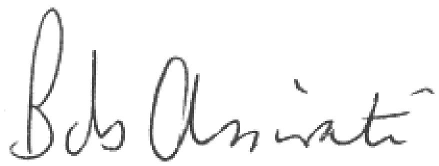
FOREWORD

Organisations are increasingly dependent on electronic delivery of services to meet customer needs. This means a requirement for high quality IT services, matched to business needs and user requirements as they evolve.

CCTA’s IT Infrastructure Library (ITIL) is the most widely accepted approach to IT Service Management in the world. ITIL provides a cohesive set of best practice, drawn from the public and private sectors internationally, supported by a comprehensive qualification scheme, accredited training organisations, implementation and assessment tools.



Bob Assirati
CCTA Chief Executive

PREFACE

The ethos behind the development of the IT Infrastructure Library (ITIL) is the recognition that organisations are increasing dependent upon IT to satisfy their corporate aims and meet their business needs. This growing dependency leads to growing needs for quality IT services – quality that is matched to business needs and user requirements as they emerge.

This is true no matter what type or size of organisation, be it national government, a multinational conglomerate, a decentralised office with either a local or centralised IT provision, an outsourced service provider, or a single office environment with one person providing IT support. In each case there is the requirement to provide an economical service that is reliable, consistent and fit for purpose.

IT Service Management is concerned with delivering and supporting IT services that are appropriate to the business requirements of the organisation. ITIL provides a comprehensive, consistent and coherent set of best practices for IT Service Management processes, promoting a quality approach to achieving business effectiveness and efficiency in the use of information systems. ITIL processes are intended to be implemented so that they underpin but do not dictate the business processes of an organisation. IT service providers will be striving to improve the quality of the service, but at the same time they will be trying to reduce the costs or, at a minimum, maintain costs at the current level.

For each of the processes described in this book, one or more roles are identified for carrying out the functions and activities required. It should be noted that organisations may allocate more than one role to an individual within the organisation (although this book indicates where specific roles should *not* be merged), or may allocate more than one individual to a role. The purpose of the role is to locate responsibility rather than create an organisation structure.

The best-practice processes described in this book both support and are supported by the British Standards Institution's Code of Practice for IT Service Management (PD0005), and in turn underpin the ISO quality standard ISO9000.

The authors

The guidance in this book was distilled from the experience of a range of authors working in the private sector in IT service management. The material was written by:

John Bartlett	Insight Consulting
Brian Johnson	Pink Elephant
David Johnston	Fox IT (Ultracomp)
Chris Keeling	Insight Consulting
Vernon Lloyd	Fox IT (Ultracomp)
Ian MacDonald	Barclays Bank PLC
Gerry McLaughlin	Fox IT (Ultracomp)
John Mather	Primax Consulting Ltd
David Wheeldon	CEC Europe
Rob Young	Fox IT (Ultracomp)

The project was managed and co-ordinated by Hilary Weston of Fox- IT (Ultracomp)

A wide-ranging national and international Quality Assurance (QA) exercise was carried out by people proposed by CCTA and itSMF. CCTA and Ultracomp wish to express their particular appreciation to the following people who spent considerable time and effort (far beyond the call of duty!) on QA of the material:

Joy Attwood-Harris	FI GROUP
Graham Barnett	DMR Consulting Group
David Bingham	DMR Consulting Group
Ian Bridge	FI GROUP
Graham Briscoe	Royal & Sun Alliance
Michael Busch	Debis Systemhaus PCM Computer AG
Steven Cain	HM Treasury

Derek Cambray	Amdahl
Martin Carr	CCTA
Duncan Carroll	Scottish Provident
Alison Cartlidge	FI GROUP
David Cronin	
Steve Daniels	Insight Consulting
Mike Davies	Proactive-sv
Kevin K Ellis BA MBA	Profission Inc. Sunbury Ontario, Canada
Jenny Elwood-Wade	Bowood Ltd (NZ)
Mary Fishleigh	*it*SMF Publications sub-committee
David Forster	UNICOM
Judith Fulcher	FI GROUP
Dave Green	Fox IT (Ultracomp)
John Groom	CCTA
Paul Habershon	DMR Consulting Group
Ken Hamilton	ManageOne
Lex Hendriks	EXIN
Marc Hodes	Fidus
Steve Houlding	Inter Access InterProm b.v.
Lesley Hughes	Hewlett Packard
Tony Jenkins	Parity Training Ltd
Chris Jones	Compaq
Colin Keeley-Huggett	IBM
Magda Kilby	ECsoft
Maggie Kneller	*it*SMF / Powergen
Arry Koetsier	Inter Access InterProm b.v.
Vladimir Kufner	Logica
Ronnie Lachniet	Quint Wellington Redwood NL
Cameron Law	WM Company
Aidan Lawes	*it*SMF International
Adrian Leach	Parity Training Ltd
Rudolf Liefers	Syntegra
Martin Lucas	Fox IT (Ultracomp)
Ivor Macfarlane	ISEB
Steve Mann	Sysop
Simon Marvell	Insight Consulting
Rob Mersel	Inter Access InterProm b.v.
Steve Moran	British Library
Alan O'Connor	HM Treasury
Paul Overbeek	KPMG Information Risk Management
Eric Overvoorde	EXIN
Dave Owen	Alliance & Leicester
Louk Peters	Pink Elephant
Tony Price	Fox IT (Ultracomp)
Rene Posthumus	Mainland Sequoia
Susan Reay	Hambleton District Council
Colin Rudd	IT Enterprise Management (itEMS)Ltd
Kathryn Rupchock	Microsoft
Frances Scarff	CCTA
Andy Smith	EDS
Richard Stone	Fox IT (Ultracomp)
Mark Sutton	Amdahl
Ed Tozer	Edwin E Tozer Ltd
Bridget Veitch	F. I. GROUP
Jenny Elwood-Wade	Bowood Ltd (NZ)
Jan M Willems	Sx Consultants B.V.B.A.
John Windebank	Computacenter
Peter Youart	Morse Group Ltd

Contact information

Full details of the range of material published under the ITIL banner can be found at www.itil.co.uk.

For further information on this and other CCTA products, please visit the CCTA website at www.ccta.gov.uk/. Alternatively, please contact:

CCTA Help Desk
Rosebery Court
St Andrews Business Park
Norwich
NR7 0HS
United Kingdom
Tel: +44 1603 704567
Email: info@ccta.gov.uk

1 INTRODUCTION TO THIS BOOK

This book is one of a series issued as part of the updated IT Infrastructure Library that documents industry best practice for the support and delivery of IT services. Although this book can be read in isolation, it is recommended that it be used in conjunction with the other IT Infrastructure Library books. Service Management is a generic concept and the guidance in the new IT Infrastructure Library books is applicable generically. The guidance is also scaleable - applicable to both small and large organisations. It applies to distributed and centralised systems, whether in-house or supplied by third parties. It is neither bureaucratic nor unwieldy if implemented sensibly and in full recognition of the business needs of your organisation.

1.1 The IT Infrastructure Library

Developed in the late 1980’s, the IT Infrastructure Library (ITIL) has become the world-wide *de facto* standard in Service Management. Starting as a guide for UK government, the framework has proved to be useful to organisations in all sectors through its adoption by many Service Management companies as the basis for consultancy, education and software tools support. Today, ITIL is known and used worldwide.

The reasons for its success are explained in the remainder of this section:

1.1.1 Public domain framework

From the beginning, ITIL has been publicly available. This means that any organisation can use the framework described by the CCTA in its numerous books. Because of this, the IT Infrastructure Library guidance has been used by such a disparate range of organisations, local and central government, energy, public utilities, retail, finance, and manufacturing. Very large organisations, very small organisations and everything in between have implemented ITIL processes.

1.1.2 Best practice framework

The IT Infrastructure Library documents industry best practice guidance. It has proved its value from the very beginning. Initially, CCTA collected information on how various organisations addressed Service Management, analysed this and filtered those issues that would prove useful to CCTA and to its Customers in UK central government. Other organisations found that the guidance was generally applicable and markets outside of government were very soon created by the service industry.

Being a framework, ITIL describes the contours of organising Service Management. The models show the goals, general activities, inputs and outputs of the various processes, which can be incorporated within IT organisations. ITIL does not cast in stone every action you should do on a day-to-day basis because that is something which will differ from organisation to organisation. Instead it focuses on best practice that can be utilised in different ways according to need.

Thanks to this framework of proven best practice, the IT Infrastructure Library can be used within organisations with existing methods and activities in Service Management. Using ITIL doesn’t imply a completely new way of thinking and acting. It provides a framework in which to place existing methods and activities in a structured context. By emphasising the relationships between the processes, any lack of communication and co-operation between various IT functions can be eliminated or minimised.

ITIL provides a proven method for planning common processes, roles and activities with appropriate reference to each other and how the communication lines should exist between them.

1.1.3 De facto *standard*

By the mid-1990s, ITIL was recognised as the world *de facto* standard for Service Management. A major advantage of a generally recognised method is a common language. The books describe a large number of terms that, when used correctly, can help people to understand each other within IT organisations.

An important part of IT Infrastructure Library projects is getting people to speak that common language. That is why education is the essential basis of an implementation or improvement program. Only when the people involved use a common language can a project be successful.

1.1.4 Quality approach

In the past, many IT organisations were internally focused and concentrated on technical issues. These days, businesses have high expectations towards the quality of services and these expectations change with time. This means that for IT organisations to live up to these expectations, they need to concentrate on service quality and a more Customer oriented approach. Cost issues are now high on the agenda as is the development of a more businesslike attitude to provision of service.

ITIL focuses on providing high quality services with a particular focus on Customer relationships. This means that the IT organisation should provide whatever is agreed with Customers, which implies a strong relationship between the IT organisation and their Customers and partners.

Tactical processes are centred on the relationships between the IT organisation and their Customers. Service Delivery is partially concerned with setting up agreements and monitoring the targets within these agreements. Meanwhile, on the operational level, the Service Support processes can be viewed as delivering service as laid down in these agreements. On both levels you will find a strong relationship with quality systems such as ISO 9000 and a total quality framework such as European Framework for Quality Management (EFQM). ITIL will support these quality systems by providing defined processes and best practice for the management of IT Services, enabling a fast track towards ISO certification. Attaining a quality standard is beneficial for organisations but it has to be recognised that this alone does not guarantee delivery of good service. There would need to be on-going review of quality of processes aligned with business requirements.

Appendix D provides more information on quality management. Generic benefits include:

- improved quality service provision
- cost justifiable service quality
- services that meet business, Customer and User demands
- integrated centralised processes
- everyone knows their role and knows their responsibilities in service provision
- learning from previous experience
- demonstrable performance indicators.

Business case for using the ISO9000 quality standards

Many companies require their suppliers to become registered to ISO 9001 and because of this, registered companies find that their market opportunities have increased. In addition, a company's compliance with ISO 9001 ensures that it has a sound Quality Assurance system.

Registered companies have had dramatic reductions in Customer complaints, significant reductions in operating costs and increased demand for their products and services.

ISO 9000 registration is rapidly becoming a should for any company that does business in Europe. Many industrial companies require registration by their own suppliers. There is a growing trend toward universal acceptance of ISO 9000 as an international standard.

Of course this applies to other standards, for example the British Standards, and in fact most European and many other standards world-wide have been consolidated in the new ISO9000-2000 standards.

1.1.5 it*SMF*

The *it*SMF (IT Service Management Forum) was set up to support and influence the IT Service Management industry. It has, through its very large membership, been influential in promoting industry best practice and driving updates to ITIL.

1.2 Restructuring the IT Infrastructure Library

The concept of managing IT services for the improvement of business functions is not new; it predates the ITIL. The idea to bringing all of the Service Management best practice together under one roof however was both radical and new. The first series of ITIL amalgamated Service Management from an IT standpoint but could have done more to capture the interest of the business. The business perspective series was published to bridge the gap between business and management and although a success, the series was published at a time when the original ITIL guidance was becoming outdated in some areas. The impact of the new series was therefore limited in the market, but a catalyst in the Service Management industry.

ITIL was originally produced in the late 1980s and consisted of ten core books covering the two main areas of Service Support and Service Delivery. These core books were further supported by 30 complementary books covering a range of issues from Cabling to Business Continuity Management. In this revision, ITIL has been restructured to make it simpler to access the information needed to manage your services. The core books have now been pulled together into two books, covering the areas of Service Support and Service Delivery, in order to eliminate duplication and enhance navigation. The material has also been updated and revised for consistency and sharpness of focus. Lastly, the material has been re-engineered to focus on the business issues of infrastructure management as well as to ensure a closer synergy with the new IS guides published by CCTA.

1.3 Target audience

This book is relevant to anyone involved in the delivery or support of IT services. It is applicable to anyone involved in the management or day-to-day practice of Service Management. It is recognised that there are several ways of delivering an IT service, such as in-house, out-sourced and partnership. Even though this book is written mainly from an in-house service provider's perspective it is generally relevant to all other methods of service provision. So those involved in out-sourced service provision or working in partnerships will find that most of this book is applicable to them. Business managers will find the book helpful in understanding and establishing best practice IT services and support. Managers from supplier organisations will also find this book relevant when setting up agreements for the delivery and support of services.

1.4 Navigating the IT Infrastructure Library

Following consultation with Service Management organisations and User groups, CCTA designed the diagram shown in Figure 1.1. This illustrates that the new library series will comprise five principal elements, each of which will have interfaces and overlaps with each of the other four. The elements are:

- the business perspective
- managing applications
- delivery of IT services
- support of IT services
- manage the infrastructure.

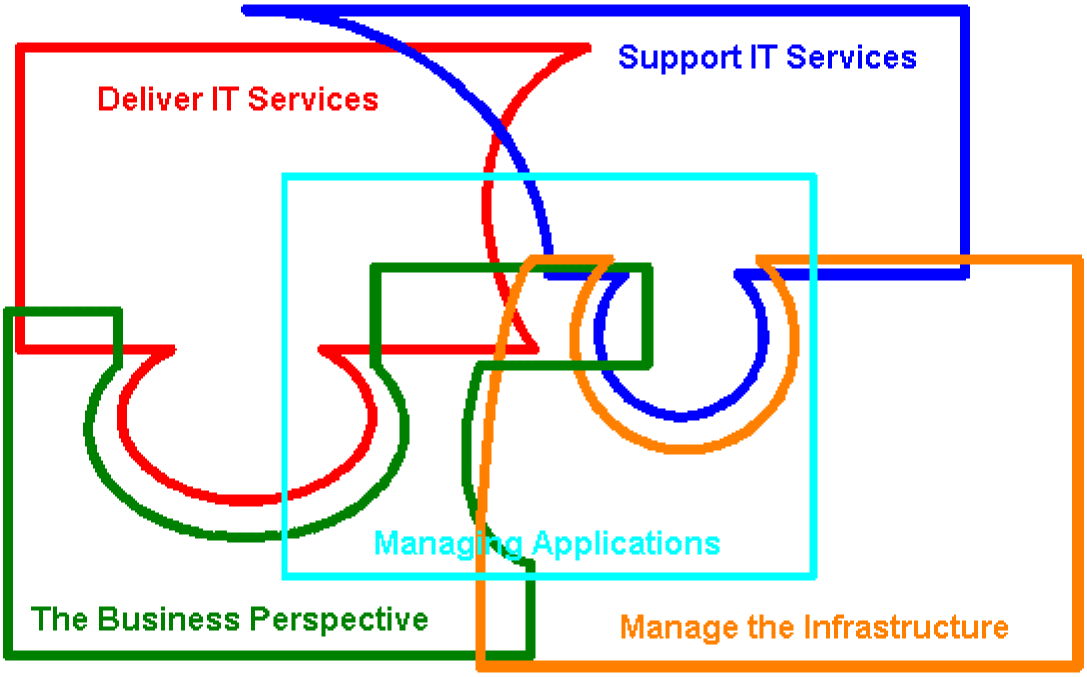


Figure 1.1 - Jigsaw Diagram

The Business Perspective book will cover a range of issues concerned with understanding and improving IT service provision, as an integral part of an overall business requirement for high quality IS management. These issues include:

- Business Continuity Management
- partnerships and outsourcing
- surviving change
- transformation of business practice through radical change.

The Service Delivery book looks at what service the business requires of the provider in order to provide adequate support to the business customers. To provide the necessary support the book covers the following topics:

- Capacity Management
- Financial Management for IT Services
- Availability Management
- Service Level Management
- IT Service Continuity Management

The Service Support book is concerned with ensuring that the User has access to the appropriate services to support the business functions. Issues discussed in this book are:

- Service Desk
- Incident Management
- Problem Management
- Configuration Management
- Change Management
- Release Management.

The ICT Infrastructure Management book includes:

- Network Service Management
- Operations Management
- Management of Local Processors
- Computer Installation and Acceptance
- Systems Management. (covered here for the first time).

Lastly, the book on Applications Management will embrace the software development lifecycle expanding the issues touched upon in Software Lifecycle Support and Testing of IT Services. Applications Management will expand on the issues of business change with emphasis on clear requirement definition and implementation of the solution to meet business needs.

The major elements of the ITIL books can be likened to overlapping jigsaw puzzle pieces (or perhaps better as tectonic plates), some of which have a precise fit, and some of which overlap or do not fit together accurately. At the highest level, there are no strict demarcation lines. Indeed, if we consider further the analogy of tectonic plates, sliding over and under one another, joining and separating, then the earthly problem of points of instability or friction caused by the imprecise nature of the pieces has an IT Infrastructure Library equivalent. It is precisely where process domains overlap or where demarcation lines cannot be clearly drawn that many management problems arise. We cannot stop all the problems from occurring (just as we cannot stop earthquakes) but we can provide advice about how to prepare for and deal with them.

1.5 Why choose a jigsaw concept?

To clarify how the concepts within ITIL work together, CCTA produced a set of process models to describe the makeup of ITIL – the process model for Service Delivery can be found at Appendix F. These process models have been used in practice and enhanced since first produced and now form the cornerstones of the ITIL core books. The process elements for management of services can be defined precisely. However, in practice, when analysing the processes in more detail, elements overlap. This situation illustrates the need for both consistency across the guidance, and advice on how to deal with management problems that may arise. The cause of these management problems may be the result of boundaries drawn that perhaps have more to do with the span of control than with logical grouping of related processes.

1.6 The Service Delivery book

Figure 1.2 expands the service delivery jigsaw puzzle pieces. The infrastructure library process elements covered in this book are also shown.

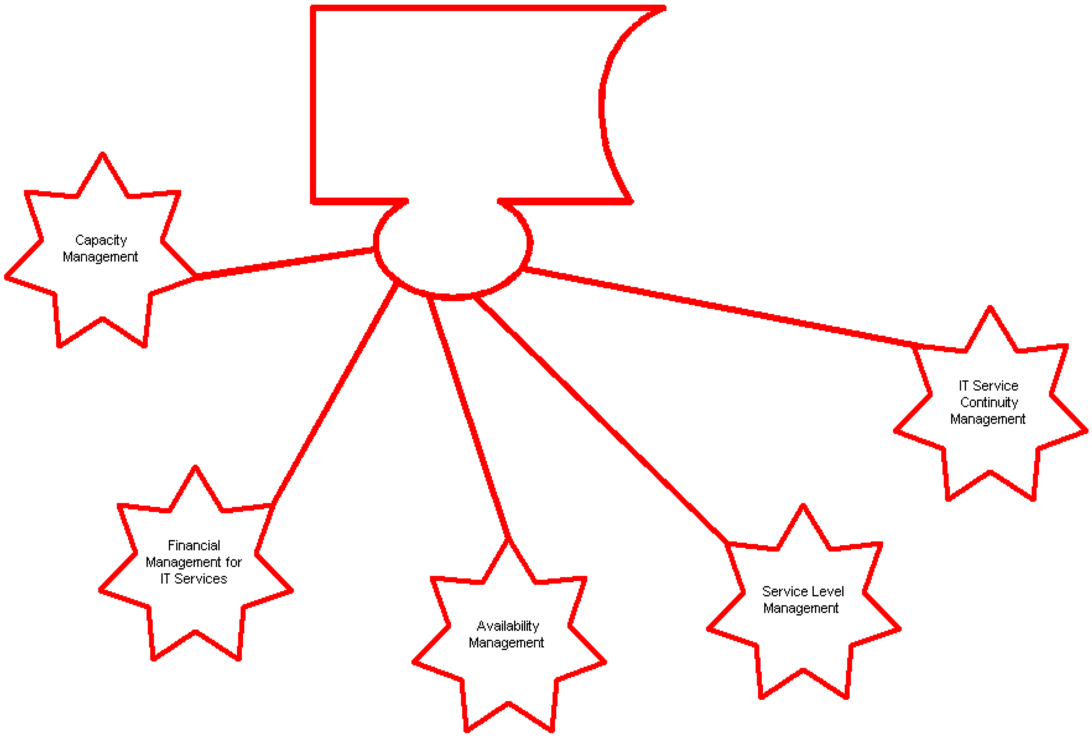


Figure 1.2 - Service Delivery; the coverage

Note that all of the chapters relate to processes.

1.7 Service Management

All the processes described in ITIL relate to each other. Half of these processes are detailed in this book, half in the book on Service Support. To better understand how these processes inter-relate, consider this example life cycle of an Incident:

1. A User calls the *Service Desk* to report response difficulties with the on-line service.

2. The *Incident Management* process deals with Incident.

3. The *Problem Management* process investigates underlying cause and calls in *Capacity Management* to assist in this process. *Service Level Management* alerted that the SLA has been breached.

4. The *Change Management* process raises and co-ordinates a Request For Change (RFC).

5. The *IT Financial Management* process assists with the business case cost justification for the hardware upgrade.

6. The *IT Service Continuity* process gets involved in the *Change Management* process to ensure recovery is possible onto current back-up configuration.

7. The *Release Management* process controls the implementation of the Change by rolling out replacement hardware and software. Release Management updates Configuration Management with details of new Releases and versions.

8. The *Availability Management* process is involved in considering the hardware upgrade to ensure that it can meet the required availability and reliability levels.

9. The *Configuration Management* process ensures the CMDB information is updated throughout the process.

1.8 Customers and Users

To avoid confusion regarding roles and terminology the terms 'Customer' and 'User' are used throughout the new books to differentiate between those people (generally senior managers) who commission, pay for and own the IT Services (the Customers) and those people who use the services on a day-to-day basis (the Users). The semantics are less important than the reason for differentiation. The primary point of contact for Customers is the Customer Relationship Manager, whilst the primary point of contact for Users is the Service Desk. A poorly functioning Incident Management process will affect the User population immediately. A service that is poor value for money will have a greater impact on the Customer.

It is therefore important that we distinguish the different, but related, needs of Users and Customers in the provision of services. Certainly, their goals may be at odds and need to be balanced; for example Users may demand high availability whereas Customers look for value for money at different levels of availability. There are information flows that should be maintained and key process elements that should be defined for use by both parties.

1.9 A Code of Practice for IT Service Management - PD0005

The British Standards Institute have published 'A Code of Practice for IT Service Management (PD0005) and the BS15000 standard which are based on the principles of ITIL; the context diagram from the standard is reproduced below. The diagram is not a process model but simply a pictorial description. It can be viewed in the same way as Figure 1.1, i.e. the main principles (of Service Management in this instance) are placed in a coherent context, providing guidance that enables the reader to make links between related process elements.

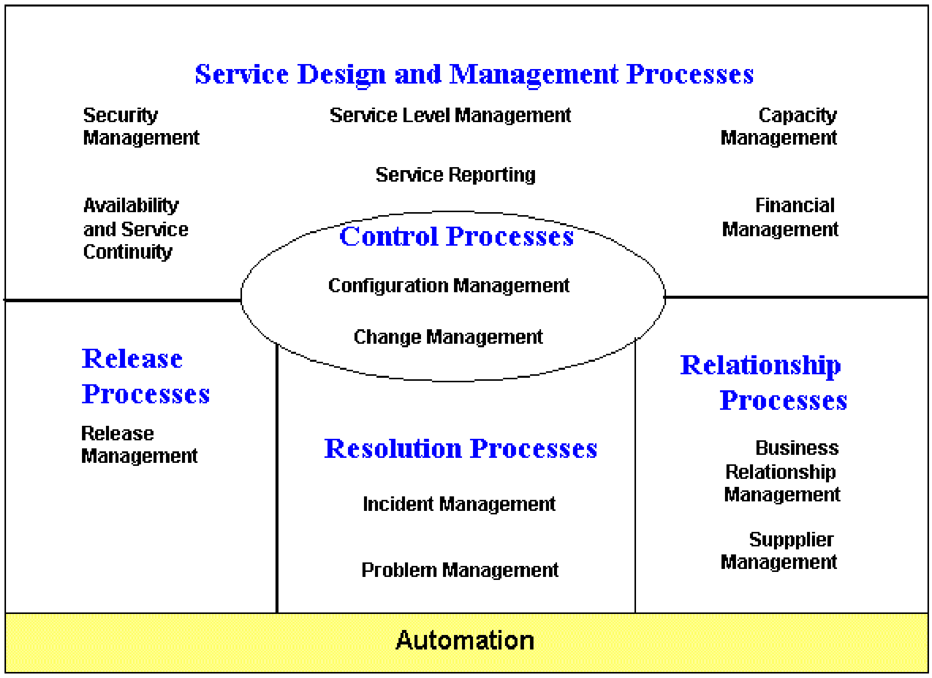


Figure 1.3 – BS15000 Service Management processes

If we now consider PD0005 in the context of Figure 1.3, it is obvious that the new IT Infrastructure Library models can be seen as an expansion of the BSI model, taking IT Service Management forward. The process elements are nearly the same, the principle of Change and Configuration Management as a linchpin is the same, the difference relates to the level of detail. Both BSI and CCTA espouse similar, if not identical, principles of best practice for IT Service Management.

1.10 Service Management: a process approach

The chapters in this book focus on Service Delivery as a set of integrated processes. Service Delivery focuses on processes in order to achieve goals; some managerial functions are required to enact the processes, but fundamentally it is the process and its suitability for purpose that is important.

More details about process theory and practice are provided in Appendix B to this book.

1.11 Recommended reading

These CCTA IS Management Guides are also related to Service Management and are recommended additional reading. They can be obtained from your ITIL bookseller or The Stationery Office:

Strategic Management of IS
Published by Format
1 90309 1 02 0
Approx 70 pages
October 1999
£25.00

This guide describes the context within which organisations in the public sector will need to think strategically about the exploitation of IS and IT, and the impact of current sector-wide policies and initiatives. It sets out the approach which senior business managers should take to the management of information systems and their IS strategy, together with guidance on managing the required changes.

Managing change
Published by Format
1 90309 1 01 2
Approx 80 pages
September 1999
£25.00

This guide is intended for senior managers who are responsible for managing complex IS-related change. It investigates the issues associated with business transformation, organisational and cultural change. It provides practical advice on working across organisational boundaries and breaking the ‘all pain, no gain’ barrier. The guidance also provides clear interfaces to Programme Management and benefits realisation.

Acquisition
Published by Format
1 90309 1 03 9
Approx 110 pages
September 1999
£25.00

This guide focuses on business objectives and outcomes in making the right IT acquisitions. It explains the key issues relating to sourcing options and partnerships, together with procurement strategies for new ways of working and new kinds of contracts.

Managing performance
Published by Format
1 90309 1 05 5
Approx 110 pages
October 1999
£25.00

This guide developed in collaboration with NAO, focuses on three levels of performance management: business management; contribution of IS/IT to the business; and performance of the IS/IT function. It examines the role of the EFQM Excellence Model ®, balanced scorecards, benchmarking and other widely adopted techniques.

Managing services
Published by Format
1 90309 1 04 7
Approx 100 pages

October 1999
£25.00

This guide explains the foundations for effective service management, appropriate contracts and good working relationships. It helps the customer of IS/IT services to understand the suppliers' perspective on service delivery. It provides practical advice on how to achieve better performance from service providers and continuing value for money.

2 RELATIONSHIP BETWEEN PROCESSES

This book refers to the need for Service Level Management, Financial Management for IT Services, IT Service Continuity Management, Availability Management and Capacity Management. Each component of service delivery is discussed separately in the book. The purpose of this section is to show the links and the principal relationships between all the Service Management and other Infrastructure Management processes.

2.1 Service Level Management

The Service Level Management (SLM) process is responsible for ensuring Service Level Agreements (SLAs) and underpinning Operational Level Agreements (OLAs) or contracts are met, and for ensuring that any adverse impact on service quality is kept to a minimum. The process involves assessing the impact of Changes upon service quality and SLAs, both when Changes are proposed and after they have been implemented. Some of the most important targets set in the SLAs will relate to service availability and thus require Incident resolution within agreed periods.

SLM is the hinge for Service Support and Service Delivery. It cannot function in isolation as it relies on the existence and effective and efficient working of other processes. An SLA without underpinning support processes is useless, as there is no basis for agreeing its content.

2.2 Financial Management for IT Services

Financial Management is responsible for accounting for the costs of providing IT service and for any aspects of recovering these costs from the Customers (charging). It requires good interfaces with Capacity Management, Configuration Management (asset data) and Service Level Management to identify the true costs of service. The Financial Manager is likely to work closely with the Customer Relationship Manager and the IT Directorate during the negotiations of the IT department budgets and individual Customer's IT spend.

2.3 Capacity Management

Capacity Management is directly related to the business requirements and is not simply about the performance of the system's components, individually or collectively. Capacity Management is involved in Incident resolution and Problem identification for those difficulties relating to capacity issues.

Capacity Management activities raise Requests for Change (RFCs) to ensure that appropriate capacity is available. These RFCs are subject to the Change Management process, and implementation may affect several CIs, including hardware, software and documentation, requiring effective Release Management.

Capacity Management should be involved in evaluating all Changes, to establish the effect on capacity and performance. This should occur both when Changes are proposed and after they are implemented. Capacity Management should pay particular attention to the cumulative effect of Changes over a period of time. The negligible effect of single Changes can often combine to cause degraded response times, file storage problems, and excess demand for processing capacity.

2.4 IT Service Continuity Management

IT Service Continuity Management is concerned with managing an organisation's ability to continue to provide a pre-determined and agreed level of IT Services to support the minimum business requirements following an interruption to the business. Effective IT Service Continuity requires a balance of risk reduction measures such as resilient systems and recovery options including back-up facilities. Configuration Management data is required to facilitate this prevention and planning. Infrastructure and business Changes need to be assessed for their potential impact on the continuity plans, and the IT and business plans should be subject to Change Management procedures. The Service Desk has an important role to play if business continuity is invoked.

2.5 Availability Management

Availability Management is concerned with the design, implementation, measurement and management of IT services to ensure the stated business requirements for availability are consistently met. Availability Management requires an understanding of the reasons why IT service failures occur and the time taken to resume service. Incident Management and Problem Management provide a key input to ensure the appropriate corrective actions are being progressed.

Availability targets specified in SLAs will be monitored and reported on as part of the Availability Management process. Additionally Availability Management supports the Service Level Management process in providing measurements and reporting to support service reviews.

2.6 Configuration Management

Configuration Management is an integral part of all other Service Management processes. With current, accurate and comprehensive information about all components in the infrastructure the management of Change, in particular, is more effective and efficient. Change Management can be integrated with Configuration Management. As a *minimum* it is recommended that the logging and implementation of Changes be done under the control of a comprehensive Configuration Management system and that the impact assessment of Changes is done with the aid of the Configuration Management system. All Change requests should therefore be entered in the Configuration Management Database (CMDB) and the records updated as the Change request progresses through to implementation.

The Configuration Management system identifies relationships between an item that is to be changed and any other components of the infrastructure, thus allowing the owners of these components to be involved in the impact assessment process. Whenever a Change is made to the infrastructure, associated Configuration Management records should be updated in the CMDB (see Figure 2.1). Where possible, this is best accomplished by use of integrated tools that update records automatically as changes are made.

The CMDB should be made available to the entire Service Support group so that Incidents and Problems can be resolved more easily by understanding the possible cause of the failing component. The CMDB should also be used to link the Incident and Problem records to other appropriate records such as the failing Configuration Item (CI) and the User. Release Management will be difficult and error prone without the integration of the Configuration Management process.

The Service Delivery processes also rely on the CMDB data. For example:

- Service Level Management needs to identify components that combine together to deliver the service so that underpinning agreements can be set up
- Financial Management for IT needs to know the components utilised by each business unit especially when charging is in place
- IT Service Continuity and Availability Management need to identify components to perform risk analysis and component failure impact analysis.

Figure 2.1 shows sample relationships between Configuration Management and other Service Management processes.

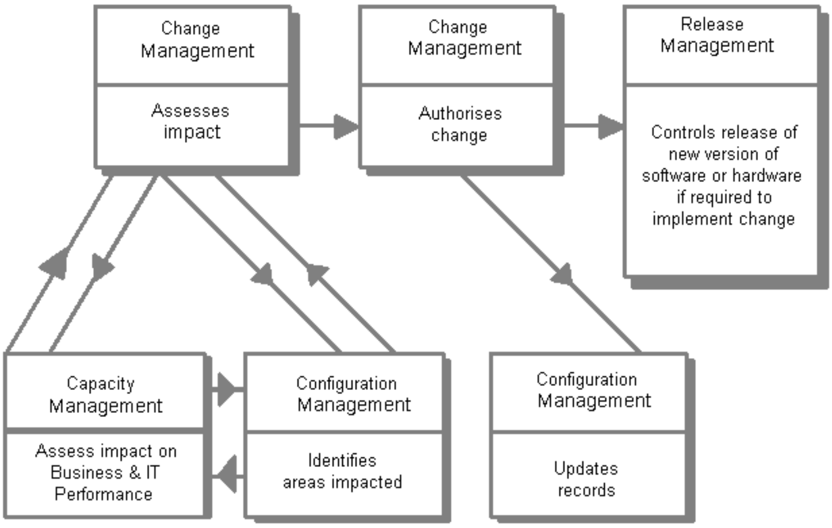


Figure 2.1 – Relationship between, Change Management, Configuration Management Capacity Management and Release Management

2.7 Change Management

The Change Management process depends on the accuracy of the configuration data to ensure the full impact of making Changes is known. There is therefore a very close relationship between Configuration Management, Release Management and Change Management.

Details of the Change process are documented in SLAs to ensure that Users know the procedure for requesting Changes and the projected target times for, and impact of the implementation of Changes.

Details of Changes need to be made known to the Service Desk. Even with comprehensive testing there is an increased likelihood of difficulties occurring following Change implementation either because the Change is not working as required or expected, or because of queries on the change in functionality.

The Change Advisory Board (CAB) is a group of people who can give expert advice to the Change Management team on the implementation of Changes. This board is likely to be made up of representatives from all areas within IT and representatives from business units.

2.8 Release Management

Changes may often result in the need for new hardware, new versions of software, and/or new documentation, created in-house or bought in, to be controlled and distributed, as part of a new ‘packaged Release’. The procedures for achieving secure, managed rollout should be closely integrated with those for Change Management and Configuration Management. Release procedures may also be an integral part of Incident Management and Problem Management, as well as being closely linked to the CMDB in order to maintain up-to-date records.

2.9 Incident Management

There should be a close interface between the Incident Management process and the Problem Management and Change Management processes as well as the function of the Service Desk. If not properly controlled, Changes may introduce new Incidents. A way of tracking back is required. It is therefore recommended that the Incident records should be held on the same CMDB as the Problem, Known Error and Change records, or at least linked without the need for re-keying, to improve the interfaces and ease interrogation and reporting.

Incident priorities and escalation procedures need to be agreed as part of the Service Level Management process and documented in the SLAs.

2.10 Problem Management

The Problem Management process requires the accurate and comprehensive recording of Incidents in order to identify effectively and efficiently the cause of the Incidents and trends. Problem Management also needs to liaise closely with the Availability Management process to identify these trends and instigate remedial action.

2.11 Service Desk

The Service Desk is the single point of contact between service providers and Users, on a day-to-day basis. It is also a focal point for reporting Incidents and making service requests. As such, the Service Desk has an obligation to keep Users informed of service events, actions and opportunities that are likely to impact their ability to pursue their day-to-day activities. For example, the Service Desk might act as the focal point for Change requests from Users, issuing Change Schedules on behalf of Change Management, and keeping Users informed of progress on Changes. Change Management should therefore ensure that the Service Desk is kept constantly aware of Change activities.

The Service Desk is in the direct firing line of any impact on the SLAs and as such needs rapid information flows.

The Service Desk may be given delegation to implement Changes to circumvent Incidents within its sphere of authority. The scope of such Changes should be predefined and the Change Management function should be informed about all such Changes. Prior approval of Change Management is essential before Changes of specification of any CI are implemented.

2.12 ICT Infrastructure Management

ICT Infrastructure Management functions are involved in most of the processes of Service Support and Service Delivery where more technical issues are concerned.

2.13 Application Management

CCTA plans to produce guidance that discusses the major processes required to manage applications throughout their lifetime. Service Management is typically concerned with a product (software/hardware) at a particular point in time to support the service requirements of the business. But it should be more than this. It delivers a maintainable service for the business - this means delivering the skills, training, and communications with the application. Applications Management considers the issues from feasibility through productive life to final demise of the application.

2.14 Security Management

The Security Management function interfaces with IT Service Management processes where security issues are involved. Such issues relate to the Confidentiality, Integrity and Availability of data, as well as the security of hardware and software components, documentation and procedures. For example, Security Management interfaces with Service Management to assess the impact of proposed Changes on security, to raise RFCs in response to security problems; to ensure confidentiality and integrity of security data and to maintain the security when software is released into the live environment.

2.15 Environmental infrastructure processes

Changes to the environment may affect the quality of service, and Changes to the infrastructure may have implications for the environmental infrastructure. It is recommended that all relevant aspects of the environmental infrastructure be brought under Configuration Management control and subjected to the Change Management procedures described in the Service Support book.

2.16 Project Management

When implementing new processes in an organisation there are benefits to running the activity as a project. Accordingly, this book refers to Service Management Projects that are implemented to introduce new processes or to improve the current processes for the delivery or support of IT services.

There are a variety of structured project management methods that can be adopted. Within ITIL when we discuss project management if we need to draw on a particular method we will use PRINCE2 which is owned by CCTA and widely adopted throughout UK Government bodies, the Public and Private sectors, and internationally.

3 GETTING STARTED

Chapter 11 discusses how to set about planning and implementing the project of introducing Service Management. Within this chapter we look at how to build commitment from management so that they will provide the necessary funding and communicate support for the project.

What you need to know first is the benefit of using the method and how to market the message of those benefits to your organisation. These issues can form part of a business case for process implementation or improvement. An important part of the business case is likely to be concerned with articulating the problems with the current position and demonstrating the benefits of the new vision. A business case should look at the benefits, disadvantages, costs and risks of the current situation and the future vision so that management can balance all of these factors when deciding if the project should proceed. Appendix E provides some costed examples for developing a business case for introducing ITIL into an organisation.

3.1 Service Management benefits

It is important to consider the benefits for the organisation of having a clear definition of the Service Management function. Some of the benefits that could be cited include:

- improved quality of service - more reliable business support
- IT Service Continuity procedures more focused, more confidence in the ability to follow them when required
- clearer view of current IT capability
- better information on current services (and possibly on where Changes would bring most benefits)
- greater flexibility for the business through improved understanding of IT support
- more motivated staff; improved job satisfaction through better understanding of capability and better management of expectations
- enhanced Customer satisfaction as service providers know and deliver what is expected of them
- increased flexibility and adaptability is likely to exist within the services
- system-led benefits, e.g. improvements in security, accuracy, speed, availability as required for the required level of service
- improved cycle time for Changes and greater success rate.

The importance and level of these will vary between organisations. An issue comes in defining these benefits for any organisation in a way that will be measurable later on. Following ITIL guidance can help to quantify some of these elements.

3.2 A process led approach

Appendix B provides detailed information about using process models and definitions to complement a process led approach to implementing IT Infrastructure Library guidance in a programme of continuous improvement.

Figure 3.1 represents a model that can be used by an organisation as the framework for process improvement.

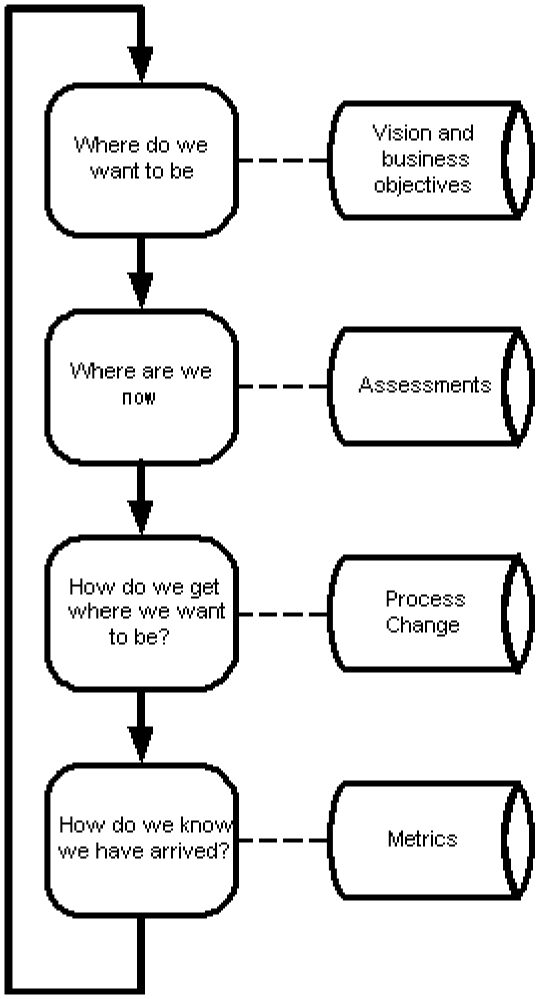


Figure 3.1 - Process improvement model

In short, if you can draw up a process model of best practice - or any practice in fact that is more effective than your current way of doing things - then you can compare that model with a description of your current practice and use it to define improvements. If you do this in the light of your business direction or critical success factors, you can define measures of how you can demonstrate improvements and achievements that are truly useful.

As a process based method, the IT Infrastructure Library is particularly suited to use in this way. Appendix B provides more information and a case study.

There are a number of different methods and notations by which processes can be defined and documented – each often associated with a specific design/modelling tool.

Taking a fairly generic view, the Figure 3.2 illustrates the typical components one might find in a process definition.

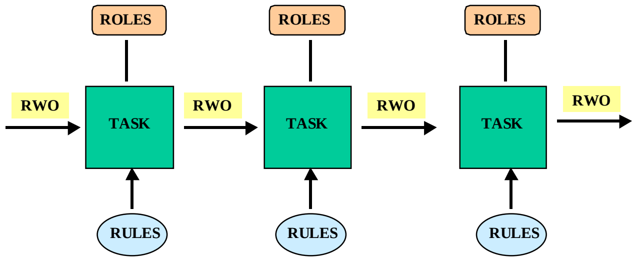


Figure 3.2 – Typical components in a process definition

Each process can be broken down into a series of **tasks**. For each task, there will be Inputs and Outputs (shown as **Real World Objects, RWO**, in the diagram). Whether these RWOs have a physical form, e.g. as a piece of paper, or are merely held as electronic information, is irrelevant.

Each task will be executed by a **role**. This may be embodied in a human being or performed by a piece of software. If human-centric, then there will be a set of competencies that an individual needs in order to perform the role.

The execution of the role is governed by a set of **rules**. These will range from the simple ("All boxes on the form should be completed") to the very complex ("Credit is only allowed if a set of criteria are met according to an algorithm").

Often, a process will span various organisational boundaries. It is important, therefore, that each process should have an **owner**. This is another **role**.

The process owner is responsible for the **process definition** itself, which should be treated as a CI, subject to all the usual Change control rigours. The process owner is responsible for ensuring that all that are involved in the execution of the process are kept informed of any Changes that occur.

3.3 Management commitment

Management commitment is about motivating and leading by example. If you or your managers do not support the use of best practice openly and demonstrably, or if you are not fully committed to Change and innovation, then staff cannot be expected to improve themselves, Service Management processes or service to Customers. Genuine management commitment is absolutely essential to 'staying the course' when implementing Service Management in an organisation.

3.3.1 Aspects of management commitment

Modern organisations require IT/ business alignment, and therefore a total quality approach to leadership is required from managers. The different aspects of management commitment can be found in commonly used Total Quality Models such as the EFQM (in Europe) or MBNQA (in North America) models.

3.3.2 Management commitment in the planning stage

Why do implementations fail?

This question has been asked many times in the past. If we examine the causes for failures at the highest level a pattern appears. Simply designing and implementing a new or updated process does not guarantee success. A number of factors could result in the process not realising its objectives. In most cases failure is caused by lack of attention to the 'process enablers'. It is not enough for management to provide the funds for the implementation process and then sit back expecting everything to work. Management should be committed during the entire *'plan-do-check-act'* cycle, and should also address all aspects of the Service Management framework. Other common reasons for failure include:

- lack of staff commitment and understanding
- lack of training
- the staff given the responsibility for implementation are not given sufficient authority to make the required decisions
- loss of the Service Management 'champion' (the person driving the implementation)
- loss of impetus after the initial hype
- lack of initial funding and lack of quantifiable long term cost benefits
- over-focus on tactical, isolated 'solutions' rather than a strategic solution, i.e., addressing individual elements of Service Management rather than the overall picture
- overly ambitious expectations of immediate benefits or try to do everything at once
- unrealistic implementation timetable
- no one accountable
- difficulties of changing the culture of the organisation
- tools unable to support the process, requiring tailoring of the process or of the tool
- inappropriate approach taken to implementation - not under Project Management controls
- inappropriate scoping of the process
- lack of appreciation for the hard work and discipline required to implement Service Management.

3.4 Cultural aspects

It would appear obvious that, in order to prosper, a business should have Customer satisfaction as its prime objective. Unfortunately, until recently, many IT departments have been too obsessed with technology and flashing lights to recognise that they have Customers at all.

In recent years pressure has been put on IT service providers to become more aware of their role in supporting the business and to be run as accountable business units. This pressure has come from many quarters:

- outsourcing
- market testing
- compulsory competitive tendering
- the Customers themselves (who are, in increasing numbers, being given their own budgets and the freedom to use them where they see fit).

The days when staff in IT departments regarded their 'Customers' as a *necessary evil* or just *difficult colleagues* have (hopefully) passed. There is a growing awareness that in order to retain the in-house IT department they have to stay close to their Customers, understand and predict their requirements and satisfy them. IT departments are now raising the priority of Customer satisfaction from being merely 'nice to have' to 'essential'.

The provision of quality IT services with high levels of availability and performance can be achieved with the correct hardware, software and underlying support processes. This level of service may satisfy, but may not delight, the Customer! Extra effort is needed for the Customer to enjoy the experience and want to come back for more. The way in which the service is delivered is dependent on the people delivering the service. Customer delight will only be achieved if the people involved are responsive to their Customers' needs, are attentive, reliable, and courteous, delivering the service in the way they themselves would like to receive it.

We all know what a Customer is - someone who deals with a trader and habitually purchases from him. Similarly, we are all familiar with service - an act performed for the benefit or advantage of a person, institution or cause. Customer service therefore is concerned with performing acts that will benefit the Customer in a way that will encourage him to purchase service again and again.

3.4.1 What is culture?

It is important that all staff involved in delivering service are committed to the concept of Customer delight. This can only be achieved if the organisation's culture demands it. The prevalent culture within any organisation is a product of a variety of factors: the age and history of the organisation, its size, the technology in use, its objectives, the market as well as geography and the personalities and backgrounds of people employed.

There are many influencing factors and it is not reasonable to expect all organisations to have, or adopt, the same culture, organisational structure or systems. Despite the growing literature on the culture of organisations there is no textbook formula for organisational culture.

The term ‘culture’ is used in this context to refer to the values and beliefs of the organisation - the normal way of doing things. Component parts of the culture include:

- the way authority is exercised and people rewarded
- methods of communication
- the degrees of formality required in working hours and dress and the extent to which procedures and regulations are enforced.

An organisation's culture can be immediately recognised by an outsider by the staff's attitudes and morale; their vocabulary - the phrases and buzz words they use and the stories and legends they tell of the organisation's heroes.

3.4.2 Responsibilities

Most often, responsibility for creating and maintaining an organisation's culture should rest with its leaders. The prevailing culture will dictate the shape of the organisation's structure and the nature of the systems and procedures used.

In their best-selling book *In Search of Excellence*, a study of the best run American companies, Tom Peters and Robert Waterman noted

...Without exception, the dominance and coherence of culture proved to be an essential quality of the excellent companies. Moreover, the stronger the culture and the more it was directed toward the market place, the less need was there for policy manuals, organisation charts, or detailed procedures and rules... people ...know what they are supposed to do in most situations because the handful of guiding values is crystal clear

Procedures, therefore, should support the culture and not govern or influence it.

3.4.3 What is meant by 'service culture'?

In order to achieve business success the culture of every organisation should embrace the concept of service and Customer care.

The satisfaction of Customer requirements should be the number one priority - for everyone, whatever their role in the organisation. The concept of service should permeate through all layers of the organisation - from staff in the front line responsible for delivering service to those in a supporting role - either as managers or in the back room.

Service delivery should ensure that Customer requirements are met in a way that makes them feel that they are valued and respected - they should be made to feel good at every transaction. The aim should be to exceed Customers' expectations (not in *what* you deliver but in the *way* you deliver it) and give them confidence in your ability to satisfy all their future needs.

By taking the time and effort to listen to the Customers it is possible to understand the service being provided from their perspective. The key point about providing service is attention to detail; to go that little bit further to delight the Customer.

3.4.4 How is this relevant to IT service provision?

Achieving a service culture in an IT organisation should, in essence, be no different from achieving a service culture in a bread shop or a carpet wholesale business; the technology used and the product delivered may be different but the end result is the same - Customers receive service. Customers of IT services have come to expect the same levels of service at work as they receive at their local shopping mall - and they use the same criteria for choosing it.

If we accept that the quality of our IT service delivery is important then we should agree that the way we treat our Customers is an important part of service delivery.

3.4.5 What do Customers want?

Having recognised that IT departments are now in the business of service provision they should now adopt a whole new way of thinking and embrace the same business concepts as those used by all service providers. There is a lot of catching up to do.

Before any business goes into production of a new product or service it should perform some market research to find out what the Customers actually want and will buy.

So, what do Customers want?

- **Specification** - they want to know, up front, what they are going to get. The problem in IT service delivery is that Customers frequently don't know what they want - the 'IT experts' should translate their business requirements into solutions. Customers don't buy products or services, they buy solutions.

- **Conformance** to specification - once the appropriate IT solution has been found then it should conform to the specification. Customers want to know when they can receive it and should be satisfied that it will fulfil their business requirement.

- **Consistency** - they want it to be the same every time they come back for more.

- **Value for money** - the price they pay should be a fair one for the product or service they receive.

- **Communication** - they want to be told what they are getting, when, how and what to do if they have a problem with it.

Popular misconceptions about Customer care
Many people are cynical about the concept of Customer care and, in many cases, they have good reason to be. The following paragraphs describe areas of Customer care.

'The Customer Is King posters' (and other clichéd campaign material)

Allegedly witty and amusing posters proclaiming the virtues of Customer care can be a decorative addition to an office or corridor, on their own they achieve very little. If the same slogans are used continually on posters then they become part of the wallpaper and are ignored - be prepared to change them and beware of empty slogans that can be misinterpreted. Ensure that all staff are aware of the nature, scope and intention behind any Customer care promotion otherwise the material will be greeted by cynicism if it arrives unannounced.

A by-product of ITIL/BS5750/ISO 9001/TQM

A service culture will not develop unaided. The adoption of the best practice in IT service delivery described in ITIL will help improve service quality. Improving the quality of procedures and obtaining quality certification will certainly focus the mind on service quality but these alone are not enough to ensure the IT department is truly Customer focused - a structured approach to cultural change is necessary.

Customer Charters and Service Level Agreements

Published guaranteed levels of service (such as SLAs) are worthless if the supplier consistently fails to meet the targets. The payment of recompense through penalty clauses is unlikely to satisfy the purchaser in the long term. These agreements tend to set targets for the tangible elements of service - availability, response times etc. and ignore the intangible elements, which are so crucial to Customers - the way in which service was delivered.

(Another) Management initiative

If Customer service is placed high on the agenda one month and then allowed to slip back, people will think that 'the management have been on another training course and have come back with some bright ideas and this actually has nothing to do with us'. Staff should be made to realise that Customer care is not management or technology driven - instead it is Customer driven, and that it makes sound business sense to adopt a Service Culture.

Someone else's responsibility

Where organisations have Customer Liaison departments or Customer Service Managers, staff may be tempted to relinquish the responsibility for Customer care, thinking it is someone else's job. The concept of Customer service should permeate throughout the whole organisation.

Something new

Although the concept of satisfying Customers' requirements is only now hitting some IT departments, it is nothing new. Man has been trading ever since he came down from the trees and learned how to make a surplus. Those men and women who produced something that wasn't marketable starved to death!

3.4.6 Common excuses for conducting 'business as usual'

Most of the Management Gurus preach what is little more than 'common sense' - to treat your Customers as individuals, to ask Customers what they want, listen to them and then provide what they want, to wander around amongst your Customers and staff and engender a feeling of trust. It sounds easy. So why do so many organisations still fail to provide consistent Customer satisfaction? Some of the common excuses include:

- ' We are increasing our market share and expanding our product portfolio'
- ' Our Customers are happy - the number of Customer complaints is down on last year and the results of our Customer satisfaction survey show that we are improving our service'
- ' We are the market leaders - we are the best'
- ' Customer Service is an important issue, that is why we have a Customer service department - Customer satisfaction is their responsibility'
- ' We are highly skilled IT people - we haven't been trained to deal with Customers'
- ' We tried Customer Care before - it cost us a lot of money for little if any return'
- ' We have always treated our Customers as Customers and always provided an excellent service'
- ' We are not in a competitive situation - our Customers are tied to internal services. We are not a business unit - we run the IT department'.

Some of these arguments show an element of arrogance, some complacency and some that someone else is responsible. These attitudes must be addressed otherwise there is a danger that customers will be dissatisfied or in some cases lost.

3.4.7 How much will all this cost?

Aiming to "exceed Customers' expectations" sounds very expensive. Indeed, if we promise to visit a Customer within four hours to fix a printer and we turn up in ten minutes with a brand-new replacement, we will certainly incur extra costs. The manpower overhead of having staff sitting around waiting for a call, the cost of having spare printers on-site etc. The Customer may very well be delighted with this level of service and will expect the same service every time - the Customer will certainly be disappointed if you take four hours next time.

Exceeding Customer expectations does not, therefore, mean 'giving away service'. In the example given above the Customer may have been quite happy with the four hour target and a ten minute response may actually in this case have been inconvenient, having arranged to do something else for the next four hours. The Customer may have been happy with the old printer and not want a new one - particularly if it is a different model or takes you some time to configure it.

The IT supplier could have delighted the Customer without incurring unnecessary cost by:

- logging and understanding the technical nature of the difficulty at the first point of contact
- responding in line with the urgency as stated by the Customer
- keeping the Customer informed of what will happen, when - and keeping the promise
- carrying out the work with the minimum level of disruption in a cheerful and professional manner
- reviewing the action taken to ensure that the difficulty has been fixed and that the Customer is happy.

Getting the job done in accordance with the Customers' wishes, correctly, first time, actually saves the cost of a potential repeat visit.

Implementing a Customer care programme *will* cost money - staff time, training, changes to operating methods, promotional material etc, but the cost of doing it should be weighed against the cost of not doing it - and the risk of losing dissatisfied Customers to a competitor. It will cost a lot more to find new Customers than to retain existing ones, so any money spent keeping existing Customers happy is going to be money well spent.

3.4.8 What are the potential benefits of Customer care?

Whatever the motivations behind the move towards a service culture the end result will be increased profitability:

- operating costs will decrease as less effort is wasted giving Customers products or services they don't want
- profits margins will improve as more repeat business is won - it is much cheaper to sell to an existing Customer than to court a new one
- efficiency will improve as staff will work more effectively as teams
- morale and staff turnover will improve as staff achieve job satisfaction and job security
- service quality will be constantly improving, resulting in an enhanced reputation for the IT department, which will tempt new Customers and encourage existing Customers to buy more
- the IT department will become more effective at supporting the needs of the business and will be more responsive to changes in business direction.

3.4.9 Service Management training

It is vitally important that the concepts of the IT Infrastructure Library are well known and understood in the Service Management function. ITIL foundation and management training is now widely available leading to an internationally recognised qualification developed by the EXIN and ISEB examination boards. At the turn of the Millennium, over 10,000 people now hold qualification in IT Service Management world wide, and many more have received training.

In addition it would be beneficial for the organisation to receive business-related training covering the broader aspects of infrastructure management. Far too often IT is accused of not knowing enough about the business and its needs; ITIL the business perspective series of books should be referred to for more information.

4 BUSINESS COMMUNICATION

4.1 Introduction

Effective and efficient communication is part of the foundation on which the success of many organisations depends. In the business world there are many pressures on time and, often, it is communication which suffers when time and resource constraints apply. In Service Management there are many areas that are involved in regular communication with their business customers to ‘ensure’ that the services delivered to the customers meet their requirements at appropriate cost. Notable examples include the Service Desk, Change Management and Service Level Management.

Everyone within IT has a responsibility to provide good communication but there is prime responsibility at Management level to ensure that communication is embedded within the Service Management processes. They also need to ensure that sufficient resource is allocated to setting up and maintaining an effective and efficient communication function.

Across the service industries, many organisations have recognised that it is necessary to manage not only communication with their customers but also the relationships with them. As a result the discipline of Business Relationship Management (BRM) has developed which deals primarily with managing the relationships between Customers and IT Service Providers, and also with the communication that takes place between the two. Full guidance on BRM will be given in the ITIL Business Perspective book.

The main aims of this chapter are:

- to show why business communication is important,
- identify who should be involved
- to list some key communication techniques
- make you think about communication in your organisation and your own contribution

4.2 Why is Communication with the Business so Important

Developing and fostering effective communication with customers and users of IT has always been an important issue. Very often, however, this was something that was almost left to chance and to the personal actions of individuals within IT departments. In the modern business world this is not sufficient.

During review exercises of organisational, departmental, or team performance, communication is one of the most common topics mentioned that needs to be improved.

Organisations can be said to operate at three key levels, strategic, tactical and operational, and there are specific corporate functions at each level. For IT to be fully integrated with their business customers, interactions must also take place at the same three levels. Table 4.1 shows some high level corporate functions and the corresponding communication activities for each of the three levels.

Level	Corporate Function	Communication Activities
Strategic	Setting policy and direction	Capture and understand business strategy Understand and influence business policy and direction. Market existing and new capabilities of IT. Inform senior staff of new technologies. Interpret high-level business needs and translate them into IT requirements.
Tactical	Planning and	Understand business requirements via Change

	implementing business change	Management, Project Management, Service Level Management, Capacity Management, IT Financial Management.
Operational	Delivering customer service	Day to day contacts between Service Desk and Users, Change Management, etc. Improve communications channels at all day to day service support and provision customer contact points.

Table 4.1. Key Communication activities at each level.

It is important to note that, at all levels, there is a management, in particular Business Relationship Management, responsibility to improve communications between the IT Service Provider and its customers.

We should remember that communication is a two way process and we should all be as proactive in requesting communication as we expect others to be in giving it to us.

There are many reasons why communication with the business is important and the following are just some prime examples.

- Generate an increased understanding of the Customer's business and its needs
- Establish and develop trusting relationships
- Provide senior customers with up to date relevant information on where the application of new technology could bring business benefit
- Manage customer expectation
- Measure customer satisfaction
- Provide a conduit for business related information into IT e.g. advance notification of a new project or forthcoming advertising campaign each of which would have a significant impact on the IT service provider
- Identify new or changed requirements on a day-to-day basis

It is also important to recognise that our customers have a responsibility to communicate effectively with us. This is another area where BRM can play a significant part either as a role within the IT Service Provider or within the customer organisation.

4.3 Who is involved?

As we have seen, communication between the IT Service Provider and the customer takes place across all levels of the organisation. As such, effective communication should be a responsibility of all those staff who come into contact with customers or users.

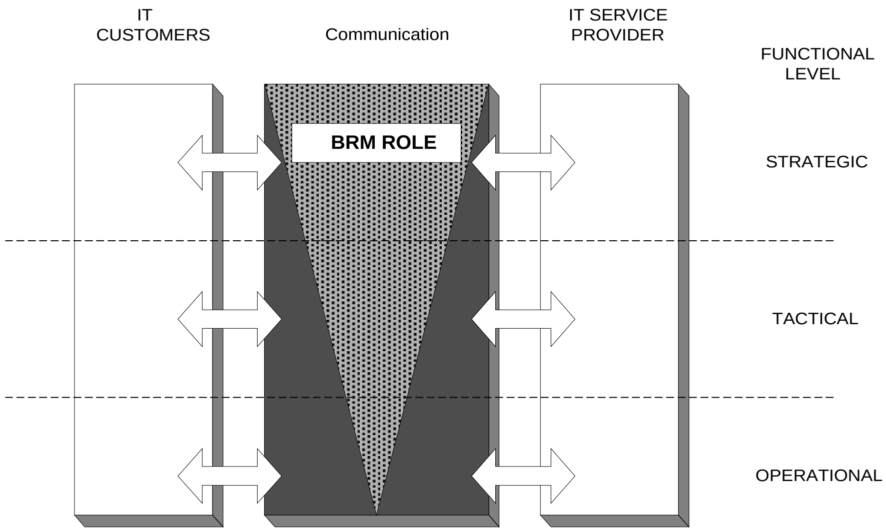


Figure 4.1. Customer communication and the role of BRM.

BRM has a key responsibility to improve communications between IT Service Providers and their customers and the role of BRM Manager is a key enabler in this area particularly at the strategic level. Figure 4.1 illustrates that communication occurs at all levels across an organisation and that the BRM Manager functions primarily at the strategic level.

Remember that the IT Service Provider communicates with each customer at all levels and therefore all IT staff need to be aware of their communication responsibilities.

Although this chapter focuses on communication with our customers, it is appropriate at this point to recall that communication within the IT service organisation, whether in a single supplier or a multi-supplier environment, is just as important. The concept that communication and relationship management could be viewed as independent of both Customer and Service Provider is illustrated in Figure 4.2 below.

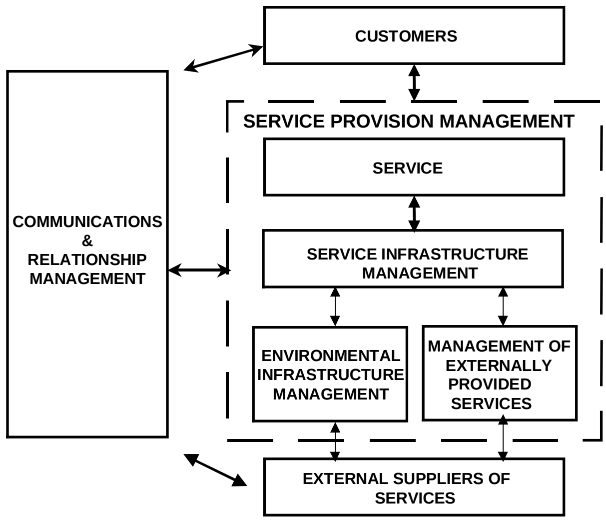


Figure 4.2. Communication across the Customer / Service Provider organisation.

At the strategic level Senior IT Managers, such as a Head of IT and Business Relationship Manager, should communicate with Business Directors and senior business managers to ensure that business strategy and policy, and business requirements are fully understood.

The ICT Steering Group is also important at the strategic level. This group provides a forum for discussions, presentations, etc on many strategic topics and is a key element in the decision making process.

An example of communication at the tactical level is the Service Level Manager discussing requirements with a Business Operations Manager and negotiating agreed service targets.

At the operational level, the Service Desk is an excellent example of a team where communication is extremely important. Being available as a first point of contact for reporting incidents, and maintaining contact with users throughout the incident lifecycle.

4.4 Benefits and Problems

There are significant benefits to be gained by both Customer and Service Provider from the implementation of effective communication. For these to be fully realised we must ensure that there is a single identifiable owner of communication who is responsible for managing the communications function.

Organisations may find the time and effort spent on communication activities, and the associated cost, is difficult to quantify.

At the operational level, and also to a large extent at the tactical level, the time and effort spent by the organisation solely on communication activities will be minimal, because it will take place mostly within existing points of contact

At the strategic level, when there may be dedicated communications staff, quantifying their time and effort will be easier.

4.4.1 Benefits

The overall benefits of improved communication come from the following:

- Business focused IT delivery - for example, effective communication will improve the ability of Service Level Management to identify and meet customer requirements.
- More efficient use of available IT resources and capacity - effective communication will improve the ability of Capacity Management to produce accurate resource plans.
- Better customer representation within IT, by ensuring that customer views are accurately reflected at all levels.
- Improved ability to actively market IT services where this is appropriate.
- Customer retention and development through higher levels of customer satisfaction.

The benefits of improved communication activities at a strategic level include:

- Business needs more accurately and speedily reflected into the IT environment
- Increased proactive communication
- Improved co-ordination of changing business needs with an evolving IT environment

At a tactical level, the benefits include:

- Improved information flow into planning activities
- Project initiation more clearly defined

At the operational level, the benefits are:

- Greater levels of customer satisfaction when dealing with service requests and minor changes.
- A better reputation for the IT service provider through a better quality of personal contact
- Greater levels of IT staff satisfaction

4.4.2 Problems

The problems associated with improving communication include:

- Switching from a ‘passive’ to an ‘active’ mode, especially at a strategic level
- Lack of commitment and support from
  - senior managers
  - internal IT sections
  - customers
- Insufficient time and/or resource.
- Lack of performance metrics, demonstrable to both IT and the customer

It is therefore important that each IT service provider decides, in conjunction with their customers, how business communication will best function within their environment.

Beware of over communication. It is imperative that you agree the level and frequency of communication with your customer. For example, nobody enjoys being deluged with a barrage of unwanted or trivial email!

4.5 Planning for Communication

The approach taken to planning the business communication function should be as rigorous as that applied to Service Management disciplines such as Capacity Management and Service Level Management. It is necessary to plan for controlled communication at the three business levels identified earlier and to agree a communication plan with the customer.

4.5.1 The Communication plan

As in all planning activities, the scope and objectives of the function should be clearly defined. Assigning specific responsibility to an individual role is advisable for the reasons mentioned earlier and resource levels allocated to the communication function should be agreed.

The communication plan should identify what is being communicated at each level, when it will be communicated and the communication technique to be used. The communication plan should not be static but be regularly updated to reflect changes in the organisation, different communication techniques, etc.

Examples of communication activities that should be detailed in the plan include the following.

Strategic:

- Business plans presented to senior IT management
- Technology strategy plans which underpin the business plan to senior Customer management

Tactical:

- Meetings between operational Business managers and SLM to update and agree the SLA and associated targets
- Project management board meetings

Operational:

- Service bulletins from the Service Desk
- Reports from Change Management on the success of managing change requests

Each of the activities covered in the plan should have a description of when the communication takes place, what format of technique will be used, who is responsible for initiating the communication, and why it is required.

It is often useful to set standard formats for different types of communication so that the recipient recognises immediately what type of communication they have received. For example, a branded quarterly newsletter, or a daily service bulletin delivered via email and which has a consistent subject heading.

4.5.2 Awareness Campaign

When used well, an awareness campaign is one of the most useful communication techniques. When launching or relaunching a communications initiative an awareness campaign must ensure that senior managers are targeted so that their influence can permeate through the organisation. The generic, structured approach to mounting awareness campaigns is described in the introductory chapter.

Organisations that already have formal Communications functions for external customers are normally more familiar with the concepts and more aware of the potential benefits. However it is often more difficult to justify internal Communications functions. Although organisations see the need for formal communications for external customers, it is not always thought necessary to apply these techniques to the provision of internal IT communications.

Remember that knowledge shared can be a powerful force, but knowledge left solely in the hands of an individual may be wasted knowledge!

4.6 Techniques

There are many communication techniques each with its own advantages. It is important to select the correct technique for the type of communication that is being delivered. For example an SLA is likely to be a paper document which formally records agreed target levels for service whereas a daily service bulletin is likely to be more effective if delivered via email.

These techniques are used regularly, sometimes without too much thought and it can be useful to remind ourselves of how to use them effectively. Using the appropriate technique for the type of communication will increase the chance of the message being both received and understood.

Many communication techniques exist and include:

Seminars	:	Seminars are useful when it is necessary to target a particular set of people on a specific topic or subject. They are often used for educational or awareness purposes and to generate discussion and feedback.
Reports	:	Reports provide a formal documented record relating to a specific topic. The scope and objectives should be defined in the introduction which should also state the background to why the report has been produced. The report writer and the recipients should be identified. It is often useful to place details of calculations or analysis in an appendix to the report. The report should not contain any technical or legal/contractual jargon and clear unambiguous language should be used. Example uses include product evaluations, regular Change Management reports and reports on potential service improvements.
Focus/special interest groups	:	The formation of a focus or special interest group is a good way of gathering a team of people who have a common interest but who may have different backgrounds and experience or who are located in different geographical areas. The group should have a clearly defined goal or set of objectives which they work towards and it is important that all members of the group actively contribute. The ICT Steering Group is a good example of a special interest group.
Listening to our customers	:	There are a number of occasions when listening is the most important communication technique. Examples include receiving feedback (either formally or informally) from our customers or when they are telling us about their plans for the future. We should ensure that we are receptive to those who are communicating to us and demonstrate by our actions that we listened to them and understood what they told us.
Circulars / Memos	:	Company circulars or memoranda can be used when there is a need to reach all staff with a communication and there is little time constraint in staff reading it.
One to one sessions	:	There are occasions when the best form of communication is a one-to-one session held in private where discussions can be open and honest. Such occasions include confidential financial or legal discussions, company plans for the future, and annual staff appraisals.

Meetings	:	When used effectively meetings can be a good forum for communication and discussion. Every meeting, whether regular or a one-off, should have a nominated chairperson who is responsible for providing the meeting agenda and objectives, and also setting/agreeing the start and finish times. The Chairperson is also responsible for managing the meeting and ensuring that all attendees have an opportunity to contribute. Example regular meetings are Change Management and Service review meetings and a Business Impact Analysis meeting is one which would be run as required .
Awareness campaigns	:	A well run Awareness Campaign is a very effective way of increasing interest in and raising awareness of a specific topic. An initial presentation, followed by supporting material such as posters, distributing labelled coffee mugs and mouse mats, etc, can be effective in generating and maintaining awareness. Examples of awareness campaigns are those which are used to inform people of forthcoming changes and plans, or to promote new products or services. It is usually important to correctly set expectation levels and not to promise something that cannot be delivered otherwise the campaign can backfire.
Presentations / roadshows	:	Use of electronic presentation aids; don't promise what you can't deliver; useful as part of an awareness campaign to reach a large number of people
Email	:	Electronic mail can be extremely useful in targeting a number of people very quickly. It is usually more informal that printed documents and can be easy, sometimes too easy, to use. Even given its informal nature it is important to check grammar and spelling and also to ensure that the message is only sent to the appropriate audience.
Internet and Intranet sites	:	Web sites can be an excellent source of reference information partly as they are (in theory) always available.

Before deciding on any of the above techniques thought must be given to whether you want the communication to be formal or informal. For example, is a formal record required for legal or contractual reasons? Another consideration is ensuring that technical jargon is not incorporated in the communication.
The type of communication will influence the medium that is most effective. For example, persuasiveness can be very effective in verbal communication where the enthusiasm and tone of voice of the speaker can be very influential.

A good example of operational level communication is major incident reporting during and after an incident. This is a critical element of customer communication that should demonstrate a true understanding of the impact and severity of the incident to the customer. It should be supported by frequent and regular bulletins detailing progress together with the expected time of service restoration, and the production and distribution of a major incident report.
It has often been the case that, when done well, good communication during a service interruption can have a positive impact on customer perception and help to increase customer satisfaction.

4.6.1 The importance of Listening

At times it seems that one of the least used communication techniques is that of listening. It is extremely important that we listen to our customers and ensure that we fully understand what they are telling us. We should also demonstrate our understanding by the actions we subsequently take.

It has often been said that each of us has two ears and one mouth and that they should be used in that proportion!

Communication is a 2 way process – it is vital to ensure that the message has been received and understood by the listener. It is often a good idea to ask the receiver to repeat back the message to confirm that what was sent has been accurately received and understood.

One of the most valuable techniques that allow us to 'listen' to our customers is the Customer Satisfaction Survey (CSS). A sample CSS can be found in the ITIL Customer Liaison book and forms a useful starting point. Before designing and issuing a CSS it is valuable to find out what our customers want to be surveyed on. This will often initiate useful discussion and can, in itself, be a useful source of customer feedback.

Once results from the CSS are returned they must be analysed and acted upon. The results should be reported back to the customer base together with any trend analysis that has been performed.

Many organisations fall into the trap of thinking that they are measuring customer satisfaction when in reality they are measuring perception. This is simply down to the type of questions that are asked as part of the survey. It is now recognised that there is a significant amount of skill required in creating meaningful, measurable and repeatable customer satisfaction surveys.

4.7 Performance Measurement

If we are to successfully demonstrate the performance of the communication function then a combination of metrics will be required which demonstrate both effectiveness and efficiency. The effectiveness of our communication is best measured by those to whom we are communicating. Our customers are prime candidates for this and we should regularly ask them how satisfied they are with the communication they receive from us.

Sample questions could include:

- How well do you feel we have communicated with you?
- Is our communication clear and easily understood?
- Do you receive enough, not enough, or too much communication from us?
- Are we communicating effectively across all levels of the organisation?
- Do you feel that you are kept informed during service incidents?

Efficiency metrics of our communication function could include question such as:

- Has the annual communication plan been produced within "n" days of the yearly production cycle?
- The overall cost of the communication function.

It can be both useful and educational to ask our customers how we should measure our communication performance and to find ways of linking our metrics to their performance metrics.
As with monitoring the performance of all functions, a cycle of review and continuous improvement should be incorporated into the annual Communication Plan.

Annex A Customer survey questionnaires

Al Purpose

Questionnaires have four main purposes:

- to collect the relevant responses
- to ensure responses are comparable
- to minimise bias
- to encourage the recipient to respond.

A2 Design

It is important when designing a questionnaire to consider what analysis is to be carried out on the results, and to ensure that all information collected is relevant to, and adequate to meet the objectives of the survey.

To encourage responses, ensure that a questionnaire is laid out clearly. Vary the format and type of question and include explanations to bridge changes of topic and explain their relevance. This is particularly important for any personal information required to put the respondents’ answers in context.

A3 Questions

A3.1 Construction

Always construct questions very carefully. Ensure that the words used are unambiguous, and have the same meaning for all respondents. Frame questions in everyday language, or the respondents’ business language, not technical computer language.

Include only questions which are essential to the objectives of the survey - resist the temptation to add other questions out of interest as they will confuse the respondent and reduce the response rate.

A3.2 Sequence

Consider the order of questions carefully. Make the early questions of interest to the respondent, and easy to answer, to encourage the respondent to continue. Put any questions that might be considered uninteresting towards the end. If any ‘filter questions’ are used, give very clear directions; eg “If ‘yes’ go on to question X”.

A3.3 Types of question

Four main types of question can be included:

- dichotomous questions (with a ‘yes’ or ‘no’ answer)
- multiple choice questions
- open-ended questions
- rating scales.

Each type of question needs different preparation, and a different form of analysis.

Yes/No questions are easy to ask and answer, but need careful preparation to ensure that they are specific enough to allow only a clear yes or no answer.

Multiple-choice questions are interesting to answer, easy to analyse, but very difficult to design - the designer must know not only what to ask, but also all the possible answers, and must divide these answers into a comprehensive series of mutually exclusive categories.

Open-ended questions are easy to ask, but require more effort from the respondent, and are time consuming to analyse. The results are qualitative rather than quantitative, but they can increase the depth of information collected, and often provide useful quotes and comments for the survey report. Strictly limit the number of this type of question in a survey.

Rating scales are useful for gathering attitudes and opinions. Two examples of scales against which statements are rated follow:

- Scale One
  1) means agree strongly
  2) means agree
  3) means neither agree nor disagree
  4) means disagree
  5) means disagree strongly

- Scale Two
  1) means you personally believe the statement applies very little or not at all
  2) means you personally believe the statement applies to a limited extent
  3) means you personally believe the statement applies to a fair extent
  4) means you personally believe the statement applies to a considerable extent
  5) means you personally believe the statement applies to a very great extent.

The results are not strictly comparable, since the statements may be interpreted variably by different respondents. The can however, be used to explore subtle differences of opinion and emphasis.

A4 Check the draft

Check all questions once written, to ensure they are specific - check the who, what, where, when and how have all been specified. For instance, rather than asking “How often is System X used” ask “How many times have you looked al the screen entitled ... in your office during the past week”. Also check for ambiguity, leading questions, loaded words and generalisations (eg on average). Consider whether what you are asking is reasonable.

A5 Produce the questionnaire

Produce the finished questionnaire as professionally as possible within the cost constraints imposed. A questionnaire which is desk-top published will look more professional than one which is hand typed.

A6 Covering letter

Issue an explanatory letter with each questionnaire stating clearly:

- by whom it has been issued and to whom it should be returned
- its title
- a brief explanation of the background to, and the objective of, the questionnaire
- to whom it is being issued
- any general information or guidance on completion
- an indication of the time likely to be needed to complete the questionnaire
- an appreciation of anticipated co-operation by the recipient.

A7 Pilot the questionnaire

Pilot the questionnaire that has been drawn up, so that faults are discovered and corrected before it is generally issued. Also do a dummy run of the analysis you propose to carry out to check finally that the questions will give the information you require. Revise the questions as necessary, or seek clarification of objectives.

A8 Useful checklist

The following is a checklist for use when producing a questionnaire:

- will the questions listed collect all the information required?
- is every question essential?
- is the mix of questions seeking fact and opinion right?
- is all the personal information really needed for analysis?
- are any questions likely to offend or be answered evasively?
- is the question sequence logical?
- are the types of question appropriate?
- are all questions simple to understand and unambiguous?
- have cushion statements (ie statement which explain the change from one subject to another) and filter statements (ie questions which enable questions that are not relevant to be skipped over) been used where necessary?
- is it reasonable to expect the respondent to answer every question?
- will the answers be easy to record?
- does the questionnaire look good?

A9 Example customer satisfaction questionnaire

The following document is an example of a small, general customer satisfaction questionnaire. It should be issued with a covering letter (see C6 above). Organisations should devise their own series of questions according to their particular objectives.

This type of questionnaire could be for completion by customer personnel in one or more branches/sections. It objective is to gather their opinions on:

- the quality of the IT services they use (which are governed by a service level agreement) and their level of satisfaction
- the general level of satisfaction with other service used
- the quality of, and customer satisfaction with, IT Services support personnel.

A9.1 Example questionnaire

If you are responding for a group of users please indicate how many, and your relation to them (manager, peer, etc).

Section 1
General

1a Which IT services are currently available to you?

Please tick those on the list below that are available to you and add up to three other services that are available and important to you. Please show the order of importance of the services available to you by putting 1 against the most important, 2 against the next most important and so on.

Service name / type Importance

Enquiry system / read only enquiries

Account management / update

etc

Other services available to you

Service name

Service name

Service name

1b Are all the IT services you receive covered by a Service Level Agreement? If not, please list those that are not covered and indicate the number of terminals with access to the service, and your general level of satisfaction with these services using the scale (1=Excellent, 5=Very poor).

Terminals
Satisfaction

1c How often are existing service level agreements, covering services you receive, reviewed and updated?

☐	☐	☐	☐	Please tick One box
6-monthly or more frequently	6-12 monthly	Less frequently than annually	Never been reviewed	

1d Are you kept informed of:

i planned changes to the services which you use

☐	☐	☐	☐	Please tick One box
Not at all	Sometimes	Regularly	Always	

ii the progress of changes?

☐	☐	☐	☐	Please tick One box
Not at all	Sometimes	Regularly	Always	

iii when they are to be implemented?

☐	☐	☐	☐	Please tick One box
Not at all	Sometimes	Regularly	Always	

Section 2
Service Levels

[NOTE: There are six sets of identical questions in this section. Please complete a set for the six (or all, if less than six are used) most important services which you use.]

Service Name: [ ]

For questions 2a to 2e, please state how satisfied you are with the following aspects of service provision giving a marking on a scale of 1 - 5, as follows:

1 Excellent

2 Good

3 Satisfactory

4 Poor

5 Very poor

2a The contracted hours of service. Please state preferred hours if a mark of 4 or 5 is given. [ ]

2b The availability of this service within contracted hours. State briefly what you consider to be the major irritants (if any) regarding downtime. [ ]

2c The current level of performance. What is your level of satisfaction with throughput rates, response times and turnround times?

i throughput rates (ability to consistently input all your daily transactions) [ ]

ii response times at the terminal [ ]

iii batch turnround time for
- an on-line submission [ ]
- manual submission [ ]

2d Please give your assessment of the reliability of the service, ie is it often affected by failures and incidents?

2e How do you regard the service provided for delivery of printed reports:

i are reports consistently delivered by the scheduled times?

ii is the quality of print consistently to the specified standard?

2f Do you know how much your requirements for the service will grow in the next year? (Yes/No)

If "Yes", are the requirements included in SLAs capacity plans? (Yes/No)

The above section is then repeated up to five more times.

Section 3
Support Personnel

For questions 3a to 3e, 3g and 3h, please state how satisfied you are with the following aspects of customer care, giving a marking on a scale of 1 -5 as for questions 2a to 2e

The Service Desk

3a How helpful are Service Desk personnel? ☐

3b How competent do you think they are? ☐

3c How satisfactory are they at keeping you informed of the status of incidents? ☐

3d How well do you rate the Service Desk for consistently dealing with incidents within the times specified in SLA? ☐

Other IT Services personnel

3e How helpful do you find customer liaison staff? ☐

3f How useful is their role to you personally?

☐	☐	☐	☐	☐	Please tick One box
Very useful	Quite useful	A little	Not at all	Don't know	

3g How satisfied are you with the service you receive from the personnel who install/maintain your terminals? ☐

[NOTE: Put a cross in the box if this is not applicable]

3h If you have received any training from IT Services, how satisfactory was it? ☐

[NOTE: Put a cross in the box if this is not applicable]

Section 4
Other aspects of service provision

4a System functionality. How well do the services and applications which you use, provide the functions that you as a customer were and are expecting?

☐	☐	☐	☐	☐	Please tick One box
Totally	Very well	Satisfactorily	Not entirely satisfactorily	Inadequately	

4b Security. Is security and privacy of your data and of access to the IT services via your terminal important to on? (Yes/No) ☐

4c If your answer is "YES", are you satisfied with the level of security and privacy?

☐	☐	☐	☐	☐	Please tick One box
Totally	Generally	Not quite	Not at all	Don't know	

4d Disaster/ recovery. Do contingency plans exist to your knowledge? (Yes/No) ☐

[Note: If you have answered "NO" to this question would you please now move on to question 5a]

4e Have you been involved in testing disaster recovery plans? (Yes/No) ☐

[Note: If you have answered "YES" to this question would you please now move on to question 5a]

4f Do you think that you should have been involved in testing the disaster recovery plans? (Yes/No) ☐

Section 5

General comments

5a What is your overall level of satisfaction with IT Services?

Excellent	Good	Satisfactory	Poor	Very poor
☐	☐	☐	☐	☐

Please tick One box

5b There is space below for any brief comments you may wish to make on any aspect of service provision or on how you feel that IT service quality could be improved. (Please feel free to use this space also for any comments you may wish to make about this questionnaire, or any of your answers.)

[If more space is required, please use another page]

If requested, would you be prepared to help further by participating in a short interview? ☐

Would you please complete the name and address details below, which are only required for checking and follow-up purposes and will be detached from the questionnaire as soon as it is received:

Name: ______ **Title:** ______

Section: ______ **Room No:** ______

Building/Location: ______

Tel No: ______ **Date completed:** ______

--- THANK YOU FOR YOUR CO-OPERATION ---

5 SERVICE LEVEL MANAGEMENT

5.1 Introduction

5.1.1 Why Service Level Management SLM?

Service Level Management is essential in any organisation so that the level of IT Service needed to support the business can be determined, and monitoring can be instigated to identify whether the required service levels are being achieved – and if not, why not.

Service Level Agreements, which are managed through the SLM process, provide specific targets against which the performance of the IT Department can be judged.

5.1.2 Mission for SLM

The mission for Service Level Management is to maintain and gradually improve IT Service quality, through a constant cycle of agreeing, monitoring and reporting upon IT service achievements and instigation of actions to eradicate poor service – in line with business or cost justification. Through these methods, a better relationship between IT and its customers should be developed.

5.1.3 Scope for SLM

Service Level Agreements (SLAs) should be established for all IT services being provided. Underpinning Contracts (UCs) and Operational Level Agreements (OLAs) should also be in place with those suppliers (internal and external) upon whom the delivery of service is dependent.

5.1.4 Basic Concept of Service Level Management

Service Level Management is the name given to the processes of planning, coordinating, drafting, agreeing, monitoring and reporting on Service Level Agreements (SLAs), and the on-going reviewing of service achievements to ensure that the required and cost-justifiable service quality is maintained or where necessary gradually improved. SLAs provide the basis for managing the relationship between the provider and the customer.

When the first ITIL Service Level Management book was published in 1989, very few organisations had SLAs in place. Today most organisations have introduced them - though not always with complete success. This version includes some coverage of the common causes of failure, and guidance on how to overcome these difficulties.

What is a Service Level Agreement?

A two-sided written agreement between an IT Service Provider and the IT Customer(s), defining the key service targets and responsibilities of both parties.

The emphasis must be on agreement and SLAs should not be used as a way of holding one side or the other to ransom. A true partnership should be developed between the IT provider and the customer, so that a mutually beneficial agreement is reached, otherwise the SLA will quickly fall into disrepute and a culture of blame will prevent any true service quality improvements from taking place.

Figure 5.1 below outlines the normal Customer/Service Level Manager relationship:

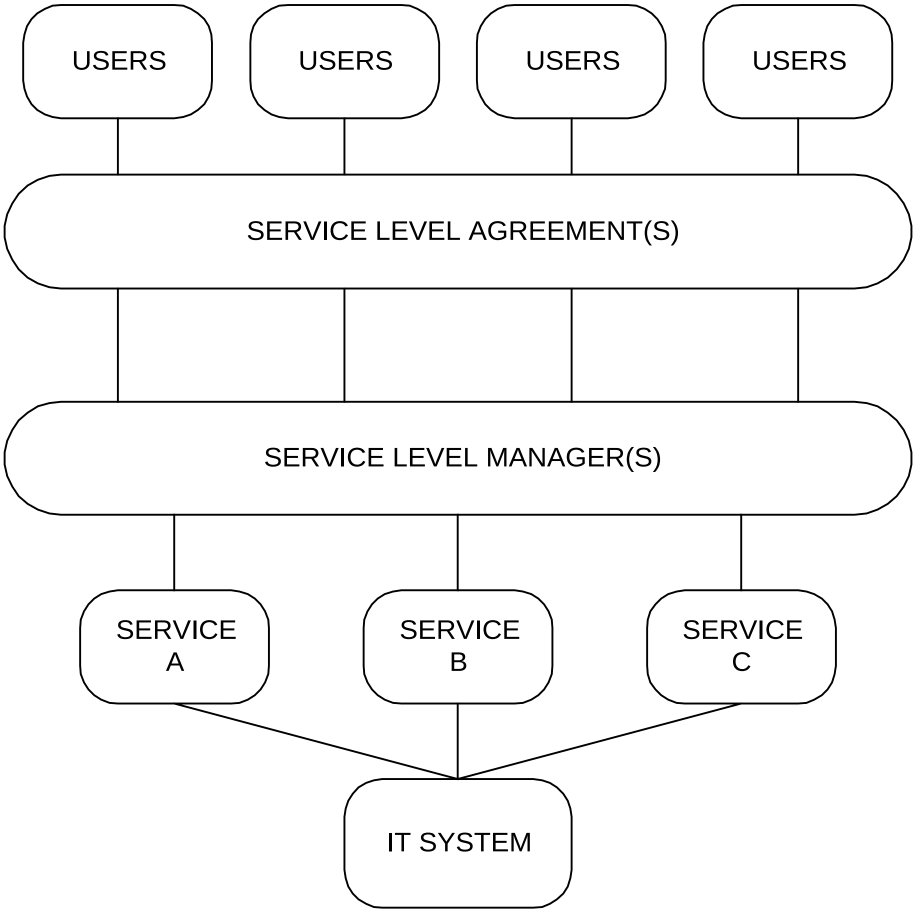


Figure 5.1 Customer / Service Level Manager Relationship

5.2 The Service Level Management Process

There is a definite Service Level Management process that must be planned, implemented, executed and controlled. Figure 5.2 below gives an overview of this:

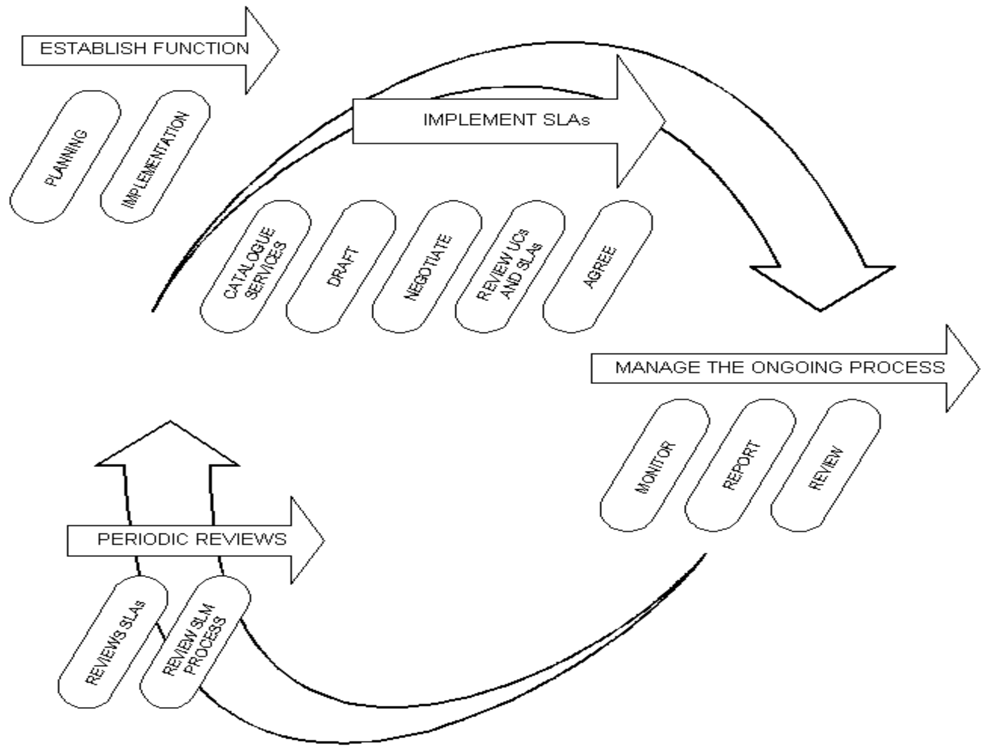


Figure 5.2 Service Level Management Process

The steps in the process are described in detail in sections 5.3 and 5.4 below.

5.2.1 Benefits of Service Level Management

The improvements in service quality and the reduction in service disruption that can be achieved through effective Service Level Management can ultimately lead to significant financial savings. Less time and effort will be spent by IT staff in resolving fewer failures and the IT customers will be able to perform their business functions without adverse impact.

Other specific benefits from Service Level Management include:

- satisfied customers
- both parties to the agreement have a clearer view of roles and responsibilities - thus avoiding potential misunderstandings or omissions
- there are specific targets to aim for and against which service quality can be measured, monitored and reported - "if you aim at nothing, that is usually what you hit"
- IT effort is focused on those areas that the business think is key
- IT and customers have a clear and consistent expectations of the level of service required (i.e. everyone understands and agrees what constitutes a "Priority One" incident, and everyone

has a consistent understanding of what response & fix times are associated with something called “Priority One”.

- service monitoring will allow weak areas to be identified, so that remedial action can be taken (if there is a justifiable business case), thus improving future service quality
- will show where customer/user actions are causing the fault and so identify where working efficiency/training needs can be improved
- Service Level Management underpins supplier management (and vice versa)- in cases where services are outsourced the SLAs are a key part of managing the relationship with the third-party - in other cases service monitoring allows the performance of suppliers (internal and external) to be evaluated and managed
- IT services will be designed to meet Service Level Requirements
- Service Level Agreements can be used as a basis for charging – and will demonstrate what customers are receiving for their money.

[Anecdote – One organisation found that one of its local offices consistently showed up as having incidents causing downtime on a regular basis every month. This was adversely affecting the National Availability figures for the entire organisation. When SLA breaches were investigated it was found that one particular set of users were incorrectly entering commands which took the server down. This highlighted a service fragility but also a training need within that particular office, which dramatically increased their effectiveness and efficiency, and brought the National Availability figures back to what were more acceptable levels.]

The cumulative effect should lead to a gradual improvement in service quality and an overall reduction in the cost of service provision.

In addition Service Level Management establishes - and keeps open - regular lines of communication between IT and its customers. The beneficial impact of this should not be underestimated.

When a true “partnership” develops then problem identification and solving becomes much easier and productive. Where a “blame” culture develops, this in itself becomes a block to Service Quality improvements, as both sides become protective and defensive in their dealings with each other.

5.2.2 Costs

The costs associated with implementing and executing Service Level Management will include:

- staff costs (salaries, training, recruitment costs, consultancy - if needed), both initial and ongoing
- accommodation costs

- support tools (monitoring and reporting, plus some element of integrated Service Management tools)
- hardware on which to run these tools
- marketing costs eg printing

This expenditure should however be seen as an investment and a value-add rather than a cost

5.2.3 Possible Problems

Possible problems that might be encountered with Service Level Management include

- monitoring of actual achievements (particularly achieving the same perception as that held by the customers) – *this is perhaps the most difficult problem that must be addressed first, as it impacts upon the next three:*
- ensuring targets are achievable before committing to them
- verifying targets prior to agreement
- SLAs that are agreed based upon desires rather than achievable targets
- inadequate resources and time - often Service Level Management is seen as something that can be done "in the margins of time" - the ongoing resources are sometimes overlooked.
- not enough seniority/authority given to the Service Level Manager to push through negotiations / improvements
- SLAs may not be supported by adequate contracts or underpinning agreements (see Paragraph 5.4.8)
- responsibilities of each party are not completely defined, creating a danger that some things fall 'between the cracks' and both parties deny responsibility for them.
- being IT based rather than business aligned, especially where the business does not know it's requirements
- SLAs may be too lengthy, not concise, not focussed
- SLAs are not properly
- for commercial companies, Service Level Management is an overhead and not a chargeable service.

Why have some attempts to introduce SLM failed?

Each of the difficulties described above could have contributed to failure. If SLAs are not consistently and accurately defined, documented and monitored, and regular reviews held, then potential service improvement will not be realised and SLAs may fall into disuse (or even disrepute). If this has occurred it will be more difficult to resurrect them or to re-launch Service Level Management. It is far better to recognise the potential difficulties in advance, to put the correct monitoring in place, to ensure that the function is properly resourced, and make sure that the initial attempt is successful.

In other cases, SLAs have been successfully agreed to at a management level, but the fact that that they exist, and details of the agreed upon targets have not been distributed down to the operational level (either on the IT or customer side, or both). In such cases the SLAs have not been actively used to drive up service quality as they should be. To avoid this difficulty it is recommended that SLAs are widely publicised and that summaries of agreed targets are held by all operational IT areas and particularly by the Service Desk. All staff should be clear on what is expected of them in order to ensure targets are met, and made aware of any penalties or consequences if targets are not met

[Anecdote: An Australian insurance company established some very basic SLAs in the late 1980's. Although very simple, they did specify the key customer requirements and the customer-supplier relationship was pretty good. After a few years a new IT Manager was appointed who felt that the SLAs were too simplistic. They were replaced by far more complex documents including many targets that could not accurately be measured and reported. Within a fairly short time, customers became dissatisfied with the SLA arrangements and IT grew disheartened because they could no longer show they were meeting requirements.

When the IT Manager departed the decision was made to return to basics in order to regain customer confidence in the SLA process. Key Message: keep things simple and concentrate on the core customer requirements.]

5.3 Planning the Process

5.3.1 Initial Planning Activities

If Service Level Management is not yet in place, there are a number of activities that must be planned. These include:

- appointment or nomination of a Service Level Manager and any necessary supporting staff
- production of a mission statement
- definition of the objectives and scope of the function
- an awareness campaign to win support for the function and to advise people how and when they might be affected
- definition of roles, tasks and responsibilities
- quantification of activities, resources, funding, quality criteria
- identification of risks
- planning of a Service Catalogue and an SLA structure
- drafting of a pilot SLA format
- identification of support tools, particularly for SLA monitoring
- setting and agreeing incident priority levels and escalation paths, with Customers, and Internal and External providers (in conjunction with Service Desk and Problem Management)

In most cases it will be appropriate for these activities to be incorporated within a formal project and managed via a recognised project management method, such as CCTA's PRINCE2.

5.3.2 Plan Monitoring Capabilities

The importance of adequate capabilities for monitoring SLAs cannot be overstressed. Current monitoring tools and techniques must be reviewed – and action must be taken to improve these where necessary. See Paragraph 5.4.7 for further details.

5.3.3 Establish Initial Perception of the Services.

Before embarking on the introduction of Service Level Management, it is worthwhile to attempt to evaluate the customers' current perception of service levels, so that later the effectiveness of your efforts might be judged. This may also assist you in determining the pace at which you want to proceed and the prioritisation of services to be addressed.

Be aware that the senior managers who are paying for the services are rarely the ones who use them – so it will be necessary to gather perceptions at all levels within the customer community.

At the same time it will also be worthwhile seeking the perception of the service providers - they may not have been asked before and may have a more jaundiced view than the customers.

If monitoring is not yet in place, production of firm metrics may not yet be possible - but some form of customer satisfaction survey or questionnaire may allow you to gain an overall impression. You should approach both at management and at user/operational levels to gain the full perception. It is often the case that the operators of a system have a totally different perception from the Manager, particularly where each is following a different agenda

Experience has shown that face-to-face discussions or telephone surveys are more successful than questionnaires alone (both in terms of percentage returns and accuracy of information), though they are more costly and take longer to complete. A combination of the two is usually the most effective approach.

It may also be worthwhile trying to 'network' and get some ideas of reasonable expectation and targets (eg attend conferences, join user groups such as the *it*SMF)

5.3.4 Underpinning Contracts and Operational Level Agreements

Plans must be made to review or implement contracts with external suppliers and Operational Level Agreements (OLAs) with internal suppliers to ensure services are provided that underpin the SLAs. See Paragraph 5.4.8 for further details.

5.4 Implementing the Process

When planning activities have been completed, the following activities must be undertaken to implement Service Level Management.

5.4.1 Produce a Service Catalogue

Many organisations' IT Infrastructure has grown and developed over many years, and there may not be a clear picture of all the services currently being provided and the users of each. In order to establish an accurate picture, it is recommended that an IT Service Catalogue is produced.

Such a catalogue should list all of the services being provided, a summary of their characteristics and details of the users and maintainers of each. A degree of 'detective work' may be needed to compile this list and agree it with the customers (sifting through old documentation, searching program libraries, talking with IT staff and customers, looking at procurement records and talking with suppliers and contractors etc). If you have a CMDB or any sort of asset database these may be a valuable source of information. (Hint - Service Desk Incident records are a good pointer to those old systems that everyone else but the user has forgotten about)

What is a service?

This question is not as easy to answer as it may first appear, and many organisations have failed to come up with a clear definition in an IT context. IT staff often confuse a 'service' as perceived by the customer with an IT system. In many cases one 'service' can be made up of other 'services' (and so on) which are themselves made up of one or more IT systems. A good starting point is often to ask your customers "which services do you use?" Customers often have a greater clarity of what they believe a service to be.

One possible definition may be: "A business process underpinned by one or more IT systems".

To avoid confusion, it may be a good idea within the Service Catalogue to define a hierarchy of services by qualifying exactly what type of service is meant (eg business service – that which is seen by the customer), infrastructure services, network service, application service (all invisible to the customer – but essential to the delivery of customer services), - or whatever is appropriate for your organisation).

When completed, the catalogue may initially consist of a matrix, table or spreadsheet. A number of organisations have progressed to integrate and maintain their Service Catalogue as part of their Configuration Management Database (CMDB). By defining each service as a Configuration Item (CI) - and where appropriate, relating them to form a service hierarchy, this allows the organisation to relate events such as incidents and RFCs to the services affected, thus providing the basis for service monitoring via an integrated tool (eg "list or give the number of incidents affecting this particular service"). This can work well and is recommended.

The service catalogue can also be used for other Service Management purposes (eg for performing a Business Impact Analysis (BIA) as part of IT Service Continuity Planning, or as a starting place for Workload Management, part of Capacity Management). The cost and effort of producing the catalogue is therefore easily justifiable. If done in conjunction with prioritisation of the BIA, then it is possible to ensure that the most important services are covered first.

An example of a simple service catalogue which can be used as a starting point is given in Annex A.

5.4.2 Expectation Management

From the outset, it is wise to try and manage customer's expectations. SLAs are just documents - and in themselves will not materially alter the quality of service being provided (though they may affect behaviour and help engender an appropriate service culture which can have an immediate beneficial effect, and longer term improvements will be possible). A degree of patience is therefore needed and should be built into expectations.

Where charges are being made for the services provided, this should modify customer demands (customers can have whatever they can cost justify – providing it fits within agreed corporate strategy - and have authorised budget for, but no more!). Where direct charges are not made, the support of senior business managers should be enlisted to ensure that excessive or unrealistic demands are not placed upon the IT provider by any individual customer group

5.4.3 Plan the SLA Structure

Using the catalogue as an aid, the Service Level Manager must plan the most appropriate SLA structure to ensure that all services and all customers are covered in a manner best suited to the organisation's needs. There are a number of potential options, including:

Service Based

Where an SLA covers one or more services, for all the users of that service. For example an SLA may be established for an organisation's Email service, covering all of the users of that service.

This may appear fairly straight forward – but difficulties may arise if the specific requirements of different user groups vary for the same service – or if characteristics of the IT infrastructure mean that different service levels are inevitable (eg head office staff may be connected via a high-speed LAN whilst local offices may have to use a dial-up). In such cases, separate targets may be needed within the one agreement.

Difficulties may also arise in determining who the signatories to such an agreement should be.

Customer Based

An agreement with an individual customer group, covering all the services they use. For example, agreements may be reached with an organisation's Finance Department covering, say, the Finance System, the Accounting System, the Payroll System, the Billing System, the Procurement System and any other IT systems that they use.

Customers often prefer such an agreement, as all of their requirements are covered in a single document. Only one signatory is normally required, which simplifies this issue.

Note. A combination of either of these structures might be appropriate, providing all services and customers are covered, with no overlap or duplication.

Multi-level SLAs

Some organisations have chosen to adopt a multiple level SLA structure. For example, a three-layer structure as follows:

Corporate Level: covering all the generic SLM issues appropriate to every customer throughout the organisation. These issues are likely to be less volatile and so updates are less frequently required.

Customer Level: covering all SLM issues relevant to the particular customer group, regardless of the service being used.

Service Level: covering all SLM issues relevant to the specific service, in relation to this specific customer group (one for each service covered by the SLA)

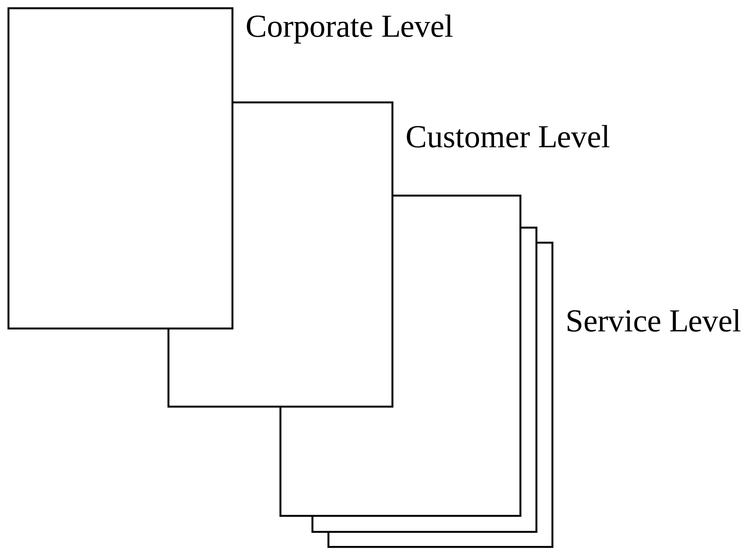


Figure 5.3 Multi-level SLAs

Such a structure allows SLAs to be kept to a manageable size, avoids unnecessary duplication, and reduces the need for frequent updates.

(The UK Employment Service successfully used a three-level SLA format. They used the titles: Framework, Standard, Specific to describe their 3 levels - but the concept was basically the same as described above)

5.4.4 Establish Service Level Requirements and Draft SLA

It is difficult to give guidance on which of these should come first – as it often requires an iterative process. Once the SLA structure has been agreed, a first SLA must be drafted. It is advisable to involve customers from the outset, but rather than going along with a blank sheet to commence with, it is often advisable to produce a first outline draft as a starting point for more detailed and in depth discussion. Be careful though not to go too far and appear to be giving the customer a fait accompli.

It can be difficult to draw out requirements, as the business may not know what they want - especially if not asked before and they may need help in understanding and defining their needs.

Be aware that the requirements initially expressed may not be those ultimately agreed, particularly where no charging is in place. Several iterations of negotiations may be required before an affordable balance is struck between what is sought and what is achievable and affordable.

Many organisations have found it valuable to produce a pro-forma that can be used as a starting point for all SLAs. The proforma can often be developed alongside the pilot SLA.

Guidance on the items to be included in an SLA is given in Section 5.6.

(Hint - make roles and responsibilities a part of the SLA. You will need to consider three perspectives, the IT provider, the IT Customer, and the actual users)

5.4.5 Wording of SLAs

The wording of SLAs should be clear and concise and leave no room for ambiguity. There is normally no need for agreements to be couched in legal terminology, and plain language aids a common understanding. It is often helpful to have an independent person, who has not been involved with the drafting, to do a final read-through. This often throws up potential ambiguities and difficulties that can then be addressed and clarified.

It is also worth remembering that SLAs may have to cover services offered internationally. In such cases the SLA may have to be translated into several languages. Remember also that an SLA drafted in a single language may have to be reviewed for suitability in several different parts of the world (ie a version drafted in Australia may have to be reviewed for suitability in the USA or the UK - and differences in terminology, style and culture must be taken into account).

5.4.6 Seek Agreement

Using the draft agreement as a basis, negotiations must be held with the customer(s), or customer representatives to finalise the contents of the SLA and the initial service level targets, and with the service providers to ensure that these are achievable. Guidance on general negotiating techniques is included in the ITIL Business & Management Skills book.

One problem you might encounter, is identifying exactly who to negotiate with. Who 'owns' the service? In some cases this may be obvious, and a single customer manager is willing to act as the signatory to the agreement. In other cases, it might take quite a bit of negotiating or cajoling to find a representative 'volunteer' (beware that volunteers often want to express their own personal view rather than represent a general consensus), or you may need to get them all to sign (messy!).

If customer representatives exist who are able to genuinely represent the views of the customer community, because they frequently meet with a wide selection of them, this is ideal. Unfortunately, all too often so-called representatives are head-office based and seldom come into contact with genuine service users. In the worst-case, the Service Level Manager may have to perform his/her own programme of discussions and meetings with customers to ensure true requirements are identified.

(Anecdote - On negotiating the current service and support hours for a large system an organisation found a discrepancy in the required time of usage between Head Office and the field offices Users. Head Office (with a limited user population) wanted Service Hours covering 8am to 6pm, whereas the field (with at least 20 times the user population) stated that starting an hour earlier would be better - but all offices closed to the Public by 4pm at the latest and so wouldn't require a service much beyond this. Head Office won the "political" argument and so the 8am to 6pm band was set. When the Service came to be used (and hence monitored) it was found that Service extensions were usually asked for by the field to cover the extra hour in the morning, and actual usage figures showed that the system had not been accessed after 5pm, except on very rare occasions. The Service Level Manager was blamed by the IT staff for having to cover a late shift, and by the Customer Representative for charging for a service that was not used (ie staff and running costs).

Hint. Care should be taken if opening discussions on service levels for the first time, as it is likely that 'current issues' (the failure that occurred yesterday!) or long-standing grievances (that old printer that we have been trying to get replaced for ages!) are likely to be aired at the outset. Important though these may be, they must not be allowed to get in the way of establishing the longer-term requirements. Beware however that you may well have to address any issues raised at the outset before gaining any credibility to progress further.

If there has been no previous experience of SLM, then it is advisable to start with a pilot SLA.

A decision should be made on which services/customers to be used for the pilot. It is helpful if the selected customer is enthusiastic and wishes to participate - perhaps because they are anxious to see improvements in service quality. The results of the initial customer perception survey may give pointers to a suitable pilot.

Hint – don't pick an area where large problems exist as your pilot. Try to pick an area where you are sure that you can show some quick benefits and develop the SLM Process. Nothing breeds acceptance of a new idea quicker than success.

One difficulty sometimes encountered is that staff at different levels within the customer community may have different objectives and perceptions. For example a senior manager may never use a service and may be more interested in issues such as value for money and output, whereas a junior member of staff may use the service throughout the day and may be more interested in issues such as responsiveness, usability and reliability. It is important that all of the appropriate and relevant customers requirements, at all levels, are identified and incorporated in SLAs.

Some organisations have formed focus groups from different levels from within the customer community to assist in successfully ensuring that all issues have been correctly addressed. This takes additional resources, but can be well worth the effort.

The other group of people that have to be consulted during the whole of this process is the appropriate representatives from within the IT provider side (whether internal or from a third-party supplier). They need to agree that targets are realistic, achievable and affordable. If they are not, further negotiations will be needed until a compromise acceptable to all parties is agreed.

The views of suppliers should also be sought and any contractual implications should be taken into account during the negotiation stages.

Where no past monitored data is available, it will be advisable to leave the agreement in draft format for an initial period, until monitoring can confirm that initial targets are achievable. Targets may have to be re-negotiated in some cases. When targets have been confirmed, the SLAs must be signed.

Once the pilot has been completed and any initial difficulties overcome, you can then move on and gradually introduce SLAs for other services/customers.

If you have decided from the outset to go for a multi-level structure, it is likely that you will need to cover the Corporate-level issues for all customers at the time of the initial pilot. It is also worth trialling the Corporate issues during this pilot phase.

Hint. Don't go for easy targets at the Corporate level, they may be easy to achieve but will not have any value in improving Service Quality. Also if the targets are set at a high enough level the Corporate SLA can be used as the standard that all new services should reach.

One point to ensure is that at the end of the drafting and negotiating process, the SLA is actually signed by the appropriate managers on the customer and IT provider sides to the agreement. This gives a firm commitment by both parties that every attempt will be made to meet the agreement by both sides. Generally speaking, the more senior the signatories are within their respective organisations, the stronger the message of commitment will be.

Once an SLA is agreed, wide publicity needs to be used to ensure that customers and IT alike are aware of its existence, and of the key targets.

It is important that the Service Desk staff are committed to the SLM process and become proactive ambassadors for SLAs, embracing the necessary service culture, as they are the first contact point for customers' incidents, complaints and queries. If the Service Desk Staff are not fully aware of SLAs in place and do not act upon them then Customers will very quickly lose faith in SLAs.

5.4.7 Establish Monitoring Capabilities

Nothing should be included in an SLA unless it can be effectively monitored and measured at a commonly agreed point. The importance of this cannot be over-stressed, as inclusion of items that cannot be effectively monitored will almost always result in disputes and eventual loss-of-faith in the Service Level Management process. A lot of organisations have discovered this the "hard way" and as a consequence, have absorbed heavy costs both in a financial sense as well as in terms of negative impacts on their culture.

Anecdote: A global network provider agreed availability targets for the provision of a managed network service. These availability targets were agreed at the point were the service entered the customer's premises. However, the global network provider could only monitor and measure availability at the point the connection left their premises. The network links were provided by a number of different national Telecommunications service providers, with widely varying availability levels. The result was a complete mis-match between the availability figures produced by the network provider and the customer, with correspondingly prolonged and heated debate and argument.

Existing monitoring capabilities should be reviewed and upgraded as necessary. Ideally this should be done ahead of or in parallel with the drafting of SLAs, so that monitoring can be in place to assist with the validation of proposed targets.

It is essential that monitoring matches the customer's true perception of the service. Unfortunately this is often very difficult to achieve. For example, monitoring of individual components, such as the network or server, does not guarantee that the service will be available so far as the customer is concerned - a desktop or application failure may mean that the service cannot be used by the customer. Without monitoring all components in the end-to-end service (which may be very difficult and costly to achieve) a true picture cannot be gained.

Similarly, Users must be aware that they should report incidents immediately to aid diagnostics, especially if performance related.

Where multiple services are delivered to a single workstation, it is probably more effective to only record downtime against the Service the user was trying to access at the time (though this needs to be agreed with the Customers). Customer perception is often that although a failure might affect more than one service all they are bothered about is the service they cannot access at the time of the reported incident – though this is not always true, so caution is needed.

A considerable number of organisations use their Service Desk, linked to a comprehensive CMDB, to monitor customer's perception of availability. This may involve making specific changes to incident/problem logging screens and will require stringent compliance with incident logging procedures. All of this will need discussion and agreement with the Availability Management function. Chapter 9 (Availability Management) gives guidance and examples of the formulae that might be used to determine service availability levels, and the amendments that may be needed to capture the required data.

The Service Desk is also used to monitor incident response times and resolution times, but once again the logging screen might need amendment to accommodate data capture, and call logging procedures may need tightening and must be strictly followed. If support is being provided by a third-party, this monitoring may also underpin supplier management.

It is essential to ensure that any incident/problem handling targets included in SLAs are the same as those included in Service Desk tools and used for escalation and monitoring purposes. Where organisations have failed to recognise this, and perhaps used defaults provided by the tool supplier, they have ended up in a situation where they are monitoring something different than that which has been agreed in the SLAs, and are therefore unable to say whether SLA targets have been met, without considerable effort to massage the data.

Some amendments may be needed to support tools, to include the necessary fields so that relevant data can be captured.

It is advisable to ensure that incident response and resolution targets are the same for all SLAs – otherwise it becomes a nightmare to try and monitor different targets via the Service Desk software.

Another notoriously difficult area to monitor is transaction response times (the time between sending a screen and receiving a response). Often end-to-end response times are technically very difficult to monitor. In such cases it may be appropriate to deal with this as follows:

(a) include a statement in the SLA along the following lines "The services covered by this SLA are designed for high-speed response and no significant delays should be encountered. If a response time delay of more than x seconds is experienced for more than y minutes this must be reported immediately to the Service Desk"

(b) agree and include in the SLA an acceptable target for the number of such incidents that can be tolerated in the reporting period.

(c) create a new incident category "poor response"(or similar) and ensure that any such incidents are accurately logged and that they are related to the appropriate service.

(d) produce regular reports of occasions where SLA transaction response time targets have been breached, and instigate investigations via Problem Management to correct the situation

This approach not only circumvents the technical difficulties of monitoring, but ensures that incidences of poor response are reported at the time they are occurring. This is very important as poor response is often caused by a number of interacting events, which can only be detected if they are investigated immediately (the 'smoking gun' syndrome!).

The preferred method however is to implement some form of automated client/server response time monitoring. These tools are becoming increasingly available and increasingly more cost effective to use. These tools provide the ability to measure or sample actual or very similar response times as those being experienced by a variety of end users.

Hint. Some organisations have found that in reality “poor response” is sometimes a User perception incident. The User will become used to a particular level of response over a period of time and will start complaining as soon as this is slower. Take the view that “if the User thinks the system is slow - then it is”.

If the SLA includes targets for assessing and implementing Requests for Charge (RFCs), the monitoring of targets relating to Change Management should ideally be monitored via whatever Change Management tool is in use (preferably part of an integrated Service Management support tool) and change logging screens and escalation processes should support this.

There are a number of important ‘soft’ issues that cannot be monitored by mechanistic or procedural means, such as customer’s overall feelings (these need not necessarily match the ‘hard’ monitoring). For example, even when there have been a number of reported service failures the customers may still feel positive about things, because they may feel satisfied that appropriate actions are being taken to improve things. Of course, the opposite may apply and customers may feel dissatisfied with some issues (eg the manner of staff on the Service Desk) when few or no SLA targets have been broken.

It is therefore recommended that attempts are made to monitor customer perception on these soft issues. Methods of doing this include:

- telephone perception surveys (perhaps at random, or using regular customer liaison representatives)
- periodic questionnaires
- satisfaction survey handouts (left with customers following installations, service visits etc.)
- user group meetings

Where possible, targets should be set for these and monitored as part of the SLA (eg an average score of 3.5 should be achieved by the service provider on results given, based on a scoring system of 1 to 5, where 1 is poor performance and 5 is excellent)”

Ensure that if users bother to provide feedback they yield some return and demonstrate that their comments have been incorporated in an action plan, perhaps a Service Improvement Programme.

5.4.8 Review Underpinning Contracts and Operational Level Agreements

Most IT Service Providers are dependent to some extent on their own suppliers (both internal and/or external). They cannot commit to meeting SLA targets unless their own supplier’s performance underpins these targets. Contracts with external suppliers are mandatory, but many organisations have also identified the benefits of having simple agreements with internal support groups, usually referred to as Operational Level Agreements (OLAs). Figure 5.4 below illustrates this.

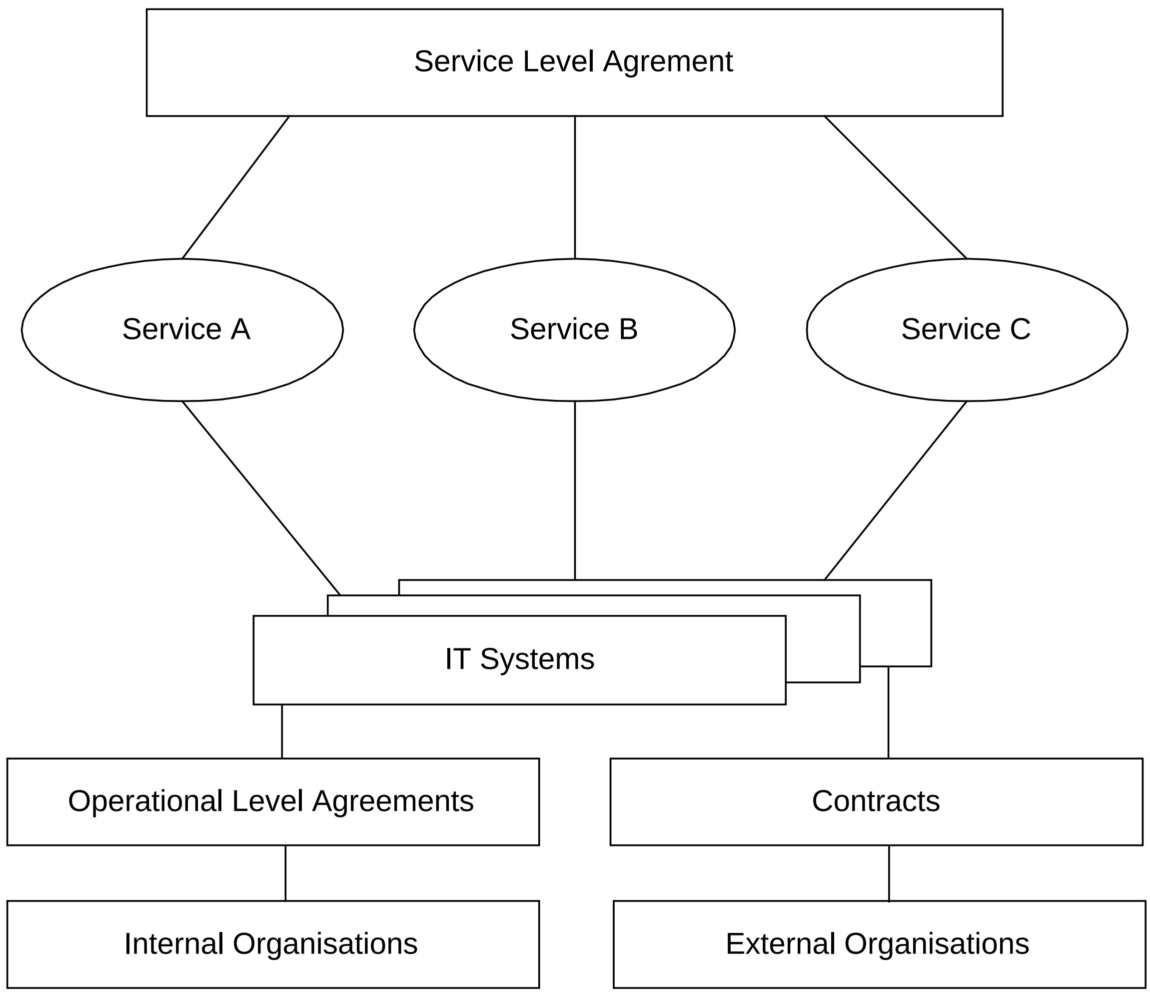


Figure 5.4 SLA Support Structure

OLAs need not be very complicated, but should set out specific back-to-back targets for support groups that underpin the targets included in SLAs. For example, if the SLA includes overall time to respond and fix targets for incidents (varying on the priority levels), then the OLAs should include targets for the each of the elements in the support chain (target for the Service Desk to answer calls, escalate etc, targets for Network Support to start to investigate and to resolve network related errors assigned to them). In addition, overall support hours should be stipulated for all groups that underpin the required service availability times in the SLA. If special procedures exist for contacted staff (eg out of hours telephone support) these must also be documented.

It must be understood however that the incident resolution targets included in SLAs will not normally <u>match</u> the same targets included in contracts or OLAs with suppliers. This is because the SLA targets must include an element for all stages in the support cycle (eg Detection time, Service Desk logging time, escalation time, referral time between groups etc, Service Desk review and closure time - as well as the actual time fixing the failure). The SLA target should cover all of this.

Before committing to SLAs, existing contractual arrangement must be investigated and where necessary, upgraded. This is likely to incur additional costs, which must either be absorbed by IT, or passed on to the customer. In the latter case the customer will have to agree to this, or the more relaxed targets in existing contracts should be agreed for inclusion in SLAs.

OLAs should be monitored against these targets and feedback given to the Managers of the support groups. This will highlight potential problem areas, which may need to be addressed internally or by a further review of the SLA

Consideration should also be given to introducing OLAs, where they do not already exist.

5.4.9 Define Reporting and Review Processes

The SLA reporting mechanisms, intervals and report formats must be defined and agreed with the customers.

The frequency and format of Service Review Meetings must also be agreed with the customers. Regular intervals are recommended. Periodic reports should fit in with the reviewing cycle.

Some organisations have chosen to incorporate SLA reviews within Customer Relationship Manager roles (with the Customer Relationship Manager(s) acting on behalf of the Service Level Manager in the SLA review process). This can be particularly beneficial in larger organisations where the customer base is widely distributed and where the Customer Relationship Management role already exists. However, there must be close co-operation between the Service Level Manager and the Customer Relationship Manager to enable feedback to flow accurately in both directions. It is advisable that the Service Level Manager attends at least some of these meetings in order to get an accurate reading of Customer issues.

The SLAs themselves must be reviewed periodically (eg annually in line with financial cycle) to ensure that they are still current and indeed still relevant - does the SLA still fit the needs of the business and the capabilities of IT? All SLAs should be under strict change management control and any changes should be reflected in an update to the Service Catalogue, if needed.

5.4.10 Publicise the existence of SLAs

Steps must be taken to advertise the existence of the new SLAs amongst the Service Desk and other support groups (with details of when they become operational). It may be helpful to extract key targets from the SLAs into tables that can be on display in support areas – so that staff are always aware of the targets they are working to.

If support tools allow it, these targets should be included as thresholds and automatically alerted against when a target is threatened or actually breached.

SLAs and the targets they contain must also be publicised amongst the user community, so that users are aware of what they can expect from the services they use, and know at what point to start to express dissatisfaction.

5.5 Executing the Process

The following ongoing activities must be undertaken to execute the process.

5.5.1 Monitoring and Reporting

Immediately the SLA is agreed, monitoring must be instigated (if this has not been done already), and service achievement reports must be produced. Operational reports must be produced frequently (daily – perhaps even more frequently), but where possible, exception reports should be produced whenever an SLA has been broken (or threatened if appropriate thresholds have been set to give an 'early warning').

Periodic reports must be produced and circulated to customers (or their representatives) and appropriate IT Managers a few days in advance of SLA reviews, so that any queries or disagreements can be resolved ahead of the review meeting. The meeting is not then diverted by such issues.

The periodic report should incorporate details of performance against all SLA targets, together with details of any trends or specific actions being undertaken to improve service quality. Annex B gives an example of an SLA Monitoring (SLAM) chart which can be used at the front of the report to give an 'at a glance' overview of how achievements have measured up against targets. These are most effective if colour coded (Red-Amber-Green, and sometimes referred to as RAG charts as a result).

Other interim reports may be required by IT management for internal performance reviews and/or supplier or contract management.

Note that this is likely to be an evolving process - your first effort will probably not be the one you end up with.

The amount of effort required to produce and verify reports should not be underestimated. It can be extremely time consuming, and if reports do not accurately reflect the customers own perception of service quality, they can add insult to injury.

The Service Level Manager should identify the specific reporting needs and automate production of these reports, so far as possible. The extent, accuracy and ease with which automated reports can be produced should form part of the selection criteria for integrated support tools.

5.5.2 Service Review Meetings

Periodic review meetings must be regularly held with customers (or their representatives) to review the service achievement in the last period and to preview any issues for the coming period. It is normal to hold such meetings monthly, or as a minimum, quarterly.

Actions must be placed on the customer and provider as relevant to improve weak areas where targets are not being met.

Particular attention should be focused on each breach of service levels to determine exactly what caused the loss of service and what can be done to prevent any recurrence for the future. If it is decided that the service level was, or has become, unachievable, it may be necessary to review and re-agree different service targets.

If the service break has been caused by a failure of a third-party or internal support group, it may also be necessary to review the underpinning agreement or OLA.

All actions must be minuted, and progress should be reviewed at the next meeting to ensure that action items are being followed up and properly implemented.

Hint. "A spy in both camps" - Service Level Managers can be viewed with a certain amount of suspicion by both the IT Service Provider staff and the Customer Representatives. This

can be particularly so in places where Customer Relationship Managers are not used. This is due to the dual nature of the job where they are acting as an unofficial customer representative when talking to IT staff, and as an IT provider rep when talking to the Customers. This is usually aggravated when you have to fight the "oppositions" point of view in any meeting etc. To avoid this the Service Level Manager should be as open and helpful as possible (within the bounds of commercial propriety), when dealing with both sides.} Ambassador for IT, lonely job, finds allies in other Service Level Managers (eg via itSMF). Never vent frustrations about colleagues with the client community – or criticise a colleague in public.

5.5.3 Service Improvement Programme (SIP)

The Service Level Management process often generates a good starting point for a Service Improvement Programme (SIP) – and the service review process may drive this.

Where an underlying difficulty has been identified which is adversely impacting upon service quality, the Service Level Manager must, in conjunction with Problem Management and Availability Management, instigate a SIP to identify and implement whatever actions are necessary to overcome the difficulties and restore service quality. Further guidance on this and the specific techniques that might be used can be found in the Chapter 6 (Problem Management) of the Service Support book and in Chapter 9 (Availability Management) of this book. SIP initiatives may also focus on such issues as user training, system testing and documentation. In these cases the relevant people need to be involved and adequate feed-back given to make improvements for the future.

At any point in time, a number of separate initiatives which form part of the SIP may be running in parallel to address difficulties with a number of services.

Some organisations have established an up-front annual budget held by Service Level Management from which SIP initiatives can be funded. This means that action can be undertaken quickly and that Service Level Management is demonstratively effective.
This practice should be encouraged and expanded to enable Service Level Management to be become increasingly proactive and predictive

If an organisation is outsourcing its service delivery to a third party, the issue of service improvement should be discussed at the outset and covered (and budgeted for) in the contract, otherwise there will be no incentive during the lifetime of the contract for the supplier to improve service targets if they are already meeting contractual obligations and additional expenditure is needed to make the improvements.

5.5.4 Maintenance of SLAs, Contracts and OLAs

SLAs, underpinning contracts and OLAs must all be maintained up to date. They should be brought under change control and reviewed periodically, at least annually, to ensure that they are still current and comprehensive, and still aligned to business needs and strategy.

These reviews should ensure that the services covered and the targets for each are still relevant – and that nothing significant has changed that will invalidate the agreement in any way (this should include infrastructure changes, business changes, supplier changes etc). Where changes are detected, the agreements must be updated to reflect the new situation.

5.6 SLA Contents and Key Targets

The specific content and the initial targets to be included in SLAs must be agreed. It is difficult to be prescriptive, as each situation will be unique, and content will vary depending upon the type of SLA, but there are a number of common features that often occur within SLAs. These include:

Introduction

- Parties to the agreement
- Title and brief description of the agreement
- Signatories
- Dates: start, end, review
- Scope of the agreement; what is covered and what is excluded
- Responsibilities of both the service provider and the customer
- A description of the services covered

Service hours

- The hours that each service is normally required (eg 7 x 24, Monday to Friday 08:00 - 18:00)
- Arrangement for requesting service extensions, including required notice periods (eg request must be made to the Service Desk by 12:00 noon for an evening extension, by 12:00 noon on Thursday for a week-end extension)
- Special Hours (eg public holidays)
- Service Calendar

Availability

- Availability target within agreed hours, normally expressed as a percentage (eg 99%) - measurement period and method must be stipulated. This may be expressed for the overall service, underpinning services and critical components or all three.

Reliability

- Usually expressed as the number of service breaks, or the Mean Time Between Failures (MTBF) or Mean Time Between System Incidents (MTBSI)

Support

- Support hours (where these are not the same as Service Hours).
- Arrangement for requesting support extensions, including required notice periods (eg request must be made to the Service Desk by 12:00 noon for an evening extension, by 12:00 noon on Thursday for a week-end extension)
- Special Hours (eg public holidays)
- Target time to respond, either physically or by other method (eg telephone contact, email), to incidents
- Target time to resolve incidents, within each incident priority

  targets will vary depending upon incident priorities

Throughput

- Indication of likely traffic volumes and throughput activity (eg the number of transactions to be processed, number of concurrent users, amount of data to be transmitted over the network). This is important to determine whether any performance issues have been caused by excessive throughput outside the terms of the agreement.

Transaction Response Times

- Target times for average, or maximum workstation response times (sometimes expressed as a percentile - eg 95% within 2 seconds – but see recommendations in Paragraph 5.4.7)

Batch turnaround times

- Times for delivery of input and the time and place for delivery of output

Change

- Targets for approving, handling and implementing RFCs, usually based upon the category or urgency/priority of the change.

Contingency & Security

- A brief mention of contingency plans and how to invoke them, and coverage of any security issues, particularly any responsibilities of the customer (eg back-up of free-standing PCs, password changes).
- Details of any diminished or amended service targets should a disaster situation occur (if no separate SLA exists for such a situation)

Charging

- Details of the charging formula and periods (if charges are being made). If the SLA covers an outsourcing relationship, charges should be detailed in an Annex as they will be often covered by commercial in confidence provisions.

Service Reporting and Reviewing

- the content, frequency and distribution of service reports the frequency of service review meetings

Performance incentives/penalties

- Details of any agreement regarding any financial incentives or penalties based upon performance against service levels - or lack of performance (more likely to be included if the services are being provided by a third-party organisation). It should be noted that penalty clauses are not entirely successful. They can prove a barrier to partnership if unfairly invoked on a technicality and can also make service provider staff unwilling to admit to mistakes for fear of penalties being imposed. This can, unless used properly, be a barrier to effective problem solving.

NOTE: The contents given above are examples only – and should not be regarded as exhaustive or mandatory in every case!

5.6.1 New Services

Service Level Requirements

Whilst many organisations have to give initial priority to introducing SLAs for existing services, it is also important to establish procedures for agreeing Service Level Requirements (SLRs) for new services being developed or procured.

The SLRs should be an integral part of the systems design criteria or functional specification, and should from the very outset form part of the testing/trialling criteria as the system progresses through the stages of development or procurement.

A draft SLA should be developed alongside the service itself, and should be signed and formalised no later than the point at which the service is introduced into live use.

Support Planning

Another area that requires attention, if new services are to be introduced in a seamless way into the live environment, is the planning and formalisation of the support arrangements. Specific responsibilities need to be defined and either added to existing contracts/OLAs, or new ones need to be agreed. The support arrangement and all escalation routes also need adding to the CMDB, so that Service Desk and other support staff are aware of them. Where appropriate, initial training and familiarisation for the Service Desk and other support groups must be instigated before live support is needed.

It should be noted that additional support resources (ie: more staff) may be needed to support new services. There is often an expectation that an already overworked support group will magically cope with the additional effort imposed by a new service!

5.7 KPIs and Metrics for SLM Efficiency and Effectiveness

The following KPIs (Key Performance Indicators) and Metrics can be used to judge the efficiency and effectiveness of the SLM processes and function.

- What number or percentage of Services are covered by SLAs?
- Are underpinning contracts and OLAs in place for all SLAs and for what percentage?
- Are SLAs being monitored and are regular reports being produced?
- Are review meetings being held on time and correctly minuted?
- Is there documentary evidence that issues raised at reviews are being followed up and resolved (eg via a SIP)

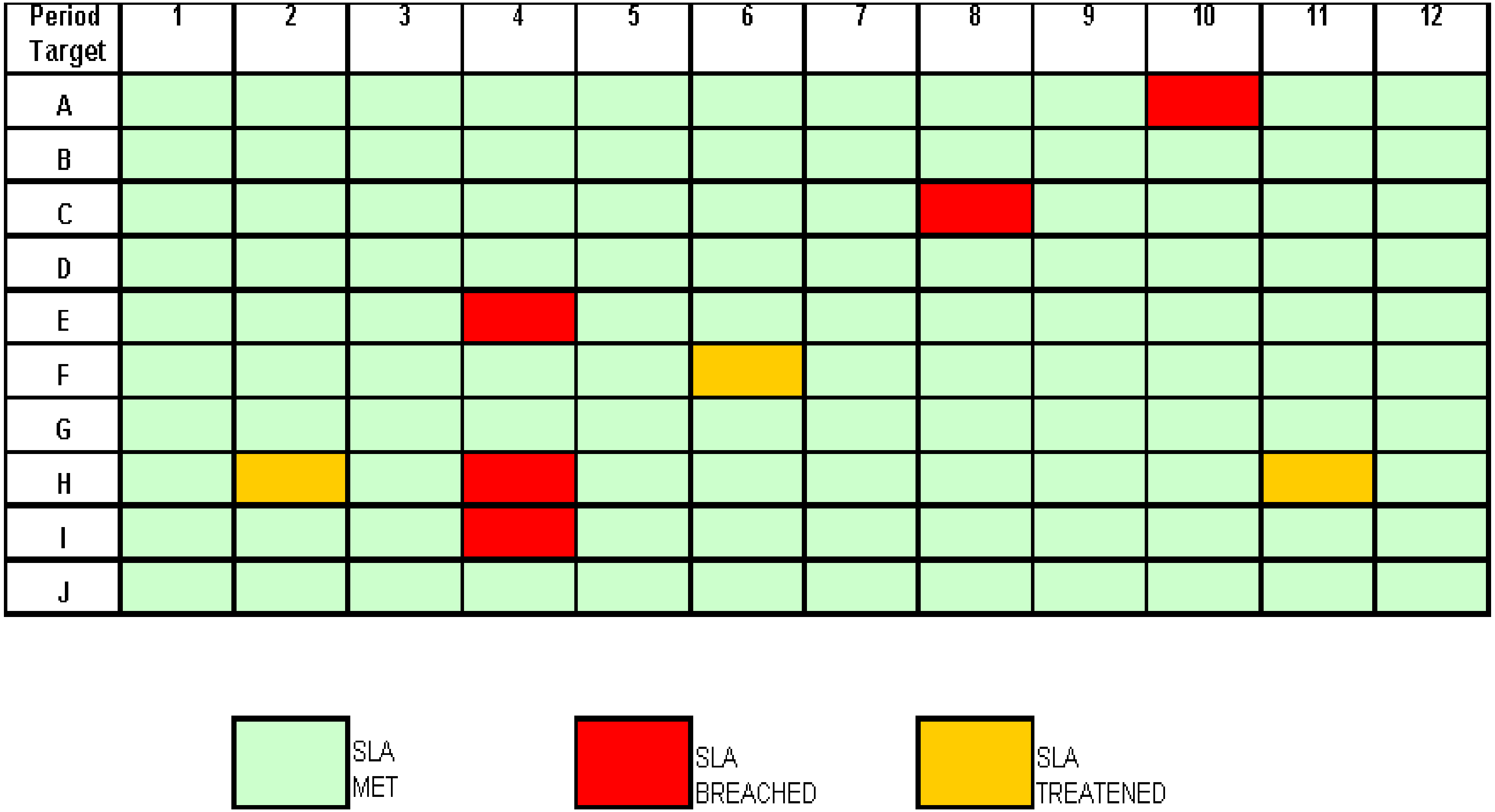
- Are SLAs, OLAs and underpinning contracts current? What percentage is in need of review and update?
- What number or percentage of service targets is being met? What is the number and severity of service breaches?
- Are service breaches being following up effectively?
- Are service level achievements improving
- Are customer perception statistics improving?
- Are IT costs decreasing for services with stable (acceptable but not improving) service level achievements?

{Hint - Don't fall into the trap of using percentages as your only metric. You will get caught out when you have a small system with limited measurement points (ie a single failure on a population of 100 is only 1%, a single failure on a population of 50 is 2% - if your target is 98.5 % then you have already breached the SLA. One should always go for a number of incidents rather than a percentage on populations of less than 100 and be careful of the targets you accept. - This is something some organisations have learned the hard way!}

Annex A - Example of a simple Service Catalogue

Customer / Service	Accounts	Sales	Marketing	Legal	Production	Retail	Warehouse	Transport	Design
Payroll System	✔			✔					
Accounts System	✔	✔	✔	✔		✔			
Invoicing	✔	✔				✔			
Customer D/Base	✔	✔	✔	✔		✔	✔	✔	
Sales D/Base	✔	✔	✔			✔			
Stock Control					✔		✔	✔	
Legal System				✔					
Factory Production					✔		✔		✔
Suppliers D/Base	✔	✔	✔	✔	✔	✔	✔	✔	
Ordering	✔	✔			✔	✔	✔	✔	
Logistics					✔		✔	✔	
Postal Addresses	✔	✔	✔	✔	✔	✔	✔	✔	
CAD/CAM					✔				✔
Intranet	✔	✔	✔	✔	✔	✔	✔	✔	✔
Internet	✔	✔	✔	✔					✔
Routemaster		✔					✔	✔	
Office Suite	✔	✔	✔	✔	✔	✔	✔	✔	✔
Email	✔	✔	✔	✔	✔	✔	✔	✔	✔

Annex B - Example SLAM Chart



Annex C - Service Level Manager - Role, Responsibilities & Key Skills

The Service Level Management Process must be "owned" in order to be effective and achieve successfully the benefits of implementation. This is not meant to imply that this should be a single post, unless that is appropriate within your organisation. However if the role is split, care should be taken to avoid giving tasks to people with "fire-fighting" type roles as they will not have the time to pro-actively manage the majority of the required tasks. The skills list given is an indication of the types of skills which would be present in an "ideal candidate in an ideal world". As long as most are present then you have potentially found a good candidate for the post.

Role

To implement and maintain the Service Level Management process to the level required by the parent organisation.

Role Positioning

The role must be of an appropriate Level to negotiate with Clients on behalf of the organisation, and to initiate and follow through actions required to improve or maintain agreed Service Levels. This will require either adequate seniority within the organisation and clearly visible management support.

Responsibilities

- Create and maintain a catalogue of existing services offered by the organisation.
- Formulates, agrees and maintains an appropriate Service Level Management structure for the organisation, to include;
  - Service Level Agreement structure (e.g. Service based, Customer based or Multi - Level);
  - Operational Level Agreements within the IT Provider;
  - Third Party Supplier / Contract Management relationships to the Service Level Management Process;
  - Accommodate any existing Service Improvement Plans / Programmes within the SLM process.
- Negotiates, agrees and maintains the Service Level Agreements with the Customer.
- Negotiates, agrees and maintains the Operational Level Agreements with the IT Provider.
- Negotiates and agrees with both the Customer and IT Provider any Service Level Requirements for any proposed new / developing services.
- Analyses and reviews Service Performance against the SLAs and OLAs.
- Produces regular reports on Service Performance and Achievement to the Customer and IT Provider at an appropriate level.
- Organises and maintains the regular Service Level review process with both the IT Customer and IT Provider which covers ;
  - Review outstanding actions from previous Reviews;
  - Current Performance;
  - Reviews Service Levels and targets (where necessary);
  - Reviews underpinning agreements and OLAs as necessary
  - Agree appropriate actions to maintain / improve Service Levels.

- Initiates any actions required to maintain or improve Service Levels
- Conduct Annual (as appropriate) reviews of the entire Service Level Process and negotiate, agree and control any amendments necessary.
- Acts as co-ordination point for any temporary changes to Service Levels required (i.e. extra support hours required by the Customer, reduced Levels of Service over a period of maintenance required by the IT provider etc.)

Skill Set (in no particular order)

- Have a good understanding of the IT Providers services and qualifying factors in order to understand how customer requirements will affect delivery.
- Have an understanding of the Customers end business and how IT can affect the delivery of that product or service.
- Excellent communication and negotiation skills.
- Have patience, tolerance and resilience.
- Knowledge and experience of Contract and/or Supplier Management type roles.
- Good People Management and Administrative skills.
- Good understanding of statistical and analytical principles and processes.
- Good presentational skills.
- Reasonable numeric skills.
- Ability to interact successfully with all levels of the Customer and IT Provider organisation.
- Reasonable technical understanding and an ability to translate technical requirements and specifications into easily understood business concepts and vice versa.
- Be innovative in respect to Service Quality and ways in which it can be improved within the bounds of the organisations limits (resource, budgetary, legal etc.)
- Be a good listener and have the ability to apply the knowledge gained effectively.
- Be seen to be even-handed and fair in their dealings with other parties.

6 FINANCIAL MANAGEMENT FOR IT SERVICES

6.1 Introduction

6.1.1 Why Introduce Formal Financial Management for IT Services ?

IT services are usually viewed as critical to the business or enterprise. The increases in user numbers, demands for new technologies and complexities of client-server systems has frequently caused IT Services costs to grow faster than other costs. As a result, enterprises are often unable or unwilling to justify expenditure to improve services, or develop new ones, and IT services may become viewed as high-cost or inflexible.

Due to the complex nature of IT Accounting for IT usage, it is rare that the actual running costs of the IT services are properly identified and this often leads to dissatisfaction with the perceived 'value for money' of the services.
"Why can't the IT department provide a better level of service ? "
"Why does the IT department budget have to be so large ?"
"How much will it cost to implement and run this new system"
"Why do we spend so much time performing redundant tasks, like reprinting large reports that are not read ?"

The above are examples of the questions asked inside and outside an IT department, often in emotive situations, such as project over-runs or during periods of loss of critical service. The answer is often :
"We're doing the best that we can with the money that we have "

but … are we ?

To understand whether we are doing the best that we can and to demonstrate this to our customers, we have to both understand the true cost of providing a service and manage those costs professionally. To do this, it is usual to implement IT Accounting and Budgeting processes & often to implement Charging processes as well.

6.1.2 Basic Concepts of Financial Management for IT Services

Financial Management is the sound stewardship of the monetary resources of the enterprise. It supports the enterprise in planning and executing its business objectives and requires consistent application throughout the enterprise to achieve maximum efficiency and minimum conflict.

Within an IT organisation it is visible in three main processes : Budgeting, IT Accounting & Charging

> **Budgeting** is the process of predicting and controlling the spending of money within the enterprise and consists of a periodic negotiation cycle to set budgets (usually annual) and the day-to-day monitoring of the current budgets.
>
> **IT Accounting** is the set of processes that enable the IT organisation to fully account for the way its money is spent (particularly the ability to identify costs by customer, by service, by activity). It usually involves ledgers and should be overseen by someone trained in Accountancy.

Charging is the set of processes required to bill a customer for the services supplied to them. To achieve this requires sound IT Accounting, to a level of detail determined by the requirements of the analysis, billing and reporting processes.

The aim of Budgeting is that the predicted costs match the budget; this budget is usually set by negotiations with the customers who are providing the funds (although this sometimes happens at a very gross level i.e. the business leaders agree proportions of their revenue which will be used to fund IT, based upon what IT have told them their costs will be).

Good Budgeting is essential to ensure that the money does not run-out before the period end. Where shortfalls are likely to occur the enterprise needs early warning and accurate information to enable good decisions to best manage the situation.

Organisations which need to account and charge to a very high level of accuracy, e.g. commercial IT service providers, will need to invest much more effort in developing IT Accounting & Charging systems than those who seek only a fair, simple apportionment of costs back to business units.

IT Accounting can be used to determine the exact costs of resource usage down to CPU, filestore and bandwidth but it is rarely advisable to use this as the basis for Charging as the costs of so doing may outweigh the benefits. It is in the interest of all parties to keep the overall cost of service low and the bureaucracy to a minimum, even at the expense of perfect precision.

Current leading practice is to use IT Accounting to aid investment and renewal decisions and to identify inefficiencies or poor value but to charge a fixed amount for an agreed capacity (determined by the level of service agreed in the Service Level Agreements or SLAs).

In this case, the IT Finance Manager works with the Service Level Manager (they may even be the same person) to ensure that the overall costs of running the agreed services should not exceed the predicted costs. Charging is then often a matter of billing for agreed periods at an agreed rate, for example $^{1}/_{12}$ of each customer's IT budget each calendar month. Additional charges are made for work above the agreed Service Levels (e.g. Office moves, major roll-out, unplanned hardware upgrade).

A commercial organisation, such as a supplier of outsourced services, is likely to have to develop more precise methods of charging and display a greater flexibility in linking charges to costs incurred, than an in-house organisation, in order that the requirements of commercial marketing and profit-making can be supported.

This chapter looks at methods for an IT department to predict and calculate the costs of service and discusses ways of estimating the proportion of costs that can be attributed to each customer where an IT service is shared. The following simple diagram, *Figure 6-1 IT Accounting, Charging & Budgeting cycle*, is used as a basis for the whole chapter.

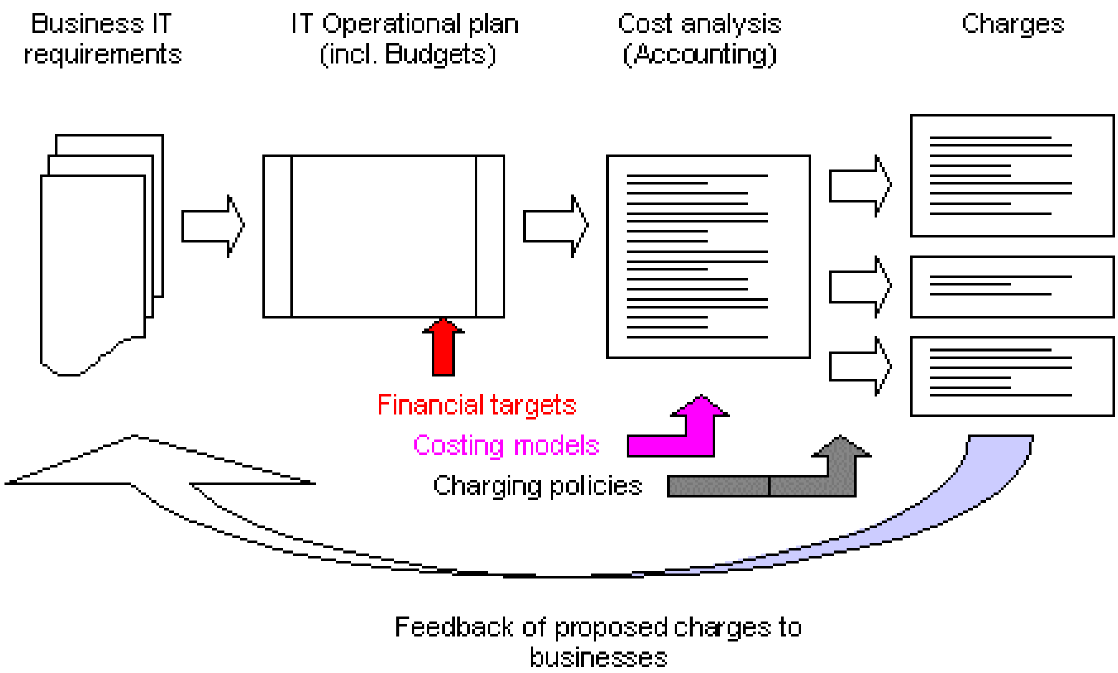


Figure 6-1 IT Accounting, Charging & Budgeting cycle

In summary :

Budgeting enables an organisation to

- predict the money required to run IT services for a given period
- ensure that actual spend can be compared with predicted spend at any point
- reduce the risk of overspending
- ensure that revenues will be available to cover predicted spend (where Charging is in place)

IT Accounting enables an organisation to

- account for the money spent in providing IT Services
- calculate the cost of providing IT services to both internal and external customers
- perform cost-benefit or Return-on-Investment analyses
- identify the cost of changes

Charging enables an organisation to

- recover the costs of the IT services from the customers of the service
- operate the IT organisation as a business unit if required
- influence user/customer behaviour (note the discussion in 6.4.2 Charging Policies)

6.1.3 Scope of IT Financial Management

The scope of IT Financial Management includes Budgeting, IT Accounting and Charging, although the responsibility for the processes and tasks may lie with the Finance department. In many enterprises the budget rules are set for all parts of the enterprise and the monitoring and reporting of budgets is performed by staff who report to the Finance department rather than to the IT department.

For the purpose of this chapter, it is assumed that Budgeting, IT Accounting and Charging for IT Services will be the responsibility of IT Services management. In some enterprises the design and implementation may be the responsibility of a Finance Division or shared with them. Even if the IT Services management assume total responsibility for the function, it is advisable to work closely with the Finance department and with qualified accountants.

6.1.4 Mission for Financial Management for IT Services

For an in-house organisation, the mission should be :

> “to provide cost-effective stewardship of the IT assets and resources used in providing IT services”.

In a commercial environment, there may be a mission statement that reflects the profit-making and marketing aims of the organisation.

The goals for any IT Services organisation should include :

> ”to be able to account fully for the spend on IT services and to attribute these costs to the services delivered to the organisation’s customers”

> “to assist management decisions on IT investment by providing detailed business cases for changes to IT services”

6.1.5 Relationship with other IT Service Disciplines

Financial Management for IT Services interacts with most IT Service disciplines and will have particular dependencies upon and responsibilities to:

Service Level Management - The SLA specifies customer expectations and IT services obligations. The cost of meeting the customer’s requirements may have major impact on the shape and scope of the services that are eventually agreed. The IT Finance Manager will liase with a Service Level Manager about the costs of meeting current and new business demands, the Charging policies for the organisation, their effects on customers and how the policies are likely to influence customer / user behaviour. The more that the SLA allows individual customers to request variations to service levels, the greater is the scope for (and potential benefits of) Charging for IT services but also the greater the overheads of Budgeting, IT Accounting & Charging.

Capacity Management – Cost information can be used to estimate the costs of the desired capacity and availability of the system. In planning the capacity it may be necessary to discuss the costs with individual customers or the enterprise as a whole. Data that is collected so that costs can be determined may also be relevant to capacity assessments, e.g. staff effort, machine usage.

Configuration Management –Financial Management requires asset and cost information that may be managed by enterprise-wide systems, such as those based on SAP, BAAN or ORACLE. Configuration Management is responsible for managing the data relating to assets (Configuration Items) and their attributes (e.g. cost).

6.1.6 Impact on the Enterprise

Changes to Budgeting and IT Accounting or the introduction of Charging for IT services are strategic business decisions. They may impact service levels, perceptions of value and usage of services. They also require an investment in planning and maintaining the processes. Business leaders throughout the enterprise should be fully aware of the changes likely from the implementation of changes to any or all of the above.

It is essential that the enterprise recognise the cost of introducing and maintaining Budgeting, IT Accounting or Charging as well as the benefits. The proposed benefits must be clear to the IT organisation and business users. Evaluation of these costs and benefits prior to implementing new systems is essential if the systems are to be quickly accepted by the users and IT services staff.

Charges for IT services must be simple, fair and accurate and this requires accurate, effective IT Accounting. The enterprise must also be clear on its overall policies on Charging e.g. whether to break-even, to subsidise or to make profits. It is essential that the enterprise is fully aware of the benefits and the pitfalls of the proposed system of Charging.

It is unlikely that IT Accounting can be introduced solely for IT services – the whole enterprise must be prepared to account in the same way for monies spent. If Charging is introduced in an enterprise where no other form of inter-departmental charges are levied, anomalies may have to be addressed when, for instance, IT ‘charges’ Personnel for running the Personnel database but Personnel cannot charge the IT department for their services.

6.1.7 Benefits of Financial Management for IT Services

Throughout this chapter, the term customer is used to refer to the department or division ‘buying’ the service. The ‘user’ is the person who makes day-to-day use of the service, e.g. a Salesperson or Customer Service Representative. Most of the benefits discussed are benefits to the enterprise as a whole, or to the customers of the IT department. The benefits to users are realised through improved service, arising from efficient use of IT spend. Financial Management can be seen as the brace that ‘locks’ IT to the business, preventing the IT organisation from drifting away from the needs of the business and preventing businesses from pursuing private deals outside the enterprise.

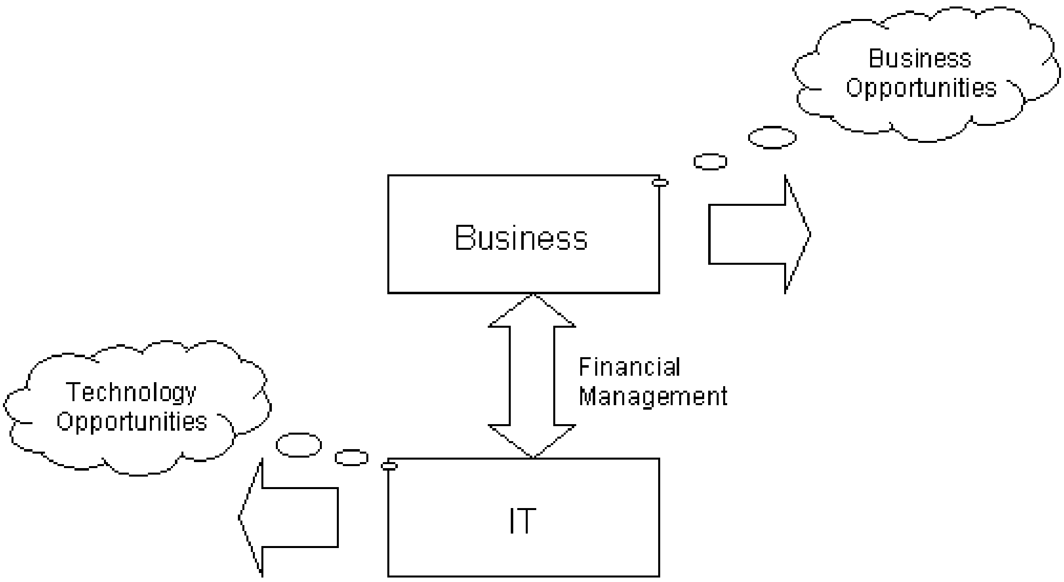


Figure 6-2 Locking IT to the business

The benefits of Budgeting, IT Accounting & Charging for IT services are discussed fully in the sections on each topic; in summary they are :

- increased confidence in setting and managing budgets.
- accurate cost information to support IT investment decisions
- accurate cost information for determining cost of ownership for ongoing services
- a more efficient use of IT resource throughout the enterprise
- increased professionalism of staff within the IT organisation

Budgeting

The benefits of Budgeting should be self-evident, but in summary are :

- Ensuring that the business provides sufficient funds to run the IT Services it requires
- Ensuring that IT Service Levels can be maintained throughout the year
- Providing early warning of under- or over-consumption of service (provided that some form of IT Accounting is in place)

Accounting for IT Services

The fundamental benefit of Accounting for IT Services (IT Accounting) is that it provides management information on the costs of providing IT services that support the enterprise's business needs. This information is needed to enable IT and business managers to make decisions that ensure the IT Services section runs in a cost effective manner.

Cost effectiveness is defined here as ensuring that there is a proper balance between the quality of service on the one side and expenditure on the other. Any investment that increases the costs of providing IT services should always result in enhancement to service quality or quantity.

IT Accounting helps the business to:

- base decisions about the services to be provided on assessments of cost-effectiveness, service by service
- make more business-like decisions about IT services and investments in them
- provide information to justify IT expenditures
- plan and budget with confidence
- demonstrate under- or over-consumption of service in financial terms
- understand the costs of **not** taking advantage of opportunities for change

Put simply, there is no prospect of IT service providers maximising value for money if the costs of providing the services are not accurately known. A key justification for investing in more IT resources is to support new or better business processes. IT Accounting provides the cost basis for cost-benefit analyses.

Charging

The fundamental benefit to the enterprise of charging customers, is that it provides a sound business method of balancing the shape and quantity of IT services with the needs and resources of the customers. Customers are charged for the services they receive and because they are paying, they have a right to influence decisions on its provision. If they do not think the services represent good value for money, they may stop using them or make formal complaints but professional IT departments will invest time in discussing the balance of charges and service levels with their customers.

Services can be improved by spending more, if there is a business justification for it. The introduction of formal Charging, often provides more evidence to support this and hence more enterprises will choose to invest in IT. Conversely, if customers believe that they can save themselves money (directly or indirectly, by reducing overall enterprise expenditure) by changing the way in which they use the IT services, they will be able to discuss this more openly with the IT department.

Charging enables the IT Services management to:

- make formal evaluations of IT services and plan for investment based on cost recovery and business benefits
- recover IT costs in a fair manner
- influence customer behaviour (where appropriate – see *6.4.2 Charging Policies*).

Notional Charging, where bills are produced but no money changes hands, is sometimes introduced to ensure that customers are aware of the costs they incur. The effectiveness of introducing Notional Charging depends on the supporting management processes: if customers ignore the information and management takes no action, there is little point in providing the information.

The introduction of real Charging (i.e. actual money changing hands), as opposed to Notional Charging, is not always necessary. Customers and enterprises that see the calculated costs may improve service provision and decision making without transferring money within the enterprise. Further, the cost of introducing Charging should be justified by better value for money for customers but overheads and system constraints may mean that the expected savings cannot be realised. For instance, it may not be possible to provide a higher quality of service to an individual customer even if that customer is prepared to pay a premium for it, nor to provide less service to a customer who expects a discount.

Real Charging is therefore desirable in principle but its introduction must actually improve effectiveness of IT spend and value for money and do so to an extent that savings outweigh the administrative costs. The level of detail available in Charging will be dependent upon the level of detail of the IT Accounting and usage information and may require a redesign of the IT Accounting systems if detailed Charging is required.

6.1.8 Costs

The costs associated with Budgeting, IT Accounting and Charging fall into 3 broad categories:

1) the administration and organisation costs for the planning, implementation, ongoing operations and management of the function (both staff directly working in these areas and Operations or other staff involved in the data collection).
2) the extra computing resources needed to automate and facilitate IT Accounting and Charging
3) the purchase and support of tools required in carrying out the processes (some of these tools are also required for other IT Infrastructure Library functions).

Once costs are visible, and particularly when real Charging is in place, the demand for some Services may fall. This results in reduced revenue but is not really a cost of implementation, as it is in the enterprise's interest to identify and reduce inefficient use of IT resource.

6.1.9 Possible problems

There are a number of possible problems in implementing IT Accounting and Charging :

- IT Accounting and Charging are often new disciplines in IT Services and there is limited understanding of leading practice in Cost Modelling and Charging mechanisms which could lead to over-complex or ineffective systems
- IT Accounting relies on planning information provided by other functions both within and outside of IT Services management which may not be routinely available, delaying the project
- staff combining accountancy and IT experience are rare, so many activities may need to be shared with staff from outside IT Services who may not have this as their priority
- enterprises' IS strategies and objectives may not be well formulated and documented and prediction of capacity requirement not accurate
- senior business managers may not recognise the benefits of IT Accounting and Charging and may resent the administrative overheads and the limitations on workload
- the IT organisation may not be able to respond to changes in users demands once costs become an influence

- the IT Accounting and Charging processes are so elaborate that the cost of the system exceeds the value of the information produced.
- The monitoring tools providing resource usage information are inaccurate, irrelevant or cost too much to develop and maintain.

The guidance contained in this chapter is intended to enable an IT Finance Manager to better manage the risks of these problems.

6.1.10 Accountancy

The chapter assumes an IT organisation serving a single enterprise or related enterprises in the same country. It does not cover taxation or accounting legislation. Guidance on IT Accounting practice should be sought from other publications or from qualified accountants.

This chapter specifically refers to Budgeting, IT Accounting and Charging in an IT organisation but the principles and advice apply to all service provision (e.g. software development, in-house consultancy, procurement, direct works departments and so on) and to the supply of services in a commercial environment i.e. outsourcing and shared service centres.

6.2 Budgeting

6.2.1 Introduction

Budgeting is the process of ensuring that the correct monies are set aside for the provision of IT services and that during the budget period they are not over-spent.

The Budgeting process is a key influence on strategic and tactical plans. It is also the means of delegating control and monitoring performance against predefined targets. It is paramount that budgets are effectively integrated within the enterprise and that managerial responsibility and accountability is matched and communicated in an efficient way.

As all spend affects profitability, it must be recognised that decisions about investment in IT Services and the integrated management IT Accounting function can help provide the competitive edge necessary for survival of an enterprise.

All enterprises have an annual round of negotiations between the business departments and the IT department covering expenditure plans and agreed investment programmes which ultimately sets the budget for IT.

These are closely linked to reviews with the businesses (individually or collectively) that cover :

- current projects and service levels
- a review of the last 12 months
- plans for the next 1 - 3 years.

The final budget agreed for an IT department may include financial disciplines imposed by the enterprise, including :

- limits on capital expenditure (see definition6.3.3 Building the Cost Model)
- limits on operational expenditure (see definition in 6.3.3 Building the Cost Model)
- limits on variance at any point in time, between actual and predicted spend
- guidelines on how the budget must be used
- an agreed workload and set of services to be delivered.
- limits on expenditure outside the enterprise or group of enterprises

Further, the business departments which provide the revenue to the enterprise, from which departmental budgets are drawn, may themselves have rigid limits on the way in which they pay for services – they may not be able to fund a change in service or service level mid-year despite IT being able to provide it.

An example, very simplified budget is shown in *Table 6-1 Example budget calculation*.

Budget item	Capital	Purchase cost	Operational cost	Spend this year	Budget next year	notes	Annualised cost
Hardware							
UNIX Server	Yes	£80,000	£8,000	£8,000	£8,000	no changes	£34,667
NT Server	Yes	£10,000	£1,000	£1,000	£1,000	no changes	£4,333
Netware server	Yes	£3,000	£300	£300	£300	no changes	£1,300
PCs (50)	Yes	£60,000	£6,000	£6,000	£6,000	no changes	£26,000
Routers (5)	Yes	£3,000	£300	£300	£300	no changes	£1,300
LAN Cabling	Yes	£40,000	£4,000	£4,000	£4,000	no changes	£17,333
Software							
General Ledgers	No		£20,000	£20,000	£24,000		
ORACLE	No		£7,000	£7,000	£8,400		
Mktg & Sales	No		£3,000	£3,000	£3,600	staff increase	
MS Windows (50-user)	No		£2,500	£2,500	£3,000	from 50 to 60	
MS Office (20-user)	No		£3,000	£3,000	£3,600		
Netware	No		£3,000	£3,000	£3,600		
NT	No		£2,500	£2,500	£3,000		
Employment							
Manager	No		£50,000	£50,000	£52,000	4% payrise	
Senior Op	No		£30,000	£30,000	£30,000	just joined	
Operator	No		£20,000	£20,000	£21,000	5% payrise	
Contractor	No		*£100,000*			paid by Mktg	
Accommodation							
Computer room	No		£10,000	£10,000	£10,200	2% rise in	
Office	No		£10,000	£10,000	£10,200	charges	
Transfer							
Wide Area connection	No		£20,000	£20,000	£20,000	fixed price,	
DR contract	No		£10,000	£10,000	£10,000	3-yr contracts	
			Total	£210,600	£222,200		

Note : the annualised cost is taken as 1/3 of the purchase cost, plus the annual maintenance cost and will be used in the Cost Model in section 3

Table 6-1 Example budget calculation

6.2.2 Estimating the cost of budget items

The categorisations above, of Hardware, Software etc, are arbitrary but help ensure that all the budget items can be identified. Transfer items are those in which a service is bought, rather than a physical product. Other categorisations can be chosen: the test is that all budget items are identified.

The cost of some budget items may not be known at the time a budget is drawn up, e.g. overtime payments, contractor payments, consumables, external network charges. These have to be estimated, usually based upon a previous IT Accounting period.

Some costs may vary from the estimates, depending upon the usage. An example of this is software licences that may increase (in steps) as further users are introduced. Other costs may need to be estimated to cover out-of-hours support, major equipment re-location.

Capital items are those which are recorded in the assets of the organisation, this is discussed further in Paragraph 6.3.3 Building the Cost Model

IT Finance Managers must be cautious in estimating costs where they do not fully control them. For example, planning a reduction of 20% computer room usage costs by removing old disc drives may not result in 20% saving in costs, as the rental for the space may be fixed by the lease.

6.2.3 Estimating the cost of workload dependent budget items

Another reason for costs changing, is when the IT workload changes. For this reason, workload estimates and forecasts should be considered when drawing up budgets. Such estimates and forecasts are also required for the preparation of Service Level Agreements and for Capacity Management.

Estimates of workload volumes are normally obtained from historical data, and forecasts are made on the basis of updated information and revised plans. A fuller approach to workload estimating can be seen in the chapter on Capacity Management.

In the simplified example below, *Table 6-2 Estimating workload-dependent budget items*, a calculation is performed for the cost (for budgetary purposes) of the Wide Area Network (WAN) service.

	current year				budget year				
Factor	**quantity**	**unit cost**	**cost**		**quantity**	**unit cost**	**cost**		Note
Users	30	-			40	-			
64k dial-up lines	5	£200	£1,000		10	£200	£2,000		
64k dial-up telephony			£3,000				£6,000		demand per user doubles each year
ISDN lines	2	£1,000	£2,000		7	£1,000	£7,000		higher proportion of customers use ISDN
Routers required	1	£500	£500		3	£500	£1,500		1 Router handles up to 8 lines
Network budget			£6,500				£16,500		

Table 6-2 Estimating workload-dependent budget items

6.3 Developing the IT Accounting System

6.3.1 Scope of IT Accounting

The basic IT Accounting principles (why do it, what to do and whom it affects) are common to all business functions. To a large extent, the implementation of IT Accounting is similar throughout the enterprise but the detail of what to cost and how to cost it, is dependent upon the type of service being provided.

To implement IT Accounting may require improved IT Accounting in many areas, for example in staff time & activity recording, supplier contracts, software licensing, resource metering or accommodation costs.

Other items which rely upon the information provided by IT Accounting and hence may dictate the shape of a IT Accounting model include :

- Budgeting guidelines
- Charging policies
- Investment guidelines

IT Accounting can be very complex and if implemented at too high a level of detail, may cost more than the benefits realised. The IT Accounting systems described here should enable an enterprise to:

- track actual costs against budget
- support the development of a sound investment strategy which recognises and evaluates the options and flexibility available from modern technology
- provide cost targets for performance and service delivery
- facilitate prioritisation of resource usage
- make day-to-day decisions with full understanding of the cost implications and hence the minimum of risk
- support the introduction, if required, of Charging for IT service

6.3.2 Business Perspective

IT Accounting and Charging are integral parts of IT infrastructure management responsibilities. The policies must be agreed by the enterprise, i.e. at board level.

The difference between IT Accounting and Charging and the responsibility for each, has to be defined and made clear by the IT organisation. They are different but have linked sets of activities.

The most visible of the two functions, **Charging**, is concerned with the recovery of the cost of IT services expenditures in a simple, fair, affordable way. **IT Accounting** is concerned with providing detailed information on where and for what reason expenditure is incurred within IT Services and is inward-looking.

The management of the IT organisation can choose to implement IT Accounting for IT Services (IT Accounting) with no Charging, or to charge for IT Services (either to break-even or to make profits). If IT Services have to make a book profit, the enterprise may even create IT Services as a separate legal entity.

IT Accounting only The benefit of this policy is that sound IT Accounting focuses awareness on costs and enables investment decisions to be better founded, without the overheads of billing and book-keeping but it is less likely to shape users behaviour and does not give the IT organisation the full ability to choose how to financially manage itself, for example in funding IT investment.

Zero-profit Zero-profit organisations are designed to fully account for all IT spend and recover it from the users. These accounts will include both cash and non-cash costs that, in effect, identify the full economic cost of running the business. The benefits of IT Accounting (before Charging is considered) include improved cost control over service provision, recognition of true costs by customers and consistency in approach by different organisations.

Profit Centre A Profit Centre is a method of managing the IT organisation in which it has sufficient autonomy to operate as a separate business entity but with business objectives set by the enterprise. A Profit Centre can be created with the business objective of making a profit, breaking even or operating with a subsidy. The key characteristics are that:

- deliverables or products are clearly identified and sold into a marketplace
- each product or service carries a price tag.

The IT Services section must be assumed to pay for its own upkeep, probably with some financial support for capital expenditures. Choosing to operate as a Profit Centre is usually the first step along the path to a truly commercial IT Services section, one in which the customers (the people who pay for the services) have some freedom of choice to go elsewhere if they are not satisfied with the quality and/or price. The Profit Centre approach provides the IT Services section with a certain amount of autonomy, but it also carries risks, such as the customer becoming aware that the IT organisation is ‘making a profit’ from them. This can be more serious if another service provider can offer a cheaper service of similar standard (corporate loyalty disappears very quickly in this situation).

6.3.3 Building the Cost Model

To calculate the costs of IT service provision, it is necessary to design a framework in which all known costs can be recorded and allocated to specific customers, activities or other category. This is called a **Cost Model**. Most Cost Models are based on calculating the cost for each customer but other models can be developed to show the cost for each service or the costs for each location. In this chapter, we concentrate on a Cost Model that will enable the calculation of Costs-by-customer, which is the usual start-point if a Charging system is to be introduced.

This Costs-by-customer Cost Model requires that we first identify all major costs elements in the current or proposed IT budget and then attribute them to the customers who ‘cause’ them. To do this we have to first identify costs as either Direct or Indirect.

Direct Costs are those clearly attributable to a single customer, e.g. Manufacturing systems used only by the Manufacturing division.

Indirect Costs (sometimes called overheads) are those incurred on behalf of all customers e.g. the network or the technical support department, which have to be apportioned to all customers in a fair manner.

Any Indirect Costs, which cannot be apportioned to a set of customers (sometimes called **Unabsorbed Overheads**), have then to be recovered from all customers in as fair way a way as is possible, usually by uplifting the costs calculated so far by a set amount. This ensures that the sum of all of the costs attributed to each customer still equals the total costs incurred by the IT organisation – in *Table 6-4 Spreadsheet example of full Costs-by-customer*, this is referred to as the ‘balance check’. This ‘balance check’ can be applied to costs divided in other ways e.g. by service or by location; always, the sum of the parts should equal the whole.

If the Cost Model is being produced for the first time the categories and Cost Elements for it, will have to be developed first, to a level of detail that meets the needs of IT Accounting and of any Charging to be performed. Hence an understanding of Charging policies (see *6.4.2 Charging Policies*) is necessary when the Cost Model is drawn up.

If costs are mainly Direct, perhaps because each customer has independent hardware and software, the method of recording and of apportioning costs can be very simple. For example, if Finance are the only users of the General Ledgers and the system on which it runs, all costs directly associated with the General Ledgers, including purchase, maintenance and support, can be attributed to Finance department’s code in the ledgers (often called a cost-centre or charge-code).

However, if resources are shared, for instance a mainframe running applications for more than one customer, the hardware costs may have to be classified as indirect and apportioned to each customer, say by CPU-seconds from workload predictions. To do this requires a model that allows these costs to be spread across a number of customers.

In the example in *Figure 6-3 Cost Model of Costs-by Customer*, it is assumed that there are 3 businesses or departments, who together are responsible for all of the IT costs. The three departments are Marketing & Sales, Manufacturing & Finance and all of the IT systems and services have been implemented on their behalf.

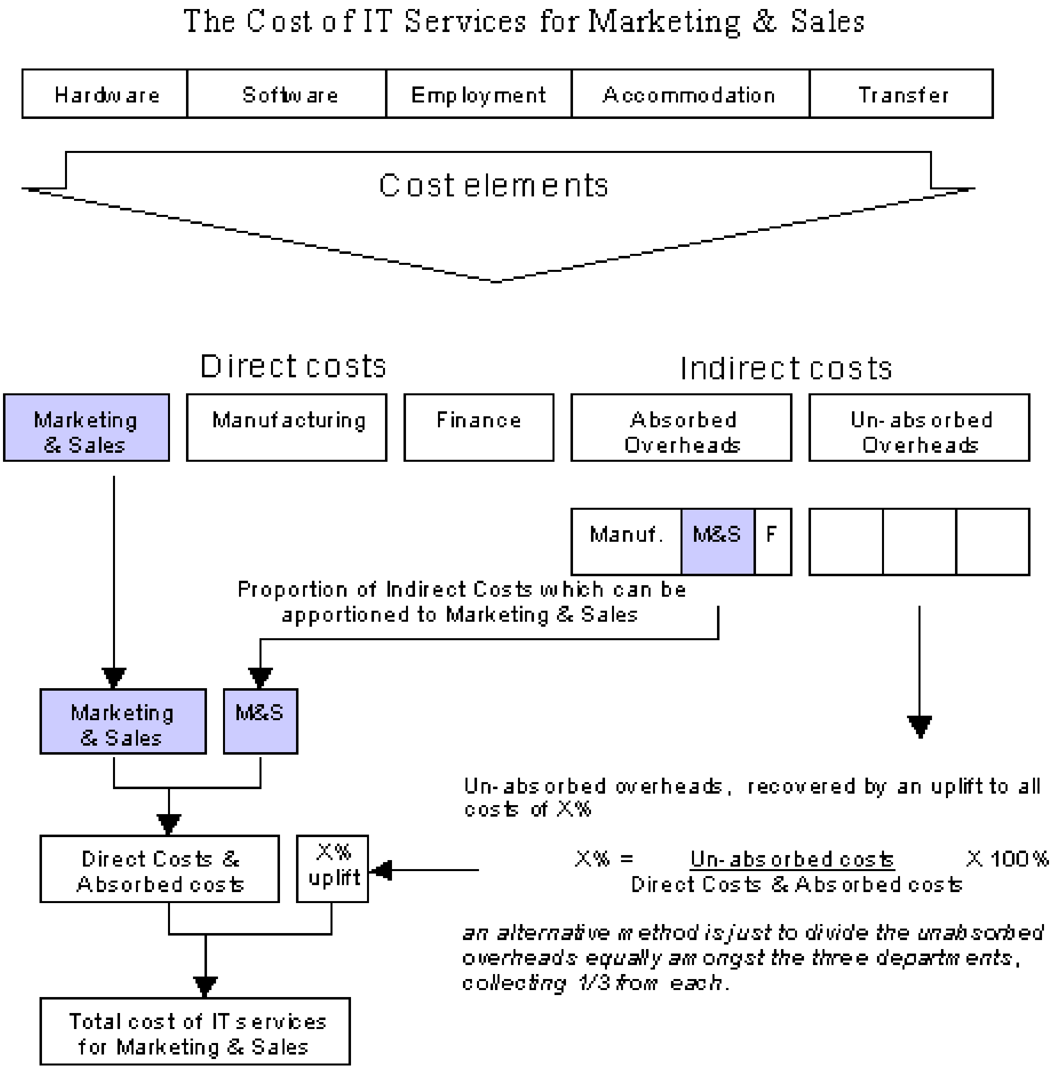


Figure 6-3 Cost Model of Costs-by Customer

An example of the calculation of Cost Model for a simple Cost-by-customer is shown in *Figure 6-3 Cost Model of Costs-by Customer*. The same principles can be applied to calculating the costs of individual application services or even parts of a service e.g. support and maintenance.

If the cost of providing some element of a service is desired, for instance producing large reports for the Marketing & Sales department, this may require measurement of resource usage to apportion indirect costs to this one activity e.g. paper used, computer time consumed, printing & operations staff and facilities. This can become very complex and would normally be treated as a separate exercise rather than as part of the standard Cost Model used for calculating Costs-by-customer.

To be able to derive cost information and report it in the formats required by different parts of the enterprise, it is necessary to ensure that all costs recorded are classified to a standard system with a level of detail that anticipates future changes, e.g. new cost centres, new equipment types, new project codes.

Cost Elements

If more detail is required in calculating costs, the chosen major Cost Types of Hardware, Software, Employment, Accommodation and Transfer can be further divided. For instance, Hardware might be divided into *Office, Network, and Central Servers*. The purpose of this is to ensure that every cost identified in the IT department can be placed within a table of costs, by type. This enables analysis to be performed by type e.g. all *Network* costs.

The decision on whether to identify more detailed cost units will often depend upon whether more detail is required to apportion charges. In general, Cost Elements will be the same as budget line items where the purpose of the model is simple recovery of costs.

If a more detailed analysis of costs is required, e.g. for organisations providing shared services, then more detailed Cost Elements will have to be identified. Typical Cost Elements within each major Cost Type are :

Major type	**Cost Elements**
Hardware	Central processing units, disk storage, peripherals, wide area network, PCs, portables, local servers
Software	Operating systems & options, scheduling tools, applications, databases, personal productivity tools, monitoring tools, analysis packages
Employment	Payroll costs, Benefit Cars, Re-location costs, Expenses
Accommodation	Offices, Storage, Secure areas
Transfer	Consultancy, Security services, Disaster Recovery services, outsourcing services, utilities,

Table 6-3 Cost Element examples

For organisations providing services based upon central mainframes, the hardware costs may be the largest proportion but it is more common to see a rough balance amongst hardware, software and employment. Increasingly, the proportion of costs attributed to networking devices and network services is becoming more significant and may be identified as a separate Cost Type.

Organisations which purchase software products, rather than developing them, will find a higher proportion for costs categorised as Software. Organisations which use outsourcing services (such as offshore development or computing services) will see Transfer costs as the largest proportion of costs.

Classification of Cost Elements

When each Cost Element has been identified, it should be classified, as a minimum, by the following :

Capital or Operational Costs should be classified as either ***Capital Costs*** or ***Operational Costs*** *(also known as Current Expenditure or Revenue Expenditure).*

For financial purposes, costs are classified into either Capital or Operational when the financial ledgers are reported (the ‘books’). This is because Capital expenditure is assumed to increase the total value of the company, while Operational expenditure does not, although in practice the value of Capital expenditure decreases over time (depreciates).

This distinction affects IT Accounting because the Cost Model needs a method of calculating the annual cost of using a capital item (e.g. mainframe) to deliver IT Service. The annual costs must make allowance for the decreasing value of capital items (assets) and make for timely renewal of Capital items e.g. buildings, servers, applications. Usually, this is taken as the annual depreciation, from a method set by the Finance department (within the boundaries of the country's laws).

Capital Costs are typically those applying to the physical (substantial) assets of the organisation. Traditionally this was the accommodation and machinery necessary to produce the enterprise's product. Capital Costs are the purchase or major enhancement of fixed assets, for example computer equipment (building and plant and are often also referred to as 'one-off' costs. It is important to remember that it is not usually the actual cost of items purchased during the year that is included in the calculation of the cost of the services but the annualised depreciation for the year as shown in *Table 6-4 Spreadsheet example of full Costs-by-customer*.

Operational Costs are those resulting from the day-to-day running of the IT Services section, e.g. staff costs, hardware maintenance and electricity, and relate to repeating payments whose effects can be measured within a short timeframe, usually the less than the 12-month financial year.

The following list gives typical examples of the Cost Elements, classified into Capital expenditure and Operational expenditure (revenue) items :

Capital

- computer equipment
- building and plant
- software packages (owned, not licensed).

Operational

- staff costs
- maintenance of computer hardware and software
- suppliers and services rental fees for equipment
- software licence fees
- accommodation costs
- administration expenditures
- electricity, water, gas, rates
- disaster recovery
- consumables.

Enterprises that measure themselves primarily on cash flow or budget adherence may have specific rules about the categorisation of costs in the accounts that differ from those who measure Return on Capital Employed (ROCE – see 6.3.11 Investment Appraisal).

The organisation's accountants will explain the rules for identifying Capital items and this will depend upon a number of business decisions. Many organisations choose to identify major expenditure as Capital, whether there is a substantial asset or not, to reduce the impact on the current financial year of such expenditure and this is referred to as 'Capitalisation'. The most common item for this to be applied to is software, whether developed in-house or purchased.

The reason for this is that a business that is investing in a major software development, that will provide service for a number of years, does not want to show all of the costs in a single year (and so, potentially, an operating loss). The board will want the value of the company and its shares to reflect the investment made but adding the cost of the item to the assets of the company without adjusting cash flow in some way would also give a false picture. The agreed method is to show Capital and Operational expenditure separately but to apply rules of depreciation, described later in this section, to provide a balance. This system allows an enterprise to spread the cost of a major purchase over a number of years although, as with all systems, many additional rules (and laws) have to be written to prevent fraud or misleading of investors.

Direct or Indirect Some costs can be directly attributed to a single user or group of users and these are referred to as ***Direct Costs***. Examples of direct costs would be a server or application used exclusively by a single cost-centre. Usually, all direct costs are attributed to the customer incurring them but there are occasions when these might be subsidised, shared, deferred or ignored, such as for the use of a new system which will be subsequently 'rolled-out' throughout the enterprise.

Other costs, such as operations staff in a Data Centre cannot be easily attributed to the running of a specific customer's service and these are referred to as ***Indirect Costs*** or *shared costs*.
To fully attribute all costs requires some form of division of indirect costs by a fair method of re-apportioning. An example would be operations staff whose total cost could be apportioned to the businesses on the basis of the number of users in each business. Often this apportionment is not completely accurate but the costs to the enterprise of a more accurate calculation, is far too high. The system of apportioning has therefore to be affordable, clear, fair and in accordance with good IT Accounting practice.

Cost Centre It is usually necessary to be able to apportion Direct Costs to a specific business group, department or external customer. Most businesses allocate ***Cost Centres*** to units within the enterprise that relate directly to the general ledger system. The number and types of Cost Centres differ from installation to installation due to size and organisational structure. Occasionally, specific projects or initiatives may be allocated a separate Cost Centre to enable costs for them to be 'ring-fenced'.

Fixed or Variable Costs that do not vary even when resource usage varies are referred to as ***Fixed Costs***. Examples of this would be a maintenance contract for a server or a corporate software licence (within agreed user limits).

Variable Costs are those that vary with some factor, such as usage or time. They are likely to be used for Cost Elements which cannot be easily predicted and which it is to the benefit of both supplier and customer to determine the costs exactly, perhaps for variable charges to be applied. Examples of charges that might vary, because the underlying cost varies, are out-of-hours cover, major equipment re-location, and the production of additional quarterly reports.

A Cost Element such as filestore may be considered to be variable. If a customer requires an additional 10Gb we may be able to calculate that the cost of this is £1000 and hence the cost per is £100 / Gb. The danger of this approach is that there are often sharp changes in costs because they cannot be continuously scaled : the next disc drive may require another cabinet, an additional process run on the server may cause queuing problems resulting in all jobs taking longer to run.

It is sometimes necessary to view a cost as having a fixed element and a variable element, for example, using filestore at all requires disc controllers and bandwidth, causing a fixed cost. The variable cost of additional disc drives can then be calculated and added to the fixed portion. This level of detail is not usually needed in calculating Service costs but can be useful when evaluating competing technologies or services.

6.3.4 Depreciation

Depreciation is the measure of the wearing out, consumption or other reduction in the useful economic life of a fixed asset, whether from use, passage of time, or obsolescence through technological or market changes. Depreciation should be allocated so as to charge a fair proportion of cost or valuation of the asset to each IT Accounting period expected to benefit from its use. This point can be a delicate balance and many IT organisations face a difficulty in funding the replacement of apparently useful items that have no capital value (i.e. are fully depreciated).

An example is a PC which is cannot perform at the level required but has not yet been fully depreciated. The organisation may not be willing to replace it because it still is an asset with a value. Similarly, a PC may be perfectly functioning but has been fully depreciated and now has no value in the business accounts – strictly, it is no longer a cost in providing services but we are unlikely to recalculate costs to take account of this.

The assessment of depreciation, and its allocation to IT Accounting periods, involves the consideration of three factors:

- the current cost (or valuation) of the asset

- the length of the asset’s expected useful economic life to the business of the enterprise, having due regard to the incidence of obsolescence
- the estimated residual value of the asset at the end of its useful economic life in the business of the enterprise.

The useful economic life of an asset may be:

- pre-determined, as in the case of a lease
- dependent on its physical deterioration through use or passage of time
- reduced by economic or technological obsolescence.

The depreciation methods used should be the ones most appropriate having regard to the types of assets and their use in the business. The Finance department will give guidance in this. The most common methods of assessing depreciation are:

straight line method - where an equal amount is written-off the value of the asset each year. Usually a fixed percentage of purchase cost, this results in the item having zero Net Book Value after a preset number of years (although it may continue to be used).

reducing balance method - where a set percentage of the capital cost is written-off the Net Book Value each year. Often this is of the form 40% in the first year, 30% in the second year and 30% in the last year. The Net Book Value is the capital cost minus the depreciation written-off to date

by usage - where depreciation is written-off according to the extent of usage during a period. It is usual to estimate the total useful ’life’ of a device and to calculate the proportion of this that has been ‘used’ during the year. For example, a laser printer may be estimated to have a useful ‘life’ of 5,000,000 pages. If the average usage is 1,000,000 pages in a year, it can be depreciated by 20% in that year. Again, an anomaly arises if after 5 years it is still in use.

The finance department may require IT assets to be ‘written-off’ before the end of their useful life, increasing the apparent cost of services but facilitating a charging system that generates revenue for the early replacement of systems.

6.3.5 Apportioning the IT Services Costs

Consider a company with three departments who require IT services – Marketing & Sales, Manufacturing, Finance. Each is asked to contribute towards the IT budget, based upon the services they require. Each Cost Element in the IT budget has to be identified; classifying them by type (Hardware, Software etc.) helps ensure that all such costs are found. It must then be decided whether these are Direct Costs or Indirect Costs and how they will be allocated to customers (in this case, other departments of the company).

In this example, all the costs of providing the shared computer infrastructure – cables, servers, routers, software, have been grouped into a single Indirect Cost Element called 'infrastructure' for which a common apportionment method can be used; in this case, the number of users of the infrastructure. This simplifies the spreadsheet and enables a simple calculation of the cost of adding new users to the network.

Cost Elements	**Capital**	**Annualised cost** (see note 1)	**Direct**	**Apportionment Method**	**Customer** Marketing & Sales	Manuf-acturing	Finance
Hardware							
UNIX Server	Yes	£34,667	No	50/50 split	£17,333	£17,333	
NT Server	Yes	£4,333	Yes		£4,333		
Netware server	Yes	£1,300	No	Infrastructure			
PCs (50)	Yes	£26,000	No	by PC	£5,200	£19,240	£1,560
Routers (5)	Yes	£1,300	No	Infrastructure			
LAN Cabling	Yes	£17,333	No	Infrastructure			
Software (see note 2)							
General Ledgers	No	£20,000	Yes				£20,000
ORACLE	No	£7,000	Yes			£7,000	
Mktg & Sales appl.	Yes	£3,000	Yes		£3,000		
MS Windows (50-user)	No	£2,500	No	by PC	£500	£1,850	£150
MS Office (20-user)	No	£3,000	No	by licence	£1,500	£1,050	£450
Netware	No	£3,000	No	Infrastructure			
NT	No	£2,500	No	Infrastructure			
Employment (see note 3)							
Manager	No	£50,000	No	Unabsorbed overhead			
Technical Support	No	£30,000	No	Unabsorbed overhead			
Assistant	No	£20,000	No	Unabsorbed overhead			
Contractor (see note 4)	*No*	*£100,000*	*No*				
Accommodation (see note 5)							
Computer room	No	£10,000	No	Unabsorbed overhead			
Office	No	£10,000	No	Unabsorbed overhead			
Transfer							
Wide Area connection	No	£20,000	No	Infrastructure			
DR contract	No	£10,000	No	Unabsorbed overhead			
Total costs		**£275,933**					
Direct & Apportioned costs		£100,500			£31,867	£46,473	£22,160
Absorbed costs (Infrastructure)		£45,433		20% / 74% / 6%	£9,087	£33,621	£2,726
Unabsorbed costs		£130,000	*89.1%*	*uplift*	£36,482	£71,349	£22,169
		£275,933			**£77,435**	**£151,443**	**£47,055**
				balance check for the 3 customers above:		***£275,933***	

Notes

1 For capital items, this is 1/3 of the purchase price (the agreed depreciation) plus the annual maintenance cost

2 The M&S application will cost £100,000 to develop (one year of contractor), and the support contract is £3,000 annually

3 Includes NI, pension & other benefits, usually adding between 30% & 50% to salary

4 Contractor is employed to develop new system for Marketing & Sales and is funded directly by them

5 Accommodation costs set by Finance department

6 Marketing has 10 PCs with all software, Operations has 37 PCs but only 7 with Microsoft Office, Finance has 3 PCs with all software

7 Infrastructure costs will be added (absorbed) based upon numbers of PCs in each department, i.e. Marketing 20%, Operations 74%, Finance 6%

8 Unabsorbed overheads are added onto each cost centre by uplifting it by 89.1%, to ensure full recovery

Table 6-4 Spreadsheet example of full Costs-by-customer

Rather than trying to determine the actual usage of the infrastructure that all departments rely upon, it is simpler to group all infrastructure Cost Elements into one and determine a fair way of recovering those costs, e.g. by number of users in each department.

In practice, the Cost Elements in each of the major Cost Types may be groups of items. For instance the UNIX server Cost Element may consist of a number of items e.g. central processor, UPS, filestore, peripherals but these rarely need to be identified individually unless the Cost Model has to show costs to that level.

6.3.6 Level of detail required in cost calculations

Some organisations will wish to calculate the cost of delivering individual services or activities. To do this requires a more complex Cost Model, with Cost Elements apportioned by **Cost Units**. These Cost Units will usually be things that can be easily counted, such as staff numbers, software licences or things easily measured, such as filestore usage, CPU usage, network packets sent.

A higher level of detail is required to fairly apportion Cost Elements that are Indirect but very few apportionment methods can ever be completely accurate. For instance, if we attribute central server usage by CPU-seconds, the issues of filestore and filestore transfers rises. Then, who 'pays' for the queuing and 'paging' of virtual store ? Large servers are rarely dedicated to a single function and the expense of apportioning their use to each customer can be both divisive and prohibitive.

Similarly, it is difficult to apportion the time of support staff to individual directorates without detailed timesheets and the benefits of this level of detail may be less than the cost of obtaining it.

At all times, when performing a detailed cost analysis it is essential to ensure that the value provided by the answers is not outweighed by the costs of data collection & analysis.

For most business-driven Cost Models, apportionment will be firstly simple, secondly fair and thirdly accurate (if possible).

6.3.7 Variable Costs or Indirect Costs requiring apportionment

Cost Unit ***Cost Units*** are the basic items of resource for which customers will be held accountable i.e. provide the method of apportioning Indirect Costs or calculating the actual cost of Variable Costs.

Cost Units should be chosen to enable simple, fair apportionment. As central hardware costs have become a small proportion of costs in many organisations, the use of CPU-seconds as a Cost Unit has decreased. It is used as an example here because it is one of the simplest measures, easy to measure on mainframe systems, accurate and apparently fair.

Staff hours are another easily identifiable cost unit, as is floor space. In the example *Table 6-5 Calculating the cost of an individual service*, PCs and software licences are both Cost Units.

Cost per Cost Unit After the Cost Unit has been set, the cost of each Cost Unit must be determined. Following the example above of CPU-seconds, if an enterprise's IT organisation knows the total cost of the necessary capacity, the average unit cost of a CPU-second can be calculated and apportioned to each customer as it is used, in the same manner that an electricity company calculates the consumption of each of its customers in KWh. In a Cost Model using Cost Units for every Cost Element, the sum of [units consumed X the cost per unit] should be demonstrated to be equal to the total cost of the IT Services (see also the 'balance check' *Table 6-4 Spreadsheet example of full Costs-by-customer*).

6.3.8 Calculating the Cost-by-service

In *Figure 6-3 Cost Model of Costs-by Customer*, we considered the Cost Model for calculating costs attributable to a customer. If we wish to calculate the Cost-by-service, our Cost Model may require more detail, as discussed in *6.3.6 Level of detail required in cost calculations*.

The basic approach is similar.

First we must identify all those costs that can be directly attributed to the service we are analysing, for instance any dedicated hardware, software, staff or contracts.

Second, we must then decide how to apportion the Indirect Costs such as infrastructure

Third, we have to adjust our total to allow for 'hidden costs' or 'Unabsorbed overheads' such as IT management or buildings. This must be the same uplift figure calculated for the whole model, or used from the Costs-by-customer Cost Model, i.e. 89.1%.

It is best to consider the model as being one layer of a set of models, one for each service, so that we can be sure that if we were to add together the costs for each Service provided, we would again reach the total of the costs incurred by the IT organisation. Even if we are only interested in Service A, we must ensure that all costs are identified and attributed to the other services, as in *Figure 6-4 Complete view of the Cost-by-service*.

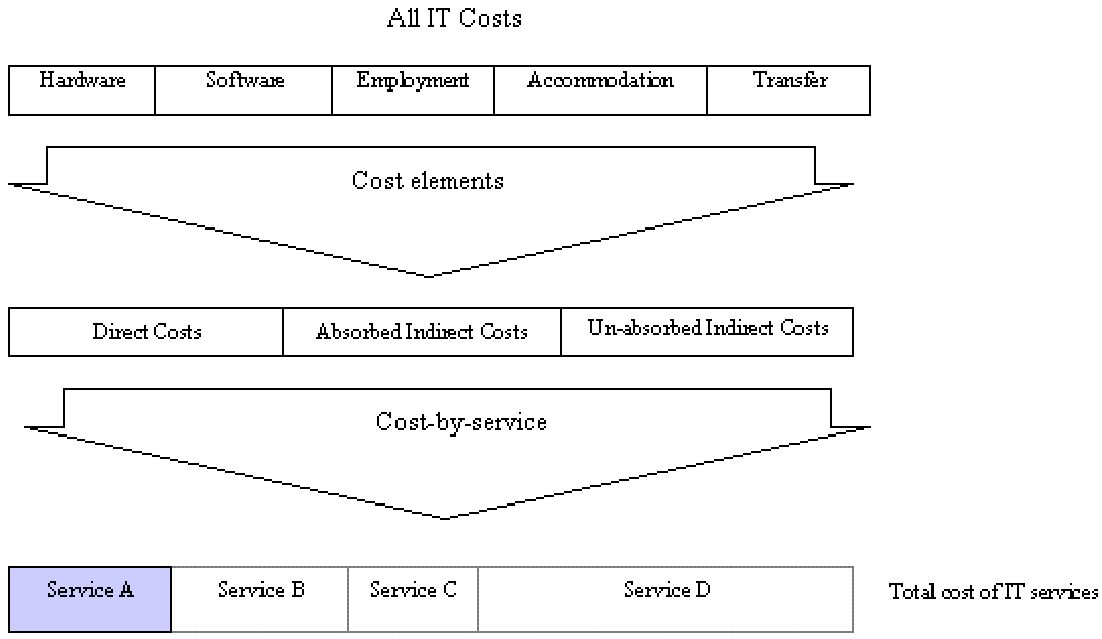


Figure 6-4 Complete view of the Cost-by-service

We can use our original spreadsheet in *Figure 6-3 Cost Model of Costs-by Customer* as a start-point. We then need to identify those costs that can fairly be attributed directly to the provision of this service and to apportion those Indirect Costs in a manner that reflects the proportion of them that result from the provision of this service.

We may not be able to identify Costs in every Category, for example there may be no Direct Accommodation costs and no Absorbed Accommodation or Staff costs – these may all be included in the Unabsorbed Costs which have to be included fairly in the costs of every service.

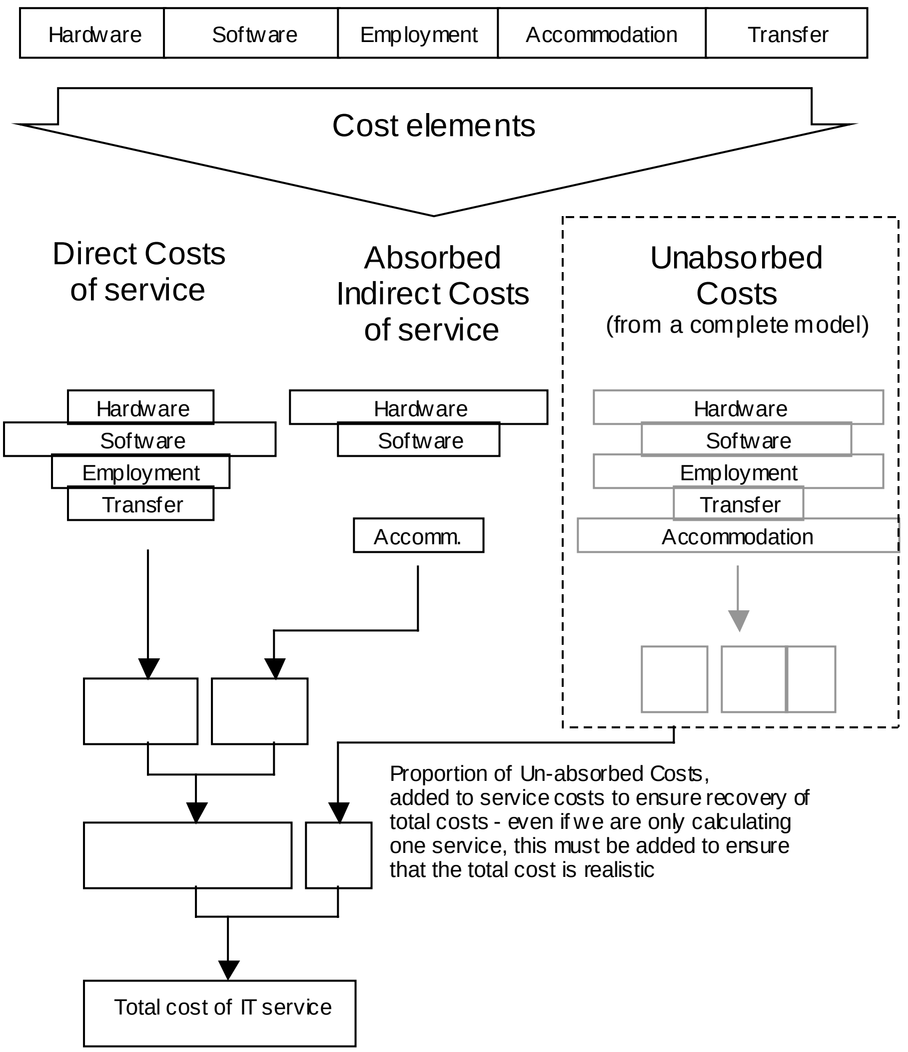


Figure 6-5 Cost Model for a Cost-by-service

The remaining complexity is how to account for the £100,000 development costs (one year of contractors time). Since this software is owned by the company (an asset), the Finance department will value it. If its net value is less than £100,000 after development is completed, there has been some misunderstanding when the business case was made for spending £100,000. If the software is to be used for the next 5 years, it could be depreciated over that period, representing £20,000 pa costs (excluding maintenance) and this cost attributed to the cost of the service to the Marketing & Sales department.

	Capital	Annualised cost	Direct	Cost Unit	Total Capacity	Cost per Cost Unit	Usage	Cost
Hardware								
UNIX Server	Yes	£34,667	No	average % CPU	100%	£34,667	37%	£12,827
NT Server	Yes	£4,333	Yes	n/a				£5,000
Netware server	Yes	£1,300	No	Infrastructure				
PCs (5)	Yes	£26,000	No	PC	50	£520	5	£2,600
Routers (5)	Yes	£1,300	No	Infrastructure				
LAN Cabling	Yes	£17,333	No	Infrastructure				
Software								
GL	No	£20,000	Yes	Users	5	£4,000	0	£0
ORACLE	No	£7,000	Yes	Users	5	£1,400	1	£1,400
Mktg & Sales appl.	No	£3,000	Yes	n/a				£3,000
MS Windows (50-user)	No	£2,500	No	PC	50	£50	5	£250
MS Office (20-user)	No	£3,000	No	licence	20	£150	0	£0
Netware	No	£3,000	No	Infrastructure				
NT	No	£2,500	No	Infrastructure				
Employment								
Manager	No	£50,000	No	Unabsorbed overhead				
Senior Op	No	£30,000	No	Unabsorbed overhead				
Operator	No	£20,000	No	Unabsorbed overhead				
Contractor (see note 1)								
Accommodation								
Computer room	No	£10,000	No	Unabsorbed overhead				
Office	No	£10,000	No	Unabsorbed overhead				
Transfer								
Wide Area connection	No	£20,000	No	Infrastructure				
DR contract	No	£10,000	No	Unabsorbed overhead				
Total costs		£275,933						
								£25,077
Apportioned costs								
Proportion of Absorbed Costs (Infrastructure)				Staff	50	£756	5	£3,780
								£28,857
			[2]	89.1 % uplift to recover Unabsorbed Costs				£25,711
								£83,425
				Application cost (annual)				**£20,000**
				Total annual cost				**£103,425**

Notes

1 The contractor cost is spread over 5 years life of the application, i.e. £20,000 pa

2 The uplift is the amount calculated from the overall Cost-by-customer model, discussed previously

Table 6-5 Calculating the cost of an individual service

When comparing the cost of internal provision of service against a bureau or packaged solution, we now have a comparative figure of £103,396. The most contentious issue this brings to light is the 89.1% uplift (£25,711) that is mainly caused by the unabsorbed overheads of Employment and Accommodation cost. If the customer feels that they could run the application themselves with the existing staff in their department administering it, they will argue strongly that the IT department is inefficient.

There is no immediate solution to this, as all analyses that break costs down to the minute, suffer from both not taking the whole picture into account (i.e. the provision of IT throughout the enterprise is most likely to be more efficient if centralised) and that single Cost Units cannot alone be used to calculate the cost of service provision but the bureaucracy involved in more detailed calculation may be prohibitive.

The exercise is valid, however, providing that these issues are understood and discussed. Ultimately, strong corporate support is required for central IT systems.

Note that if we ignore the 89% uplift, our service costs would not be a true reflection of the cost of service provision and could mislead decision-making.

6.3.9 Calculating the costs of Cost Units

When calculating how to apportion Indirect Costs or how to assess the actual cost of a Variable Cost, we need to select a Cost Unit that is a factor that directly and fairly represents the cost incurred. For example, using the PC as a Cost Unit for calculating the apportionment of infrastructure costs.

When we are calculating the cost of a Cost Unit such as usage of a central mainframe, we may find that the precise calculation results in an amount that is dependent upon when the item is used or upon how many customers use the resource. For instance, if we choose processor-second as a Cost Unit for determining the apportionment of costs across a number of users, we find that this varies with the total usage of the server. That is, the cost per Cost Unit varies unless the server runs at full capacity all of the time.

To prevent absurdities arising, such as attributing all the costs during overnight running of an unattended server to the sole user, it is usual to set fixed rates based upon the calculation of the capacity available in a fixed period. Hence a Standard Cost for a processor-second can be derived **but this must be regularly recalculated** to prevent anomalies from creeping in.

The following table demonstrates the calculation of some simple Costs of Cost Units but even this is beyond that which would normally be performed by an in-house IT organisation.

Cost Unit	Total available in 1 month	Total cost in one month	Cost per Cost Unit	Note
CPU-seconds	792,000	£100,000	£0.126 / CPU-sec	Based upon 10 hours x 22 days i.e. excluding out-of-hours
Filestore	4,000 Gb	£12,000	£3/Gb	Based upon allocated filestore (not actual used space) and annual cost of £36/Gb
Tapes	10,000	£2,500	£.25/tape	Based upon tapes allocated, £30 each with a life of 10 years, hence annual budget for new tapes is £30x1000
Operator hours	528	£30,000	£56/hour	3 operators, working, on average, 22 days x 8 hours

Table 6-6 Calculating the Cost of Cost Units

The cost of a given activity that requires 10,000 CPU-seconds of processing, 100Gb of filestore, 40 tapes and 3 operators for 10 hours can then be calculated by adding the costs of each Cost Element, for example :

Cost Unit	Quantity	cost per Cost Unit	
CPU-secs	10,000	x £.126	£1,260
Gb	100	x £3	£ 300
Tapes	40	x £.25	£ 10
Operator hours	30	x £56	£1,680
		Total	**£3,250**

Table 6-7 Calculating Activity Costs

We can further simplify this by using only the CPU-seconds as the factor and by diving the £3,250 by 10,000 we can base future cost calculations on £0.325 / CPU-second, ignoring the other factors This, of course, assumes that the other items are always used in the same ratios as the example above.

To ensure true costs, "Tapes" might have to include the cost of tape decks and maintenance, Operator hours may have to include Service Desk and Support staff. In current practice it is rare to perform these calculations 'from the bottom up, i.e. by first breaking down all resource usage to its component parts. Most businesses will be content with choosing a fair factor that is agreed to be the measure of usage and use this.

Even after calculating these detailed costs, many installations reduce all measures of usage to two Cost Units : **processor time** & **filestore occupancy**. By taking the total costs (including accommodation, hardware, staff etc) of storage and dividing this by the storage capacity, a Cost per Cost Unit can be determined. Similarly, the cost of providing the processor time (including accommodation, staff, DR etc.) can be divided by the cumulative processor-seconds of processing time to arrive at a cost for each CPU-second. Costs are then allocated to cost-centres solely on the basis of processor time and filestore occupancy.

6.3.10 Changes Affecting Costs

When systems change, underlying costs also change. To avoid having to recalculate Costs for Cost Units too frequently, the IT Finance Manager may choose to anticipate changes or to defer the inclusion of the cost in the Cost Model until stable usage has been established.

An example would be the addition of extra filestore. The purchase cost of the extra filestore can be easily determined but this may affect the calculated Cost per Cost Unit of filestore. This is because some overhead costs used in calculating the cost of filestore Cost Units may not change e.g. accommodation, employment cost, and some factors may increase the overall costs for filestore, such as the need for additional disc controllers or accommodation.

When performing detailed cost analysis, a cost recovery plans is needed to provide a picture of under (or over) calculation. The IT Finance Manager may choose to leave the Cost per Cost Unit unchanged until the error is greater than, say, 5% to avoid the trap of recalculating after every sneeze.

There are a number a methods employed to reduce the effect on costs of changes to equipment and services. They are included here for completeness as they were widely used in the past by mainframe-based organisations and computer bureaux.

Average cost

One common method of involves establishing only the average cost of machine time and using this average as the sole Cost Unit. To do this, the total costs for providing the machine during a given period are divided by the total number of hours it is used and the resultant cost per hour of service time is used to calculate costs incurred by each customer or usage.

The main advantage of the method is that all costs are attributed and a simple rate per hour of machine time is available for budgetary purposes.

The main disadvantages to a customer, however, are:

- the rate for machine time can vary from period to period, depending upon the number of other users

The main disadvantages to an IT organisation are:

- as everything is based upon machine time only, changes in other costs, such as staff are not properly recovered
- if usage decreases, revenues decrease immediately
- if usage increases (and capacity was available) an excess profit may be made

Standard Rate

Standard Rate is calculated by estimating the costs of the resources necessary to meet an estimated capacity, based upon the forecast workloads and service requirements of all the customers, and dividing this by a usage factor, such as number of customers files or system users. The resultant rate is then fixed for an agreed period.

The advantage of this approach is that a consistent rate is used for a whole period, allowing customers to predict costs simply.

The disadvantage is that the calculation differs from true cost if the workload changes, for instance a customer changes their business process or submits increased volumes of work. Providing that Charging is not directly based on this rate calculation, the impact is minimal and can be addressed when the rate is next calculated.

The current installation life of the equipment has a direct bearing on the rate. Thus, recently installed equipment tends to be lightly loaded giving a high average cost, while a machine just about to be upgraded and, therefore, heavily loaded, gives rise to a low average rate. The major flaw in this method, therefore, is that as usage increases so the cost of machine time per hour decreases. The effect is to show reducing costs when, in fact, no such decrease is actually taking place.

Standard Rate using a standard machine

The use of Standard Rate on a standard machine is a refinement of the Standard Rate described above and is an attempt to overcome the disadvantages that are produced by that method. The rate is not based on the actual forecast machine usage but rather on use of a machine at an average load. The effect is to fix the costs of machine time at what should reflect a "the market rate" – that is the rate charged by external organisations.

Estimates can be made of a probable optimum configuration and the resultant additional costs and throughput potential added to the calculation.

The advantage of using this method is that a constant rate is obtained, which is not influenced by the actual equipment installed and allows customers to match their usage to the anticipated costs.

The disadvantage is that the IT organisation has to be capable of delivering service, at all times, to this Standard Rate. It may not be able to sustain this as business environments and customer usages change.

6.3.11 Investment Appraisal

Basic Principles

Sound IT Accounting methods, including the Cost Models described previously, enable businesses to determine the costs of IT service provision. These costs can be used in Investment Appraisal, a process of determining whether the business will benefit from changes to IT service quantity and quality. Techniques of appraisal have been developed mainly in the context of decisions on capital spending, but the general principles apply to any proposal for spending or saving money that involves changes in the use of resources. Systematic appraisal entails:

- being clear about objectives
- thinking about different ways of meeting them
- estimating and presenting the costs and benefits of each potentially worthwhile option.

Used properly, appraisal leads to better decisions by policy makers and managers; it encourages both groups of people to question and justify what they do and it provides a framework for rational thought about the use of limited resources. It establishes the link between business operations and the cost of underpinning them; it enables the IT organisation to support the business in its cost/benefit analyses.

Investment Appraisal has long been recognised as an essential prerequisite to sound financial management both in the public and private sectors. The importance of Investment Appraisal has grown in recent years as an aid to decision making in its broadest sense, as a means of identifying efficiency savings and controlling investment expenditure to maximum effectiveness.

There is often a trade-off between capital investment and running cost expenditure: i.e. between maximising effectiveness in the long-term and the risk of failing to achieve short-term goals. Capital investment decisions are essentially longer-term decisions, and thus it is more difficult to hold management responsible and accountable for such decisions. However, because the performance of a manager is often measured on the efficient and effective use of allocated resources within a budget period (of one year), there are only limited ways of holding managers responsible for investment decisions : this is why sound Investment Appraisal procedures are essential to an organisation.

The purchase of IT equipment and software is one of the most important investment decisions an enterprise must make. It is therefore crucial that such decisions are properly included in the enterprise's strategic planning. In all well-run businesses, there is a requirement to prepare and update a business IS Strategy, thus IT managers both at the strategic and operational levels have a central role in helping to achieve organisational aims and objectives.

From the viewpoint of the business manager, IT investment and the supply of IT services is the same as any other planned expenditure or allocation of resource in that it is measured in terms of its contribution to the effective, efficient and economic achievement of business goals. We must enable the business to determine whether, for example, the returns would be better from a new IT system or from increased advertising.

Return On Investment (ROI)

Many enterprises now insist that IT projects, in line with other business projects, calculate a Return On Investment. This enables decision-making to be based on common business standards and allows comparisons between investment in IT and non-IT projects.

The factors classed as a "Return" vary from enterprise to enterprise and from year to year. In concept, a return is a revenue or benefit, or the prevention of a lost benefit or revenue (opportunity cost), which is attributable to the project. If the return is divided by the expenditure required to complete the project, a figure is derived which can be compared with the returns on doing nothing (often taken as the bank base rate) or the returns from other projects or initiatives.

$$\text{ROI} = \frac{\text{average increase in profits}}{\text{Investment}} \quad \text{(average taken over an agreed number of years)}$$

The Investment figure may take into account the cost of borrowing money

The accuracy of the estimates for returns on investment is often challenged but that is the nature of estimating benefits. To improve the acceptance of these calculations, it is important to involve the business in determining the returns and how they are to be measured. It is essential that the Capital (one-off) and Operational (on-going) costs for a given project or initiative are identified and can be recorded in the manner previously shown.

Frequently, the benefits will lag behind the expenditure and the Return has to be calculated over an agreed period. For some financially aggressive companies this could be 12 months. For others, it could be between 5 and 10 years. The exact expectations of the financial managers may be dictated by board-level strategies, in turn responding to shareholder demands.

Return On Capital Employed (ROCE)

Shareholders and potential investors in a company will look very closely at a number of "ratios". The actual ratios felt to be important are subject to fashion but a primary one is the Return On Capital Employed (ROCE).

$$\text{ROCE} = \frac{\text{Net Profit Before Tax and Interest}}{\text{Total assets less current liabilities}}$$

This ratio, of the Returns over the Capital Employed, is one frequently used by business analysts to judge the effectiveness of the enterprise as a whole. Any changes to services or products would normally be expected to improve this figure and hence the ROCE calculated for a proposed projects should be higher than the overall enterprise ROCE.

The use of ROCE enables comparisons of different investment opportunities in a fair way and represents how effectively an enterprise generates revenue from its assets (the capital employed). At its most basic, this can be compared with bank base rate that offers a Return on Capital Employed which has minimal risk.

Adventurous organisations will usually seek to employ their capital in furtherance of their business while enterprises that are undergoing a period of stabilisation will reserve the capital for future use.

6.3.12 Total Cost of Ownership (TCO)

The Gartner Group pioneered a method of calculating the costs of a product or service with the title of "Total Cost of Ownership". This referred to assessing the lifecycle costs of an item rather than just the visible capital expenditure.

The most widely known example was that for Personal Computers. In an era where the price of a PC on a desk had fallen to $2,000, Gartner demonstrated that the 5-year cost of a PC, when taking into account purchasing overheads, upgrades, maintenance, a proportion of support staff and Service Desk costs, disposal etc. was closer to $35,000.

They argued that the cost per PC per year was thus $5,000. This is actually the same process as that described in *6.3.8 Calculating the Cost-by-service* but applied to a single product / activity rather than a customer or service.

It remains a vivid demonstration of the difference between purchase price and ownership cost which many customers of IT departments still find surprising, particularly when IT budgets are discussed.

The current extension of this is to include Social costs such as recycling, environmental damage or work injury to produce Total Life Cost.

6.3.13 Budgeting, IT Accounting and Charging cycles

If the Cost Model is being developed during the IT financial year, the figures used will be based upon the costs already incurred and those predicted for the remainder of the year.

If the Cost Model is being developed to enable the calculation of future IT budget, it will depend upon workload, capacity and technology requirements from the businesses. It is likely that there will be some iteration or adjustment to both the model and the resulting costs during the budget preparation period (see *Figure 6-1 IT Accounting, Charging & Budgeting cycle*.

There are two distinct cycles associated with Budgeting, IT Accounting and Charging:

- a planning cycle (annual) where cost projections and workload forecasting form a basis for cost calculations and price setting
- an operational cycle (monthly or quarterly) where costs are monitored and checked against budgets, bills are issued and revenue collected

	Budgeting	IT Accounting	Charging
Planning (annual)	Agree overall expenditures	Establish standard unit costs for each IT resource	Establish pricing policy Publish price list
Operational (monthly)	Take actions to manage budget exceptions or changed costs	Monitor expenditure by cost-centre	Compile and issue bills

Table 6-8 Budgeting, IT Accounting and Charging cycles

It is important to plan the expenditure and cost recovery profiles annually. The plans should cover each month in the financial year to facilitate monthly monitoring and cover peaks and troughs. Actual expenditures and recoveries are monitored and compared at the end of each month: the expenditures and recoveries are also compared with what was planned. It is also appropriate to plan changes to the IT Accounting system to coincide with SLA reviews and any change should be reflected in the SLAs.

6.3.14 Charging in Profit Centres

In a Profit Centre the objective is to recover, through Charging, an amount greater than the costs incurred. To manage this, it is usual to be able to report the profitability of individual customers and services at various points in time. The revenues and expenditures are controlled through the Budgeting cycle and operational management of the agreed budgets. The performance of the IT organisation and its managers may be evaluated by measuring both profitability and expenditure.

This may require the overall Financial Model for the IT Organisation to include profit and margin information for individual services and the Cost Model must have sufficient detail to enable these calculations.

6.4 Developing the Charging System

6.4.1 Scope of Charging

The concept of Charging for incurred internal costs is not new but is often seen as too bureaucratic & too difficult to implement fairly.

The Charging systems described should enable an enterprise to:

- determine the most suitable Charging policies for their enterprise
- recover fairly and accurately, the agreed costs of providing the services
- shape customer / user behaviour to ensure optimal return on IT investment by the enterprise

Such a system controls IT service costs and influences the proper use of IT resources, so that these scarce resources are used in the manner which best reflects business need.

To calculate the charge for providing IT services internally, or between (or to) subsidiaries, an enterprise must decide, <u>prior to implementation</u>, what it is hoping to achieve. One key factor will be to analyse the motivational aspects of Charging, considering both the effects upon the provider and the user of the service. The objective is to optimise the behaviour of both parties in achieving the enterprise's aims.

6.4.2 Charging Policies

Traditional, centrally funded IT Service organisations are under pressure from many sides. They are expected to reduce overall costs while maintaining or improving service in an increasingly complex environment. Business divisions may make unrealistic, competing and unjustifiable demands on the fixed resource available. Within the organisation, staff may feel trapped into a slave role with little opportunity to manage workload or to develop new skills.

Charging for IT Service is seen as a method of

- Forcing the business divisions to control their own users' demands
- Reducing overall costs and highlighting areas of service provision which are not cost effective
- Allowing the organisation to match service to justifiable business need, through direct funding

" Users will only value what they have to pay for " : the corollary is that once users have to pay, **they will demand value for their money**.

To implement charging requires a management commitment to resolving the issues that this will bring to the organisation as a whole. **Unless the IT Service organisation has the support of the whole company in introducing charging, it will fail.** It has to be *simple*, *fair* and *realistic*.

Simple - *" I can see three more administrators but two less IT professionals "* - the overheads of cost management must deliver the benefits of an improved overall cost-effectiveness without the bureaucracy commonly associated with IT Accounting for costs.

Fair - *" I can obtain the services cheaper elsewhere and that's what I'll do "* - the system must be fair and realistic, services which are not cost-effective will have to be reviewed and hard decisions taken. Each business will have to pay the same money for the same service.

Realistic - *" I'm saving money, even though it must be costing the company more "* - anomalies in the charging system will be exploited by businesses. The charging mechanisms must be designed to achieve optimal behaviour.

The image of the IT Service organisation is likely to change; they may be seen initially as demanding money without providing the required service, as having become bureaucratic and focussed on trivial accounting. To limit this risk, IT organisations considering the implementation of charging should

- Publicise the programme and work with the businesses to define charging policy
- Ensure that Service Level Agreements are in place and representative of actual service
- Ensure that the benefits are quantifiable and demonstrable

Four factors govern the requirements of a Charging system in the enterprise. For example, full commercial Charging requires that costs can be forecast and collated in a manner that provides profitable revenues in the chosen marketplace. In a simple environment, the aims of a Charging system may be solely to make costs more visible and to cause customers to pay more attention to them. The four factors are as follows.

Level of recovery of expenditure required If the IT organisation opts for full recovery of all costs, then it is opting to function as an autonomous unit, financially self-sufficient. This then requires that costs can be forecast and a Charging system selected that is rational, easily understood and very accurate (although not necessarily based on business unit Charging).

Desire to influence customer / user behaviour Customers and users are encouraged to make more efficient use of IT resources through levying charges that vary with usage. This can be applied to :

- reducing the inefficient use of IT resources e.g. reprinting large reports
- reducing the peak capacity required e.g. running lower priority work in periods of lower demand.

There is a conflict between the aim of simple, fixed charges referred to in *6.1.2 Basic Concepts of Financial Management for IT Services*, and the levying of variable charges (which makes the customer's task of Budgeting more difficult). Leading practice would be to help the customer identify where poor process or lack of knowledge was increasing the costs of providing their service and work with them to reduce this excess cost to zero over a number of years.

Ability to recovery according to usage Recovering costs according to usage requires that the selected Chargeable Items have a reasonable correlation with the amount of resources required to produce them, thereby promoting the perception of a fair pricing and Charging structure.

Control of the internal market Introducing market-priced services requires an efficient and effective IT infrastructure management with capacity properly managed, costs well controlled, and services delivered according to expectations.

Pitching services at market price in turn leads to being able to provide quality services consistently, and at reasonable prices, thereby establishing a professional interface with customers. Ideally, the Charging will be based on business deliverables, recognisable to the customers e.g. business transactions, monthly reports.

6.4.3 Leading practice

Charging customers by resource usage, particularly where they have no control over that usage, is one of the most hotly debated topics. Most enterprises which seek to charge for service, particularly those who are Charging internal business divisions or departments, levy the charges as a method of forcing sound IT service procurement discipline in both customers and suppliers (the latter usually being the in-house IT department).

In this situation, the business division benefit from being able to evaluate the best options for the supply of IT service while the in-house IT organisation benefits by the increased awareness of costs resulting in greater efforts by the customer to reduce or remove inefficient use of IT.

The actual charges made are usually negotiated annually, with the customer forecasting service quantity and quality requirements and the IT organisation attempting to balance the requirements from all of its customers with a corporate strategy to provide the lowest cost core services and to support the IT-enabled directions that the businesses and the enterprise need to take.

In order to prevent over-capacity and under-Charging, or under-capacity which results in over-budget or insufficient service to meet business needs, the majority of enterprises separate the provision of core services from the more optional and variable services, such as installing, moving, adding & changing hardware and software infrastructure components. This excludes applying ‘fixes’ to hardware and software which should be seen as part of the cost of providing the agreed IT services (defined in the Service Level Agreements).

Additional capacity is often made available through the use of contractors or third –party service agents. The role of a modern IT organisation can often include being an ‘honest broker’ by assisting the customers in the selection of the most suitable service provider and reducing costs by ‘bulk-buying’ on behalf of all customers.

It is still necessary to fully understand the Cost Model for the in-house IT organisation if the core charges are to be fairly assessed and most customers will need to see this broken down by Chargeable Item in order to perform full comparisons with alternative suppliers of service. Any comparison should take into account the different environment and scope of services provided by external suppliers.

One of the greatest dangers for an IT organisation beginning to apply charges to internal divisions or departments is that, in the absence of a breakdown by element of service, the service provided appears more expensive than outsourcing or third-party suppliers. This is often due to levels of service being offered that could not be duplicated by an external organisation. It is therefore critical that the level of service being provided is discussed with customers at the same time as proposals to introduce Charging.

6.4.4 Deciding Chargeable Items

The key decisions for Charging are the choice of what to charge for (Chargeable Items) and how much to charge for them. Chargeable Items should be understandable and controllable by the customer. Suitable Chargeable Items would be PCs connected to the network or number of batch enquiry jobs submitted. The customer can then manage their budget by controlling their demand for these items.

The more closely the Chargeable Items relate to the enterprise’s business deliverables the better the interface to the users. Only a lack of information should force Charging to be directly based on resource usage; this lack of information must be overcome and it is important that in the analysis phase, steps are taken to ensure the future availability of information.

Consider the following example: airlines sell tickets which allow you to make a journey; they do not issue an bill covering usage of plane, fuel, food, proportional crew costs and so on. The flight or journey is the chosen Chargeable Item.

Often, business deliverables are not suitable as Chargeable Items because they require too detailed a measurement of the resources consumed, for example a customer may run require a service to produce financial analyses but calculating the individual cost of one analysis or even of many types of analysis could add to the total costs unacceptably. In such cases, a structure may have to be established in which the service is charged for as a whole.

Another example is work performed by a statistical analysis program that on one occasion may run through very quickly and on another, consume vast amounts of resource. Although over a period of time, Charging by the business deliverable (the report) may collect the appropriate money for resources consumed, more accurate Charging for the resource-intensive runs may influence the customer to alter the way in which the program runs enabling a more effective use of IT.

Often, business deliverables cannot be easily attributed to single processes or applications. Programs are rarely written to produce single logical business deliverables: often many users utilise portions of a multitude of programs, each of which contributes to the production of parts of many business deliverables. Batch systems in particular are characterised by this trait and identifying the costs of running a large suite of batch jobs may not be possible.

Where a customer requires charges to be variable, dependent upon usage, the Chargeable Items have to be more specific to that customer and easily attributable to that customer. The more freedom the customer has to define their own service, the more detailed the Charging structure required.

6.4.5 Variable costs and charges

Other costs that are known to vary, for instance computer consumables or overtime hours, have to be managed carefully. In the budget, it may be necessary to estimate the likely annual total and manage the cost by checking monthly that the usage is in proportion to the period measured.

The customer will want an estimate of the likely charges and possible upper and lower limits, in order that they can budget. If a service depends upon staff costs and the effort required cannot be predicted accurately, the charge calculation might be in the form of :

e.g. first 1000 hours @ £60 /hour, subsequent hours @ £100/hour (reflecting the need to pay contractors)

When estimating Costs-by-customer or Costs-by-service for a new or changed service, the potential variation from estimate could be large. For example if a major project is due to be implemented, a large variable cost may be the overtime necessary to install equipment outside normal working hours or additional processor time to re-run work.

Often, apparently variable costs will not decrease with decreasing usage, for example decreased usage of a mainframe server may not result in any decreased costs, unless we can reduce licence, maintenance or hardware charges.

Hence, the IT Finance Managers must be cautious in identifying a cost as variable if this figure is to be used in calculating Costs-by-customer or Costs-by-service. At the very least, it will need to be re-checked whenever there is a change to the system or its usage.

6.4.6 Pricing

Pricing is just one element of the marketing quartet "product, pricing, promotion, place". Deciding upon the appropriate charge/price is, therefore, not merely a question of cost recovery but also of its impact upon the demand for the product. If an organisation charges for its product in an open marketplace (i.e.

no constraints upon the businesses to buy the service) there must be clear understanding of whether the product (i.e. IT service) is attractive to the marketplace.

The pricing of any product or service involves :

- the determination of a pricing objective
- understanding the true (not perceived) demand for the service
- accurate determination of direct and indirect costs
- the level of control of the internal market
- understanding the services available externally if customers have a choice
- legal, regulatory & tax issues

Achieving an anticipated rate of return figures greatly in many managers' considerations. For the IT Services manager selling services to internal divisions, the price to be charged is often based on what has to be charged to recover costs, with any cost above market prices explained through the benefits of internal spend, the assumed flexibility of internal organisations or the value-added aspects of working for a common enterprise.

Whatever pricing decision is taken it will be essential for the IT Services Manager to first know the actual costs of providing the service.

Examples of pricing methods include:

- cost
- cost-plus
- going rate
- market rate
- negotiated contract

Cost plus The cost can be defined in several ways, for example:

- full cost (calculated as a total cost of ownership, including depreciation / planned renewal)
- marginal cost (the cost of providing the service now, based upon the investment already made).

Cost plus There are a number of cost-plus pricing models. The basic form is:

Price = cost + x%

The mark-up (x%) can either be set by the enterprise as a standard Target Return, comparable with returns on other business investments, or varied by the IT organisation to meet strategic business needs e.g. encouraging the use of strategic applications but discouraging the use of legacy applications.

Going rate The price is comparable with other internal departments within the enterprise or with similar organisations.

Market rate The price is the same as that charged by external suppliers. Care should be exercised in asking external suppliers to quote prices – they may well disguise a discount to gain the business.

Fixed price The IT organisation sets a price based upon negotiation with the customer for a set period, based upon a predicted consumption.

However the prices are determined, it is essential that they are visible to the customers so that customers can tailor their budget forecasts and service level requirements to match the likely costs.

6.4.7 The internal market

Tied users In many IT organisations, users are tied to using the internal IT services. Where the IT organisation is intending to operate as an independent business unit, the stages usually followed are:

- tied, services provided “free”
- tied, notional Charging introduced
- tied, actual charges introduced (sometimes subsidised, initially)
- untied, real Charging.

Where no direct charge is levied, customers do not take account of the cost of producing a service and tend to treat it as free. As a consequence they may be extravagant and uneconomical in their use of IT services.

Untied users In an untied situation Charging becomes particularly important because it enables customers to choose between using the in-house IT Services section or outside suppliers, based on the relative quality and price of services offered.

If IT provision is to be untied (i.e. users can buy in from other sources), the IT providers should operate on a commercial basis with the aim of recovering full cost, including return on capital employed. When users are first freed from being tied to internal IT services, it is not uncommon to find some enterprises subsidising the cost of the internal IT services, to discourage users from switching suppliers. In practice, subsidising IT services should be discouraged, since prices will appear to be unrealistically low.

Charging at full cost entails recovering all revenue spend costs associated with providing an IT service (including a share of overheads), insurance premium, depreciation of fixed assets and interest on capital employed. It is important that the charge for each service provided by an IT Services section should, wherever possible, be charged out at full cost.

Cross-subsidisation between different customers of the same service should be avoided whenever practicable. Where surpluses and deficits accumulated by a number of services are pooled so that they break-even in total, the effect is to subsidise the cost of one service with the income from another. Such cross-subsidies may lead to a department being criticised for taxing one group of customers for the benefit of another group.

An IT organisation might not charge out a particular service at full cost when:

a) it is possible to distinguish between internal and external customers and where the external customer is the ones with the primary reason for establishing the Services. In this case the external users could be charged a premium rate in order to make a profit but internal users would not be charged.

b) there is excess capacity created for specific customers that can be used on a short-term basis by other customers at no cost or at marginal cost.

6.4.8 Differential Charging

Setting different charges for different usage of the same or similar services enables an organisation to reward some usage patterns over others. For instance the use of Differential Charging to increase charges for batch work during peak daytime processing periods may encourage changes in customer behaviour which reduce overall costs, often without the IT organisation being blamed for poor value of service. This can result in the smoothing demand for capacity and reduce the overall capacity required.

However, if users are not tied to the IT department, great care has to be taken in applying differential charges that are not seen in the open market or cannot be justified on a business basis. For instance, trying to discourage the use of daytime development work on a mainframe because of the impact on the service levels of the production service, may cause a customer to move their development workload, decreasing the revenues for the IT organisation without reducing their need for capacity (i.e. revenue drops, costs stay constant)

6.4.9 Pricing flexibility

The IT Finance Manager in a Profit Centre may wish to set prices for an annual period to guarantee revenues; however, where new resource requirements are committed to meet the needs of particular customers it may be necessary to build into any contract for the provision of such services a clause to permit changes to the capacity available during the contract period. This reduces the risk of having excess capacity. The alternative is to be able to sell excess capacity to other customers but it is rare for this to be effective as other customers' budgets will also have been fixed.

The IT Finance Manager in a Profit Centre may wish to consider the stability and duration of a customer requirement for IT services and seek regular updated forecasts if demand is likely to change. This may be particularly important if there are large peaks and troughs in customer demand especially when these are not synchronised to the availability of additional capacity. It may be that the Manager should consider setting specific prices for particular times of usage; but this will depend upon size, circumstance, and objectives of the Profit Centre.

6.4.10 Billing

Three objectives are key to Billing:

- The bills must be simple, clear and matched to the ability to pay (in amount, time & method)
- Chargeable Items must be understood by the user, with reasonable correlation to usage of resources, including hardware, software, accommodation, and organisation
- IT Accounting data must be available to provide details on, and justification for, bills

Bills Charging information is passed to customers to make them aware of the cost of the resources used by their business. This can be done by :

- calculating and circulating to managers the full details of the cost of providing each business's IT or services (No Charging)
- as above, but including details about how much the IT organisation would charge, should a charge-back system be operated, without applying transactions to the financial ledgers (Notional Charging)
- as above, but applying transactions to the financial ledgers (Full Charging)

Often the first two options above lead ultimately to the introduction of Full Charging. Whichever approach is followed, the presentation of the information to the customer must be simple, understandable and honest.

Notional Charging is useful when a Charging system is being introduced for the first time. Notional Charging allows the IT Services section to gain experience and time to correct errors in the Charging formulae or cost recovery plans and familiarises customers with the concept of being charged for using IT resources.

Notional Charging is not recommended for long-term use unless the enterprise does not intend to move to a real Charging system, because the incentive to become cost conscious is lessened when money does not change hands.

Billing cycle If Full Charging is introduced, the enterprise may also expect the IT organisation to manage cash flow. If a bill is only paid annually, the organisation may have to account for the gap between revenues and expenditure in the same way that a separate company would have to manage its cash flow.

In the simplest example, an IT organisation which pays licence fees annually in advance, has monthly Operational expenditure and some Capital expenditure but does not bill its customers until the end of the year would operate at a loss throughout the year and only break-even when all bills are paid. This is recognisable to businesses as a cash flow problem, which requires funding.

For most enterprises, the budgets are agreed and monitored monthly by the Finance department, which takes care of any cash flow issues for the enterprise as a whole and can include those of the IT organisation.

In the following simplified example, the overall cost of implementing a service has to be borne within the financial year ending August 31 but the customer will only begin paying once installation is complete.

				Month				
Expenditure	January	Febraury	March	April	May	June	July	August
Software	£20,000							
Development		£10,000	£10,000	£5,000	£2,000			
Testing				£10,000				
Hardware			£30,000					
Installation			£5,000	£2,500				
Support				£2,000	£5,000	£4,000	£1,000	£1,000
Total	£20,000	£10,000	£45,000	£19,500	£7,000	£4,000	£1,000	£1,000
Charge					£26,875	£26,875	£26,875	£26,875
Revenue-expenditure	-£20,000	-£30,000	-£75,200	-£95,002	-£75,882	-£53,965	-£28,858	-£3,530
Cumulative cashflow	-£20,000	-£30,200	-£75,502	-£95,757	-£76,840	-£54,733	-£29,405	-£3,824
Interest @ 1%	£200	£302	£755	£958	£768	£547	£294	

Table 6-9 The effects of interest on late billing

The result of interest applied to the cash flow is to show a £3,824 shortfall in recovery. Even if the Finance department take responsibility for managing this, they will often want to see the Cash Flow profile to aid their planning.

6.4.11 Conditions for Charging to be effective

In considering whether it is worth investing time and resources in setting up a Charging system, regard should be given as to whether customers:

a) have meaningful budgets in which internal IT Services charges are regarded as real costs to be set against other items in the budget. If this is not the case, many of the advantages of Charging are reduced; though it is still worthwhile doing the IT Accounting to provide cost data as management information

b) have an element of choice in their level of usage of the services provided by the IT Service organisation, both overall and at particular times.

If demand is not sensitive to pricing, Charging is effective only as a mechanism for funding IT service provision: alternative arrangements such as inter-departmental reviews must take place to verify and assure the cost effectiveness of IT service provision (see also *6.7.11 Auditing the systems*).

Any considerations, however, must take into account the fact that customers perception of IT is constantly changing and that they increasingly expect their IT services supplier to be a thoroughly professional organisation, whether in-house or external.

6.4.12 Case Studies

The intent of this section is to provide the reader with a few examples of how Charging may become a real hindrance for an organisation if the initial strategy is not clearly defined or the implementation imperfect.

Case Study One : Internal money versus external money

A company has an underground parking lot. This parking lot is for the use of all employees but is over-subscribed. The company decides to introduce Cost Accounting and calculates that the £200,000 yearly cost of the car parking facility will have to recovered by an internal charge to each department of £50 / month for each car park pass issued.

The company parking lot is :

- underground, at the office
- provided as a benefit to staff
- with electronic remote control to open/close the main door
- cameras and security guards.

A hotel nearby provides parking facilities :

- outside, 200m away
- with a hotel-operated barrier
- no security or guarantee of safety

Because the hotel has the space free during working hours, it need only recover costs of administration plus any profit it wishes to make. It decides to charge £30 / month for business car parking. The result of this is that the internal parking lot is only used at 20% of its capacity while the hotel parking lot is full.

Using the external hotel parking facility costs less to the department manager but more overall to the enterprise, as it has to pay the costs of the building and would usually prefer to minimise external spend. They cannot reduce their overall costs as the company parking is an integral part of the offices and grounds.

Resolution

The company agreed that it should subsidise parking to reduce the charge to £40 / month while also banning managers from purchasing outside services when suitable internal services were available.

Bottom line

- Do not allow businesses to buy external services when suitable internal services are available.
- Avoid higher internal prices than the market prices.

Case Study Two : Exceeding business need

One company is charging £860 / year for the provision of an infrastructure to which users workstations are connected. This infrastructure is available 24 hours a day, monitored 16 hours a day and includes a number of facilities such as office automation tools, shared printers, email, external gateways (faxes, Internet, access to large servers).

Some small departments (from 5 to 20 users) were used to very simple peer-to-peer infrastructure that suited their needs. Moving to the new, more expensive common infrastructure did not appear to provide any benefits and even seemed to add to overheads.

Several of these departments conducted a study and found that it would be less expensive for them to continue with their dedicated infrastructure and just have one shared workstation connected to the common infrastructure for email purpose.

Resolution

The charges were adjusted by reducing the standard charge and adding an additional charge for those requiring the additional facilities. The standard charges were raised above true cost, under the direction of the enterprise management, to encourage users to adopt the systems that were felt to be strategic and to wean them from outdated systems.

Bottom line

A **small, dedicated infrastructure** tailored to specific needs **will always be cheaper than a company wide infrastructure**.

However, **several small, dedicated infrastructures will always cost much more than a global one** and are also unlikely to meet strategic aims.

Case Study Three : Discouraging use of services

A company provided its users with a dedicated, outsourced Service Desk facility. The vendor charged the company on a per-call basis, the price varying in bands, depending upon the total number of calls during the month. The charging policy was to recharge all IT spending to the business on the basis of true cost.

Once the Service Desk was in place, customers realised that they could reduce their costs by not placing Service Desk calls. Some business managers instructed their users not to use the Service Desk, or to route all issues through a single, local support person.

Decreasing the total number of calls decreased the calculated charges to the customer but did not reduce overall price of the service by the same amount. It also resulted in :

- increased wasted time for users
- reduced effectiveness of IT systems
- poor perception of the IT services and the IT organisation
- additional work IT organisation to discover problems
- reduced leverage in negotiating service costs with the outsourcing vendor

Resolution

The charging method was changed to one in which a fixed fee per user was negotiated, based upon an estimated call rate taken from previous years volumes and business predictions. This charge was reviewed quarterly to check that call levels were within agreed thresholds.

Bottom line

Some cost should be fixed and should not depend on actual usage

Case Study Four : Hidden costs

In the same Service Desk context as the previous example, business managers may be tempted to set-up their own Service Desk facility by appointing on staff dedicated to this. It may cost less to the business department than a centralised desk and allows the business to direct the efforts of the staff toward the incidents and problems that concern them. However, the total cost to the organisation of allowing one or more businesses to do this is :

- the quality and level of service will obviously not be the same
- knowledge will not be shared with other departments
- studies will be undertaken on a department basis where they should have been on the company or group basis
- costs for this dedicated facility will not be monitored or at least not monitored centrally

This results in an under-used central facility and an increase in hidden costs of IT as the business Service Desks are not accounted for in the IT budget and probably not fully costed by the businesses.

Resolution

The enterprise forced all business units to measure effectiveness and introduce cost analysis. This demonstrated that the internal Service Desk and support services were actually costing more *in total* than would be the case if the central facility were correctly used.

Bottom line

If the businesses are allowed to develop private infrastructure or services, it may result in :

- a decreasing influence for the IT organisation
- an increasing IT cost to the enterprise
- inefficient use of resources and IT
- arguments where the department and the central Service Desk are asked to resolve problems or maintain service in areas of overlapping responsibility

6.5 Planning for IT Accounting & Charging

This section covers the evaluation and planning activities necessary to prepare for the implementation of sound IT Accounting and Charging systems within IT organisation. It is not exhaustive and is intended to provide specific activities within a formal project framework such as PRINCE-2.

The planning phase can be expected to take 3-6 months, mostly dependent on the size and complexity of the IT organisation and the availability of data. The preparation for IT Accounting can be carried out at any time but it is recommended that any new or amended IT Accounting system be brought into use at the start of a new financial year. See also *6.3.13 Budgeting, IT Accounting and Charging cycles.*

6.5.1 The project

The first step in planning the introduction of, or changes to, IT Accounting and Charging is to confirm the scope and requirements with agreed Terms Of Reference (TOR)

The responsibility for IT Accounting and Charging in the operational environment does not need to be decided at this point although in many enterprises, the person most suitable for this role would take the project management responsibility for its introduction.

The Project Board should comprise the following :

- the director of IS in the enterprise
- a senior manager in the Financial Services division
- one or more of the customers to be billed
- the IT Services Manager.

The project team It is vital that the team has a fundamental appreciation of both the enterprise's business and the IT Services section in order to understand the options available to management. Team members must understand the principles of IT Accounting and the following sections are intended to provide the proper context and background for the work.

The board will make all of their business decisions based on the recommendations of the project team's feasibility study and will set appropriate budgets for the development and implementation project.

6.5.2 Feasibility study

Decisions about the scope and objectives of any implementation will be based on the findings of the feasibility study, although decisions may be imposed because of management adoption of Profit Centre, or Shared Services initiatives.

The feasibility study objectives should include :

- evaluation and quantification of the likely benefits and costs of the introduction of IT Accounting and / or Charging
- recommendations on Charging policy
- outline implementation plans
- identification of the resources, staff and costs needed to introduce and run IT Accounting and Charging
- specification of the financial interfaces to the rest of the enterprise
- recommendations on customers' control of budgets : are they simply for information purposes ? can customers spend the money externally ?
- assessment of scope for customisation of billing and purchasing systems to support the IT Accounting and Charging system
- evaluation of any support tools (e.g. monitoring, IT Accounting) that are already available, or can be obtained
- identification of Charging or IT Accounting activities already being carried out within the IT organisation
- the likely business components (for example, timesheets, cheques, enquiries, pay-slips) upon which the recommended Charging system is to be based
- any risk areas that might militate against successful introduction of IT Accounting and Charging
- performance indicators for the successful implementation of IT Accounting and Charging

The following information should be included in the TOR for the study:

- the enterprise’s management requirements (what IT Accounting and Charging are intended to achieve, and whether the IT organisation wish simple cost recovery or the creation of a profit centre)
- specification of how any new systems are expected to interface with any existing systems

It is important for the project team to gain a clear understanding of the business of the enterprise and highlight areas of doubt for a senior management decision, for example how to charge an intermediate customer such as Personnel or Finance, who are using the systems on behalf of another business department

When management approval has been given, a project plan must be produced for the development and implementation phase (based upon the outline plan produced by the feasibility study).

6.5.3 The project plan

Planning for the introduction of, or changes to, IT Accounting and Charging must be carried out in an integrated fashion. The project may require discrete stages and the project plan must cover :

- Analysis of costs and potential Charging policies
- design & development of IT Accounting and Charging systems

- implementation
- post-implementation and review

The project team must identify the:

- establish and agree the organisational and budget boundaries of IT Services
- business objectives of the enterprise in introducing IT Accounting to IT services
- align the review of IT Accounting and Charging with other IT organisation planning
- determine whether Charging will be implemented, with what objectives and to what level of detail
- develop a spreadsheet to capture the projected expenditures allocated to projected usage of major resource types
- establish the Cost Units to be used and the costs for each Cost Unit
- develop a feed-back system to verify that the actual expenditure and usage are, within acceptable limits, comparable with cost and workload projections
- ensure the mechanisms for recovering under- & over-charge are in place (if it becomes necessary to correct mismatches)
- examine the cost of maintaining existing quality levels and how much the enterprise loses through waste, inefficiency or poorly specified service.

The initial analysis can be carried out by developing stand-alone spreadsheets or computer applications. On-going analysis and reporting is best carried out with the assistance of a proprietary packages, designed for IT organisations.

6.5.4 Designing the Cost Model and related processes

To design the Cost Model, it will be necessary to determine the way in which the businesses set out their IT requirements and even to assist in setting this in place if no formal process exists. The IT organisation budget information and policies for IT Accounting and Charging will also be needed.

The following steps should be followed in developing the Cost Model

Identify the planning cycle It is important to analyse the planning cycles in the enterprise so that the source, timing and content of required plans are identified.

Business divisions must produce business plans that include IT requirements (i.e. service requirements and workload forecasts) for their expected use of IT services. These plans provide the information that enables the IT Executive to design the target Technical Architecture and produce the Operational Plan for the IT organisation. The production of these business plans is iterative and the IT organisation will often develop a key role in advising on their production

A simple model of this process was discussed in 6.1.2 Basic Concepts of Financial Management for IT Services

Establish current costs by identifying the IT services that will be provided for the coming financial year, then estimating the total cost of resources needed to provide these services.

Determine the Cost Units that are to be used for apportioning Indirect Costs and calculate a Cost per Cost Unit for each, as shown in 6.3.5 Apportioning the IT Services Costs. Use this data to produce a first estimate of the Cost-by-customer or Cost-by-service, whichever is required.

The data necessary to identify costs is not always readily available in most IT service organisations. It may be necessary to know, for each customer, the proportionate use of resources that are not currently accounted for, e.g. staff, accommodation, hardware, and software. Ideally, monitoring of resource usage should be automatic, using one or more software tools but some data is likely to be collected from paper records and reports or from stand-alone systems.

If costs are not properly identified it is likely that the initial calculations will be shown to be inaccurate during the monitoring periods. This could entail a redesign of the system, at the very least, complex recalculation of the spreadsheets will be needed.

To paint a more complete picture of cost estimates, the Project Team must attempt to quantify the anticipated growth (or other change) in workloads expected in each financial year. This information might be present in the SLA, may be obtained from the Capacity Manager or from the User community. This workload information is necessary to establish the amount of work by business units (i.e. the business units that will be charged, if Charging is to be implemented) that is to be undertaken by the IT organisation and from this, the unit costs for the future years.

Check the data and the model The identified costs should calculated and checked from more than one to collection to validate the data. The model should be checked by performing a full balance check. This check should be made in more than one way, e.g. by totalling all the individual customers' costs and proving that this equals the budget figures as shown in *Figure 6-3 Cost Model of Costs-by Customer* and also by totalling the individual service costs and proving that this comes to the same figure.

6.5.5 Performance indicators

The obvious measures of the success of the IT Accounting and Charging systems are that:

- cost recovery profiles and expenditure profiles prove to be accurate
- charges, where applied, are seen to be fair
- the IT organisation is provided with the expected income/level of profits
- IT customers and users behaviours and perceptions change

The business objective of either break-even or profit whichever is the objective of the enterprise, should therefore be met.

However, even the most accurate Cost Model may become invalid because of changes in customer behaviour perhaps resulting from changes in the behaviour of their clients. Whilst we aim for accuracy in the calculations of the costs and associated charges, there are additional aspects of success that can be measured and used to determine the effectiveness of the IT Accounting system. Examples of other performance indicators that could be used include:

- plans and budgets produced on time
- specified reports produced at the required time
- the inventory schedules are kept up-to-date
- all costs are accounted for
- timeliness of annual audits
- meeting of monthly, quarterly and annual business objectives
- the number (and severity) of changes required to the IT Accounting system
- accuracy of monthly, quarterly and annual profiles
- number of changes made to the Charging algorithm (where appropriate).

6.5.6 Dependencies

Project dependencies the three most important dependencies are *:*

- Senior management must fully understand the implications and costs of the introduction of IT Accounting and Charging and fully support it.
- Service Level Management must provide information on the impact on service levels of different usage patterns that may result from Charging (and also to help to ensure delivery of service in accordance with user expectations).
- IT Infrastructure Management should be set in place to ensure that IT services provided are efficient and effective, at reasonable cost.

The most important dependency is that there is senior management commitment to the introduction of IT Accounting and or Charging systems. Senior managers must also be prepared to specify what they require from the systems and how these systems should interface with other systems in the enterprise. (e.g. capacity management)

A fully functioning system will also depend on the availability of detailed and accurate information.

Professional IT Accounting skills are required to help ensure the IT Accounting and Charging systems are well designed. Where accountants are not available, company auditors may be able to advise.

6.5.7 Organisation

IT Accounting is an integral part of the IT Services management structure and must be permanently staffed. The IT Services Manager is responsible for the IT Accounting function and whoever performs the Finance Manager role should report directly to the IT Services Manager.

The IT Finance Manager is responsible for the day-to-day operation of the IT Accounting and Charging systems and ensuring that the planned structure is operational The Project Management Office and other IT Services managers also have responsibilities for preparing and checking budgets and for assisting the IT Finance Manager in determining suitable IT Accounting methods.

It is likely that the IT Finance Manager needs support to carry out the work. In a small enterprise the functions of Service Level Management and IT Accounting and Charging could probably be combined so long as sufficient assistance is available to support the day-to-day activities. It is also common to share the responsibility for the IT Accounting aspects of IT Accounting and Charging with the Finance department.

Responsibility for IT Accounting and Charging for IT services should belong to the IT organisation, since it is their task to ensure that the business objectives underpinned by IT are met; responsibility for IT Accounting of IT services cannot be placed anywhere else.

6.5.8 Tools

Except for very small IT Services sections, it is not practical to attempt IT Accounting without the availability of a suitable application or software tool. Facilities required include the management of ledgers, reporting of resource usage, and the calculation of costs according to the chosen Cost Model.

Tools are available to provide the data that IT Accounting and Charging require but most IT organisations use standard spreadsheets for developing budgets and Cost Models.

Frequently, the General Ledgers and Purchasing systems are already in place, and the tools for producing Cost and Charging analysis will require methods of importing data from these. Service Desk and integrated ITIL management systems are likely to be able to provide data that will be input to Cost calculation but care must be exercised that the development and implementation of such facilities does not lead to systems which are too expensive to maintain.

6.5.9 Training

Everyone involved with IT Accounting should be provided with appropriate training about IT Accounting. It is also recommended that they understand the fundamentals of capacity, change, configuration and service level management issues.

Many training organisation offer courses with titles similar to “Financial Management for non-Finance Managers” which cover the understanding of budgeting, annual reports and balance sheets necessary to undertake the implementation of IT Accounting for IT Services.

6.6 Implementation

The time necessary to implement IT Accounting and Charging depends heavily on the tools and information already available but can take six or more months even with good tools and information readily available. It will be impossible to introduce full Charging other than at the start of a financial year, although Notional Charging can be introduced at any time, providing that the impact on behaviours has been fully anticipated (see *6.6.3 Awareness*)

IT Accounting can be introduced at any time although it is usual to use a full year’s data in calculating Costs per Costs Unit. Where a full year’s data is not available or is inconsistent (perhaps as the result of organisational changes) the extrapolations made should be clearly identified and appraised for risk.

6.6.1 Documentation

Document the operational procedures covering normal daily, weekly, monthly and yearly operation of all the aspects of IT Accounting, including the planning, cost reporting, cost recovery reporting and billing activities. In addition to the documents covering the IT operational procedures, develop user manuals covering learning, references and user guide for the IT Finance Manager and his staff.

To attempt to implement a system without proper and complete documentation to help the customers and support personnel and to complement training is inadvisable, particularly when the system involves accounts procedures that may be unfamiliar to staff outside the Finance department. The complexity of the support documentation varies depending on whether the user is the IT Accounting team, the Service Desk or a business customer.

It is vital to cover:

- How & when IT management and customer management will be informed of costs
- How IT Accounting data is to be collected
- How the Charging system works and what pricing structures are to be used
- How budgets are monitored between bills
- How accounts are to be settled
- Who is responsible for policing the IT Accounting system and producing reports, bills, and so on
- How & when auditing takes place
- What contingency option(s) are in place
- How errors in billing will be handled
- What change control is applied to the IT Accounting system.

The analysis and design deliverables should be retained as detailed support documents.

6.6.2 Preparation

Set up the Cost centres & Cost Units

IT Accounting is dependent upon the analysis of cost data collected from a variety of sources. Some is collected from automatic meters (e.g. processor-seconds, Service Desk calls) other data may be from semi-automatic processes (e.g. timesheets, purchasing systems, equipment movement forms) or from manual calculation (e.g. support out-of-hours records): coding should therefore be simple to ensure an easy fit from these different sources.

The data has to be consolidated in a way that enables the total costs to be easily checked against financial ledgers and for sub-sets of cost information to be easily derived.

Frequently, the Finance department has already allocated the cost centres and the main tasks for the Finance Manager in coding Cost Units are :

- Separating capital & operational costs
  (in concert with Finance)
- classifying costs (using agreed Cost Types)
- identifying the Direct Costs & Indirect Costs
- agreeing a fair system of apportioning Indirect Costs (some iteration may be necessary as outline requirements for Charging become clearer and practical considerations of measurement affect intellectual requirements)
- checking the above with technical, financial and customer staff

Set up data recording

In parallel with implementing the system design, examine the existing data recording mechanisms for information about planning, reporting and (eventually) Charging. Information will be needed about:

- Workloads / schedules
- Services / customers
- Costs / cost centres
- Resource inventories (hardware and software).

Workload and service data should be available from the Capacity Management database : cost information (including future acquisitions) should be available from accounts departments or from the Configuration Manager. The Configuration Manager should also be capable of providing inventories of all resources.

The data-recording task does not end once the initial data is available. IT Accounting is based on an annual cycle and it is important to ensure that all required data is made available on time, every time. It is not possible to perform accurate cost analysis unless the required data is routinely available.

Select, install and test software tools

Many organisations will use spreadsheets, both to work out the Costs per Cost Unit (and the Price) and to calculate the costs and charges to each customer.

Data is often typed into the spreadsheets which may become very complex but due to the wide range of sources of data, very few integrated packages exist other than for the specific Charging for mainframe usage.

A number of Data Mining, Data Warehouse or OLAP products can be configured to use the data embedded in other applications e.g. Time Management, Work Management or Systems Management products.

Where software tools have been developed or purchased, install and test them at this stage to ensure that they function as expected.

Extend measuring

Once data collection has begun and early analyses performed, it may become necessary to extend existing monitoring systems or develop new ones. Timesheets frequently have to be automated to be used in IT Accounting and procedures need to be put in place to enforce use of them.

Usage data from the network or distributed systems may have to be collected if apportionment of costs is required at this level.

6.6.3 Awareness

Mounting a cost awareness campaign in the enterprise prior to project initiation will have significant benefits in reducing the resistance to new processes and beginning the change in customers and users to a fuller understanding of how they can affect costs. Short presentations should be given about identifying costs and why costs must be controlled.

The presentation must cover plans for cost recovery and the timescale over which it is intended to recover the identified costs. If it is the policy of the enterprise to profit from IT, or indeed to plough back profit into IT, then some information about how the profits are to be used is also appropriate.

Capacity and cost Improving the level of services may increase costs because it is necessary to increase capacity. Similarly, cutting costs may alter the capacity, the quality and variety of services provided. It is the balance between costs and capacity that must be maintained if services are to be provided which are matched to agreed business need.

Customer reaction Identifying all of the costs necessary to provide IT services and establishing a fair means of Charging for those services places IT service provision on a business-like footing. Customers will become aware that they pay of the cost of their IT services and, if charged, may change the amount of service they use or the pattern of their usage.

Even without explicit charges, awareness of costs will focus attention on the effectiveness, efficiency and management capability of the IT department and it is important to be as open and honest as is possible about the true costs uncovered. Where costs are felt to be unacceptable, a forward path will have to be negotiated between IT and the business.

Flexibility of charges Customers may want flexible Charging, where lower use of a service results in lower charges and vice versa. While this is reasonable, there are key reasons for an IT department to limit flexibility in Charging :

a) if the IT organisation cannot reduce its own costs in tandem with any reduced consumption, it risks having to subsidise customers (e.g. if there is a fixed contract for mainframe or network capacity)
b) it is not usually in the overall enterprise or business interest to encourage short-term or paper cost-saving, e.g. if charges are related to Service Desk calls, managers may discouraging the use of the Service Desk, resulting in inefficient use of the IT systems and hidden user dissatisfaction

Current leading practice for IT organisations running a mainly in-house service, is to determine the desired capacity of each customer and provide core services that meet, in concert with the requirements of other customers, the bulk of the needs of the customers. Flexibility is then reserved for the additional services that either cannot be predicted or are used differently by different customers. Some element of flexibility may be retained using external service suppliers to assist during periods of peak demand and by setting thresholds outside of which the Service Levels are suspended, removing the need for over-capacity.

Perceptions The fact that charges are made is likely to cause customers and users to recognise that the resources are not a low-cost commodity that can be wasted. However, services that were previously perceived as good value may suddenly be perceived as poor value when charges are made. It is the responsibility of all IT Services Managers to ensure that the true cost of providing IT services is published
.
Adopting sound IT Accounting practices should establish in people’s perceptions that the IT Services section is business-oriented and cost conscious.

Implementation resource Staff with accountancy and practical IT experience are rare. IT Accounting and Charging may be new disciplines in IT Services and will rely on expertise outside of IT Services management that must be made easily available to the Finance Manager.

6.6.4 Pilot the system

If the IT Accounting information is new or involves Charging users for the first time, it is recommended that the first year of running the system should be viewed as a pilot scheme, perhaps using Notional Charging. This will enable the IT organisation to influence the demands of the customer before Charging is fully implemented if that particular customer is using resources at a higher rate than was first measured or estimated.

If the system is to be piloted, one approach is to begin with mission critical services and to leave less important services for later.

A new or changed system should begin at the start of the financial year, providing that there is sufficient data available to estimate costs and budgets for the first year of full operation.

If monitoring of the costs or charges identifies discrepancies or anomalies, the credibility of the IT Accounting process may be compromised. Providing that this occurs during a pilot, the Charging algorithms and/or the cost projections can be re-examined and corrected before introducing the system fully, thus maintaining credibility.

Customers should be notified at the earliest possible stage about events that are likely to cause changes to their bills. It is important that customers are aware of any likely alterations to charges that may result from changes to:

- the IT infrastructure
- the IT organisation
- the business environment outside customer control

6.6.5 Monitor the system

Monitor the actual resources being consumed (regardless of whether a Charging system is to be introduced) three to six months in advance of system implementation in order to gain experience. It is essential to make sure that this data can be correlated with that from financial planning and capacity management.

For example, if support staff costs are now to be attributed to the customer using the application or system requiring support, it is essential to be able to check that the total time being logged is true and reasonable before using it in any calculation of cost (and particularly of charge).

6.6.6 Plan IT Accounting and Charging Continuity

It is possible that, in the event of a major disaster, the contingency plan is to withdraw the IT Accounting and Charging systems entirely. If not, then IT Accounting and Charging must be tested as part of the Disaster Recovery (DR) testing.

The Contingency Plan should reflect the enterprises use of IT Accounting and Charging. Simply holding paper copies of spreadsheets & bills will not provide sufficient contingency in the event of, say, a spreadsheet corruption. The dependency of the enterprise on IT Accounting and Charging will vary with the type of enterprise, e.g. a bureau would find it difficult to operate without some form of Charging.

The IT Finance Manager should determine which reports and plans are indispensable to the enterprise and ensure that they are able to be produced in the event of a major incident affecting the IT Accounting or Charging systems. The contingency arrangement for IT Accounting and Charging needs to be regularly reviewed.

The impact of a disaster is minimised by keeping off-site back-up of all required data and by maintaining standby or spare equipment upon which to run the IT Accounting and Charging systems but the data sources on which they rely may not be available in the event of a disaster. Any integration that has facilitated the transfer of data from these sources to the Cost Model or Charging application may be impossible to use in the event of a disaster. See also Chapter 8 (IT Service Continuity Management).

6.7 Ongoing management and operation

This section is concerned with the review of the IT Accounting and Charging systems and with ongoing operation of the IT Accounting and Charging system, including dealing with variance from forecasts.

6.7.1 Ongoing operation

The day-to-day operation of the IT Accounting system includes both the production of cost and Charging data and the checking of its accuracy. Checks to be performed are broadly covered in 6.3.13 Budgeting, IT Accounting and Charging cycles; below is a more detailed schedule:

Daily/Weekly	*Collect cost data and check accuracy and completeness*
	Instigate changes, if necessary through Change Management
	Attend Change Advisory Boards (CAB) as required
Monthly	*Run the cost reporting system*
	Check that costs are in line with predictions and explain any variances
	Produce cost analyses
	Produce charges per customer and compare to budgets
	Circulate a monthly balance sheet

	Review cost recovery metrics against IT business targets
Quarterly / half-yearly	*Prepare for annual budget cycle by checking quarterly, half-year or interim forecasts (inward from businesses and upward from IT)*
	Assess the accuracy of the Charging algorithms by balancing the actual revenue against expected revenue (similarly for actual costs against predicted)
	Assess the accuracy of forecasts as a means of improving them in the future
	Verify the price lists
	Plan the changes necessary for next year's staff and resources and any alterations to cost projections and cost recovery projections.
Annually	*Review and audit the IT Accounting system and produce final accounts for the previous financial year*
	Produce annual cost analyses
	Circulate a final balance sheets
	Review standard Cost Units (amending only if critical to accuracy as changes will make any year-on-year comparisons difficult)
	Recalculate Costs per Cost Unit to check for conformance with predicted results
	Review Charging policies and IT Accounting methods
	Assist customers and other IT managers to set IT budgets for new financial year
	Review the IT Accounting and Charging systems to ensure that the IT organisation business objectives are being met
	Review the IT Service Continuity plans and ensure that the dependency on IT Accounting or Charging systems is clear

6.7.2 Change management

The IT Accounting system, when running in production, is based on an annual cycle for which plans and budgets are produced. Expenditure and revenue can be compared against these plans and budgets. If fundamental changes to the system are needed they should be aligned to this annual cycle although it should be noted that to be in place for the start of a financial year, they have to be a part of the budget cycle and may require planning 3 – 6 months earlier.

If changes cannot wait to be implemented until the beginning of a new financial year, there may be severe consequences involving users (prices may have to go up, costs must be curtailed, demand reduced and so on). It may be possible to arrange for packages of changes to be aligned to the changes made to SLAs between IT Services and its customers.

Where changes are unavoidable, their introduction must be subject to Change Management procedures. Small-scale changes must also be subject to change management. Such small-scale changes can, however, be made through short-term special arrangements, but must be included in the IT Accounting system for the coming financial year. Agree all decisions about changes to the system under the control of Change Management, with representatives from:

- IT Financial Management
- Service Level Management
- Finance department
- IT Services Management.

If customers are affected, customer management must also participate in the decision-making.

6.7.3 Reasons for changing the IT Accounting & Charging system

With a complex IT Accounting system there may be many reasons for changes needing to be made. Examples of these are:

- Changes to the enterprise or the IT organisation
- mathematical or base data errors in the IT Accounting or associated systems leading to incorrect forecasts, budgets or calculations
- workloads varying from forecast., leading to recovery above or below predictions
- customers demanding volumes of service which would or has exceeded their budgets

Careful monitoring of the system during the pilot phase reduces the likelihood that major changes will become necessary and where changes are identified, certainly lessens their impact.

6.7.4 Dealing with variances

It is recommended that any variances between calculated Costs and current Customer budgets are reduced or eliminated prior to implementing Charging.

Any errors must be rectified at source to prevent recurrence; where forecasts are at variance with actuals, the appropriate expenditure plans, cost recovery plans or proposed charges, must be altered. At all times, openness with customers, even in a commercial environment is advisable.

It is recommended that where Charging is introduced (real or notional) the charges should not be altered even during the pilot unless serious errors are discovered which have caused very large variance in recovery forecasts leading to severe over-Charging (and of course, excessive revenue) or under-Charging leading to severe losses.

Where price lists must be altered and charges raised or lowered, give advance notice to provide a warning to users about the impact of the impending changes on their budgets.

6.7.5 Workload variance

Explicit in the organisation’s SLAs should be statements about what happens should a customer workload exceed (or fall below) agreed levels. The statements should cover this eventuality where:

- spare capacity is available
- spare capacity is not available.

In any enterprise it is likely that the importance of the business in question will directly affect the contingency action for this type of problem. In every case however, the first question to answer is “Who pays the bill ?”, the second is “How ?”.

IT organisations usually have a margin of capacity in addition to forecast. It is likely that increased capacity would be provided, if available and the IT Services section would negotiate an uplift in charges for the next IT Accounting period if the usage continued.

If spare capacity were not available, it might be necessary to engage a bureau or even to upgrade machine capacity. The cost would probably be met centrally, but it is very important to ensure that SLAs are not compromised and are specific on these issues.

The enterprise as a whole may need to consider whether a customer is allowed to continue to use excess resources to the detriment of other customers(whether staff time or computing resource). Every case, however, is likely to require a management decision and mechanisms must be in place to cater for such problems.

When spare capacity is plentiful, its ad hoc use must not set a precedent. Spare capacity is usually earmarked for future use (indeed it is profligate to maintain plentiful spare capacity). Customers wishing to increase their use of resources on a long-term basis must provide a business case and obtain authorisation via the normal channels. It is vital that the SLAs made with other customers are adhered to and not compromised to accommodate a single customer.

6.7.6 Cost variance

Costs for Cost Units (see *6.3.9 Calculating the costs of Cost Units*) are the basis of the budgeted costs of IT services. If they have been miscalculated, they affect the overall recovery of costs (i.e. Charges) and this has to be addressed by recalculation as soon as is possible (see *6.7.4 Dealing with* above)

Where there has been some change in the expenditure required to purchase or maintain the IT services, it may be possible to pass this on to the customer, with their agreement, particularly if the costs are clearly attributable to a single customer; an example of this might be increases in software or hardware maintenance.

Usually, such changes should have been predicted and it is seen as a responsibility of the IT organisation to have managed any changes in the price of the commodities they purchase to provide IT service.

6.7.7 Revenue variance

Although the enterprise may decide to run the IT services at a profit, it is possible that monitoring reveals either too great a profit (leading to user dissatisfaction) or a shortfall. Where too much revenue is generated the cause must be identified and justified or corrected; unless absolutely necessary the charges must not be altered. Some of the reasons for generating excessive revenue include:

- workload has grown faster than forecast and flexible Charging has resulted in increased charges although sufficient capacity was available without the IT department incurring additional costs.
- customers are being forced to buy more IT resource than they anticipated possibly because there is insufficient capacity or inefficient service
- charges really are excessive.

Where a poor service is the difficulty, the underlying causes must be addressed. When workload growth is a problem, forecasts must be revised (otherwise trouble is being stored) and it may become necessary to make a case to procure additional capacity.

If the IT services are sustaining a loss, possible causes include:

- running insufficient work (perhaps forecasts of workload were too high)
- running costs are too high (perhaps because of too much overtime, because staffing is above complement or underlying costs of operation have increased)
- charges really are too low.

Once more, forecast revision may become necessary. Where running costs are the problem, the IT Services section will have to examine the possibility of making economies.

6.7.8 Service Level issues

Demands for changes in Service Levels are similar to workload changes and should be handled through Service Level Management.

Demand for extra resource (which is incurring cost) is often due to attempts to resolve poor service or due to additional work to complete budgeted tasks where the estimates are now shown to be inaccurate. Customers will usually expect IT Services to manage this without additional cost although there may be no IT budget available. The likely effect is a stretching of resource, which will impact all Service Levels; hence the Service Level Manager must be involved in any decisions to provide chargeable service in excess of agreed budgets.

Since no IT organisation can actually 'turn off the tap', it is essential that the IT Finance Manager tracks customer budgets to be able to identify over-runs before they occur.

6.7.9 Ongoing planning

Any changes to IT capacity (or indeed any other IT or human resources) that were not present in the tactical plans must be reported to the Finance Manager for inclusion in revised cost plans, projections and so on.

Typically, workload changes as agreed at SLA reviews must also be reported. Cost implications and Charging implications must be explicit in SLAs. SLA changes are normally aligned to annual budget negotiations but IT Services management may delegate to the Service Level Manager the powers to vary SLAs within agreed bounds during the year (see Chapter 5 - Service Level Management).

If there is a long lead-time between planning and bringing into service (e.g. due to delaying to the start of the financial year) it may become necessary to revise the cost plans, projections and so on. Cost planning should use data that is no more than 3 months old. .

6.7.10 Management reporting

Regular monthly reports for each supported business are required. A monthly summary of costs and revenue should be provided to the IT Services Manager, together with a balance sheet. Senior IT management committees may also require a report, probably each quarter.

Although the format of management reports is largely dependent on the standards set by the enterprise it is suggested that reports to customers are kept simple but include details of:

- how much they have spent on IT during the financial year
- whether the charges made match the predicted profile
- the current Charging policies and IT Accounting methods
- how the IT department is investing any profits (e.g. in infrastructure or service improvements)
- any variances, what caused them and what actions are being taken

Reports to the IT Services Manager should cover:

- cost total, broken down by business
- cost analyses by service line, equipment domain or other relevant view
- revenues total, broken down by business
- costs and cost recovery against profile
- outlook on costs and cost recovery
- problems and costs associated with IT Accounting and Charging systems
- recommendations for changes

- future investments required

Senior IT executive committee reports should contain similar information. Reports from any external efficiency and effectiveness reviews or from auditors are also of interest and these should include details of any recommendations that were followed up from previous reports.

6.7.11 Auditing the systems

The purpose of the checks during an audit or review is to demonstrate that :

- the systems are delivering the expected benefits
- the systems are efficient, effective and cost-justified
- procedures are formal and adhered to

The post-implementation review and subsequent periodic reviews of the IT Accounting and Charging systems, check that they are working effectively. They also check that the IT Finance Manager is continually reviewing the system and that all identified deficiencies are corrected at source to prevent recurrence.

Audits may be performed internal staff or by external auditors. In both instances audits are intended to confirm that the IT Finance Manager and the supporting personnel are adhering to defined procedures.

Management should however operate its own controls and not rely solely on audit, whether internal or external. The IT Finance Manager will therefore perform their own audits and checks to reassure management that the system runs properly and is policed effectively. Independent audit should confirm this and also provide comparison with other organisations.

Budgeting

- budgets are provided for all functions
- budgets are monitored and reported regularly
- budget projections are reviewed at the end of the budget period
- escalation procedures exist for budget over-runs

IT Accounting

- cost projections are accurate each month and at the year end (overall and for each business supported)
- the sum of (Cost per Cost Unit x Cost Unit) is equal to the total cost of providing IT services
- all costs, including unexpected costs, are accounted for (Hardware, Software, Employment, Accommodation & Transfer)
- regular and accurate reports are produced for management (including the production of price lists)

- the IT Accounting system is understood and users are satisfied with the manner in which it operates
- the IT Accounting system identifies clearly & easily where IT Services business objectives are being met (e.g. profit margins, cost recoveries and so on)
- interfaces to Configuration Management, Capacity Management and Service Level Management are working and provide the necessary workload information
- the appropriate configuration management items (CIs) are up-to-date and accurate
- change and problem management procedures are adhered to

Charging

- bills are simple, clear, accurate and issued on time
- charges are considered fair
- income is collected on time
- price lists are available and any changes to the charges or price lists are implemented within target timescales.
- customers are neither under-charged nor overcharged for their IT services (this is in terms of correctness according to price lists
- discrepancies in charges are identified quickly and resolved with the customers
- senior managers (across the enterprise and within IT Services) are satisfied with the reports produced
- cost recovery plans are on target (cost recovery provides a barometer of how well prices have been set in relation to predicted costs of IT usage).
- interfaces to Service Level Management are effective

The IT Finance Manager should perform an audit or review on a quarterly basis initially and six-monthly after the project has been in successful operation for one year. Where possible, reviews should be synchronised with those for Service Level Management & Capacity Management.
An audit should also:

- check that regular reviews are carried out regularly and non-conformances followed up
- randomly select bills to test for clarity, accuracy and timeliness
- examine cost recovery projections and revenue to assess the accuracy of the system
- ensure that audit trails are provided
- ensure that revenues are collected and properly accounted for
- check that all documentation is accurate, up-to-date and complete.

Annex A - IT Finance Manager - Role, Responsibilities & Key Skills

Some organisations will have a dedicated IT Finance Manager, others may share the tasks amongst the Finance department, senior IT managers, especially those with responsibility for Service Level or Customer relationships and the head of IT.

The IT Financial Management process must all have an "owner" , that is someone responsible for developing and reviewing them. The IT Finance Manager may be able to agree that some of these are "owned" by the main Finance Department, however if this responsibility is split, care should be taken to avoid giving them to people with primarily administrative roles as they are unlikely to have the time or seniority required to manage the required tasks.

The skills list given is an indication of the skills that would be present in an ideal candidate for a post that takes full responsibility for all the topics discussed in this chapter.

Role

To work, at an appropriate kevel, with representatives of the enterprise management and the Finance department, to develop the policies of Budgeting, IT Accounting and Charging.

To implement and maintain the IT Financial Management process, covering Budgeting, IT Accounting and Charging.

To assist in developing account plans and investment cases for the IT organisation and its customers.

Responsibilities

Budgeting

- Manage the IT organisation budget
- Prepare budget forecasts and assist customers in preparing IT elements of their budgets
- Report regularly to IT managers and customers on conformance to budgets

IT Accounting

- Select suitable tools and processes for gathering cost data
- Develop suitable Cost Models
- Agree suitable IT Accounting policies, e.g. depreciation
- Assist in developing cost-benefit cases for IT investments
- Advise senior management on the cost-effectiveness of IT solutions

Charging

- Develop charging policies
- Provide justifications and comparisons for charges
- Prepare regular bills for customers

- Prepare a price list of services, if required

Other

- Provide close support to the Service Level Manager, Capacity Manager and Customer Relationship Manager, especially during budget and IT investment planning
- Recommend scope for internal audits
- Assist external auditors

Skills

Sound numerical and financial skills
Ability to interact successfully with all levels of Customer and IT organisation management
Thorough approach to documentation and schedules
Excellent communication and negotiation skills
Good presentational skills

Relevant knowledge or experience

Understanding of the Customers' businesses and how IT can affect the delivery of their products or services.
Accountancy and company financial reporting
Contract or Supplier Management.
Statistical and analytical principles and processes

7 CAPACITY MANAGEMENT

7.1 Introduction

Capacity Management is responsible for ensuring that IT processing and storage capacity matches the evolving demands of the business in the most cost-effective and timely manner. The process encompasses:

- the monitoring of performance and throughput of IT services and the supporting infrastructure components
- undertaking tuning activities to make the most efficient use of existing resources
- understanding the demands currently being made for IT resources and producing forecasts for future requirements
- influencing the demand for resource, perhaps in conjunction with Financial Management
- the production of a Capacity Plan which will enable the IT service provider to provide services of the quality defined in Service Level Agreements (SLAs)

As shown in Figure 7.1, Capacity Management is essentially a balancing act; balancing

- cost against capacity – i.e. the need to ensure that processing capacity that is purchased is not only cost justifiable in terms of business need, but also the need to make the most efficient use of those resources, and
- supply against demand – i.e. making sure that the available supply of processing power matches the demands made on it by the business, both now and in the future. It may also be necessary to manage or influence the demand for a particular resource.

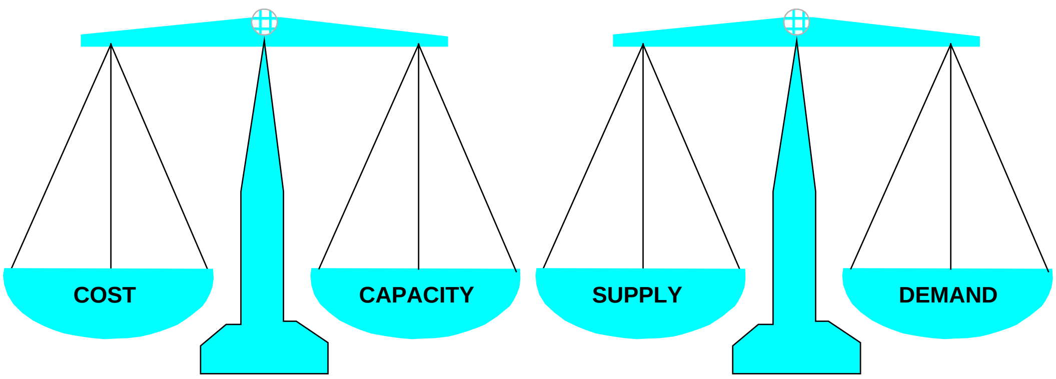


Figure 7.1 Capacity Management - a balancing act

This chapter provides guidance on

- how to justify and establish a Capacity Management process
- the activities and tasks that need to be undertaken
- planning and implementing Capacity management in an IT organisation
- the responsibilities of the Capacity Management process

7.1.1 Why Capacity Management?

Capacity Management is often viewed as on old-fashioned, mainframe-oriented discipline. IT Services managers in charge of distributed computing facilities have argued that Capacity Management takes more time and effort, and therefore cost, than it is worth, and that it would be better to 'pay for upgrades as required'. IT organisations with this view tend to exhibit the following symptoms:

- Procurement of IT equipment is justified on an individual capital return basis, rather than the overall corporate requirement
- There are no corporate Capacity Plans
- Network Capacity Management is done reactively
- As is Capacity Management of servers, albeit less often
- Little or no Capacity Management is performed on desk-top equipment

This insular, short-term attitude is the very antithesis of Capacity Management which is based on the following points:

- Capacity Management of large networks of distributed equipment is more complex than in the 'good old days' of the mainframe, and for all thriving organisations the financial investment in IT is increasing. Therefore it makes even more sense to plan for growth. While the cost of the upgrade to an individual component in a distributed environment will be less than the upgrade to a component in a mainframe environment, there are often many more components in the distributed environment that need to be upgraded. Also there could now be economies of scale, because the cost per individual component could be reduced when many components need to be purchased. So Capacity Management should have input to the procurement process to ensure that the best deals with suppliers are negotiated.
- A corporate Capacity Management process ensures that the entire organisation's capacity requirements are catered for. The cost of upgrading all the desktop equipment in an organisation could easily exceed the cost of a mainframe upgrade. Capacity Management should have responsibility for the 'refresh policy', ensuring that desktop equipment has sufficient capacity to run the applications that the business requires for the foreseeable future.
- Capacity Management will provide the necessary information on current and planned resource utilisation of individual components to enable organisations to decide, with confidence,
  - which components to upgrade (i.e more memory, faster storage devices, faster processors, greater bandwidth)
  - when to upgrade. Ideally this will not be too early, resulting in expensive over-capacity, nor too late, resulting in bottle-necks, inconsistent performance and ultimately customer dissatisfaction and lost business opportunities.
  - how much the upgrade will cost. The forecasting and planning elements of Capacity Management feed into budgetary lifecycles, ensuring planned investment.

Many of the other Service Management processes will be less effective if there is no input to them from the Capacity Management process. For example

- can the Change Management process properly assess the effect of any change on the available capacity?
- when a new service is implemented, can the Service Level Management process be assured that the Service Level Requirements of the service are achievable, and that the SLAs of existing services supported?
- can the Problem Management process properly diagnose the underlying cause of incidents caused by poor performance?
- can the Service Continuity process accurately determine the capacity requirements of the key business processes?

Capacity Management is one of the forward-looking processes, which, when properly carried out, can forecast business impacts before they happen.

Key message : Good Capacity Management ensures NO SURPRISES!

7.1.2 Mission for Capacity Management

Capacity Management needs to understand the business requirements (the required service delivery), the organisation's operation (the current service delivery) and the IT infrastructure (the means of service delivery), and ensure that all the current and future capacity and performance aspects of the business requirements are provided cost-effectively.

However Capacity Management is also about understanding the potential for service delivery. New technology needs to be understood and, if appropriate, used to deliver the services required by the business. Capacity Management needs to recognise that the rate of technological change is very unlikely to slow down, and will continue at its current rate, if not speed up. This technology should be harnessed to ensure that the IT services required by the business can be delivered.

The two laws of Capacity Management:
Moore’s Law
In 1965 Gordon Moore, one of the founders of Intel observed that each new memory chip produced contained about twice as much processing capacity as its predecessor, and that new chips were released every 18 – 24 months. This trend has continued ever since, leading to an exponential increase in processing power.
Parkinson’s Law of Data
We all know that work expands to fill the time available to complete it, but a variation on that law is that “data expands to fill the space available for storage”.
.

While these two laws hold true then effective Capacity Management is even more important as supply and demand grow exponentially.

Mission Statement

The Capacity Management process’s mission is
“To ensure that cost justifiable IT capacity always exists and that it is matched to the current and future identified needs of the business”

7.1.3 Scope of Capacity Management

The Capacity Management process should be the focal point for all IT performance and capacity issues. Other technical domains, such as Network Support, may carry out the bulk of the relevant day-to-day duties but overall responsibility lies with the Capacity Management process. The process should encompass, for both the operational and the development environment:

- all hardware – from PCs, through file servers, up to mainframes and super-computers
- all networking equipment (LANs, WANs, bridges, routers etc.)
- all peripherals (bulk storage devices, printers etc.)
- all software - operating system and network software, in-house developments and purchased packages,
- human resources, but only where a lack of human resources could result in a delay in end-to-end response time (e.g. overnight data backups not completed in time because no operators were present to load tapes). In general human resource management will be a line management responsibility.

However the driving force for Capacity Management should be the business requirements of the organisation.

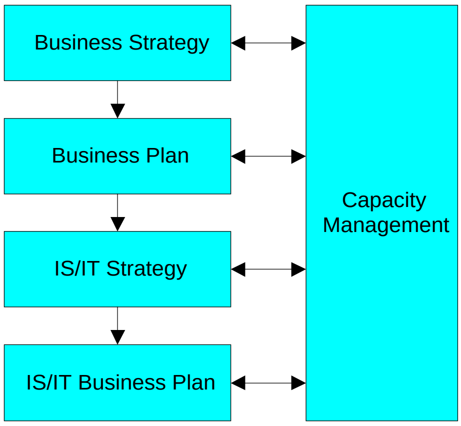


Figure 7.2 Capacity Management and the business

Figure 7.2 shows that Capacity Management has a close, two-way relationship with the business strategy and planning processes within an organisation. On a regular basis, the long-term strategy of an organisation will be encapsulated in an update of the business plans. The business plans will be developed from the organisation's understanding of the external factors such as the competitive marketplace, economic outlook and legislation, and its internal capability in terms of manpower, delivery capability etc.

Capacity Management needs to understand the long-term strategy of the business while providing information on the latest ideas, trends and technologies being developed by the suppliers of computing hardware and software.

The organisation's business plans will dictate the specific IT/IS strategy and business plans, the contents of which Capacity Management needs to be familiar with, and to which Capacity Management will have had a large input. In the IT/IS specific business plans, particular technologies, hardware and software will be identified, together with some indication of the timescale in which they will be implemented.
For more information on the interface between the business plans and strategy, and how they need to be reflected in IS/IT specific business plans and strategy see the ICT : Infrastructure Management manual.

Anecdote:

On November 29th, 1999 an American Internet Service Provider specialising in on-line mail order sales of children's toys issued a press release. The recent system crashes had been due to an unexpected 30% increase in customer demand during the previous week.
"Unexpected increase" in the demand for toys just before Christmas? I'm sure the sales and marketing departments produced some forecasts. Were these communicated to the Capacity Manager?
The same day another press release from a research company reported that the US internet audience is growing significantly slower than it used to, forcing ISP marketing departments to focus more attention on attracting customers from competitors and on improving customer service to existing customers.
Moral:
Uncertainty in business demand will always exist - even the experts can't agree
Capacity Management needs to take account of this uncertainty, keep abreast of marketing and business plans and provide cost-effective, timely processing capacity.
Capacity Management is a key enabler for business success.

7.2 The Capacity Management process

Figure 7.3 shows the inputs to, the sub-processes within and the outputs from the Capacity Management process.

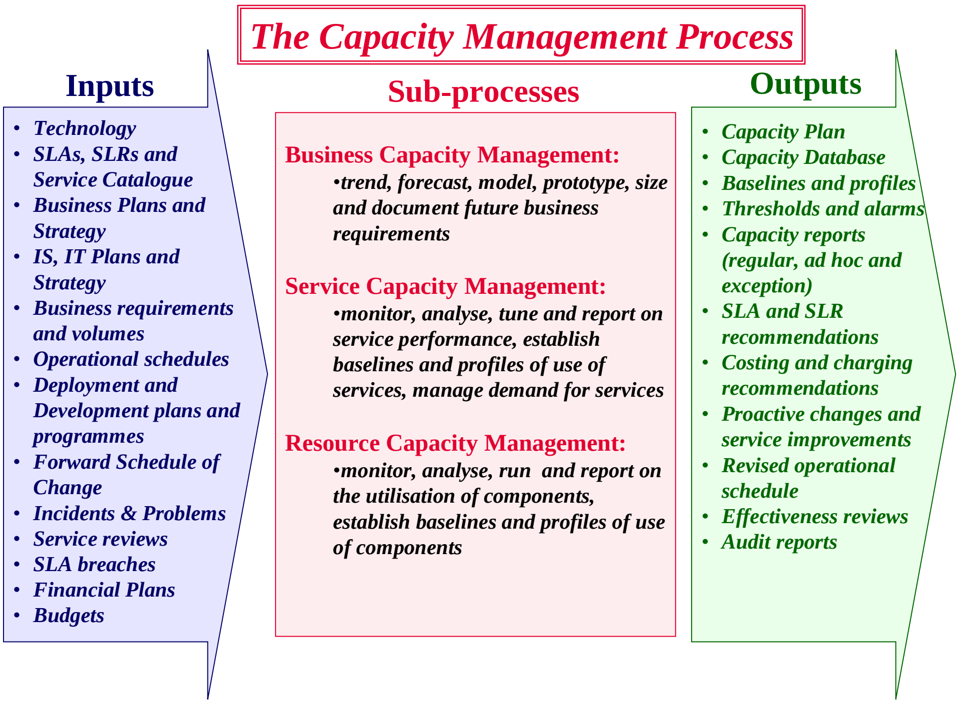


Figure 7.3 The Capacity Management Process

The inputs

There are a number of sources of information that are relevant to the Capacity Management process. Some of these are as follows:

- external suppliers of new technology
- the organisation's business strategy and plans, and financial plans
- the IT strategy and plans and current budgets

- the Incident and Problem Management processes with incidents and problems relating to poor performance
- the Service Level Management process with details of the contents of the Service Level Agreements and Service Level Requirements, and possibly from the monitoring of SLAs, service reviews and breaches of the SLAs
- the Change Management process with a forward schedule of changes and a need to assess all changes for their impact on the capacity of the infrastructure
- the IT Operations team with schedules of all the work that needs to be run and information on the dependencies between different services, and the interdependencies within a service.

The sub-processes

Capacity Management consists of a number of sub-processes, within which there are various activities. The sub-processes of Capacity Management are:

- Business Capacity Management: This sub-process is responsible for ensuring that the future business requirements for IT services are considered, planned and implemented in a timely fashion. This can be achieved by using the existing data on the current resource utilisation by the various services to trend, forecast or model the future requirements. These future requirements will come from business plans outlining new services, improvements and growth in existing services, development plans etc.
- Service Capacity Management: The focus of this sub-process is the management of the performance of the IT services used by the customers. It is responsible for ensuring that the performance of all services, as detailed in the targets in the SLAs and SLRs is monitored and measured, and that the collected data is recorded, analysed and reported. As necessary, action will be taken to ensure that the performance of the services meets the business requirements. This is performed by staff with knowledge of all the areas of technology used in the delivery of end-to-end service, and will often involve seeking advice from the specialists involved in Resource Capacity Management.
- Resource Capacity Management: The focus in this sub-process is the management of the individual components of the IT infrastructure. It is responsible for ensuring that all components within the IT infrastructure that have finite resource are monitored and measured, and that the collected data is recorded, analysed and reported. As necessary, action will be taken to manage the available resource to ensure that the IT services that it supports meet the business requirements. In carrying out this work, the Capacity Management process will be assisted by individuals with specialist knowledge in the particular areas of technology.

The activities that are carried out by the sub-processes are described in Section 7.3. Each of the sub-processes carry out many of the same activities, but each sub-process has a very different focus: Business Capacity Management is focussed on the current and future business requirements, while Service Capacity Management is focussed on the delivery of the existing services that support the business and Resource Capacity Management is focussed on the technology that underpins all the service provision.

The outputs

The outputs of Capacity Management are used within other parts of the process, by other Service Management processes and by other parts of the organisation, as follows:

- within other parts of the Capacity Management process. For example the data monitored and collected as part of Resource and Service Capacity Management will be used in Business Capacity Management to determine what hardware or software upgrades are needed, and when. The Capacity Database will hold the information needed by all the sub-processes within Capacity Management.
- by other Service Management processes. For example the Capacity Management process will verify new Service Level Requirements, and will assist the Financial Management process by identifying when money needs to be budgeted for hardware or software upgrades, or the purchase of new equipment.
- by other parts of the organisation. For example IT Operations will need to implement any changes that Capacity Management may recommend to the schedule of when services are run, to ensure that the most effective and efficient use is made of the available resource. The Capacity Plan will need to be acted upon by the management of the IT service provider and the senior management of the organisation.

Capacity Management - reactive or proactive?

Some activities in the Capacity Management process are reactive, while others are proactive. The proactive activities of Capacity Management will

- pre-empt performance problems by taking the necessary actions before the problems occur
- produce trends of the current resource utilisation and estimate the future resource requirement
- model the predicted changes in IT services, and identify the changes that need to be made to the component parts of the IT infrastructure to ensure that appropriate resource is available to support these services
- ensure that upgrades are budgeted, planned and implemented before SLAs are breached or performance problems occur
- actively seek to improve the service provision

However there will be occasions when the Capacity Management process will need to react to specific performance problems. For example the Service Desk may refer incidents of poor performance to Capacity Management for resolution.

Key message
The more successful the proactive activities of Capacity Management, the less need there will be for the reactive activities of Capacity Management.

The remainder of this section describes the sub-processes of Capacity Management in more detail, and Section 7.3 explains the activities that constitute these sub-processes.

7.2.1 Business Capacity Management

A prime objective of the Business Capacity Management sub-process is to ensure that the future business requirements for IT services are considered and understood, and that sufficient capacity to support the services is planned and implemented in an appropriate timescale.

The Capacity Management process must be responsive to changing requirements for processing capacity. New services will be required to underpin the changing business. Existing services will require modification to provide extra functionality. Old services will become obsolete, freeing up spare capacity.

As a result, the ability to satisfy the customers' service level requirements will be affected. It is the responsibility of Capacity Management to predict these changes and cater for them.

These new requirements may come to the attention of Capacity Management from many different sources and for many different reasons. They may be generated by the business or may originate from the Capacity Management process itself. Such examples could be a recommendation to upgrade to take advantage of new technology, or the implementation of a tuning activity to resolve a performance problem.

Capacity Management should endeavour to become included in as many of the planning processes as possible – particularly Change Management and Project Management.

Key message
Capacity Management should not be a last minute 'tick in the box' just prior to Operations Acceptance and User Acceptance.

The Business Capacity Management sub-process is illustrated in Figure 7.4 which shows how the new requirements for capacity are dealt with by this sub-process. In dealing with these requirements, Capacity Management has to work closely with many of the other Service Delivery and Service Support processes.

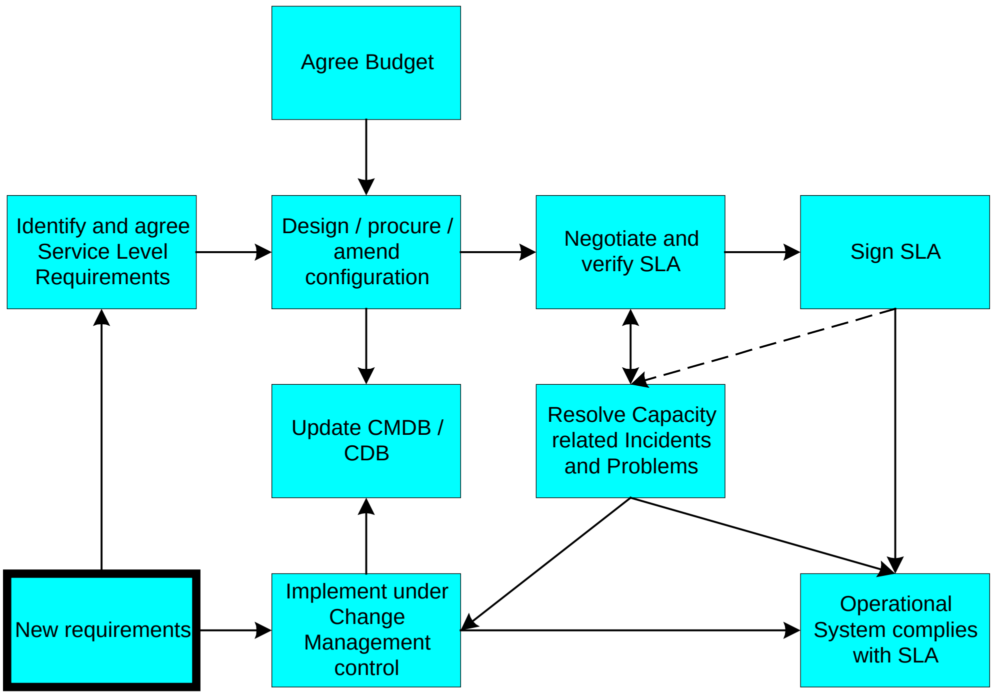


Figure 7.4 The Business Capacity Management sub-process

Identify and agree Service Level Requirements

Capacity Management should assist Service Level Management in determining the customers' capacity requirements, for example in terms of required response times, expected throughput and pattern of usage, terminal population. These requirements should cover the whole life of the SLA.

Capacity Management should help in the negotiation process by providing possible solutions to a number of scenarios. For example, if the terminal population is less than 20 then response times can be guaranteed to be less than two seconds. If more than 20 users connect then extra network bandwidth is needed to guarantee the required response time. Modelling or Application Sizing may be employed here, see Paragraphs 7.3.7 and 7.3.8.

Design, procure or amend configuration

Capacity Management will be involved in the design of new services and will make recommendations for the procurement of hardware and software, where performance and / or capacity are factors. In some instances Capacity Management will instigate the implementation of the new requirement through

Change Management, where it will also be involved as a member of the CAB.
In the interest of balancing cost and capacity, the Capacity Management process will obtain the costs of proposed solutions.

Update CMDB and Capacity Database

The details of the new or amended infrastructure configuration items should be recorded in the CMDB under Change Management. The Change Management process must identify the anticipated throughput that will then be translated into requirements for specific resources. In addition, the Change Management process will identify the performance requirements of the planned change, for example response times. This information should be held as configuration item attributes within the CMDB.

The Capacity Database will be updated to include the technical specification of the procured or amended configuration items, eg disc space, speed of processor, the service performance requirements and expected workloads and demands that are to be placed on the IT resources. From this information thresholds can be identified and monitored. Any threshold breaches and near misses should be addressed by some of the iterative activities of the Capacity Management sub-processes. See Paragraph 7.3.5 for a full description of the Capacity Database.

The CMDB and Capacity Database may be the same database, but even if not a single database, there is data set which will be common between the two databases, for example the CMDB CIs and the attributes that are relevant to the Capacity Management process will be held in both the CMDB and CDB.

Verify SLA

The Service Level Agreement should include details of the anticipated service throughputs and the performance requirements. Capacity Management will provide Service Level Management with targets that have the ability to be monitored and upon which the service design has been based. Confidence that the service design will meet the service level requirements and provide the ability for future growth can be gained by using modelling.

Sign SLA

The results of the modelling activities will provide the verification of service performance capabilities. There may be a need for Service Level Management to renegotiate the service agreement based upon these findings. Capacity Management will provide support to Service Level Management should renegotiations be necessary, by recommending potential solutions and associated cost information. Once assured that the requirements are achievable, it is the responsibility of Service Level Management to agree the service levels and sign the SLA.

Operational system complies with SLA

When the Service Level Agreements have been signed and implemented it is important that the performance and capacity of the system is monitored to ensure it meets the agreed service levels.

Capacity Management should provide Service Level Management with information relating to the system performance as identified in the targets specified in the SLA. At this point the focus of the work in Capacity Management passes to the Service Capacity Management sub-process.

7.2.2 Service Capacity Management

A prime objective of the Service Capacity Management sub-process is to identify and understand the IT services, their use of resource, working patterns, peaks and troughs, and to ensure that the services can and do meet their SLA targets, ie to ensure that the IT services perform as required. In this sub-process, the focus is on managing service performance, as determined by the targets contained in the SLAs or SLRs.

When the business requirements for a service have come through the Business Capacity Management sub-process as described above in Paragraph 7.2.1, and the service has become operational, then the Service Capacity Management sub-process is responsible for ensuing that it meets the agreed service targets. The monitored system will provide data that can identify trends from which normal service levels can be established. By regular monitoring and comparison with these levels, exception conditions can be defined, identified and reported upon. Therefore the Capacity Manager will inform Service Level Management of any service breaches or near misses.

There will be occasions when incidents and problems have been referred to Capacity Management from other Service Management processes, or it has been identified that a service could fail to meet its SLA targets. On some of these occasions the cause of the potential failure may not be resolved by Resource Capacity Management. For example, when the failure is analysed it may be found that there is no lack of resource, or no individual component is over-utilised. However the design or programming of the application is inefficient, and so the service performance needs to be managed, not the individual hardware or software resources.

The key to successful Service Capacity Management is to pre-empt difficulties, wherever possible, before they occur. So this is another sub-process that has to be proactive and anticipatory rather than reactive. However there will be times when it has to react to specific performance problems.
From a knowledge and understanding of the performance requirements of each of the services being run, the effects of changes in the use of services can be estimated, and actions taken to ensure that the required service performance can be achieved.

The activities that need to be carried out as part of this sub-process are described in Section 7.3.

7.2.3 Resource Capacity Management

A prime objective of Resource Capacity Management is to identify and understand the utilisation of each of the component parts in the IT infrastructure. This ensures the optimum use of the current hardware and software resources in order to achieve and maintain the agreed service levels. All hardware components and many software components in the IT infrastructure have a finite capacity, which, when exceeded, has the potential to cause performance problems.

This sub-process is concerned with resources such as processors, memory, discs, network bandwidth, network connections etc. So information on resource utilisation needs to be collected on an iterative basis. Monitors should be installed on the individual hardware and software components, and then configured to collect the necessary data.

As in Service Capacity Management the key to successful Resource Capacity Management is to pre-empt difficulties, wherever possible, before they occur. Therefore this sub-process has to be proactive and anticipatory rather than reactive. However there will be times when it has to react to specific problems that are caused by a lack of resource, or the inefficient use of resource.

From a knowledge and understanding of the use of resource by each of the services being run, the effects of changes in the use of services can be estimated. Then hardware or software upgrades can be budgeted and planned. Alternatively, services can be balanced across the existing resource to make most effective use of the resource currently available.

New technology

Resource Capacity Management also involves understanding new technology and how it can be used to support the business. It may be appropriate to introduce new technology to improve the provision and support of the IT services on which the organisation is dependent. This information can be gathered by attending:

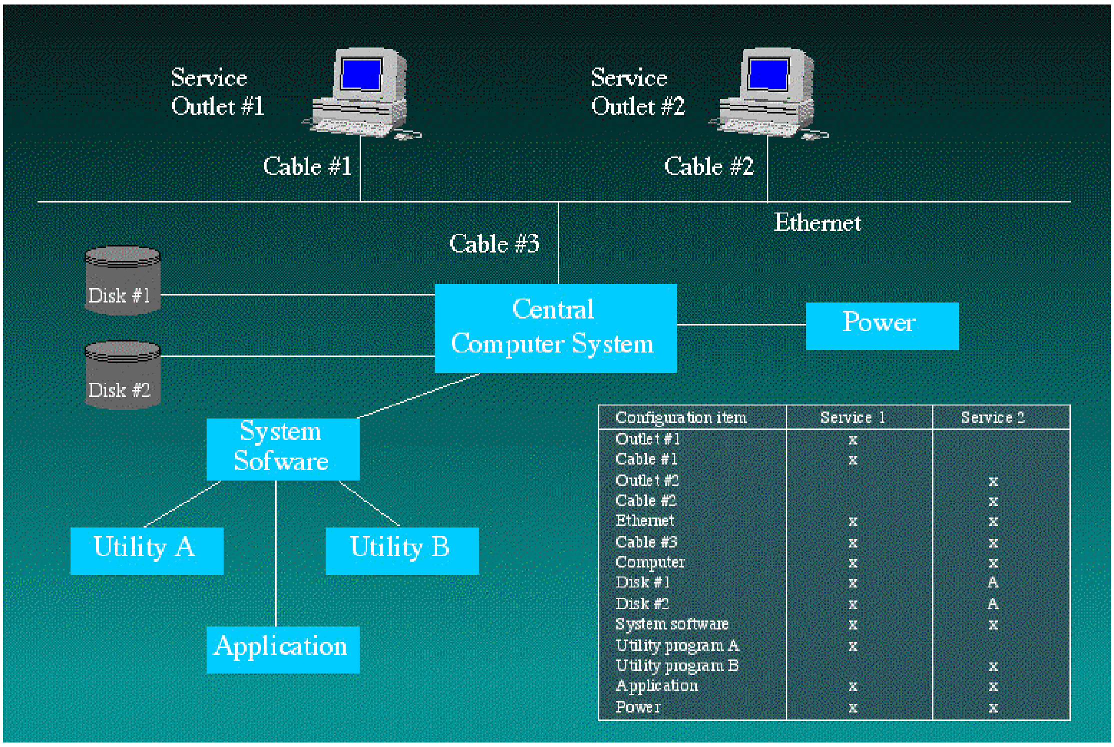
- promotional seminars by hardware and software suppliers
- user group meetings of suppliers of potential hardware and software
- user group meetings for other IT professionals involved in Capacity Management

Each of these forums provide sources of information relating to potential technology, hardware and software, which it might be advantageous for IT to implement for the benefit of the business. However at all times Capacity Management should recognise that the introduction and use of this new technology must be cost-justified and it should be required by the business.

Resilience

The Resource Capacity Management sub-process is also responsible for identifying the resilience inherent in the IT infrastructure or any subset of it. In conjunction with the Availability Management process, Capacity Management should use techniques such as to highlight how susceptible the current configuration is to the failure of individual components and make recommendations on any cost-effective solutions. See Chapter 9 (Availability Management) for a full description of Component Failure Impact Analysis.

Capacity Management will be able to identify the impact on the available resources of particular failures, and the potential for running the most important services on the remaining resources.



Configuration item	Service 1	Service 2
Outlet #1	x	
Cable #1	x	
Outlet #2		x
Cable #2		x
Ethernet	x	x
Cable #3	x	x
Computer	x	x
Disk #1	x	A
Disk #2	x	A
System software	x	x
Utility program A	x	
Utility program B		x
Application	x	x
Power	x	x

Figure 7.5 Sample configuration and basic CFIA grid

Figure 7.5 shows a sample configuration and a basic CFIA grid that identifies the components that are critical in the provision of the services. In the above example the central computer is critical to the provision of both Service 1 and Service 2. For resilience reasons it may be decided to have a number of processors in the central computer, so that if one processor fails the computer can continue to operate with the remaining processors. This may satisfy the resilience requirement, but is there sufficient capacity in the remaining processors to provide the two services to the levels documented in the SLAs?

In the final solution the requirements of both Availability Management and Capacity Management must be addressed.

Ideally the requirements for resilience in the IT infrastructure will be considered at the time of the application design, see Paragraph 7.3.8 Application Sizing. However for many services, the resilience of the service is only considered after it is in live operational use.

7.3 Activities in Capacity Management

The activities described in this section are undertaken when carrying out any of the sub-processes of Capacity Management and these activities can be done reactively or pro-actively.

The major difference between the sub-processes is in the data that is being monitored and collected, and the perspective from which it is analysed. For example the level of utilisation of individual components in the infrastructure will be of interest in Resource Capacity Management, while the transaction throughput rates and response times will be of interest in Service Capacity Management. For Business Capacity Management, the transaction throughput rates for the on-line service need to be translated into business volumes, for example, in terms of sales invoices raised or orders taken.

A number of the activities need to be carried out iteratively and form a natural cycle as illustrated in Figure 7.6.

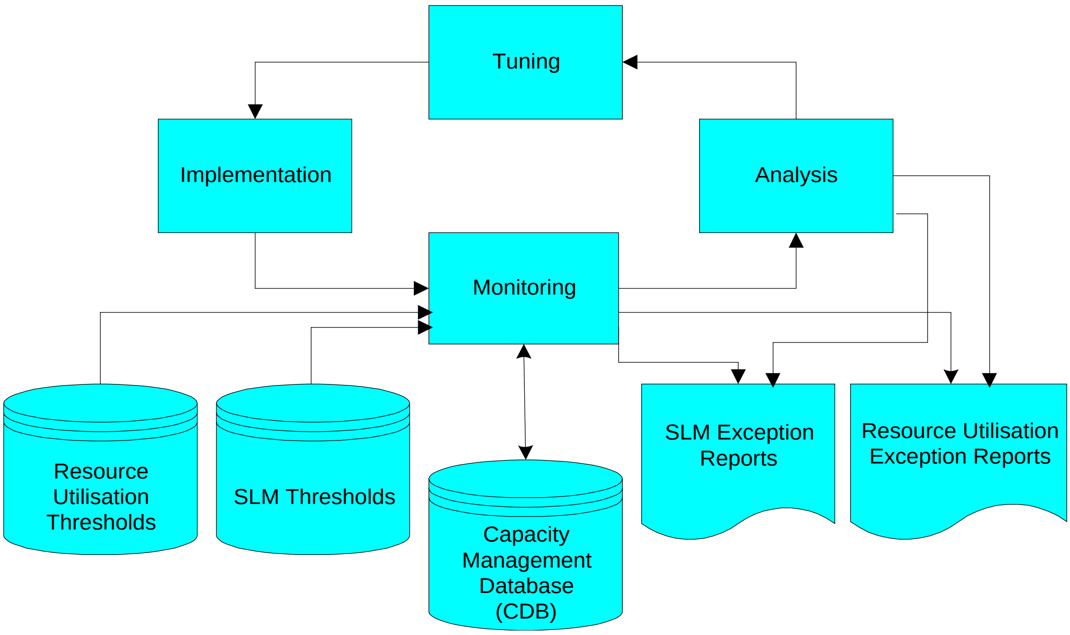


Figure 7.6 Iterative activities in Capacity Management

Monitors should be established on all the components and for each of the services. The data should be analysed, using wherever possible, expert systems to compare usage levels against thresholds. The results of the analysis should be included in reports, and recommendations made as appropriate. Some form of control mechanism may then be put in place to act on the recommendations. This may take the form of balancing services, changing concurrency levels, and adding or removing resource. The cycle then begins again, monitoring any changes made to ensure they have had a beneficial effect and collecting the data for the next day, week, or month.

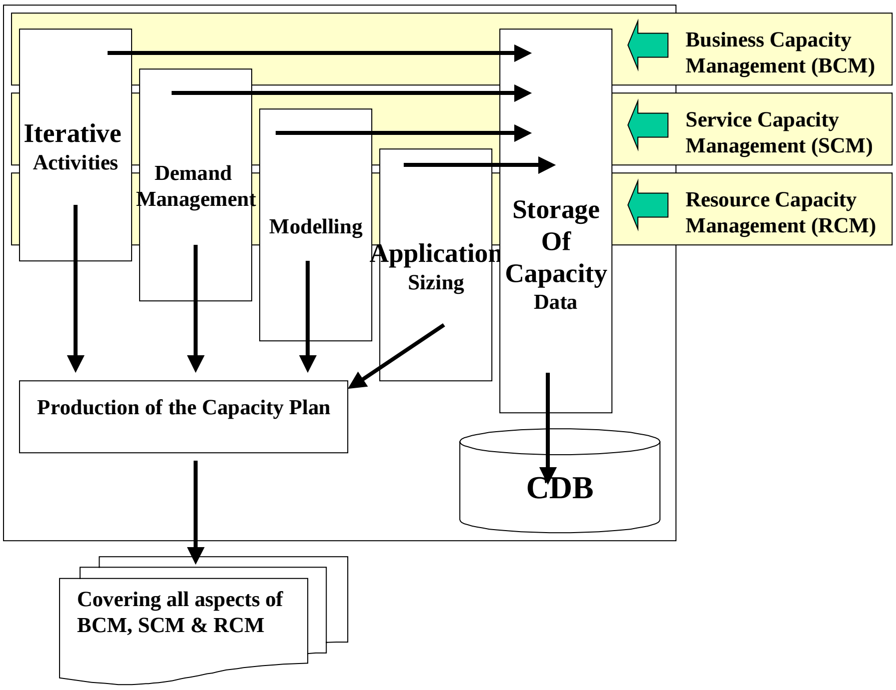


Figure 7.7 Activities in Capacity Management

Figure 7.7 shows the iterative activities as shown is Figure 7.6, together with the other activities of Capacity Management that need to be carried out:

- on a on-going basis: Iterative activities, Demand Management and the storage of data in the Capacity Database
- ad hoc: Modelling and Application Sizing
- regularly: The production of the Capacity Plan

Any one of the sub-processes of Capacity Management may carry out any of the activities, with the data that is generated being stored in the Capacity Database. This section describes all the activities of Capacity Management in detail and shows how the various sub-processes of Capacity Management use each of them as required.

7.3.1 Monitoring

Objective

It is important that the utilisation of each resource and service is monitored on an on-going basis to ensure the optimum use of the hardware and software resources, that all agreed service levels can be achieved, and that business volumes are as expected.

Description

The monitors will be specific to particular operating systems, hardware configurations, applications, etc. Some of the monitors will be free utilities within a hardware or software product, while others form part of a larger systems management tool set and need to by purchased independently. It is important that the monitors can collect all the data required by the Capacity Management process, for a specific component or service.

Typical monitored data will include:

- CPU utilisation
- memory utilisation
- % CPU per transaction type
- IO rates (physical and buffer) and device utilisation

- queue length (maximum and average)
- file store utilisation
- transactions
- transaction per second (maximum and average)
- transaction response time
- batch duration profiles
- number of hits
- number of logons and concurrent users
- number of network nodes in use (eg network devices, PCs, servers etc)

In considering the data that needs to be included, a distinction needs to be drawn between the data collected to monitor capacity (eg throughput), and the data to monitor performance (eg response times). Data of both types is required by the Service and Resource Capacity Management sub-processes.
The data should be gathered at total resource utilisation level and at a more detailed profile for the load that each service places on each particular resource. This needs to be carried out across the whole infrastructure, host or server, the network, local server and client or workstation. Similarly the data needs to be collected for each service.

Part of the monitoring activity will be of thresholds and baselines or profiles of the normal operating levels. If these are exceeded, alarms will be raised and exception reports produced. These thresholds and baselines will have been determined from the analysis of previously recorded data, and can be set on

- individual components, for example monitor that the utilisation of a CPU does not exceed 80% for a sustained period of one hour
- specific services, for example monitor that the response time in an on-line service does not exceed 2 seconds and that the transaction rate does not exceed 10000 transactions per hour.

All thresholds should be set below the level at which the resource is over-utilised, or below the targets in the SLAs. When the threshold is reached, there is still an opportunity to take corrective action before the SLA has been breached, or the resource has become over-utilised and there has been a period of poor performance.

Note:
Many monitors do not report the whole picture, and beware of the monitor that requires so much resource that a processor or memory upgrade is required to get the monitor to run.

Often it is more difficult to get the data on the current business volumes as required by the Business Capacity Management sub-process. These statistics may need to be derived from the data available to the Service and Resource Capacity Management sub-processes.

Response time monitoring

Many SLAs have end-user response times as one of the targets to be measured, but equally many organisations have great difficulty in supporting this requirement. End user response times of IT and network services can be monitored and measured in several ways, as follows:

- By incorporating specific code within client and server applications software: These can be used to provide complete 'end-to-end' service response times or intermediate timing points to break down the overall response into its constituent components. The figures obtained from these tools will give the actual response times as perceived by the end users of a service.

- By using “Robotic scripted systems” with terminal emulation software: These systems consist of client systems with terminal emulation software (eg: browser or VT100 systems) and specialised scripted software for generating and measuring transactions and responses. These systems will generally provide “end to end” service response times and are useful for providing representative response times particularly for multi-phase transactions or complex interactions. These will only give sample response times not the actual response times as perceived by the real users of the system.
- By using distributed agent monitoring software: Useful information on service response times can be obtained by distributing agent systems with monitoring software at different points of a network (eg within different countries on the Internet). These systems can then be used to generate transactions from a number of locations and give periodic measurements of an Internet site as perceived by international users of an Internet web site. However, again the times received are only indications of the response times and are not the real end user response times.
- By using specific passive monitoring systems tracking a representative sample number of client systems: This method relies on the connection of specific network monitoring systems, often referred to as “sniffers” being inserted at appropriate points within the network. These can then monitor, record and time all traffic passing a particular point within the network. Once recorded this traffic can then be analysed to give detailed information on the service response times. Once again however these can only be used to give an approximation to the actual end user response times, although these are often very close to the real world situation but this will depend upon the position of the monitor itself within the IT infrastructure.

In some cases a combination of a number of systems may be used. The monitoring of response times is a complex process even if the service is an in-house service running on a private network. However if this is an external Internet service, the process is much more complex because of the sheer number of different organisations and technologies involved.

Anecdote: A private company with a major Internet web site implemented a web site monitoring service from an external supplier that would provide automatic alarms on the availability and response time of their web site. The availability and speed of the monitoring points were lower than those of the Internet web site being monitored. Therefore the figures produced by the service were of the availability and response time of the monitoring service itself, rather than those of the monitored web site.

Hint: When implementing external monitoring services, ensure that the service levels and performance commitments of the monitoring service are in excess of those of the service(s) being monitored.

7.3.2 Analysis

Objective

The data collected from the monitoring should be analysed to identify trends from which the normal utilisation and service level, or baseline, can be established. By regular monitoring and comparison with this baseline, exception conditions in the utilisation of individual components or service thresholds can be defined, and breaches or near misses in the SLAs can be reported upon. Also the data can be used to predict future resource usage, or to monitor actual business growth against predicted growth.

Description

Analysis of the data may identify issues such as:

- contention (data, file, memory, processor)
- inappropriate distribution of workload across available resource
- inappropriate locking strategy
- inefficiencies in the application design
- unexpected increase in transaction rate
- inefficient use of memory

The use of each resource and service needs to be considered over the short, medium and long-term, and the minimum, maximum and average utilisation for these periods recorded. Typically, the short-term pattern will cover the utilisation over a 24-hour period, while the medium term may cover a one-week to four-week period, and the long term, a year-long period. Over time the trend in the use of the resource by the various IT services will become apparent.

It is important to understand the utilisation in each of these periods, so that changes in the use of any service can be related to predicted changes in the level of utilisation of individual resources. The ability to identify the specific hardware or software resource on which a particular IT service depends, will be improved greatly by an accurate, up-to-date and comprehensive CMDB.

When the utilisation of a particular resource is considered, it is important to understand both the total level of utilisation and the utilisation by individual services of the resource.

Example:
If a processor that is 75% loaded during the peak hour is being used by two different services, A and B, it is important to know how much of the total 75% is being used by each service. Assuming the system overhead on the processor is 5%, the remaining 70% load, could be split evenly between the two services. If a change in either Service A or B is estimated to double its loading on the processor, then the processor would be overloaded.
However if Service A uses 60% and Service B uses 10% of the processor, then the processor would be overloaded if Service A doubled its loading on the processor. But if Service B doubled its loading on the processor, then the processor would not necessarily be overloaded.

7.3.3 Tuning

Objective

The analysis of the monitored data may identify areas of the configuration that could be tuned to better utilise the system resource or improve the performance of the particular service.

Description

Tuning techniques that will be of assistance include:

- balancing workloads - transactions may arrive at the host or server at a particular gateway, depending where the transaction was initiated. Balancing the ratio of initiation points to gateways could provide tuning benefits.

- balancing disc traffic – storing data on disc efficiently and strategically, e.g. striping data across many spindles may reduce data contention
- definition of an accepted locking strategy that specifies when locks are necessary and the appropriate level, e.g. database, page, file, record, and row. Delaying the lock until an update is necessary may provide benefits.
- efficient use of memory – may include looking to utilise more or less memory depending upon the circumstances. A process may utilise resources more efficiently if data is read into memory and manipulated there rather than a sequential read through files. Alternatively, many processes may be contending for memory resource. The excessive demands may lead to increased CPU utilisation and delays while pages are swapped in and out of memory.

Before implementing any of the recommendations arising from the tuning techniques, it may be appropriate to consider using one of the on-going, or ad hoc activities to test the validity of the recommendation. For example, 'Could Demand Management be used to avoid the need to carry out any tuning?' or 'Can the proposed change be modelled to show its effectiveness before it is implemented?'.

7.3.4 Implementation of tuning activities

Objective

The objective of this activity is to introduce to the live operation services, any changes that have been identified by the analysis and tuning activities.

Description

The implementation of any changes arising from these activities must be undertaken through a strict, formal Change Management process. The impact of system tuning changes could have major implications on the customers of the service. The impact and risk associated with these types of changes are likely to be greater than that of other different type of changes. Implementing the tuning changes under formal Change Management procedures will result in:

- less adverse impact on the users of the service
- increased user productivity
- increased productivity of IT personnel
- a reduction in the number of changes that need to be backed-out, but the ability to do so more easily
- greater management and control of business critical application services.

It is important that further monitoring takes place, so that the effects of the change can be assessed. It may be necessary to make further changes or to regress some of the original changes.

7.3.5 Storage of Capacity Management Data - the CDB

The Capacity Management Database (CDB) is the cornerstone of a successful performance management system. Data in the CDB will be stored and used by all the sub-processes of Capacity Management because it is a repository that contains business, service, technical, financial and utilisation data. However the CDB is unlikely to be a single database and probably will exist in several physical locations.

The information in the CDB is used to form the basis of performance and capacity management reports that are to be delivered to management and technical personnel. Also the data will be utilised to generate future capacity forecasts and allow the Capacity Management to plan for future capacity requirements. Capacity and performance data from the necessary system components should be identified or extracted and delivered to a central location where a capacity database is housed. Data from components that make up the service can then be combined for analysis and provision of technical and management reporting.

The inputs to the CDB

There are a number of sources of the data that will be input to the Capacity Database, as follows:

- *Business Data*

  To understand exactly what drives the capacity and performance of an IT system it is essential to have quality business data. The future business plans of the organisation need to be considered, and the effects on the IT services understood. The business data is used to forecast and validate how changes in business drivers affect the capacity and performance of the system. Typically business data includes:

  - Number of accounts supported
  - Number of calls into call centres
  - Number of branches
  - Number of registered users of a system
  - Number of PCs
  - Anticipated workloads
  - Seasonal variations of anticipated workloads
  - Revenue from an e-commerce site
  - Number of web site business transactions

- *Service data*

  It is essential that the Capacity Management process considers at all times the effect that the technical infrastructure has on the work of the end user. To achieve this service-orientated approach to Capacity Management, service data should be stored within the Capacity Management Database. Typical service data is transaction response times that are consistent with the perceived level of service delivered to the end user.

  Other service data is the times taken for batch jobs to be processed. In general the targets in the SLA s and SLRs will provide the service data that the Capacity Management process needs to recorded and monitor against. To ensure that the targets in the SLAs are achieved, SLM thresholds should be included, so that the monitoring activity can measure against these thresholds and raise exceptions reports. By setting the thresholds below or above the actual targets, action can be taken and a breach of the SLA targets avoided.

  The Capacity Management process as well as the Service Level Management process is interested in service data. This will enable the IT provider to be more focused on the delivery of SLA-compliant services. To do this the service data should be correlated with technical and business data to forecast future breaches of service targets.

- *Technical data*

  Most of the components in the IT infrastructure have limitations on the level to which they should be utilised. Beyond this level of utilisation the resource will be over-utilised and the performance of the services using the resource will be impaired. For example the maximum recommended level of utilisation on a CPU could be 80%, or the utilisation of a shared Ethernet LAN segment should not exceed 40%.

Also components will have various physical limitations beyond which greater connectivity or use is impossible. For example the maximum number of connections through a network gateway is 100, or a particular type of disc has a physical capacity of 15Gb.

The technical limits and constraints on the individual components can be used by the monitoring activities as the thresholds at which alarms are raised and exception reports are produced. However care must be exercised when setting thresholds, because many thresholds are dependent on the work being run on the particular component and should not be applied unilaterally.

- *Financial data*

The Capacity Management process requires financial data that can be obtained from a number of sources, for example

- Financial plans, which may indicate long term plans to reduce the costs of IT service provision
- IT budgets, which may include specific budgets for hardware and software expenditure in the next year
- External suppliers, for the cost of new hardware and software component upgrades
- The CMDB, for the purchase or rental costs of current hardware and software components

Most of this data may be available from the Financial Management for IT Services process but Capacity Management needs to consider this information when managing the future business requirements.

- *Utilisation data*

Potentially there is a vast amount of utilisation data available. Ideally data is required that shows the current utilisation of all components of the IT Infrastructure, minute-by-minute, hour-by-hour, day-by-day etc. However after a period of time, for example one week, the minute-by-minute utilisation data will no longer be required. Similarly after one month, the previous month's hour-by-hour data will not be required. However even one year later the average daily utilisation of each component, or service transaction throughput rate and average response times may be required.

So in the collection of the utilisation data, there need to be facilities to enable data to be consolidated or refined, so that data can be deleted as it becomes out-of-date.

The utilisation data needs to be recorded for each component and service, and examples of utilisation data are shown in Table 7.1

Technology	Example metrics available for collection		
Mainframe	CPU utilisation	Paging rates	I/Os per second
Application	No of transactions	Response times	
UNIX server	CU utilisation	Memory utilisation	No of processes
Middleware	Average queue lengths	No of transactions serviced	
Network	Bandwidth utilisation	No of connections	Error rates
Database	Shared memory utilisation	No of queries per second	
PC Client	CPU utilisation	Memory utilisation	

Table 7.1 Example utilisation data

Capacity Management will store data that is relevant to the IT systems, services and the customers. There are many hardware and software tools that monitor systems and store performance data across all technology types and infrastructure components. Some are free utilities within a hardware product, while others form part of a larger systems management tool set. It is important that that any tool is chosen based on the data requirements of the Capacity Management process to be implemented.

The outputs from the CDB

The aim of a Capacity Management Database is to provide the relevant capacity and performance information to the appropriate sub-processes of Capacity Management. In addition, these reports could be used by a number of the other Service Management processes. This information is provided through various reports.

- *Service and component based reports*

  For each infrastructure component there will be a team of technical staff responsible for its control and management, and management staff who are responsible for the overall service. Reports must be produced to illustrate how the service and its constituent components are performing and how much of its maximum capacity is being used.

- *Exception reporting*

  Reports that show management and technical staff when the capacity and performance of a particular component or service becomes unacceptable are also a required output from a CDB. Exceptions can be set for any component, service or measurement that is stored within a CDB. An example exception may be that CPU percentage utilisation for a particular server has breached 70 % for three consecutive hours.

  In particular, exception reports will be of interest to the Service Level Management process in determining whether the targets in SLAs have been breached. Also the Incident and Problem Management processes may be able to use the exception reports in the resolution of incidents and problems.

- *Capacity forecasts*

  To ensure the IT Service Provider continues to provide the required service levels, the Capacity Management process must predict future growth. To do this, future component and service capacity must be forecast. This can be done in a variety of ways depending on the technology used by the component. Changes to workloads by the development of new functionality must be considered alongside growth in workload that is driven by business growth. A simple example of a capacity forecast is a correlation between a business driver and a component utilisation, eg CPU utilisation against the number of accounts supported by the company. Then this data can be correlated to find the effect that increased numbers of accounts will have on the utilisation of particular components of the configuration.

  If the forecasts on future capacity requirements identify a requirement for increased resource, this requirement needs to be input to the IT budget cycle.

7.3.6 Demand Management

Objective

The prime objective of Demand Management is to influence the demand for computing resource and the use of that resource.

Description

This activity can be carried out as a short-term requirement because there is insufficient current capacity to support the work being run, or, as a deliberate policy of IT management, to limit the required capacity in the long term.

Short-term Demand Management may occur when there has been a partial failure of a critical resource in the IT infrastructure. For example, if there has been a failure of part of the memory on a processor, it may not be possible to run the full range of services. However a limited subset of the services could be run. Capacity Management should be aware of the business priority of each of the services, know the resource requirements of each service (in this case, the amount of memory required to run the service) and then be able to identify which services can be run while there is a limited amount of memory available.

Long-term Demand Management may be required when it is difficult to cost-justify an expensive upgrade. For example, many processors are heavily utilised for only a few hours each day, typically between 10: 00 - 12:00 and 14:00 - 16:00. Within these periods, the processor may be over-loaded for only one or two hours. For the hours between 18:00 - 08:00 these processors are only very lightly loaded, and the resource is under-utilised. Is it possible to justify the cost of an upgrade to provide additional resource for only a few hours in 24 hours, or can you influence the demand and spread the requirement for resource across the 24 hours, thereby avoiding the need for the upgrade?

Demand Management needs to understand which services are utilising the resource and to what level, and needs to know the schedule of when they must be run. Then a decision can be make on whether it will be possible to influence the use of resource, and if so, which option is appropriate.

The influence on the services that are running could be exercised by:

- physical constraints. For example, it may be possible to stop some services from running at certain times, or to limit the number of customers who can use a particular service, for example by limiting the number of concurrent users. The constraint could be implemented on a specific resource or component, for example by limiting the number of physical connections to a network router or switch.
- financial constraints. If charging for IT services is occurring, reduced rates could be offered for running work at certain times of the day, that is the times when there is current less demand for the resource.

Demand Management could be carried out as part of any one of the sub-processes of Capacity Management. However Demand Management must be carried out sensitively, without causing damage to the business customers or to the reputation of the IT Department. It is necessary to understand fully the requirements of the business and the demands on the IT services, and to ensure that the customers are kept informed of all the actions being taken.

7.3.7 Modelling

Objectives

A prime objective of Capacity Management is to predict the behaviour of computer systems under a given volume and variety of work. Modelling is an activity that can be used to beneficial effect in any of the sub-processes of Capacity Management.

Description

The different types of modelling range from making estimates based on experience and current resource utilisation information, to pilot studies, prototypes and full scale benchmarks. The former is cheap and a reasonable approach for day-to-day small decisions, while the latter is expensive but may be advisable when implementing a large new project.

- *Trend analysis*

  Trend analysis can be done on the resource utilisation and service performance information that has been collected by the Service and Resource Capacity Management sub-processes. The data can be held in a spreadsheet and the graphical and trending, and forecasting facilities used to show the utilisation of a particular resource over a previous period of time, and how it can be expected to change in the future.

Typically trend analysis will only provide estimates of future resource utilisation information. Trend analysis is less accurate in producing an accurate estimate of response times in which case either analytical or simulation modelling should be used.

- *Analytical modelling*

  Analytical models are representations of the behaviour of computer systems using mathematical techniques, eg multi-class network queuing theory. Typically a model is built using a software package on a PC, by specifying within the package the components and structure of the configuration that needs to be modelled, and the utilisation of the components, eg CPU, memory and discs, by the various workloads or applications. When the model is run, the queuing theory is used to calculate the response times in the computer system. If the response times predicted by the model are sufficiently close to the response times recorded in real life, the model can be regarded as an accurate representation of the computer system.

  The technique of analytical modelling requires less time and effort than simulation modelling, but typically it gives less accurate results. Also the model must be kept up-to-date. However if the results are within 5% accuracy for utilisation and 15 - 20% for on-line application response times, the results are usually satisfactory.

- *Simulation modelling*

  Simulation involves the modelling of discrete events, eg transaction arrival rates, against a given hardware configuration. This type of modelling can be very accurate in sizing new applications or predicting the effects of changes on existing applications, but can also be very time-consuming and therefore costly.

  If you are simulating transaction arrival rates, you can do this by having a number of staff entering a series of transactions from prepared scripts, or by using software to input the same scripted transactions with a random arrival rate. Either of these approaches will take time and effort to prepare and run. However it can be cost-justified for organisations with very large systems where the cost (millions of pounds) and the associated performance implications assume great importance.

- *Baseline models*

  The first stage in modelling is to create a baseline model that reflects accurately the performance that is being achieved. When this baseline model has been created, predictive modelling can be done, ie ask the 'what if?' questions that reflect planned changes to the hardware and / or the volume/variety of workloads. If the baseline model is accurate, then the accuracy of result of the predicted changes can be trusted.

  Effective Service and Resource Capacity Management together with modelling techniques will enable Capacity Management to answer the 'What if' questions. 'What if the throughput of Service A doubles?' 'What if Service B is moved from the current processor onto a new processor - how will the response times in the two services be altered?'

7.3.8 Application sizing

Application Sizing has a finite life-span. It is initiated at the Project Initiation stage for a new application or when there is a major change of an existing application, and is completed when the application is accepted into the operational environment.

Objective

The primary objective of Application Sizing is to estimate the resource requirements to support a proposed application change or new application, to ensure that it meets its required service levels. To achieve this Application Sizing has to be an integral part of the applications lifecycle.

Description

During the initial systems analysis and design the required service levels must be specified. This enables the application development to employ the pertinent technologies and products, in order to achieve a design that meets the desired levels of service. It is much easier and less expensive to achieve the required service levels if the application design considers the required service levels at the very beginning of the application lifecycle, rather than at some later stage.

Other consideration in Application Sizing is the resilience aspects that it may be necessary to build into the design of the new application. Capacity Management will be able to provide advice and guidance to the Availability Management process about the resources required to provide the required level of resilience.

The sizing of the application should be refined as the development process progresses. The use of modelling can be used within the application sizing process.

The service level requirements of the planned application developments should not be considered in isolation. The resources to be utilised by the application are likely to be shared with other services and potential threats to existing SLA targets must be recognised and managed.

When purchasing software packages from external suppliers it is just as important to understand the resource requirements needed to support the application. Often it can be difficult to obtain this information from the suppliers, and it may vary, depending on throughput. Therefore, it will be beneficial to identify similar users of the product and to gain an understanding of the resource implications from them. It may be pertinent to benchmark trial the product prior to purchase.

Key message
Quality must be built in
Some aspects of service quality can be improved after implementation (additional hardware can be added to improve performance, for example). Others – particularly aspects such as reliability and maintainability of applications software – rely on quality being "built in", since to attempt to add it at a later stage is in effect redesign and redevelopment, normally at a much higher cost than the original development. Even in the hardware example quoted above it is likely to have cost more to add capacity after service implementation rather than as part of the original project".

from Quality Management for IT Services, ITIL

7.3.9 Production of the Capacity Plan

Objective

The prime objective is to produce a plan that documents the current levels of resource utilisation and service performance, and after consideration of the business strategy and plans, forecasts the future requirements for resource to support the IT services that underpin the business activities. The plan should also include any recommendations quantified in terms of resource required, cost, benefits, impact etc.

Description

The production and update of a Capacity Plan should occur at pre-defined intervals. Ideally the plan should be published annually, in line with the business or budget lifecycle, and should be completed before the start of negotiations on future budgets. A quarterly re-issue of the updated plan may be necessary to take into account changes in business plans, to report on the accuracy of forecasts and to make or refine recommendations.

The typical contents of a Capacity Plan are described in Annex A - Contents of a Capacity Plan.

7.4 Costs, benefits and possible problems

7.4.1 Costs

The costs associated with establishing a Capacity Management process include:

- procurement of required hardware and software tools including:
  - monitoring tools (performance and utilisation) covering hardware, operating system and application capacity. These tools should be capable of monitoring and collecting data from all required host, network and client environments
  - capacity database for holding a historic record of all service, technical, utilisation, financial and business data
  - modelling tools for performing simulation modelling and statistical analysis
  - graphical and textual reporting tools (often web-enabled)
- project management – implementation of Capacity Management should be treated as a project
- staff costs – recruitment, training and consultancy costs associated with the set-up of the Capacity Management process
- accommodation – provision of an adequate working environment and facilities, which may be distributed

The costs associated with maintaining a Capacity Management include:

- annual maintenance and required upgrades of all hardware and software tools
- on-going staff costs including salaries, further training and ad-hoc consultancy
- recurring accommodation costs such as leasing, rental and energy costs

As discussed earlier Capacity Management responsibilities may be distributed either geographically or functionally within the organisation. Where this is the case, additional costs will be incurred for the central co-ordination and reporting of capacity information.

7.4.2 Benefits

Reduced risk

Reduced risk is a major benefit of Capacity Management. Effective Capacity Management reduces the risk of performance problems and failure in the following ways:

- for existing applications the risk will be minimised through managing the resources and service performance
- the risk to new applications will be reduced through application sizing. As new applications can have an adverse effect upon existing applications, the risk to those applications will also be minimised
- the Capacity Management process will be included on the Change Advisory Board (CAB) to assess the impact of changes upon existing capacity, thus reducing the risk of capacity problems caused by changes
- the number of urgent changes to increase capacity will be reduced, and hopefully eliminated, through effective capacity planning

More confident forecasts

Capacity planning will improve over time. By establishing normal operating baselines and monitoring usage over time, capacity requirements for existing services will become more accurate. Through application sizing and modelling for new services more accurate forecasting and greater confidence will result.

Value to applications lifecycle

Throughout its lifecycle, application development will be influenced by Capacity Management. Additional capacity requirements will be identified during the early development stages and built into the Capacity Plan. This is in contrast to the more usual approach of thinking about capacity just before go-live. Therefore benefits will result in terms of reduced risk and more economic provision of new services.

Increased efficiency and cost savings

Capacity Management will lead to increased efficiency and cost savings in several areas, including:

- deferred expenditure – If you can defer the cost of new equipment to a later date, then the money that is currently in the budget can be spent in other ways. It may be possible to permanently defer the expenditure, so that the money need never be spent. Also, with the pace of technological change, the later you leave a purchase, the more capacity you will get for your money.
- economic provision of services – capacity will be matched to business need. Unnecessary spare capacity is not being maintained and therefore cost savings result. Use of existing capacity will be optimised as far as possible, again resulting in cost savings through not paying for unwanted capacity during quiet usage periods
- planned buying is always cheaper than panic buying !

7.4.3 Possible problems

Over expectation

Customer expectations often exceed technical capability. Therefore it is essential that customer expectations for new applications be managed from the outset. Capacity Management needs to educate customers on the technical feasibility and more importantly, the cost implications of meeting over-ambitious requirements. The opportunity to achieve this is provided as part of the application sizing activity, where requirements and expectations should be discussed and agreed between all parties.

The improvements that can be made through regular tuning may not be significant. However if a service or application has been badly designed or implemented, a large performance gain may possible.
Also Demand Management is often constrained by the requirement for constant on-line access to corporate information. It is not always easy or possible to re-schedule the use of services to quieter off-peak periods. For example, it would be difficult for an internet site selling flowers to influence the demand for flowers on or around St. Valentine’s Day!

Vendor influence

Where budget and sales target deadlines coincide, it is not uncommon to be offered what seems to be the deal of a lifetime i.e. “purchase sufficient capacity for today and tomorrow at yesterdays prices”. On face value, cost efficiencies can be realised, however, before purchasing, remember:

- the pace of change is rapid: Today’s special offer is often tomorrow’s end-of line product
- technological advancement: Tomorrow’s technology will perform faster and have more built-in capacity than today’s technology
- the overall reducing cost of technology: Today's performance and capacity will always be cheaper tomorrow and considerably cheaper in six months time

Manufacturer’s quoted performance figures are often not achievable within a production environment. Care should be taken when negotiating with vendors for additional performance. Where possible, performance figures should be verified before purchase through reference site visits and by simulation testing where appropriate.

Lack of information

Time-to-market demands are ever decreasing and as a result business planning cycles are shorter. This is not an excuse, but a statement of fact about the environment within which Capacity Management must function. Traditionally, it has always been difficult to obtain accurate business forecasts, in order to predict increases and decreases in demand for IT capacity.

However, even the best business planning function cannot always accurately predict demand. There have been numerous recent examples of unprecedented and unpredicted consumer demand for either the latest product or latest information. Internet sites that are suddenly popular, are good examples of consumer demand far outstripping supply and causing failed or drastically delayed delivery of information or products. The Internet is a prime example where customers literally “at the click of a button” are lost forever as they go elsewhere to receive a service, never again to return to a temporarily unavailable or slow site.

It is not possible to provide consistently high quality service levels, cost effectively, without timely, accurate business planning information being made available. However Capacity Management can work effectively, even with crude business estimates. Also it helps if the Capacity Management process understands the business and can talk to the customer in their language. Customers are much more likely to know how many chocolate bars they are going to sell, rather than the amount of CPU seconds they will use in the process. The Capacity Management process will always improve over time.

Capacity Management in a distributed environment

Capacity Management is often considered only as a requirement within the host environment. The network and client environments are not included as part of the Capacity Management process.

Anecdote:
A group of Network Managers were asked “When is the first time you get involved in capacity planning for new applications to be utilised across the network?”. Unfortunately the most common answer was “Usually during the first week of live running, when users complain about poor performance”. Another common answer was “During user acceptance testing, when the application is being tested by users just before transition to live running”.
However when also asked “If you had been involved from the outset, when the new application was just a user requirement, could you have provided accurate figures on spare network capacity within the target environment?” only a handful were confident in their ability to do so.
Result:
There are a number of results and lessons to be learned from the above:

- The network is not within the scope of Capacity Management, but it should be.
- Huge potential for new applications to perform poorly or fail totally.
- Customer perception that IT fails to deliver (again).
- Potential financial costs incurred by the business due to application failure.
- Unpredicted and unbudgeted IT costs to resolve the performance problems, which usually require to be addressed urgently and thus cost even more

Level of monitoring to be implemented

Tools available today provide extremely comprehensive monitoring capabilities. It is possible to monitor most aspects of most components in the IT infrastructure. However careful consideration should be given to the level of monitoring to be undertaken, and the decision should be based upon:

- business impact of component failure - can monitoring be justified on the basis of potential impact of failure ?

- utilisation volatility – is utilisation subject to a steady growth or decline over time, or do highly volatile changes in utilisation occur ?
- ability to monitor components - can monitoring and reporting be automated ?
- cost of component monitoring and reporting – are costs greater than the potential use that can be made of the data collected ?

Question:
Should an organisation with 40,000 desktops in over 1,000 locations monitor spare capacity and performance for every desktop?
Answer:
It is very unlikely that this will prove cost effective, especially when it is realised that none of the desktops store any local data and they all run a standard application set that cannot be added to by users. However there would be some benefit in monitoring at least a sample (if not all) of the servers located in each of the 1,000 locations used to store local data. The level of monitoring to be implemented would need to be based upon business criticality and volatility of the server based applications. The requirement is for exception report monitoring, with reports being produced only when predefined threshold have been reached or exceeded.

7.5 Planning and implementation

The introduction of a Capacity Management process will need to be planned carefully and implemented. Probably some of the activities, such as monitoring and tuning are already taking place. However with many different hardware and software components in the IT infrastructure, it is unlikely that all the necessary platforms will be monitored, and it is probable that some activities in the sub-processes of Capacity Planning are not being carried out.

7.5.1 Review what exists already

It is possible that some parts of the Capacity Management process are already in place in the organisation, so it will be necessary to establish details of the existing procedures and tools. These details will be required to identify the gap between the current functionality and the required functionality. In particular this review should identify:

- current responsibility for any Capacity Management activities
- the tools already in use
- current and desired requirements by other Service Management processes, especially Service Level Management, Availability Management and Financial Management
- current budget and cost-effectiveness
- the management commitment to the introduction of Capacity management

When the review has been completed, it will be possible to produce a report on:

- the assessment of the current situation re Capacity Management
- the improvements that need to be made
- the need for Capacity Management and the benefits that can be expected
- how the improvements can be implemented
- example output from the Capacity Management process
- a project plan , showing timescales, staffing levels, costs and specifying the objectives, main tasks, ongoing activities and output from each part of the process.

When there is a documented statement of what already exists, it will be possible to plan for a full, ITIL-compliant Capacity Management process.

7.5.2 Planning the process

Structure of the Capacity Management process

One of the first decisions that has to be taken is how the Capacity Management process will operate in an environment where typically, the component parts of the IT infrastructure are widely distributed.

If the Capacity Management process is centralised in one location, does the centralised process have sufficient access and control over the components in remote locations that it has to manage? Conversely if the staff who are carrying out the Capacity Management process are co-located with the components, and so widely distributed, can sufficient control be exercised to ensure that the Capacity Management process is working effectively.

The Resource Capacity Management sub-process works best if it is platform based, ie the same members of staff carry out the analysis, tuning and implementation work for a particular technology resource. In general, the staff with technical knowledge to carry out the activity on one hardware platform, operating system and set of applications, will not have the necessary skills to do the work on another platform and different set of operating system and applications. So in many organisations Capacity Management will only have responsibility for the management of these sub-processes. Capacity Management should be identifying the monitoring requirements of this sub-process, while the technical staff carry out the actual implementation.

However the Service and Business Capacity Management sub-processes are best performed by staff who can take an end-to-end view of service performance and business requirements across all the technology resources involved.

This ensures that the necessary data from all the sub-processes is available for Capacity Management to produce all reports that are required, in particular the Capacity Plan.

Monitoring

The review of the current Capacity Management capability will have identified the existing monitors to which the Capacity Management process will have access. Some of the required monitors will have been supplied with the operating systems and applications that are in use, and will only need to be switched on. Other monitors and monitoring equipment will need to be purchased and installed.
Monitors will be required for

- hardware – from PCs, through file servers, up to mainframes and super-computers
- networking equipment (LANs, WANs, bridges, routers etc.)
- peripherals (bulk storage devices, printers etc.)
- software - operating system and network software, in-house developments and purchased packages

However the monitoring of all equipment can be difficult and costly, so the monitoring should be focussed on the components of the infrastructure that support the business critical services. The potential danger of this approach is that there will be no monitoring of some component in the IT infrastructure that nonetheless is critical to the provision of a business critical service, and that subsequently has capacity or performance problems.

The next decision is the data items to monitor, record and store and the intervals over which the recorded data is to be retained. The Resource Capacity Management sub-process will require data on the levels of utilisation of each of the appropriate components of the IT infrastructure. The Service Capacity Management sub-process will require data on the throughput of applications, on-line response times etc. The Business Capacity Management sub-process will require data on the business volumes, such as number of orders processed, deliveries scheduled etc. The data will need to be collected at a variety of intervals, namely per second, per minute, per hour, per day, for measurements such as transactions per second, average utilisation over an hour, etc.

All the required data items need to be identified, and the monitors checked to ensure that they can collect the data.

The CDB

The Capacity Management process will need to collect data from a variety of hardware platforms and software applications that could be widely distributed. Also many different monitor and recording products will be required, which may be incompatible with each other. So how should the CDB be designed?

In essence the Capacity Management process has the following options:

- to get all the data onto a single platform for analysis and storage, or
- to keep the raw data and much of the analysed data on each of the host platforms, and only transfer to a central location, a very limited amount of analysed data.
- To adopt a hybrid of these two approaches above

The monitors will provide the utilisation data for the CDB, but the other potential data will be provided as follows:

- the service data from the SLAs and SLRs of the Service Level Management process
- the business data from the business plans and strategy
- the technical data from the manufacturers' specification of the hardware and software components
- the financial data from the Financial Management process and the CMDB

Other factors that will influence the decision include

- the availability of suitable hardware and software for holding the data centrally, and
- whether a distributed Capacity Management process being implemented

It is probable that the final implementation will result in the CDB being a conceptual database, with a number of physical storage locations. This is likely to occur because not all the various hardware and software monitors will record their data in a format that is easily transferable to another platform.

Also business plans and strategies, and SLAs and SLRs will be written documents containing the information that the Capacity Management process needs. These documents will exist already, but not within the CDB. The necessary data could be extracted from them and then held separately in the CDB, but the problem of keeping the CDB data up-to-date then arises.

Wherever the data is held, disc space needs to be allocated for the CDB, probably in a number of locations. The size of the CDB needs to be considered in each location, together with any requirements for specific software.

It is important to maintain the integrity of the CDB. Therefore it is recommended that identified personnel have responsibility for updating the CDB and other staff have read-only access. It is vital that comprehensive operating instructions are available for all the regular monitoring activities. Recovery instructions should be available to cater for activities that fail to complete. In general, monitoring activities should be automated as much as possible.

Whatever implementation is chosen for the CDB, it is important that the CDB is covered by the Change Management process, so that changes are considered for their impact on the CDB.

The final consideration is for a regular audit of the contents of the CDB. Such an audit should be carried out regularly, and check the contents of the CDB for completeness, ie all the necessary data items are being recorded, and the length of time that information is held for each data item is appropriate.

Integration into the other planning processes

The Capacity Management process will be most effective if it is linked closely with the other IT Service Management planning processes, such as Availability Management and development activities such as

Application Development. This enables Capacity Management to be proactive in the work that it does, which is the key to a successful Capacity Management process. In particular the need to be an integral part of the planning processes is most apparent on the work that needs to be done to produce the Capacity Plan.

Production of the Capacity Plan

A major output from the Capacity Management process is the Capacity Plan. Its regular production needs to be planned, and will depend on the following:

- the need to input to the budgeting cycle
- the accuracy or volatility of the business plans
- the volatility of the IT infrastructure and IT service supported

The Capacity Plan should be produced in a timescale that allows its recommendations to be considered as part of the budget planning cycle, because the Plan may have recommendations that require the cost of the upgrade to be included in budgets etc. Also if the business plans are shown to be inaccurate, or are being updated regularly, it will be necessary to update the Capacity Plan more regularly. Finally if there are frequent, unexpected changes in the IT service provision or in the components in the IT infrastructure, because new business is won or lost, the Capacity Plan may need to be updated more frequently.

All Capacity Plans should be updated at least once a year, and for organisations with very static IT service provision and no unexpected changes in the business plans, this will be sufficient. For organisations with rapidly changing business plans and IT Service provision, the Capacity Plan needs to be updated every three months, and in extreme cases, every month.

7.5.3 Implementation of the process

In the implementation of the Capacity Management process it is necessary to draw a distinction between the activities that will be carried out as part of the implementation of the process, the activities that will be carried out on regularly, and the activities that will carried out as necessary. In the remainder of this section the first two activities of training staff and establishing monitoring and the CDB are initial implementation activities. The activities to be carried out regularly or as necessary, are the activities that are described in Section 7.3, and are carried out as part of the Business, Service and Resource Capacity Management sub-processes.

Train staff

The IT services in most organisations are supported by an IT infrastructure that consists of a large variety of hardware and software resources from many different suppliers. Probably many different operating systems and programming languages are in use, on a wide range of hardware platforms. So a wide range of technical skills will be required to install, set-up and run all the necessary monitors, to analyse the information, and to make and implement tuning recommendations.

The Capacity Management process needs access to staff who have the necessary technical skills of the hardware platforms, operating systems and applications that all the sub-processes of Capacity Management require.

Establish monitoring and the CDB

The monitoring facilities need to be established on each of the hardware platforms and for each of the services. Also the CDB needs to be established, if necessary across a range of platforms. These activities will be carried out as part of the set-up of the Capacity Management process, although there is then an on-going requirement to ensure that the monitoring is occurring at all times and is at the level required. Also it is necessary to ensure that adequate security copies of the CDB are always being taken.

Business Capacity Management

When the necessary data is being produced by each of the Resource, Service and Business Capacity Management sub-processes, the Business Capacity Management sub-process can use the data to produce a series of resource and service utilisation reports and graphs for IT management and the business. Some of the reports will be linked to the achievement of the targets in the SLAs, while others will be confirmation that SLRs can be met. The main document produced by the Business Capacity Management sub-process is the Capacity Plan. All the reports and the Capacity Plan are produced on a regular basis.

In the preparation of any of the reports, Modelling or Application Sizing may have been used to determine the future resource requirements of new services or changes to exiting services or their SLA targets.

Service Capacity Management

When the necessary service monitoring has been established, the data produced needs to be analysed, and if necessary, action taken to tune the service performance. For example, the activities of monitoring and analysis may indicate that the design and implementation of a particular service needs to be tuned for improved on-line response times. Any amendments to the design of the service should be implemented under Change Management and then further monitoring undertaken. These activities are on-going from the first day that each service became operational.

The Service Capacity Management sub-process may need to consider the use of Demand Management if there are short or long-term requirements to influence the demand for resource. This activity will be carried out as necessary.

Resource Capacity Management

When the necessary resource monitoring has been established, the data produced needs to be analysed, and if necessary, action taken to tune the resources used. For example, the activities of monitoring and analysis may produce exception reports for the Service Level Management process that indicate that the tuning of a particular resource is required. Any tuning recommendations should be implemented under Change Management and then further monitoring undertaken. These activities are on-going from the first day that each resource was installed and became operational.

The Resource Capacity Management sub-process will consider new technology and the resilience of the IT infrastructure as required by the Business Capacity Management sub-process.

7.6 Review of the Capacity Management process

The Capacity Management process should be reviewed for effectiveness and efficiency at regular intervals to ensure that

- it is producing the required output at the required times for the appropriate audience, and
- its activities are cost effective

Metrics

The ideal indicator for the success of the Capacity Management process is that sufficient IT capacity exists at all times, to provide customers with the agreed level of service. However more specific targets and metrics should be identified, of which the following are some examples:

- The utilisation of all components and services is being recorded in the CDB.
- The correct amount of utilisation data is being collected in the CDB. Too much data and the collection overhead becomes unacceptable and filestore is wasted, too little data and investigations into incidents and problems may be unsuccessful, and Capacity Plans may be inaccurate.
- All recommendations for tuning actions are accompanied with predictions of their likely effect. The actions are reviewed for their success and the results documented.

- The Service Level Management process is informed of any potential or actual breaches of the targets in the SLAs.
- Constraints that are imposed on demand occur with the understanding of the customers and do not seriously affect the credibility of the IT Service Provider.
- All regular management reports are produced on time and all ad hoc reports are produced within the agreed timescales.
- The annual Capacity Plan is produced on time and is accepted by the senior IT management.
- Recommendations for hardware or software upgrades that are identified in the Capacity Plan are accurate, both in terms of the financial cost and the timescale in which they are required.

Critical Success Factors

Success in Capacity Management is dependent on a number of factors:

- accurate business forecasts
- knowledge of IT strategy and plans, and that the plans are accurate
- an understanding of current and future technologies
- an ability to demonstrate cost effectiveness
- interaction with other effective Service Management processes
- an ability to plan and implement the appropriate IT capacity to match business need

Key Performance Indicators (KPIs)

Success of the process can be measured by producing the following KPIs in support of the CSFs

- resource forecasts
  - timely production of forecasts of resource requirements
  - accurate forecasts of trends in resource utilisation
  - incorporation of business plans into Capacity Plan
- technology
  - ability to monitor performance and throughput of all services and components, as appropriate
  - implementation of new technology in line with business requirements (time, cost and functionality)
  - the use of old technology does not result in breached SLAs due to problems with support or performance
- cost-effectiveness
  - a reduction in panic buying
  - no significant over-capacity that can’t be justified in business terms
  - accurate forecasts of planned expenditure

- plan and implement the appropriate IT capacity to match business need
  - reduction in lost reduction in thb incidents due to poor performance
  - reduction in lost business due to inadequate capacity
  - new services are implemented which match Service Level requirements (SLRs)
  - recommendations made by Capacity Management are acted upon

7.7 Interfaces with other SM processes

Capacity Management is closely related to business requirements. As such it is a vital element of the planning process and hence has close links with all aspects of Service Delivery. However Capacity Management maintains close links with Service Support to provide support for all operational performance and capacity issues. The major Service Management interfaces are defined below.

7.7.1 Incident Management

Incident Management keeps Capacity Management informed of incidents related to capacity and performance. Capacity Management supports the Incident Management process by resolving and documenting capacity related incidents.
Co-ordinated through Problem Management, Capacity Management provides the Incident Management with diagnostic scripts and diagnostic tools to assist with Incident Management. For example, real-time performance and capacity monitoring tools will be accessible on the Service Desk. As a result Capacity Management will keep the Incident Management and Problem Management processes informed of any potential performance or capacity problems through automatic alerts or recording known errors.

7.7.2 Problem Management

Capacity Management provides a specialist support role to identify, diagnose and resolve capacity related problems. As discussed above, Capacity Management utilises and makes available, diagnostic scripts and tools to support both the Incident and Problem Management processes. Capacity problems and known errors are documented and made available through Problem Management.
Capacity Management supports the pro-active role of Problem Management through the analysis of performance and capacity information to identify any significant trends.

7.7.3 Change Management

Capacity Management will be represented on the Change Advisory Board, to assess the impact of changes on existing capacity and to identify additional capacity requirements. The cumulative effect of changes upon capacity will need to be closely monitored by Capacity Management.
Additional capacity requirements will need to be included in the Capacity Plan and as such treated as Requests For Changes (RFCs) in their own right. Capacity Management will raise RFCs for any planned upgrades, tuning activities and additional use of monitoring tools throughout the infrastructure.

7.7.4 Release Management

Capacity Management helps determine the distribution strategy, particularly where the network is used for distribution. Factors such as network bandwidth, host and target capacity, distribution window and number of targets need to be considered as part of the distribution strategy. Capacity Management provides the necessary planning data and technical expertise to support the strategy on an ongoing basis and for individual distribution requirements.
The Release Management checklist includes the following check:

> “Have all the Capacity Management issues been considered and planned? For example,
>
> response times, storage requirements and LAN traffic”

Clearly Capacity Management would provide the response to this request.

Immediately prior and post distribution, automated tools can be used to perform capacity audits and to delay distribution and implementation if there is insufficient capacity. Where automated tools are not used, these checks should still be performed manually to prevent capacity related incidents being raised.

7.7.5 Configuration Management

Conceptually the Capacity Database forms a subset of the Configuration Management Database (CMDB). The technical, service, utilisation, financial and business data all pertain to attributes of Configuration Items (CIs) maintained in the CMDB. Without this information on individual CIs, Capacity Management cannot function effectively. Information provided by Capacity Management will be made available to other processes via the CMDB. For example, Problem Management will require access to current and historic, service utilisation data to resolve capacity problems. Service Level Management will require access to current and historic performance and workload data in order to set achievable SLA targets.

7.7.6 Service Level Management

Capacity Management supports Service Level Management to ensure that performance and capacity targets for new or changed requirements can be achieved. Performance is dependent on a given workload, therefore both are required in an SLA with specific performance targets. Similarly there may be a requirement for Capacity Management to assist Service Level Management in drafting and reviewing Operational Level Agreements and External Contracts where capacity or performance issues are involved. For operational services, Capacity Management monitors and reports on performance and throughput to Service Level Management. Where weak areas are identified, Capacity Management provides technical input to the overall Service Improvement Programme to improve service performance.

7.7.7 Financial Management for IT Services

Capacity Management is concerned with economic provision of services. Financial justification and efficiency gains will be demonstrated with the assistance of Financial Management. A cost summary will be provided as part of the Capacity Plan. Procurement recommendations included in the Capacity Plan will be compared with budget forecasts and actuals before purchase.

In organisations where customers are charged according to resource utilisation Capacity Management provides Financial Management with the usage profiles that form the basis for charging. Capacity Management will provide assistance with the definition and execution of capacity related charging calculations. The use of modelling to predict demand will also provide revenue estimates.

7.7.8 IT Service Continuity Management

Capacity Management determines the capacity required for all recovery options used. The minimum hardware and software configurations required will be defined to provide the required performance and throughput levels following an invocation.

It is essential that the capacity identified for recovery is maintained in line with that used in the live environment. In one recent intermediate recovery test, one organisation arrived at the third party recovery site to find the pre-negotiated 9 disk packs available for recovery. Their current live operation was configured with 36 disk packs! The Capacity Plan should incorporate IT Service Continuity requirements and RFCs should be assessed for their impact on any recovery options.

7.7.9 Availability Management

Performance and capacity problems result in unavailability i.e. unacceptably slow performance is effectively the same as unavailability. Therefore Capacity and Availability Management share common goals and complement each other. Ideally Availability and Capacity Management should be aligned as there are many interdependencies. For example, implementing additional resilience will require consideration of the associated capacity requirements. Capacity Management needs to be aware of availability techniques deployed, such as mirroring or duplexing, in order to plan accurately for capacity. The tools used by Availability and Capacity Management will often be shared. Both will require access to common planning, monitoring and alerting tools along with the CMDB. Both processes commonly use techniques such as Component Failure Impact Analysis (CFIA) and Fault Tree Analysis (FTA). See Chapter 9 (Availability Management) for details of these techniques.

7.7.10 Application Design and Development

If the Service Level Requirements of new applications are to be met, the Capacity Management process has a crucial role to play in the application design and development lifecycle. Ideally Capacity Management will be involved at the earliest stages of application design, ensuring that the proposed hardware, software and application design will enable the operational service to perform as required. As the development progresses, Capacity Management can check that predicted performance is be achieved, and if not, can recommend changes to ensure that the required performance can be achieved. Similarly if the application is bought-in, the Capacity Management process should ensure that capacity and performance issues are considered when the evaluation phase is being carried out.
It is much easier and less expensive to achieve the Capacity Management requirements of any new application, if these requirements are considered at the beginning of the application lifecycle, rather than at some later stage.

Annex A – Contents of a Capacity Plan

The Capacity Plan should be published annually in line with the budgetary cycle. Ideally it should be updated quarterly. This will take extra effort, but, if it is regularly updated, the Capacity Plan is more likely to be accurate and to reflect the changing business need.

Introduction

This section should briefly explain the background to this issue of the Capacity Plan. For example:

- the organisation's current levels of capacity
- problems being experienced or envisaged due to over or under capacity
- the degree to which service levels are being achieved
- what has changed since the last issue of the plan.

The introduction should also include the following sections:

Scope of the plan

Ideally, the Capacity Plan should encompass all IT resources. This section should explicitly name those elements of the IT infrastructure that are included.

Methods used

The Capacity Plan uses information gathered by the sub-processes. This section therefore should contain details of how and when this information was obtained, for example business forecasts obtained from business plans, workload forecasts obtained from users, service level forecasts obtained by the use of modelling tools.

Management summary

Much of the Capacity Plan will, by necessity, contain technical detail that will not be of interest to all readers of the plan. The management summary should highlight the main issues, options, recommendations and costs. It may be necessary to produce a separate executive summary document that contains the main points from each of the sections of the main plan.

Business scenarios

It is necessary to put the plan into the context of the current and envisaged business environment. For example, a British airline planned to move a large number of staff into its headquarters building. A ratio of 1.7 people per desktop terminal was forecast. Capacity Management was alerted and was able to calculate the extra network traffic that would result.
It is important to explicitly mention all known business forecasts so that readers can determine what is in and what is outside the scope of the plan.

Service summary

The service summary section should include the following:

Current and recent service provision

For each service that is delivered, provide a service profile. This should include throughput rates and the resulting resource utilisation, for example of memory, storage space, transfer rates, processor usage, network usage. Short term, medium term and long-term trends should be presented here.

Service forecasts

The business plans should provide the Capacity Manager with details of the new services planned, and the growth or contraction in the use of existing services. This section should report on new services and the demise of legacy systems.

Resource summary

The resource summary section should include the following:

Current and recent resource usage

This section concentrates on the resulting resource usage by the services. It will report, again, on the short, medium and long-term trends in resource usage, broken down by hardware platform. This information will have been gathered and analysed by the sub-processes of management of resources and service performance and so should be readily available.

Resource forecasts

This section forecasts the likely resource usage resulting from the service forecasts. Each business scenario mentioned above should be addressed here. For example, a carpet wholesale business in the North of England could accurately predict what the peak and average processor usage would be before they decided to take over a rival business. It was proved that an upgrade would not be required. This was fed into the cost model, leading to a successful take-over.

Options for service improvement

Building on the results of the previous section, this section outlines the possible options for improving the effectiveness and efficiency of service delivery. It could contain options for merging different services on a single processor, upgrading the network to take advantage of technological advances, tuning the use of resource or service performance, rewriting legacy systems, purchasing new hardware or software etc.

Cost model

The costs associated with these options should be documented here. In addition the current and forecast cost of providing IT services should be included. In practice the Capacity Manager will obtain much of this information from the Financial Management process and the IT Financial Plan.

Recommendations

The final section of the plan should contain a summary of the recommendations made in the previous plan and their status, for example rejected, planned, implemented. Any new recommendations should be made here ie which of the options mentioned in the plan is preferred.
The recommendations should be quantified in terms of

- the business benefits to be expected
- the potential impact of carrying out the recommendations
- the risks involved
- the resources required
- the cost, both set up and on-going

Annex B- Capacity Management's role

Responsibilities

Capacity Management has overall responsibility for ensuring that there is adequate IT capacity to meet required levels of service and for ensuring that senior IT management is correctly advised on how to match capacity and demand, and to ensure that use of existing capacity is optimised.
Capacity Management is also responsible for advising the Service Level Management process about appropriate service levels or service level options.

Tasks

Capacity Management should

- ensure that appropriate levels of monitoring of resources and system performance are set, and the information recorded in a CBD that is kept up-to-date and used by all parts of the Capacity Management process
- produce capacity plans in line with the organisation’s business planning cycle, identifying capacity requirements early enough to take account of procurement lead times.
- document the need for any increase or reduction in hardware based on service level requirements and cost constraints
- produce regular management reports which include current usage of resources, trends and forecasts
- size all proposed new systems to determine the computer and network resources required, to determine hardware utilisation, performance service levels and cost implications
- assess new technology and its relevance to the organisation in terms of performance and cost
- assess new hardware and software products for use by Capacity Management that will improve the efficiency and effectiveness of the process
- carry out performance testing of new systems
- report on performance against targets contained in SLAs
- maintain a knowledge of future demand for IT services and predict the effects of demand on performance service levels
- determine performance service levels that are maintainable and cost justified
- recommend tuning of systems and make recommendations to IT management on the design and use of systems to help ensure optimum use of all hardware and operating system software resources
- recommend resolutions to performance-related incidents and problems
- recommend to IT management when to employ Demand Management, to dampen customer demands on systems
- carry out ad-hoc performance and capacity studies on request from IT management
- ensure requirements for reliability and availability are taken into account in all capacity planning and sizing activity
- be represented on the CAB, assessing and authorising changes
- ensure that regular and ad hoc audits are carried out on the Capacity Management process

8 IT Service Continuity Management

8.1 Introduction

Since the IT Infrastructure Library produced its book on 'Contingency Planning', there have been significant changes in technology and the way in which technology is used within business. The dependencies between business processes and technology are now so inter-twined that Contingency Planning (or Business Continuity Management as it is now sometimes referred) incorporates both a business element (Business Continuity Planning) and a technology element (IT Service Continuity Management Planning). Their dependencies on each other determine that one is a sub-set of the other, depending on the nature of the business and the extent to which technology has pervaded the organisation. In this chapter it is assumed that business continuity is the main driver and that IT Service Continuity Management is a sub-set of the Business Continuity Management process.

8.1.1 Why IT Service Continuity Management?

In today's highly competitive and service oriented business environment, organisations are judged on their ability to continue to operate and provide a service at all times. IT Service Continuity Management is concerned with managing an organisation's ability to continue to provide a pre-determined and agreed level of IT Services to support the minimum business requirements following an interruption to the business. This may range from an application or system failure, to a complete loss of the business premises. As such, IT Service Continuity Management forms an integral part of the Business Continuity Management process to ensure that IT services and facilities can be provided.

8.1.2 Mission for IT Service Continuity Management

The mission for IT Service Continuity Management is to support the overall Business Continuity Management process by ensuring that the required IT technical and services facilities (including computer systems, networks, applications, telecommunications, technical support and service desk) can be recovered within required, and agreed, business timescales.

8.1.3 Scope for IT Service Continuity Management

IT Service Continuity Management focuses on the IT Services required to support the critical business processes. The impacts of a loss of a business process, such as financial loss, damage to reputation or regulatory breach, are measured through a Business Impact Analysis, which determines the minimum critical requirements. The specific IT technical and service requirements are supported by IT Service Continuity Management. The scope of IT Service Continuity Management within an organisation will be determined by the organisational structure, culture and strategic direction (both business and technology) in terms of the services provided and how these will develop and change over time.

8.1.4 Basic Concepts of IT Service Continuity Management

As organisations become more dependent upon technology, which is now a core component of most business processes, continued availability of IT is critical to their survival. This availability is accomplished through a balance of risk reduction measures such as resilient systems, and recovery options including back-up facilities. Successful implementation of IT Service Continuity Management can only be achieved with visible Senior Management commitment and the support of all members of the organisation. Ongoing maintenance of the recovery capability is essential if it is to remain effective. This is achieved through:

- a rigorous configuration, change management and review process;
- education and awareness for the whole organisation;
- Utilising the latest technology and software supporting tools;
- specific training for personnel involved in the process;
- regular testing.

8.1.5 Benefits of IT Service Continuity Management

ITSCM supports the BCM process and delivers the required IT supporting infrastructure to enable the business to continue to operate following an service disruption.

ITSCM must be an integral part of the overall business process especially when a business is highly IT dependent.

The annual spend on ITSCM can be likened to an insurance premium and, like insurance, the optimum spend will be determined by the circumstances and risks that could influence the organisation's business. The variety and frequency of events that pose a threat to businesses in today's technology age have forced organisations to regard the continuing requirement for ITSCM as part of their corporate and management philosophy and culture.

This allows an organisation to identify, assess and take responsibility for managing its risks, thus enabling an organisation to understand better the environment in which it operates, decide which risks it wishes to counteract and act positively to protect the interests of all stakeholders (staff, customers, shareholders, third parties, creditors). ITSCM can complement this activity and help to deliver business benefit. For example:

Management of risk: The IT department can actively manage the infrastructure and systems to reduce the impact of component, multiple component or site failure (e.g. RAID implementation on disks, multiple processors/power supplies, mirroring or clustering of systems, etc.). The consequence of not managing this and other risks effectively is likely to impact on the ability of the organisation to meet its client's expectations, potentially resulting in lost customers, revenue and market share.

- **Potential lower insurance premiums:** The IT department can help the organisation demonstrate to underwriters or insurers that they are pro-actively managing down their business risks. Therefore the risk to the insurance Organisation is lower and the premiums due should reflect this. Alternatively, the organisation may feel comfortable in reducing cover or self-insuring in certain areas as a result of limiting potential losses.
- **Regulatory requirements:** In some industries a recovery capability is becoming a mandatory requirement (for example, regulators stipulate that financial organisations require sufficient continuity and security controls to meet the business requirements) and failure to demonstrate tested business and ITSCM facilities could result in heavy fines or the loss of trading licenses. Within the service community, there is an obligation to provide

continuous services, e.g., hospitals, emergency services (police, fire and ambulance services) and prisons.

- **Business relationship:** The requirement to work closely with the business to develop and maintain a continuity capability fosters a much closer working relationship between IT and the business areas. This can assist in creating a better understanding of the business requirements and the capability of IT to support those requirements.
- **Positive marketing of contingency capabilities:** Being able to demonstrate effective ITSCM capabilities will enable an organisation to provide high service levels to clients and customers and thus win business.
- **Organisational credibility:** There is a responsibility on the directors of organisations to protect the shareholders' interest and those of their clients. Contingency facilities increase an organisation's credibility and reputation with customers, business partners, stakeholders and industry peers. For example, the growth of call centres in many organisations has meant that the need to maintain customer communications at all times is vital to the ability to retain customer confidence and loyalty.
- **Competitive advantage:** Service organisations are increasingly being asked by business partners, customers and stakeholders to demonstrate their contingency and may not be invited to tender for business unless they can demonstrate appropriate recovery capabilities. In many cases this is a good incentive for customers to continue a business relationship and becomes a part of the competitive advantage used to win or retain customers.

In many industries, consumers and business customers are requesting an ability to make enquiries, place orders, monitor progress and settle bills with much less human intervention than is currently the case. Technology allows products to be tailor-made to customers' exact requirements and advanced just-in-time production systems allow inventories to be reduced to a minimum.

These changes will make business even more dependent on technology and ITSCM than ever before. Failure of a centralised electronic order capture system or an inventory management system will stop a business in its tracks. In the digital world, "no systems" will mean "no business". In order to address these risks, resilience, security, Business Continuity and the ITSCM components will have to be designed into electronic business services from the outset in a way that is rarely done today.

For example:

(1) If a mobile phone provider were to lose the messaging capability on its service, this would be unacceptable to many customers who may transfer to another service provider either immediately or when the contract renews.

(2) Supermarkets rely on just-in-time to provide stock control and ordering. A failure at any stage in the supply chain management, such as bar code scanners, could mean that stocks of perishable items such as milk are not re-ordered with the effect of no milk on the shelves within a very short period of time and a consequent loss of customer income, confidence and loyalty.

(3) Profit for a financial institution involved in money market trading is dependent upon decisions taken based on split second availability of pricing information delivered through

the computer infrastructure. If this information was not available, the organisation would immediately lose its ability to make informed decisions and could potentially begin to lose money. Similarly, if IT services fail in a heavily automated manufacturing plant utilising just-in-time processing, the plant will almost certainly lose production time and therefore money.

(4) In the electronic commerce environment, if an organisation were to lose the availability of its Internet site then this could lead to an immediate loss of business and customers as they will immediately find alternative sites and organisations offering similar products and services. In addition, customer loyalty cannot be relied on and once customers are lost, they are unlikely to return.

8.1.6 Business Continuity Management and IT Service Continuity Management

Business Continuity Management (BCM) is concerned with managing risks to ensure that at all times an organisation can continue operating to, at least, a pre-determined minimum level. The BCM process involves reducing the risk to an acceptable level and planning for the recovery of business processes should a risk materialise and a disruption to the business occur.

IT Service Continuity Management (ITSCM) must be a part of the overall BCM process and is dependent upon information derived through this process. ITSCM is focused on the continuity of IT Services to the business. BCM is concerned with the management of Business Continuity that incorporates all services upon which the business depends, one of which is IT.

There is a need for the minimum business requirements to be determined to a level of detail and agreed between the business and the IT service providers (internal or external) prior to the ITSCM scope being defined. These requirements may define a need to establish an immediate transfer of the service to an alternative location or a requirement to recover elements of the service over a longer period of time (e.g. a week). It is vital that these prerequisites are fully understood, defined and agreed by the business as part of the BCM process to ensure ITSCM is specifically used in the most effective and efficient manner to deliver the IT portion of these requirements.

Business requirements drive IT Service Continuity Management provision. Provide what the business wants, not what the IT community think they need.

8.1.7 Management Commitment

The implementation of ITSCM mechanisms, identified through the BCM process, must be based on business requirements in order to reduce the risk from interruption or failure of critical business processes and ensure business benefit. The process is concerned with the performance and, at the most elementary level, the survival of the business and is, therefore, a top management issue. The process must be understood and visibly supported by all senior managers and directors.

Get senior business management on board – without them effective IT Service Continuity Management is impossible to achieve.

However, management commitment does not end when the initial project has been completed. Ongoing commitment is necessary to maintain the effectiveness of the recovery capabilities and there is a requirement for a continual review, update and test process for the ITSCM plans, facilities and service provision. Providing this is achieved, ITSCM will react to changes in the business world that it is there to protect. Business continuity and recovery must be considered as part of the culture of the organisation and indeed be part of business as usual. To help achieve this, responsibility for ITSCM should be a key objective in a business manager’s job description.

Business-as-usual pressures tend to focus management time on the business itself and ITSCM typically takes a back seat. It is essential that the necessary resources are focussed through a strategic directive from the highest levels of management to prevent business as usual pressures being used as an excuse for the non-delivery of the recovery facilities.

To this end, ITSCM becomes one of the multitude of corporate management activities performed in operational and executive roles, enabling ITSCM to be embedded, either explicitly or implicitly, within business strategy, business objectives and key performance indicators through all levels of the organisation. Similarly, as the IT strategy is generally a derivative of the business strategy, this should reflect those changes in terms of ITSCM.
Failure to include the ITSCM process as part of the corporate planning activities will result in the ITSCM mechanisms becoming inaccurate or ‘out of date’ and failing to meet business requirements.

8.1.8 Relationship with other IT Service Disciplines

IT Service Continuity Management interacts with the other IT Service disciplines such as:

- Service Level Management - understanding the obligations of IT service delivery;
- Availability Management - delivering risk reduction measures to maintain ‘business as usual’;
- Configuration Management – defining the core infrastructure;
- Capacity Management – ensuring that business requirements are fully supported through appropriate IT hardware resources;
- Change Management – ensuring the currency and accuracy of the continuity plans through established processes and regular reviews;
- Service Desk/Incident Management – utilising historical data (statistics).

8.2 Scope of IT Service Continuity Management

ITSCM primarily considers those IT assets that support the key business processes. However, the installation of mechanisms to deliver ITSCM will not necessarily be sufficient to keep those business processes operating after an service disruption. Should it be necessary to relocate to an alternative working location, provision will also be required for items such as office and personnel accommodation, copies of critical paper records, courier services and telephone facilities to communicate with customers and third parties.

The ITSCM management process should identify the required and agreed minimum level of business operation following an service disruption, along with a requirements definition covering systems, facilities and service requirements. The process then examines the risks and threats to these requirements and develops an IT risk reduction or mitigation programme. This programme will implement mechanisms delivering the continuity requirements necessary to provide the required optimum level of business operation. These mechanisms may include splitting a data centre over more than one location, implementing disk mirroring (or other levels of RAID as required), system replication to a standby site, dual routing of communications links, installing secondary communications links, or provision of standby power supply (UPS and/or generator).

An organisation's structure, culture and strategic direction (both business and technology) are key drivers in determining the scope of ITSCM. Significant benefit can be derived from the involvement of someone with good specific business and infrastructure knowledge and experience who can ensure that these are considered.

Scope considerations will include:

- the organisation's dependence on technology, its infrastructure and any external providers of support services;
- the number and location of the organisations offices and the services performed in each;
- the number of critical business processes and the level of integration between them;
- the level of services that need to be provided to the business to support those critical business processes;
- any limitations in the provision of ITSCM mechanisms (e.g. cost of downtime);
- the organisation's attitude towards risk.

At the broadest level, the scope of ITSCM is usually defined in terms of the:

- business processes to be covered and their IT support requirements (e.g. systems, networks, communications, support staff skills, data and documentation etc.);
- risks that need to be addressed.

8.2.1 Risks 'in scope'

Organisations continually face risks, ranging from a localised service disruption in a single department to major service disruptions that affect multiple organisations and communities. As the business activities and infrastructure of an organisation change, so do the business processes and the organisation's risk assessment profile. The risks covered by ITSCM tend to be those that could result in serious disruption to business processes, for example the loss of, or denial of access to, IT systems or networks. These risks are discussed in more detail later in the chapter.

The likelihood of events such as these happening has been proven over the years. Talking to Building Services and IT Operations Support managers will provide an insight into the frequency, types and nature of regular service disruptions. In addition, the press has provided substantial coverage of major service disruptions, from terrorist activities to natural disasters and infrastructure problems. Below is a brief list of high profile events that have caused significant problems to organisations over the years:

Poison Gas	Tokyo Underground System, Japan (March, 1995)
Power Loss	Auckland, New Zealand (December, 1997)
Earthquake	Kobe, Japan (January, 1995)
	Los Angeles, USA (January, 1994)
Bomb	World Trade Centre, New York, USA (February, 1993)
	Oklahoma City, Oklahoma, USA (April, 1995)
	Docklands, London, England (February, 1996)
	Bishopsgate, London, England (April, 1993)
	Manchester, England (June, 1996)
Flood	Bangladesh (July, 1996)
	Pakistan (August, 1996)
Technical Failure	London Stock Exchange (2000)
	Web site denial of service attacks e.g. Yahoo (2000)

In the late 1970s a major earthquake in Romania left 1000 dead in Bucharest alone. This is a phenomenon which recurs approximately every 40 years. There were no major IT networks in Romania 20 years ago. It is almost certain that there will be a major earthquake by 2020 that will destroy all existing IT infrastructure within the fault area. Contingency for them is not a potential luxury, it is absolute survival which affects not only existing IT infrastructure but every significant IT change.

8.2.2 Risks 'out of scope'

ITSCM does not usually directly cover longer-term risks such as those from changes in business direction, diversification, restructuring, and so on. While these risks can have a material impact on IT service elements and their continuity mechanisms,management usually has some time to identify and evaluate the risk and include risk mitigation through changes or shifts in business and IT strategies, thereby becoming part of the change management programme.

Similarly, ITSCM would not usually cover minor technical faults (for example, non critical disk failure), unless there is a possibility that the impact could have a material impact on the business. These risks would be expected to be mainly covered through the Service Desk and Incident Management process, or resolved through the planning associated with the disciplines of Availability Management, Problem Management; and to a lesser extent through Change Management, Configuration Management and 'day to day' operational management.

8.2.3 Roles and Responsibilities

The initial implementation of ITSCM will typically be progressed as a project and will need to be considered as a part of other projects, whereas, on an ongoing basis ITSCM will evolve into operational responsibilities. Without the roles and responsibilities being addressed, and openly endorsed and communicated from a senior level within the organisation at an early stage in the project, it will struggle to generate the support, resourcing and 'buy-in' required for it to deliver the business requirements. This could have a number of consequences resulting in the recovery capability being delayed, running over budget, the scope being reduced or work not being completed to the required standard.

Due to the understanding required of the business and the management of risks within the project, there is a logical fit within the Security Management and Business Analysis areas of most IT departments. Quite often these roles posses similar skills/knowledge and often have to undertake risk and impact assessments, and implement risk reduction measures as part of their usual role. For this reason it is not unusual to find an ITSCM responsibility within an Information Security function.

The roles and responsibilities of individuals within the initial project and ongoing support of the ITSCM facilities are discussed later in this chapter.

Roles and responsibilities need to be endorsed and communicated from a senior level to ensure respect and commitment for the process.

8.3 The Business Continuity Lifecycle

It is not possible to develop an effective ITSCM plan in isolation, it must fully support the requirements of the business. This section considers the four stages of the business continuity lifecycle with particular emphasis on the IT aspects. A full understanding of the business continuity process can be obtained through the CCTA ITIL publications, 'An Introduction to Business Continuity Management' and 'A Guide to Business Continuity Management'.

The process is illustrated in the diagram below:

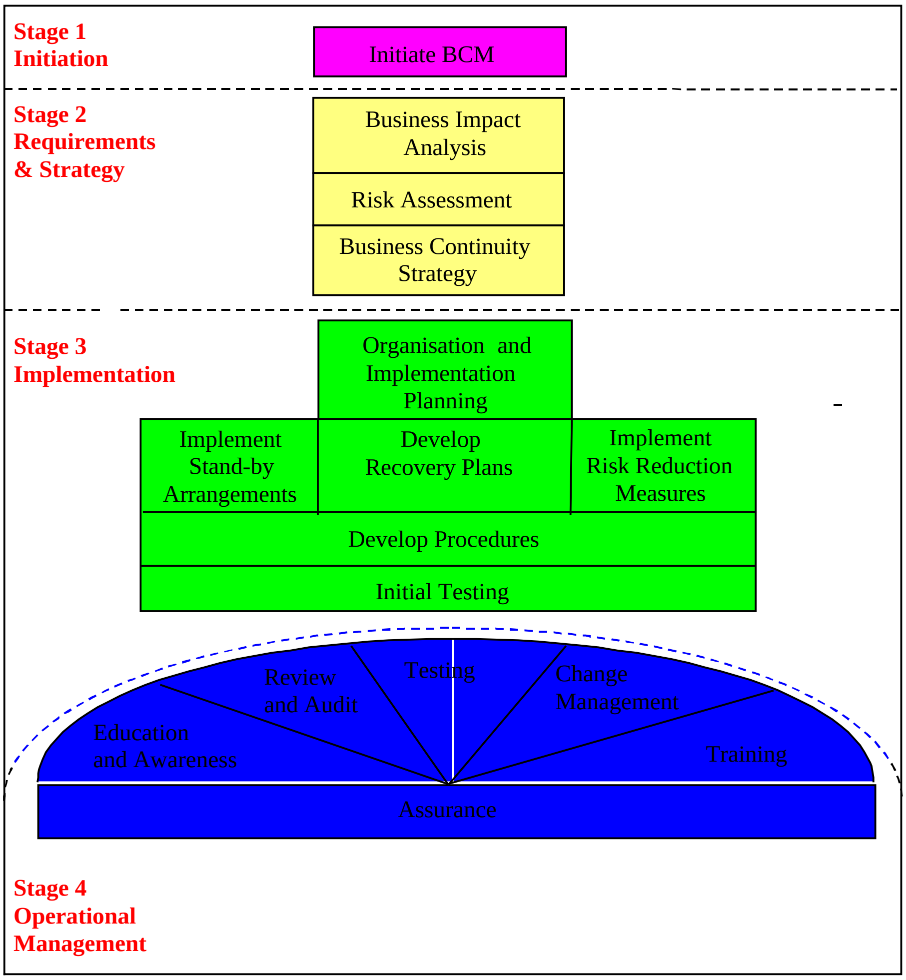


Figure 8.1 Business Continuity Management Process Model

8.3.1 Stage 1 – Initiation

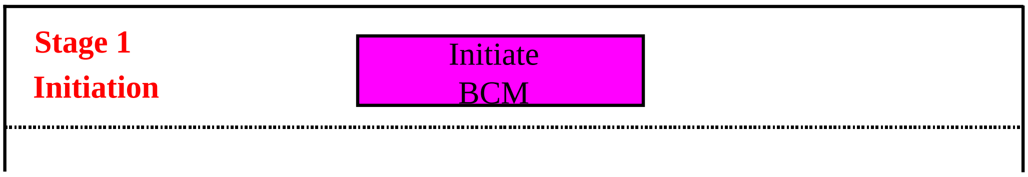


Figure 8.2 - Initiation

The activities to be considered during the initiation process depend on the extent to which contingency facilities have been applied within the organisation. Some parts of the business may have established individual continuity plans based around manual workarounds and IT may have developed contingency plans for systems perceived to be critical. This is good input to the process, however, effective ITSCM is dependent on supporting critical business functions and ensuring that the available budget is applied in the most appropriate way.

The only way of implementing effective ITSCM is through the identification of critical business processes and the analysis, and co-ordination, of the required infrastructure and IT supporting services.

The initiation process covers the whole of the organisation and consists of the following activities:

1. Policy setting – this should be established and communicated as soon as possible so that all members of the organisation involved in, or affected by, business continuity issues are aware of their responsibilities to comply with and support ITSCM. As a minimum the policy should set out management intention and objectives.
2. Specify terms of reference and scope – this will include defining the scope and responsibilities of managers and staff in the organisation, the method of working and will cover such tasks as undertaking a risk assessment and Business Impact Analysis and determination of the command and control structure required to support a business interruption. There is also a need to take into account such issues as outstanding audit points, regulatory or client requirements and insurance Organisation stipulations, and compliance with standards such as BS 7799, the British Standard on Information Security Management (which also addresses service continuity requirements).
3. Allocate resources – the establishment of an effective business continuity environment requires considerable resource both in terms of money and manpower. Depending on the maturity of the organisation, with respect to ITSCM, there may be a requirement to familiarise and/or train staff to accomplish the Stage 2 tasks. Alternatively, the use of experienced external consultants may assist in completing the analysis more quickly. However, it is important that the organisation can then maintain the process going forward without the need to rely totally on external support.
4. Define the project organisation and control structure – ITSCM and BCM projects are potentially complex and need to be well organised and controlled. It is advisable to use a standard project planning methodology such as PRINCE complemented with a project-planning tool such as Project Manager Workbench or Microsoft Project. The appointment of an experienced project manager reporting to a steering committee and guiding the

working groups is key to success. As IT is an important component of the overall process, it may be that the project will be driven from the IT area reporting through to the highest levels of management.

5. Agree project and quality plans – plans enable the project to be controlled and variances addressed. Quality plans ensure that the deliverables are achieved and to an acceptable level of quality. They also provide a mechanism for communicating project resource requirements and deliverables, thereby obtaining 'buy-in' from all necessary parties.

A well planned project initiation will enable ITSCM work to proceed smoothly with the necessary sponsorship, 'buy-in' and awareness, with all contributing members of the organisation aware of their responsibilities and commitment.

8.3.2 Stage 2 – Requirements Analysis and Strategy Definition

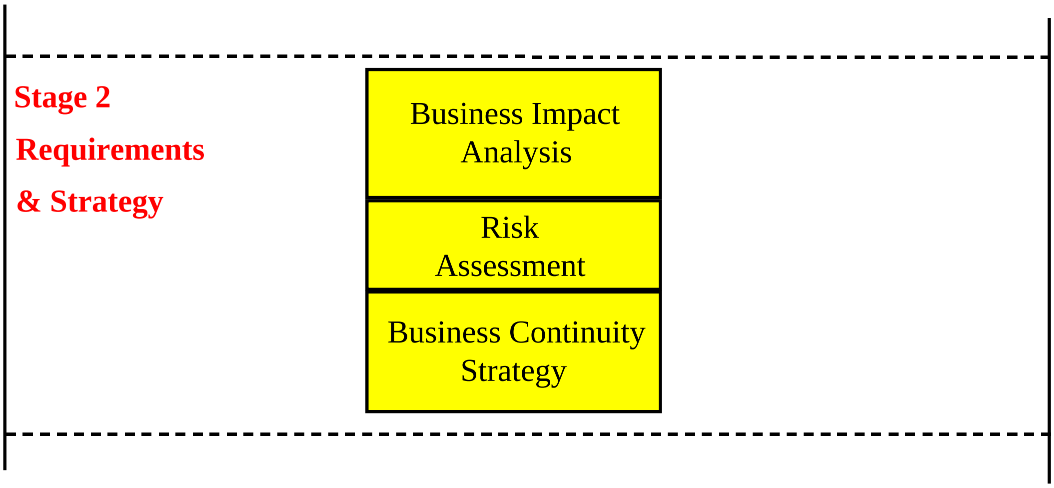


Figure 8.3 – Requirements and Strategy

This stage provides the foundation for ITSCM and is a critical component in order to determine how well an organisation will survive a business interruption or disaster and the costs that will be incurred.

If the requirements analysis is incorrect or key information has been missed, this could have serious consequences on the effectiveness of ITSCM mechanisms.

This stage can effectively be split into two sections:

- Requirements – perform Business Impact Analysis and risk assessment;
- Strategy – determine and agree risk reduction measures and recovery options to support the requirements.

Requirements

Business Impact Analysis

A key driver in determining ITSCM requirements is *how much the organisation stands to lose* as a result of a disaster or other service disruption and the speed of escalation of these losses. The purpose of a Business Impact Analysis (BIA) is to assess this through identifying:

- critical business processes;

- the potential damage or loss that may be caused to the organisation as a result of a disruption to critical business processes.

The BIA also identifies:

- the form that the damage or loss may take including lost income, additional costs, reputational damage, loss of goodwill, loss of competitive advantage;
- how the degree of damage or loss is likely to escalate after an service disruption;
- the staffing, skills, facilities and services (including the IT services) necessary to enable critical and essential business processes to continue operating at a minimum acceptable level;
- the time within which minimum levels of staffing, facilities and services should be recovered;
- the time within which all required business processes and supporting staff, facilities and services should be fully recovered.

The latter three items will provide the drivers for the level of ITSCM mechanisms that need to be considered or deployed. Once presented with these options, the business may decide that lower levels of service or increased delays are more acceptable based upon a cost/benefit analysis.

Key inputs into the Business Impact Analysis will include any service or application definitions for business areas or business processes.

These definitions and their components will enable the mapping of critical service, application and infrastructure components to critical business processes, thus helping to identify the ITSCM elements that will need to be provided. The business requirements are ranked and the associated ITSCM elements confirmed and prioritised in terms of risk assessment/reduction and recovery planning.

Impacts are measured against particular scenarios for each business process such as an inability to settle trades in a money market dealing process, or an inability to invoice for a period of days.

The impact analysis will concentrate on the scenarios where the impact on critical business processes is likely to be greatest.

Impacts will be measured against the scenarios and will typically fall into one or more of the following categories:

- failure to achieve agreed internal service levels
- financial loss
- additional costs
- immediate and long term loss of market share
- breach of law, regulations, or standards
- risk to personal safety
- political, corporate or personal embarrassment
- breach of moral responsibility
- loss of goodwill
- loss of credibility

- loss of image and reputation
- loss of operational capability, for example in a command and control environment

This process enables a business to understand at what point the unavailability of a service would become untenable. This in turn allows the types of ITSCM mechanisms that are most appropriate to be determined to meet these business requirements.

In a money market dealing environment, loss of market data information could mean that the organisation starts to lose money immediately as trading cannot continue. In addition, customers may go to another organisation which would mean a potential loss of core business. Loss of the settlement system does not prevent trades from taking place, but if trades already conducted cannot be settled within a specified period of time, the organisation may be in breach of regulatory rules or settlement periods and suffer fines and reputational damage. This may actually be a more significant impact than the inability to trade because of an inability to satisfy customer expectations.

It is also important to understand how impacts may change over time. For instance, it may be possible for a business to function without a particular process for a short period of time, for example invoicing, but over a longer period re-establishment will become critical, i.e. in order to maintain cash flow to pay bills and staff. This can be effectively represented using a graphical illustration of how business impacts vary with length of disruption as shown in Figure 8.4.

In a balanced scenario, impacts to the business will occur and become greater over time, however, not all organisations will be affected in this way. In some organisations, impacts will not be apparent immediately such as a consultancy Organisation where the need to issue reports can be deferred and impacts to the business will not begin to accrue for a period of time. In this case the ‘contingency’ line of Figure 8.4 will be applicable. In other organisations, such as investment banks, a loss of service for even a short period of time will cause major impacts to accrue immediately and the ‘preventative’ line applies. At some point however, for any organisation, the impacts will accrue to such a level that the business can no longer operate. ITSCM will ensure that contingency options are identified so that the appropriate measure can be applied at the appropriate time to keep business impacts from service disruption to a minimum level.

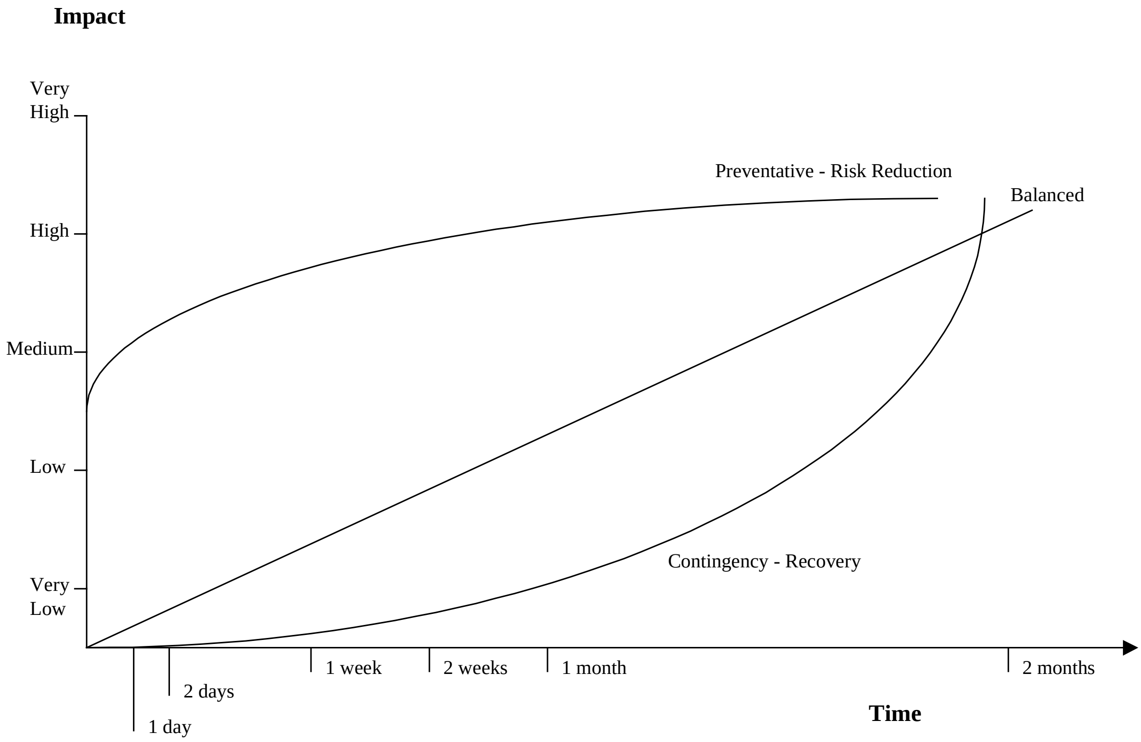


Figure 8.4 – Graphical representation of business impacts

In the majority of cases, business processes can be re-established without a full complement of staff, systems and other facilities, but still maintain an acceptable level of service to clients and customers. The business recovery objectives should therefore be stated in terms of:

- the time within which a pre-defined team of core staff and stated minimum facilities must be recovered; and
- the timetable for recovery of remaining staff and facilities.

It may not always be possible to provide the recovery requirements to a detailed level. There is a need to balance the potential impact against the cost of recovery to ensure that the costs are acceptable. The recovery objectives do, however, provide a starting point from which different business recovery and ITSCM options can be evaluated.

The Business Impact Analysis identifies the minimum critical requirements to support the business.

Risk Assessment

The second driver in determining ITSCM requirements is *the likelihood that a disaster or other serious service disruption will actually occur*. This is an assessment of the level of threat and the extent to which an organisation is vulnerable to that threat. This is expressed in Figure 8.5 below:

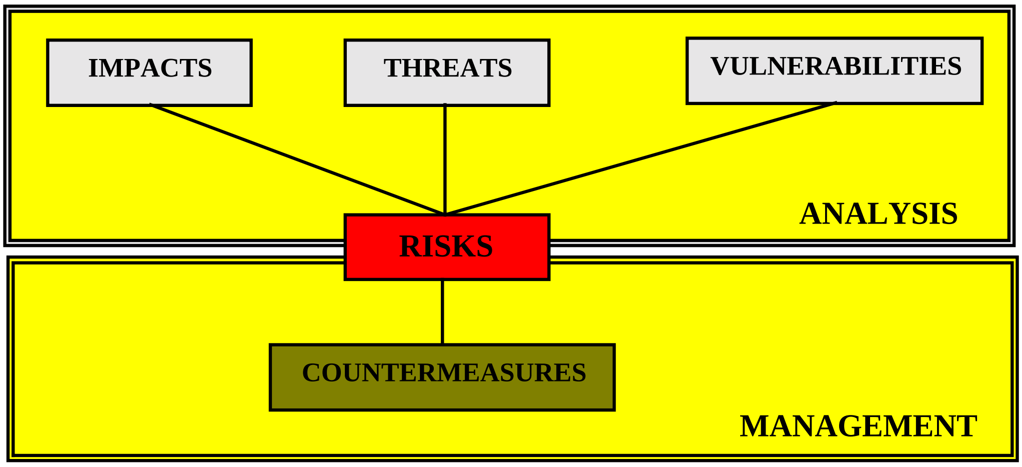


Figure 8.5 – Risk Assessment Model

The top section of figure 8.5 refers to assets - if an organisation's assets are highly vulnerable **and** there is a high threat **and** the impact is high, there would be a high risk. Countermeasures are then applied to manage the business risks by protecting the assets.]

As a minimum, the following risk assessment activities should be performed:

1) Identify risks – i.e. risks to particular IT service components (assets) that support the business process which will cause an interruption to service. Typical risks for IT include:
   a) damage or denial of access to premises;
   b) loss of IT systems, networks, PABX, Automatic Call Distribution systems, firewalls, cryptographic systems, Public Key Infrastructure (PKI), etc;
   c) loss of data or loss of integrity to data;
   d) loss of network services including telecommunications providers;
   e) unavailability of key staff, e.g. only one person knowing how to maintain a particular critical network server or business application and no documentation existing;
   f) failure of partner or service providers, e.g. outsourcing organisations providing IT systems or services (e.g. support, development or maintenance);
   g) loss of service from a partner due to excessive demand on services (e.g. excessive hits or volumes from web-site);
   h) breach of security (e.g. fraud, sabotage, computer viruses or malicious software);
   i) loss of environment (e.g. air-conditioning);
   j) loss of critical paper records or media (e.g. manuals, documents, backups, etc.);
   k) loss of utilities (e.g. power, gas, water).
2) Assess threat and vulnerability levels – the threat is defined as "how likely it is that an service disruption will occur" and the vulnerability is defined as "whether, and to what

extent, the organisation will be affected by the threat materialising". A threat is dependent on such factors as:

a) likely motivation, capability and resources for deliberate service disruptions such as malicious damage to computer systems, commercial failure of a key technology provider, attack against a Organisation's web servers and corruption of Internet sites;

b) for accidental service disruptions, the organisation's location, environment, and quality of internal systems and procedures;

c) business processes are vulnerable where there are single points of failure for the delivery of IT services (for example, a travel agent relies on information feeds for flight bookings, if the link were to fail and no backup is available, flights cannot be sold).

3) Assess the levels of risk – the overall risk can then be measured. This may be done as a measurement if quantitative data has been collected, or qualitative using a subjective assessment of, for example, low, medium or high. An example of a tabular format used to express the level of risk is illustrated in figure 8.6. Each risk can be assessed in terms of the associated threat and vulnerability. Using the table in figure 8.6 it is be possible to determine the probability of specified risks occurring (e.g. a high threat and high vulnerability implies a high probability of occurrence).

Vulnerability				
	High			
	Medium			
	Low			
		Low	Medium	High
			Threat	

Figure 8.6 – Risk Measurement table

There are many tools and methodologies available to assist in the measurement of risks including the CCTA Risk Analysis and Management Method (CRAMM). Further guidance can be found in the CCTA Management of Risk Library.

ITSCM needs to consider and assess potential risks and reduction measures across the whole infrastructure.

Following the risk analysis it is possible to determine appropriate countermeasures or risk reduction measures (ITSCM mechanisms) to manage the risks, i.e. reduce the risk to an acceptable minimum level or mitigate the risk.

In the context of ITSCM there are a number of risks that need to be taken into consideration. Table 8.1 provides a checklist of some of the risks and threats to be considered by the IT Manager:

Risk	**Threat**
Loss of internal IT systems/networks PABXs, ACDs, etc	Fire
	Power failure
	Arson and vandalism
	Flood
	Aircraft impact
	Weather damage, e.g. hurricane
	Environmental disaster
	Terrorist attack
	Sabotage
	Catastrophic failure
	Electrical damage e.g. lightning
	Accidental damage
	Poor quality software
Loss of external IT systems/networks e.g. e.commerce servers, cryptographic systems, etc	All of the above
	Excessive demand for services
	Denial of Service attack e.g. against an Internet firewall
	Technical failure e.g. cryptographic systems
Loss of data	Technical failure
	Human error
	Viruses, malicious software e.g. attack applets
Loss of network services	Damage or denial of access to network service providers' premises
	Loss of service provider's IT systems/networks
	Loss of service provider's data
	Failure of the service providers
Unavailability of key technical and support staff	Industrial action
	Denial of access to premises
	Resignation
	Sickness/Injury
	Transport difficulties
Failure of service providers e.g. outsourced IT	Commercial failure e.g. insolvency
	Denial of access to premises
	Unavailability of service provider's staff
	Failure to meet contractual service levels

Table 8.1 – Risks and Threats to be addressed by the IT Manager

The risk assessment identifies specific risks to the organisation. Many of these risks concern the ability to provide continuity of IT services. Failure to assess all the relevant risks will result in an incomplete risk assessment and will leave the business exposed to disruption.

Business Continuity Strategy

The information collated in the impact analysis and the risk assessment, and the associated ITSCM mechanisms chosen, will enable an appropriate strategy for the organisation to be developed with an optimum balance of risk reduction and recovery or continuity options. This will include consideration of the relative service recovery priorities and the changes in relative service priority for the time of day, day of the week, and monthly and annual variations.

As businesses become more dependent and driven through the use and availability of technology (e.g. e.commerce developments), ITSCM elements become a more integral part of the overall Business Continuity Strategy. Referring back to Figure 8.4, the graphical representation of business impacts, a Organisation that identifies high impacts in the short term will want to concentrate efforts on preventative risk reduction methods e.g. through full resilience and fault tolerance, while a Organisation that has low short term impacts would be better suited to comprehensive recovery options.

Risk Reduction Measures

Most organisations have to adopt a balanced approach where risk reduction and recovery are complementary and both are required. This entails reducing, as far as possible, the risks to the continued provision of the IT service and will usually be achieved through Availability Management. However well planned, it is impossible to completely eliminate all risks – for example, a fire in a nearby building will probably result in damage, or at least denial of access, as a result of the implementation of a cordon. As a general rule, the invocation of a recovery capability should only be taken as a last resort. Ideally, an organisation should assess all of the risks to reduce the potential requirement to recover the business and/or IT services.

A financial institution dealing in the equities market relies on high availability of market information and computer systems to analyse that information. Failure of the market data feeds would mean that the business process fails with an immediate financial impact to the organisation. The failure may result from a failure of the information provider (in which case competitors may suffer the same loss so the impact is lessened) so, to prevent this, the organisation takes feeds from multiple providers. Alternatively there may be a technical failure of the equipment or damage to the location where the feeds enter the building (in which case competitors are unaffected and the impacts are greater) so the organisation establishes at least two entry points and alternative equipment on immediate availability.

Typical risk reduction measures include:

- a comprehensive backup and recovery strategy, including off-site storage;
- the elimination of single points of failure such as a single power supply into a building or power supply from a single utility organisation;
- outsourcing services to more than one provider;

Outsourcing mainframe processing to a third party who provide the service remotely means that a service disruption affecting the organisation's building will not necessarily affect the availability of the host system. Outsourcing to different third parties will have the benefit of reducing the risk of a major failure as component parts of the service will always be available. This can be likened to a bookmaker 'laying off' bets to reduce the exposure on a particular gamble. This does, of course, assume the availability of resilient networks to maintain continuity of service and the fact that the third party has itself an effective and tested service continuity plan.

- resilient IT systems and networks constantly change-managed to ensure maximum performance in meeting the increasing business requirements;
- greater security controls such as a physical access control system using smartcards;
- better controls to detect local service disruptions such as fire detection systems coupled with suppression systems;
- improving procedures to reduce the likelihood of errors or failures such as change control.

The above measures will not necessarily solve an ITSCM issue and remove the risk totally, but all or a combination of them may significantly reduce the risks associated with the way in which services are provided to the business. As with recovery options, it is important that the reduction of one risk does not increase another. The risk of availability of systems and data may be reduced by outsourcing to an off-site third party, however, this potentially increases the risk of compromise of confidential information unless rigorous security controls are applied.

It is important that organisations check that recovery and ITSCM options selected are capable of implementation and integration at the time they are required, and that the required service recovery can be achieved.

An organisation's ITSCM strategy is a balance between the cost of risk reduction measures and recovery options to support the recovery of critical business processes within agreed timescales.

Recovery Options

Recovery options will be considered for:

1) People and accommodation – including alternative premises either owned, leased or through agreement with a third party; reciprocal arrangements with other organisations; and rapid procurement of alternative premises or refurbishment of existing premises. Consideration should also be given to the respective location of the proposed premises, the mobility of the staff who will be supporting the recovered business operations including IT staff and the total number of staff required to support the business process.
2) IT systems and networks – these options need to be identified and agreed by the IT Manager responsible for ITSCM and include recovery of IT systems, hardware, applications, software and networks, and the data used within these systems and facilities. This relies on the availability of effective backups to enable restoration of the service and needs to be performed in collaboration with Availability Management. This strategy should also include the implementation of continuity mechanisms to support local

disruption/interruption of IT Services supporting critical business processes, such as disk mirroring, UPS or dual power supplies, dual communication links, etc.

3) Critical services such as power, telecommunications, water, couriers and post;
4) Critical assets such as paper records and reference material.

There may be a need to consider different options for short-term and long-term recovery. Where business processes are highly dependent on external service providers, there is a need to consider the options to address failure of, or peak contention for, the services.

The costs and benefits of each option need to be analysed. This will involve a comparative assessment of the:

- ability to meet the business recovery objectives;
- likely reduction in the potential impact;
- costs of establishing the option;
- costs of maintaining, testing and invoking the option;
- technical, organisational, cultural and administrative implications against the risk of disruption or disaster and the potential impact if no action is taken.

When undertaking the analysis there is a need to consider whether the introduction of an option will adversely affect other risks.

Do not forget to check the organisation's insurance provision to determine whether adequate cover is provided.

IT Recovery Options

There are a number of options that can be considered by IT to provide contingency:

1) **Do nothing** – Few, if any, organisations can function effectively without IT services. Even if there is a requirement for stand-alone PC processing, this is still a need for recovery to be supported.
2) **Manual workarounds** – IT facilities enable organisations to process information much more quickly and efficiently. Indeed the justification for much IT spend is made on the basis of a reduced headcount. In some organisations, such as the finance, banking and insurance industries, complex calculations are undertaken by applications which would be difficult to reproduce manually in a short period of time, are dependent upon a succession of calculations by different systems with information dfed between them or are dependent upon information being fed to them from external sources. However, manual workarounds can be an effective interim measure until the IT service is resumed wherever they are practical and possible.
3) **Reciprocal arrangements** – Entering into an agreement with another organisation using similar technology used to be an effective contingency option when the computing workload was essentially batch processing. Today, the distributed computing environment means that there is a much greater requirement for individual processing power and high availability and suggests that this is not a practical solution and may not support an effective resumption

of service. In addition, there are maintenance difficulties in keeping reciprocal arrangements in step and increased need for security. Benefits can exist, however, in maintaining some reciprocal arrangements, for example, in the off-site storage of backups and other critical information.

4) **Gradual Recovery –** (sometimes referred to as 'cold standby'). This option is applicable to organisations that do not need immediate restoration of business processes and can function for a period of up to 72 hours, or longer, without a re-establishment of full IT facilities. This may include the provision of empty accommodation fully equipped with power, environmental controls and local network cabling infrastructure, telecommunications connections, and available in a disaster situation for an organisation to install its own computer equipment.

   The accommodation may be provided commercially by a third party, for a fee, or may be private, (established by the organisation itself) and provided as either a fixed or portable service.

   A fixed facility may be located at the premises of the third party that provides the service, or specially built at a location owned by the subscriber. There is a need to ensure that all services including telecommunications, market data feeds, etc. are established and adequate accommodation is available to house staff involved in the recovery process.

   A portable facility is typically a prefabricated building provided by a third party and located when needed at a predetermined site agreed with the organisation. This may be in a car park or another location some distance from the home site, perhaps, another owned building.

   The organisation will call on contracts for the supply of required computer equipment including PCs, servers, and mini computers. The organisation or the contractor (whichever has been formally pre-agreed) will then configure the equipment to the organisational requirements and load all data before a service can be provided.

   Third parties rarely guarantee replacement equipment within a fixed deadline, but would normally do so under their best efforts.

   When opting for a gradual recovery consideration must be given to highly customised items of hardware or equipment that will be difficult, if not impossible, to replace if no spares are kept securely by the organisation. Other contingency measures may be needed to cope with having to use different equipment. The same difficulties would apply to items supplied by organisations that have since gone out of business and alternatives would need to be identified, possibly putting the service delivery at risk due to delays or potential problems.

5) **Intermediate Recovery –** (sometimes referred to as 'warm standby'). This option will be selected by organisations that need to recover IT facilities within a predetermined time to prevent impacts to the business process. This will typically involve the reestablishment of the critical systems and services within a 24 to 72 hour period.

   Most common is the use of commercial facilities which are offered by third party recovery organisations to a number of subscribers, spreading the cost across those subscribers. Commercial facilities often include operation, system management and technical support.

The cost varies depending on the facilities requested such as processors, peripherals, communications, and how quickly the services must be restored (invocation timescale).

The advantage of this service is that the customer can have virtually instantaneous access to a site, housed in a secure building, in the event of disaster. It must be understood, however that the restoration of services at the site may take some time as delays may be encountered whilst the site is re-configured for the organisation that invokes the service, and the organisation's applications and data will need to be restored from back-ups.

There is a disadvantage in that the site is almost certainly some distance from the home site, which presents a number of logistical problems. The positions are shared (usually up to 20 to 30 times) with other organisations so there can be no guarantee of availability if an service disruption were to affect two organisations at the same time. There is a need to ensure that a recovery Organisation is not providing the same services for firms within an immediate geographical area. This is well understood by the recovery organisations who apply good risk management to the sale of the positions to reduce the risk of multiple invocations. It is also a fairly expensive option and can be likened to insurance. What is being paid for is peace of mind. In recent years the number of recovery centres has increased considerably and, together with the falling cost of computer hardware, good deals can be negotiated for 3, 5, or 7 year contracts.

If the site is invoked, there is often a daily fee for use of the service in an emergency, although this may be offset against additional cost of working insurance. Most commercial agreements limit invocation access to a pre-determined length of time, typically between 6 to 12 weeks and therefore longer term options are also required.

It is important that any arrangements of this sort include adequate opportunity for testing at the contingency site.

Commercial recovery services can be provided in portable form where an agreed system is delivered to a customer's site, within a certain time, typically 24 hours. The computer equipment is contained in a trailer and transported to the site by truck. The trailer is fitted out as a computer environment with the necessary services and only needs power and telecommunications links from the site to the trailer for the service to be established. Special measures may need to be taken to make the site secure.

The service provider normally charges an annual fee for such a service, and there is often a 'call-out' charge if the service is invoked. However in some circumstances, such as when there is damage to the site or when an exclusion zone is applied by the emergency services to the site, this option cannot be used

An advantage of this approach is that the trailer can be installed close to the main site subject to the necessary parking consents having been obtained. Parking a trailer on a busy road in a city is likely to draw the unwelcome attention of the police who may insist on removal.

Organisations with alternative locations may opt for a mutual fallback arrangement where accommodation is provided through displacement of non-critical staff at the unaffected building and computer facilities provided via mobile recovery.

6) **Immediate Recovery** – (sometimes referred to as 'hot standby'). This option provides for immediate restoration of services and is usually provided as an extension to the intermediate recovery provided by a third party recovery provider. The instant recovery will be supported by the recover of other critical business and support areas during the first 24 hours following an service disruption. Instances where immediate recovery may be required would be where the impact of loss of service has an immediate impact on the organisation's ability to make money, such as a Bank's dealing room.

   Where there is a need for a fast restoration of a service, it is possible to 'rent' floorspace at the recovery site and install servers or systems with application systems and communications already available and data mirrored from the operational servers. In the event of a system failure, the users can immediately switch to the backup facility with little or no loss of service.

   In the case of building loss or denial of access an organisation can pay for a limited number of exclusive positions at a recovery centre. This is a highly expensive option and would not be appropriate for the majority of organisations. However, these positions will always be available and ready for immediate occupation and use.

   Some organisations may identify a need for their own exclusive immediate recovery facilities provided internally. This again is an expensive option but may be justified for a certain business process where non-availability for a short period could result in a significant impact. The facility will need to be located separately and far enough away from the home site that it will not be affected by a disaster affecting that location.

   For highly critical business processes, a mirrored service can be established at an alternative location, which is kept up to date with the live service, either by data transfer at regular intervals, or by replications from the live service. Such a service could be used merely as a backup service, but it might also be used for enquiry access (such as reporting) without affecting the live processing performance. This is also useful if there are legal or legislative obligations to safeguard the completeness and integrity of all financial records. As this is essentially spare capacity, under normal circumstances this spare capacity can be used for development, training or testing, but could be made available immediately when a Service Continuity situation demands it.

   The ultimate solution is to have a mirrored site with duplicate equipment as part of the live operation. However, these mirrored servers and sites options, should be implemented in close liaison with Availability Management.

   It is important to distinguish between the previous definition of 'hot standby' and 'immediate recovery'. Hot standby typically referred to availability of services within a short timescale such as 2 or 4 hours whereas immediate recovery implies the instant availability of services. A recovery plan for an organisation will include a combination of some or all scenarios. Instant recovery for critical business processing, 4 hour recovery for additional business processes, 8 hour recovery for key support services and the other business areas being recovered as and when required.

8.3.3 Stage 3 – Implementation

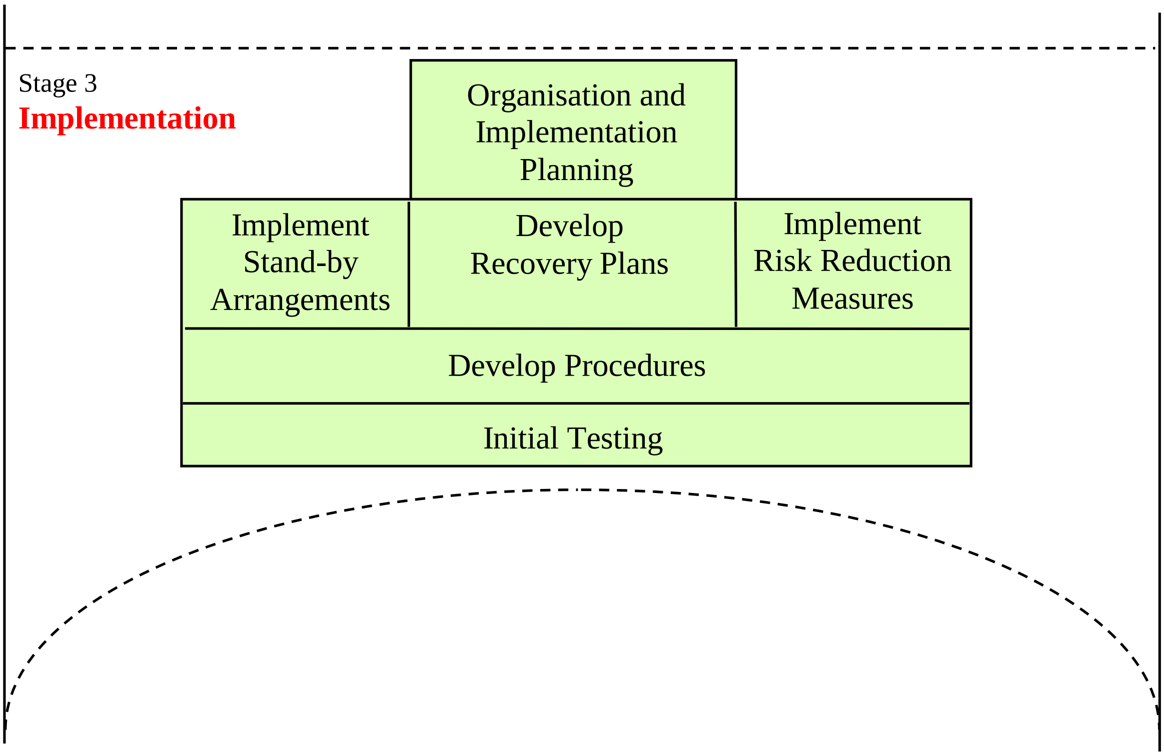


Figure 8.7 – Implementation

Once the strategy has been agreed the Business Continuity lifecycle moves into the implementation stage, involving IT at a detailed level. The implementation stage consists of the following processes:

- establish the organisation and develop implementation plans
- implement stand-by arrangements
- implement risk reduction measures
- develop IT recovery plans
- develop procedures
- undertake initial tests

Each of the above is considered with respect to the specific responsibilities that IT must action.

Organisation Planning

The IT function is responsible for the provision of IT services to support the business requirements identified during the Business Impact Analysis and requirements definition. However, for recovery purposes, IT in itself only forms part of the overall command, control and communications structure. The structure is based around three tiers:

Executive – including senior management / executive board with overall authority and control within the organisation and responsible for crisis management and liaison with other departments, divisions, organisations, the media, regulators, emergency services etc.

Co-ordination – typically one level below the Executive group and responsible for co-ordinating the overall recovery effort within the organisation.

Recovery – a series of business and service recovery teams representing the critical business functions and the services that need to be established to support these functions. Each team is responsible for executing the plans within their own areas and for liaison with staff, customers and third parties. Within IT the recovery teams should be grouped by IT service and application, for example the infrastructure team may have one or more people responsible for recovering external connections, voice services, local area networks, etc., the support teams may be split by platform, operating system or application. In addition, the recovery priorities for the service, application or its components identified during the Business Impact Analysis should be documented within the recovery plans and applied during their execution.

Implementation Planning

Plan development is one of the most important parts of the implementation process and without workable plans the process will certainly fail. At the highest level there is a need for an overall co-ordination plan that will include:

- Emergency Response plan;
- Damage Assessment plan;
- Salvage plan;
- Vital Records plan;
- Crisis Management and Public Relations plan.

These plans are used to identify and respond to an service disruption, ensure the safety of all affected staff members and visitors and determine whether there is a need to implement the business recovery process. If so, then the next level of plans are invoked which include the key support functions such as:

- Accommodation and Services plan;
- Computer Systems and Network plan;
- Telecommunication plan;
- Security plan;
- Personnel plan;
- Finance and Administration plan.

Finally, each critical business area is responsible for the development of a plan detailing the individuals who will comprise the recovery team and a detailed task list to be undertaken on invocation of recovery arrangements. The owners of each plan must ensure that they have identified and agreed support and services from other parties upon whom they have a reliance for a service or resource.

The IT contingency plan must contain all the information needed to recover the computer systems, network and telecommunications in a disaster situation once a decision to invoke has been made and then to manage the business return to normal operation once the service disruption has been resolved. There is a need to consider the various sources of information that are required in the development of the plan and these include the minimum requirements identified through the Business Impact Analysis, Service Level Agreements, security requirements, operating instructions and procedures, and external contracts. This plan will be complemented by the other plans such as the personnel plan that will address the need for transport and accommodation of key IT recovery personnel to the recovery site, or the use of overnight accommodation for critical staff. This is especially important if the recovery site has to be used for extended periods of time (e.g. weeks).

As part of the implementation planning process, it is vitally important to review key and critical contracts required to deliver business critical services. These contracts should be reviewed to ensure that, if appropriate, they provide a BCM service, there is a defined Service Level agreed and that the contracts will still be valid and in-force if operations have to switch to the recovery site (either wholly or partially). If contracts do not include these details, then the service criticality should be reviewed and the risks associated with the service not being provided should be assessed.

Check all key and critical contracts to ensure they provide a BCM service if required, have SLA's defined for business as usual and BCM and will still be valid if the recovery site has to be invoked.

Implement Risk Reduction Measures

The risk reduction measures detailed in the previous section need to be implemented. This is often achieved in conjunction with Availability Management as many of these reduce the probability of failure affecting the availability of service. Typical risk reduction measures will include such things as:

- installation of UPS and potentially back-up power to the computer;
- fault tolerant systems for critical applications where even minimal downtime is unacceptable, for example, a bank dealing system;
- offsite storage and archiving;
- RAID arrays and disk mirroring for LAN servers to prevent against data loss and to ensure continued availability of data;
- spare equipment / components to be used in the event of equipment or component failure, for example, a spare LAN server already configured with the standard configuration and available to replace a faulty server with minimum build and configuration time.

Some of the ITSCM measures that can be implemented to maintain the availability of services due to a localised disruption are described in more detail in the Chapter 9 (Availability Management).

Implement Stand-by Arrangements

The recovery options were detailed in the preceding section. It is important to remember that the recovery is based around a series of stand-by arrangements including accommodation as well as systems and telecommunications. Certain actions will be necessary to implement the stand-by arrangements, for example:

- negotiating for third party recovery facilities and entering into a contractual arrangement;
- preparing and equipping the stand-by accommodation;
- purchasing and installing stand-by computer systems;
- negotiating with external service providers on their contingency plans and undertaking due diligence if necessary.

In a manufacturing Organisation, there may be a reliance on a critical component produced by a third party without which the product cannot be manufactured. Unavailability of this component would severely disrupt, if not stop the business process so the Organisation must satisfy itself that appropriate contingency arrangements are in place to maintain the manufacture and delivery of the component.

A call centre that relies on telecommunications services should ensure that they do not rely on a single provider. Failure of that provider would mean that there would be no capability to make or receive telephone calls which would have a major reputational impact.

Training and new procedures may be required to operate, test and maintain the stand-by arrangements and to ensure that they can be initiated when required.

Develop IT Service Continuity Management Plans

IT Service Continuity Management plans need to be developed to enable the necessary information for critical systems, services and facilities to either continue to be provided or to be re-instated within an acceptable period to the business. Generally the Business Continuity plans rely on the availability of IT systems and facilities. As a consequence of this ITSCM plans need to address all activities to ensure that the required systems and facilities are delivered in an acceptable operational state and are 'fit for purpose' when accepted by the business. This will entail not only the restoration of systems and facilities, but also the understanding of dependencies between them, the testing required prior to delivery (performance, functional, operational and acceptance testing), and the validation of data integrity and consistency.

The criticality and priority of services, systems and facilities needs to be communicated by the Business Continuity planners for inclusion in the ITSCM plans. This will ensure that disruptions are dealt with in the priority order required by the business and subject to system inter dependencies. In addition, mutual agreement of the provision of the service infrastructure and the completeness of the ITSCM planning will be possible.

Management of the distribution of the plans is important to ensure that copies are available to key staff at all times. The plans should be controlled documents (with formalised document control, change control and distribution) to ensure that only the latest versions are in circulation and each recipient should ensure that a personal copy is maintained off-site.

In addition, ensure that:

- There are sufficient details to enable a technical person not familiar with the system to be able to follow the procedures. Involve people who are not familiar with the system to perform a recovery test.
- The recovery plans include key detail such as the data recovery point, a list of dependent systems, the nature of the dependency and their data recovery points, system hardware and software requirements, configuration details and references to other relevant or essential information about the system.
- A check-list is included that covers specific actions required during all stages of recovery for the system, for example after the system has been restored to an operational state, connectivity checks, functionality checks or data consistency and integrity checks should be carried out prior to handing the system over to the business,

Develop Procedures

The ITSCM plan will be dependent on specific technical tasks being undertaken. It is necessary that these are fully documented and comprehensive so that any literate IT person could undertake the recovery. Procedures will need to be developed to include the:

- installation and testing of replacement hardware and networks;
- restoration of software and data to a common reference point which is consistent across all business processes;
- different time zones in a multinational Organisation;
- business cut-off points.

Many procedures may already exist, such as the procedures for restoring systems and data in the event of equipment failure, and these should be refined and attached to the IT contingency plan as an Appendix.

Initial Testing

Testing is a critical part of the overall ITSCM process and is the only way of ensuring that the selected strategy, stand-by arrangements, logistics, business recovery plans and procedures will work in practice.

IT is responsible for the provision of the technical components and testing that these function effectively. An initial technical test can usually be done without the need to involve the business. However, for subsequent tests it is prudent to get the business to be involved to ‘prove’ the capability and to aid mutual understanding of the activities and resources needed to achieve the common goal of business recovery.
A full test will replicate as far as possible the invocation of all stand-by arrangements, including the recovery of business processes and the involvement of external parties. This will test completeness of the plans and confirm:

- time objectives, e.g. to recover the key server applications within a certain number of hours;
- staff preparedness and awareness;
- staff duplication and potential over commitment of key resources, e.g. a system administrator being required to support a number of modular plans (service desk, operations, networks and communications);
- the responsiveness, effectiveness and awareness of external parties.

Tests may be announced or unannounced, however, in the latter case it is necessary to ensure that senior management approval is obtained in advance otherwise it may be difficult to achieve commitment.
All tests will be undertaken against defined test scenarios which will be described as realistically as possible. It should be noted, however, that even the most comprehensive test will not cover everything. For example in an service disruption where there has been injury or even death to colleagues, the reaction of staff to a crisis cannot be tested and the plans need to make allowance for this. In addition, tests must have clearly defined objectives and critical success factors which will be used to determine the success or otherwise of the exercise.

8.3.4 Stage 4 – Operational Management

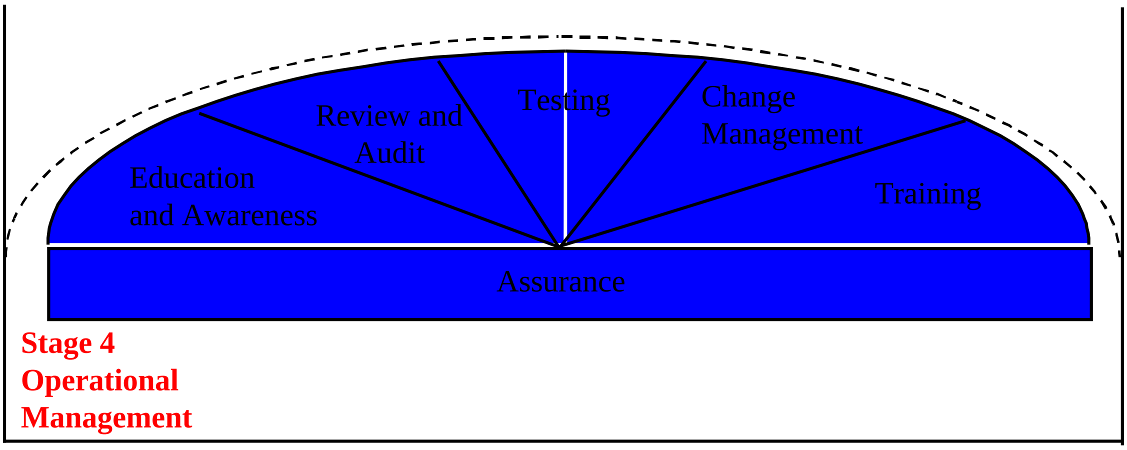


Figure 8.8 – Operational Management

Once the implementation and planning has been completed there is a need to ensure that the process is maintained as part of business as usual. This is achieved through operational management and includes:

1) Education and awareness throughout the organisation and in particular, the IT department, for service continuity specific items. This will ensure that all staff are aware of the implications of Business Continuity and of Service Continuity and consider these as part of their normal working routine and budget.
2) Training – IT may be involved in training the non-IT literate business recovery team members to ensure that they have the necessary level of competence to facilitate recovery.
3) Review – regular review of all of the deliverables from the ITSCM process needs to be undertaken to ensure that they remain current. With respect to IT this will be required whenever there is a major change to the IT infrastructure, assets or dependencies such as new systems or networks or a change in service providers, as well as when there is a change in business direction, business strategy or IT strategy. As organisations typically have rapid change, it is necessary to invest in an ongoing review programme and incorporate ITSCM into the organisational business justification processes. New requirements will be implemented in accordance with the change control process.
4) Testing – following the initial testing it is necessary to establish a programme of regular testing to ensure that the critical components of the strategy are tested at least annually or as directed by senior management or audit. It is important that any changes to the IT

infrastructure are included in the strategy, implemented in an appropriate fashion and tested to ensure that they function correctly within the overall provision of IT services.

5) Change control – following tests and reviews and in response to day to day changes, there is a need for the ITSCM plans to be updated. ITSCM must be included as part of the change management process to ensure that any changes in the infrastructure are reflected in the contingency arrangements provided by IT or third parties. Inaccurate plans and inadequate recovery capabilities may result in the failure of ITSCM. Further guidance is provided in the Chapter 8 (Change Management) in the ITIL Service Support book.
6) Assurance – the final process in the ITSCM lifecycle involves obtaining assurance that the quality of the ITSCM deliverables is acceptable to senior business management and that the operational management processes are working satisfactorily.

8.3.5 Invocation

Invocation is the ultimate test of the business continuity and ITSCM plans. If all the preparatory work has been successfully completed and plans developed and tested then an invocation of the business continuity plans should be a straightforward process.

Invocation is a key component of the plans which must include the invocation process and guidance. It should be remembered that the decision to invoke, especially if a third party recovery facility is to be used, should not be taken lightly. Costs will be involved and the process will involve disruption to the business. This decision is typically made by a 'crisis management team' comprising senior managers from the business and support departments (including IT) using information gathered through damage assessment and other sources.

A disruption could occur at any time of the day or night, therefore it is essential that guidance on the invocation process is available. Plans must be available both in the office and at home and key members of the crisis management team should be issued with a short aide memoire which they must keep with them at all time detailing key actions and decision points.

The decision to invoke must be made quickly, as there may be a lead-time involved in establishing facilities at a recovery site. In the case of a building fire, the decision is fairly easy to make, however, in the case of power failure, where a resolution is expected within a short period, a deadline should be set by which time if the problem has not been resolved, invocation will take place. This deadline will be established by the crisis management team working back from the critical point by which the business processes must be re-established to prevent an unacceptable impact to the organisation.

Whenever there is a situation where invocation may be required, put the recovery service provider on alert immediately so that facilities can be made available as quickly as possible if a decision to invoke is made.

The decision to invoke will take into account a number of factors:

- the extent of the damage and scope of the potential invocation;
- the likely length of the disruption and unavailability of premises and/or services;

- the time of day/month/year and the potential business impact. At year end the need to invoke may be more pressing to ensure that year end processing is completed on time;
- specific requirements of the business depending on work being undertaken at the time.

Once the crisis management team has decided to invoke business recovery facilities, there is a need to communicate this within the organisation. This is typically done through the use of call trees, a mechanism for communicating quickly and efficiently with identified recovery personnel throughout the organisation. The crisis management plan should include details of key personnel to be contacted to initiate the business and ITSCM recovery plans. Within each of these plans contact details for essential personnel (and their deputies will be included) to enable the plans to be initiated.

It is vital to ensure that the message has been passed to all essential personnel involved in the recovery process. The last person to receive the message should 'close the loop' by contacting the initiator and confirming the action to be taken.

The ITSCM plan will include details of activities that need to be undertaken including:

- retrieval of backup tapes or use of data vaulting to retrieve data;
- retrieval of essential documentation, procedures, workstation images, etc. stored off-site;
- mobilisation of the appropriate technical personnel to go to the recovery site to commence the recovery of required systems and services;
- contacting and putting on alert telecommunications suppliers, support services, application vendors, etc. who may be required to undertake actions or provide assistance in the recovery process.

Throughout the initial recovery, it is important that all activities are recorded. These will be used following the service disruption to analyse what went well and identify areas for improvement. The plans should include blank logs that must be given to all personnel to record activities (such as telephone conversations, timings for activities, etc.) and issues experienced.

The invocation and initial recovery is likely to be a time of high activity involving long hours for many individuals. This must be recognised and managed by the recovery team leaders to ensure that breaks are provided and prevent 'burn-out'. Planning for shifts and handovers must be undertaken to ensure that the best use is made of the facilities available. The commitment of staff (especially technical staff who will typically spend in excess of 24 hours ensuring a successful recovery) must be recognised and potentially rewarded once the service disruption is over. It is also vitally important to ensure that the usual business and technology controls remain in place during invocation, recovery and return to normal to ensure that Information Security is maintained at the correct level and that Data Protection is preserved.

It is vital to ensure that Information Security and Data Protection controls and mechanisms are maintained and enforced during the invocation, recovery and return to normal stages of Service Continuity.

Once the recovery has been completed, the business will be able to operate from the recovery site at the level determined and agreed in the business continuity strategy. The objective, however, will be to build up the business to normal levels and vacate the recovery site in the shortest possible time. The recovery period will depend on the original service disruption. In the case of a power failure, return to normal may be achieved fairly quickly, whereas in the case of a fire, reoccupation of the affected building may be impossible and alternative accommodation should be sought. Whatever the period, a return to normal must be carefully planned and undertaken in a controlled fashion. Typically this will be over a weekend and may include some necessary downtime in business hours. It is important that this is managed well and that all personnel involved are aware of their responsibilities to ensure a smooth transition.

8.4 Management Structure

As with most IT issues, ITSCM crosses organisational boundaries and consumes management time and financial resources, these should however, be in proportion to the risks being addressed. Sponsorship at the highest level and integration into the IT and ITSCM management structure is paramount to its success. Without this level of sponsorship, ITSCM risks a:

- misalignment with the business and IT strategies, thereby failing to address the true values and risks as perceived by senior management;
- lack of momentum, profile or resources to develop into a successful management and operational discipline;
- lack of extensive co-operation and input required from senior and junior management.

Without senior management sponsorship investment in ITSCM may be wasted and fail to address the requirements of the business, resulting in the organisation failing to survive a major disruption to its business due to IT systems failure.

In order for ITSCM to be successful within an organisation, a suitably capable management structure should be implemented. These roles should be integrated into the existing suite of IT management responsibilities. The optimum management structure will:

- allow responsibilities for ongoing ITSCM to be clearly defined and allocated;
- integrate into existing management structures, hierarchies and responsibilities;
- avoid concentrating responsibility in an isolated centre of excellence;
- allocate responsibilities to functions or individuals that have the necessary presence, credibility, skills, knowledge and expertise within the IT department;
- ensure that the management structure to manage ITSCM on an ongoing basis, closely resembles the structure that will control and facilitate continuity or recovery should the ITSCM mechanisms have to be utilised;
- ensure the ITSCM strategy and requirements are integrated with the business and IT strategies.

A frequent mistake is for the IT department to implement recovery mechanisms for business processes that it considers to be the most critical, without proper consultation with the business, and in isolation from the overall process.

Invariably, these mechanisms will be, at best, only partially right and result in some business interruption and lead to a lack of faith in the IT department and the ITSCM process. Conversely, the business may assume that IT and service infrastructures will be available and provided. Unless these have been notified to and agreed with the IT providers, this is unlikely to be the case.

If ITSCM is driven by business need and is fully integrated into the management structures of the organisation, significant benefit will be achieved in terms of awareness, education, and the effectiveness of the mechanisms implemented.

Ensure that the business and the ITSCM process are fully integrated to determine the mechanisms to be implemented – otherwise there is a strong likelihood they will fail to fully meet the business requirements.

8.4.1 Management Roles

A typical management structure for large organisations that will support both ongoing management and invocation for ITSCM is shown in Figure 8.9.

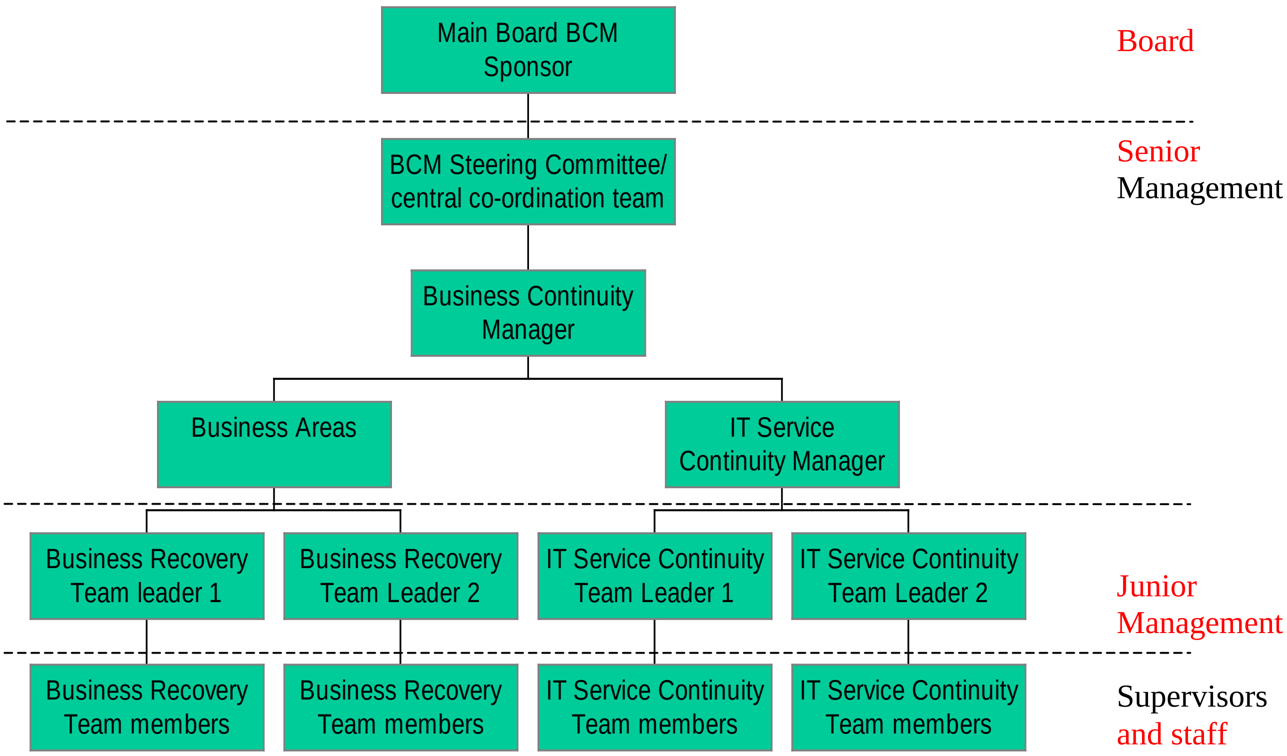


Figure 8.9 – Typical management structure for business and IT Service Continuity Management

Since management sponsorship at Board level is often given to a person whose responsibilities already encompass most of the organisation (for example Facilities, Finance or IT) it is not unusual for an IT Director to have responsibility for ITSCM.

Day to day responsibility for Business Continuity, often falls to a senior manager (shown in Figure 8.9 as the Business Continuity Manager) in an operational management role, who will:

- advise the Board representative or Board on ITSCM strategy and policy, and ensure that these are in-line with the Business and IT strategies and policies;
- ensure that the ITSCM strategy and mechanisms remain up-to-date and effective;
- ensure that change control, testing, auditing, awareness and training programmes are established.

In larger organisations, the Business Continuity Manager may be supported by a junior manager for day to day activities. The ITSCM Manager will support the Business Continuity Manager in achieving these objectives within the IT department. Within small and medium organisations, there may not be the resources available to allocate individual roles as illustrated. In organisations such as these it may be necessary to allocate one or more roles to an individual. For example, you may find that the Business Continuity Manager is also a Business Area Manager or the IT Service Continuity Manager. In the latter case, being responsible for both Business Continuity and ITSCM.

Many organisations establish steering committees at senior management level to co-ordinate business continuity activities across the organisation and support the Business Continuity Manager. The steering committee should meet regularly to:

- confirm the ITSCM strategy is still valid and that deliverables are still being maintained;
- ensure that any changes that could effect the strategy, e.g. business or IT strategy changes are addressed and reflect in the ITSCM strategy;
- review and agree test programmes for ITSCM mechanisms;
- initiate any actions necessary to address upcoming changes or to resolve difficulties.

Senior management within the IT department will typically be given ownership of the ITSCM deliverables that relate to their area of expertise or responsibility, e.g. the ITSCM mechanisms for external communications such as voice and data links will be owned by the External Communications IT manager. Ownership will not only involve responsibility for ensuring deliverables are met, but also for ensuring that they remain up to date and fit for purpose.

IT may establish a working party who will, typically, fill key roles in the IT recovery process and will fill operational management roles to deal with the continuity and availability management issues. The ITSCM Manager will typically chair the working group and lead the co-ordination team during recovery.

Invocation of continuity mechanisms and recovery options will usually be undertaken by one or a series of ITSCM teams each led by junior management and focused on specific areas of the IT department (e.g. external communications, local area networks, servers, mainframes, etc.). During periods of operational stability, the service continuity teams will play a vital role in the implementation, testing, maintenance and support of these continuity or recovery and risk reduction mechanisms and associated plans.

The approach described above correlates the management structure with the ITSCM structure and allows the team in times of operational stability to easily translate into an IT service command and control structure during times of problems and crisis (e.g. hardware or network failure, or major service disruption to the operational environment). This in turn will facilitate commitment, education, and awareness and allow cross-functional integration to take place with business activities.

Responsibilities for ITSCM should be integrated with corresponding operational responsibilities to maximise synergy and capitalise on existing knowledge, skills and expertise in the operating environment.

8.4.2 Responsibilities

Normal Operation

Figure 8.10, outlines the typical responsibilities for ITSCM during times of normal operation. These layers of responsibility also correlate with the typical management structure for ITSCM shown in Figure 8.9 and if properly implemented will allow these responsibilities to be carried out in an effective and efficient manner.

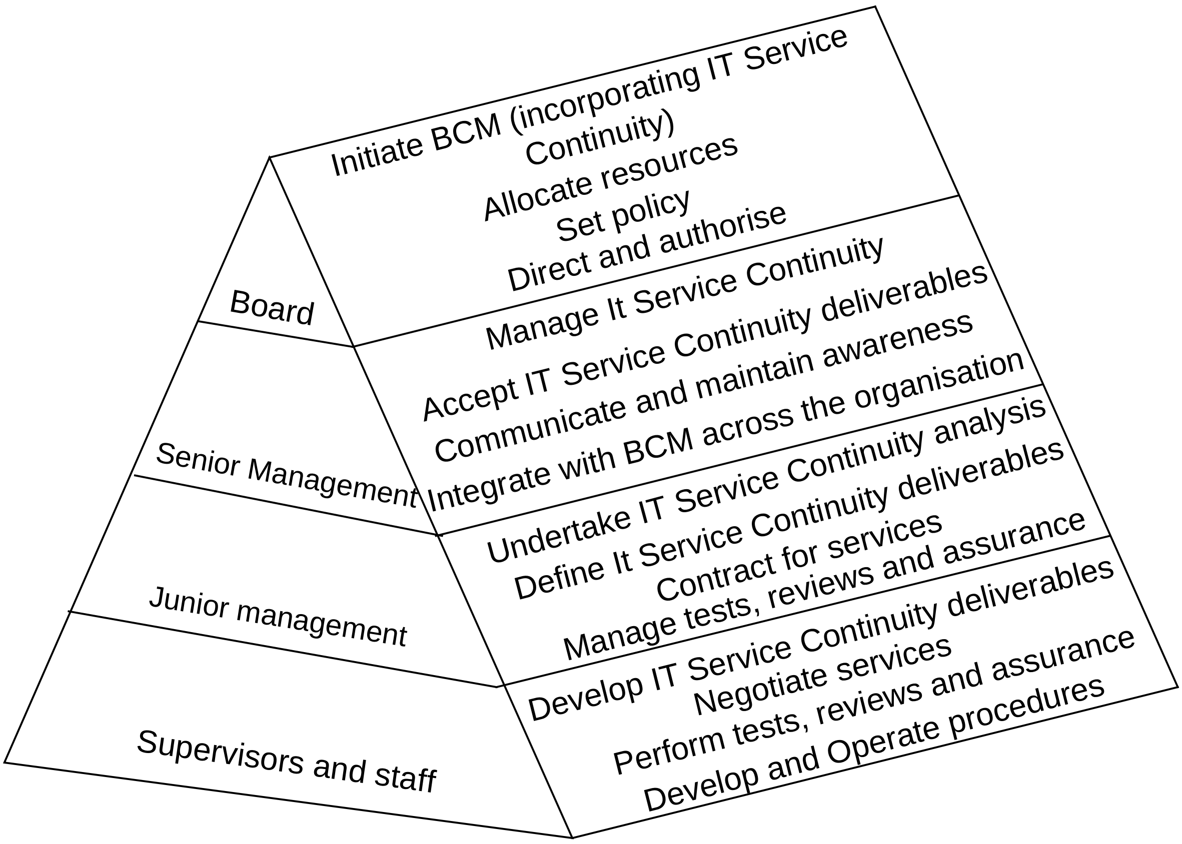


Figure 8.10 – Typical responsibilities for IT Service Continuity Management during normal operation

These responsibilities should be clearly defined, communicated to the managers concerned and documented in appropriate role or job descriptions. To ensure continual management of ITSCM at an operational level (for instance change control of the ITSCM mechanisms), best results have been achieved by incorporating specific deliverables into individual staff objectives and responsibilities into job descriptions.

Invocation Responsibilities

Following a disruption to the normal operating environment and the invocation of crisis control, management responsibilities will change in line with command, control and operational roles and responsibilities outlined in the crisis control and recovery plans. These will include responsibilities for taking corrective action to minimise impact and contingency or recovery facility invocation. Typical responsibilities for the various management layers during times of crisis or disruption to normal operations are shown in Figure 8.11.

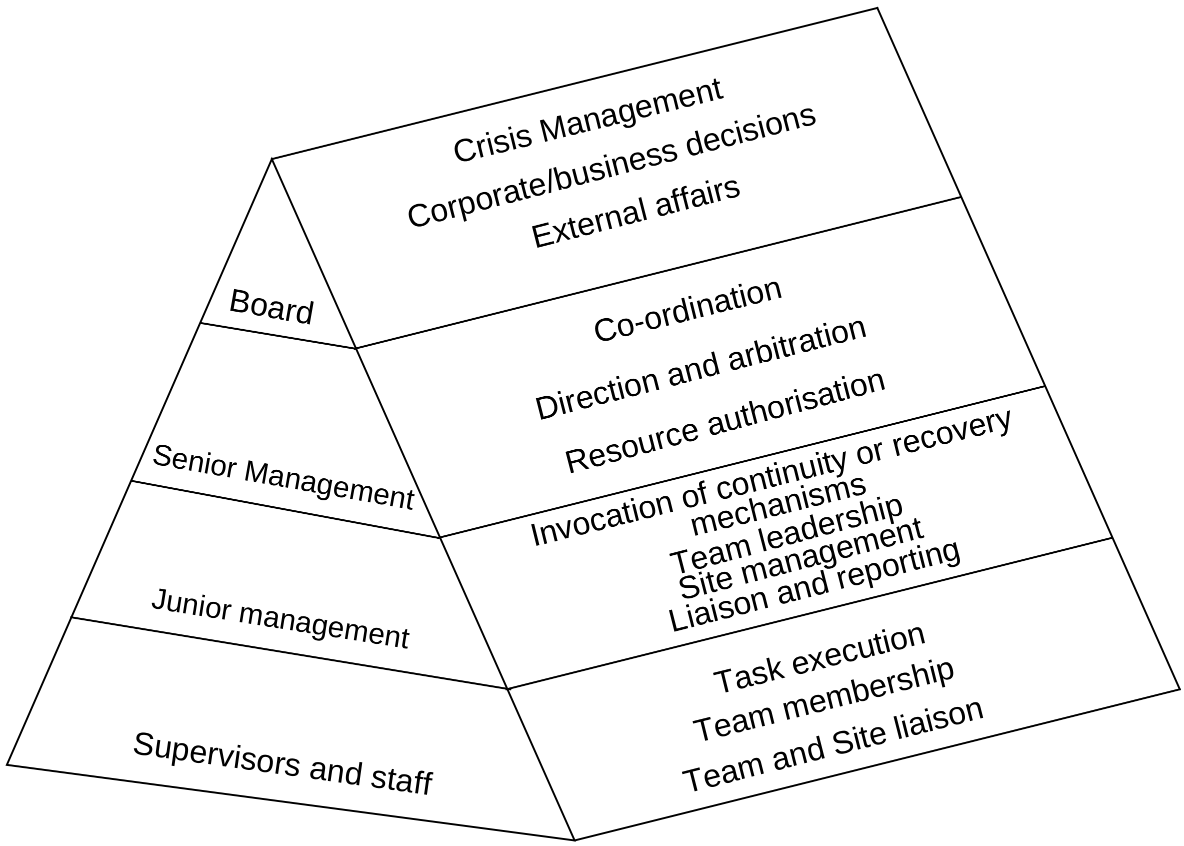


Figure 8.11 – Typical responsibilities for IT Service Continuity Management during times of crisis

For individuals associated with ITSCM, ensure high-level responsibilities are documented in job descriptions, deliverables are detailed in personal objectives and performance is monitored and reviewed during the appraisal process.

Specific responsibilities during the invocation will include recording all actions taken – to be used in subsequent debriefs, agreeing additional costs of working, liaison with the media and regulatory bodies. It is also important to understand that there may be casualties involved, such as after a terrorist bomb explosion, which may require a need for counselling during and after the service disruption and subsequent claims on the organisation for mental and physical disability.

8.5 Generating Awareness

In order to establish a successful ITSCM initiative within an organisation, it is imperative that there is awareness and commitment from the Board and senior management. The initiative has to be fully endorsed and sponsored in order to obtain the acceptance and commitment of key managers and staff, and ensure the quality of contribution that they will be able to make.

Similarly, continuing success of ITSCM will depend on a continuing commitment at all levels in the organisation and on people's awareness of their respective responsibilities. Management will need to continually monitor this and prioritise business continuity work against operational activities.

ITSCM can deliver other business benefits through the implementation of mechanisms that also improve availability management. However, in order to meet and support the true business requirements, ITSCM needs to be fully integrated into the corporate management activities for which awareness is vital.

8.5.1 Awareness of the Need for ITSCM

Awareness of the need for ITSCM will come from:

- the range of risks facing an organisation and its vulnerability;
- the potential business impacts that could result should any of the risks materialise;
- the likelihood of each of the risks materialising;
- personal responsibilities and liabilities, e.g. of directors;
- external pressures, e.g. from regulators, shareholders or clients.

In London, and in particular the City, a number of high profile bombings helped focus senior executives minds on the risks from terrorism and many new Business and ITSCM initiatives followed. Other major events have had a similar effect in other major cities in the world, including the bombing of the world trade centre in New York, the power loss in Auckland, New Zealand, earthquakes in Los Angeles and South America, and so on.

Although major events and service disruptions such as terrorism have often been the primary concern over recent years, there are many other risks that could result in serious disruption to the business and/or services on which businesses depend. Service disruptions such as computer or telecommunications failure, failure of key third parties, localised fires and floods have, collectively, caused much greater damage than high profile terrorist bombings.

Experience has demonstrated that the best way to raise awareness of the need for ITSCM is to highlight the potential risks and business impacts facing an organisation. Wherever possible, these should be stated in relation to the failure of business processes and key performance indicators for those processes, for example, customer service levels, costs, turnover, profitability, market share, etc. If the business strategy and objectives are known, comparison of risks against these also provides a useful mechanism to identify key risks.

The best and most effective way to raise senior management awareness is to highlight the potential risks and business impacts facing an organisation in terms of business failure to meet key performance indicators or corporate objectives.

If the potential impacts and risks are unclear or if there is some resistance to the concepts of ITSCM at senior levels, then a useful first step may be to undertake a scoping study, in order to bring the risks and impacts to senior management attention. The scoping report can then be used to raise awareness of the need for ITSCM, identify some of the high-level business benefits that could be realised, generate an initial senior management commitment and act as the starting point for more detailing project plans.

8.5.2 Initial Awareness

Successful awareness cannot be achieved without obtaining ‘buy-in’ from key IT and business staff (e.g. IT Support staff and Operational Business Managers) early on during this process. This is especially difficult as they can often be busy and sceptical.

The way in which the ITSCM message is communicated to those involved and the way in which the project is conducted are critical to its success.

Initial briefing sessions for the project team, stakeholders and the business areas, are always worthwhile to raise awareness, prompt support and manage expectations. Ideally these will develop into an ongoing awareness campaign along with other communication methods (a section on the staff noticeboard, inclusion in the Organisation newsletter, etc.). Regular feedback to participants will also serve to demonstrate that progress is being made as a result of their contributions.

8.5.3 Ongoing Awareness

As an organisation progresses through to the final stages of the business continuity project, emphasis shifts from awareness of the need for ITSCM mechanisms towards the responsibilities and actions necessary to implement, test and maintain those mechanisms in an operational environment.

Involvement of people from across the IT department and business areas will have helped to train and educate personnel and generate a good level of awareness.

The investment made by organisations in the ITSCM deliverables, both in terms of time and money, warrants that organisations protect this investment through an ongoing programme of awareness, education and training. Used correctly, the testing programme is an ideal way of educating and raising awareness for key individuals. For example technical and functional tests will raise awareness and educate staff about potential problems and issues to be faced. It will also confirm the operation of ITSCM mechanisms that have been installed. The use of different staff during each test will help to educate and train deputies, in order to provide cover in case of absence of the primary staff.

Use a business or ITSCM testing programme to help deliver the training and education to staff.

The following aspects of education and awareness will also be required:

- ongoing programme of briefings to all staff on the need for vigilance, on emergency procedures and security guidelines and procedures;
- regular desk top walkthroughs of the plans – especially following significant change;
- demonstration of stand-by facilities to staff that will use them;
- use of an organisation’s newsletter, noticeboards, or an intranet to maintain the profile of ITSCM;
- inclusion of an overview of the organisation’s business and ITSCM mechanisms in the staff induction process;
- regular progress reports to the Board and regular agenda items on other management and IT committees.

It is important that wherever possible senior management personnel are involved, in order to demonstrate commitment from the top and so that they can see and understand how the money has been spent. The overall aim must be to get to a stage where management considers Business and ITSCM issues in relation to, and prior to, making key business decisions. This will allow a balanced assessment of the risks to be considered in the decision making process.

The primary awareness aim is to get management to consider business and ITSCM issues in relation to and prior to making key business decisions.

8.6 The Future of IT Service Continuity Management

There are numerous challenges that will face all organisations in the future, this section outlines some of the influencing factors and highlights how they are likely to affect ITSCM.

As the corporate world develops through the use of technology and an ever increasing demand on optimisation, globalisation and an ever increasing competitive environment, the dependence upon technology in the work place and within business processes increases. This direct correlation instils the need for a co-ordinated ITSCM and change management programme that spans business processes and their supporting infrastructure. This has already been discussed at various stages throughout this chapter. However, when this need is directly referenced against ongoing initiatives within many organisations today, the need becomes even more apparent.

8.6.1 The Corporate Challenge

The influence of technology in the workplace today cannot be underestimated. The rapid improvements in the cost/benefit of processing power, disk storage and global connectivity now allows organisations to achieve speeds of information processing and transfer that could not even be dreamed of previously.

This evolution has enabled organisations to both centralise and de-centralise functions to suit the particular needs of a business stream, division or office. Where technology may be centralised in a data centre in one part of the world, business activities may take place across the globe using those systems in order to provide access to local markets and resource. Conversely, technology and IT processes may be de-centralised where business activities are centralised, making use of inexpensive labour in other countries.

This new reality allows organisations to choose how they operate their business, with few limitations, and corporate management focus can now be on the most efficient method. Consequently the impact on IT service is immense.

As businesses merge, or change their operating structure or strategy, in many cases this has a direct impact on which business processes continue to be the most critical and which elements of those business processes are more at risk. Unless organisations undertake to re-assess their Business Impact Analysis and risk assessment as part of this process, the continuity or recovery mechanisms put in place and being maintained to support the business will not support all the elements of a critical business process (and it therefore fails) or will not support new critical business processes (therefore they run the risk of failure or disruption).

People, processes, organisation structures and information systems are focused on helping the organisation achieve its objectives. In order to do this, management essentially has to do three things:

- use its resources effectively and efficiently;
- monitor and exploit new opportunities that could help it achieve or better its targets, e.g. expand into new markets, acquire new businesses or merge with other organisations, or implement improved technology;
- monitor and manage risks that could stop it from achieving its targets.

The modern business environment is characterised by *risk*. ITSCM aids an organisation in reducing its exposure to risks. Effective and efficient management of key risks to the organisation through ITSCM is therefore essential if business objectives and targets are to be achieved and maintained. Decisions to exploit new opportunities must also balance benefit against cost and *risk* and incorporate potential changes to ITSCM initiatives as part of the decision making process.

ITSCM addresses risks which could cause a *sudden and serious* impact, such that they could *immediately threaten the continuity of the business*. These typically include the following subset of the risks described earlier:

- loss, damage or denial of access to key infrastructure services;
- failure or non-performance of critical providers, distributors or other third parties;
- loss or corruption of key information;
- sabotage, extortion or commercial espionage;
- deliberate infiltration or attack on critical information systems.

Where these risks affect critical business processes the consequences can be severe and include substantial financial loss, embarrassment and loss of credibility or goodwill for the organisation concerned. The consequential damage can extend much wider impacting on staff welfare, customers, third parties, shareholders and the general public. Ultimately they will obstruct organisations from achieving their objectives and targets.

Organisations that practice ITSCM effectively will have assessed the risks to business continuity, identified minimum acceptable levels of business and put in place tested plans to ensure that these can be maintained.

If an organisation does not manage business continuity effectively, the chances are that it does not even understand its minimum acceptable levels of business. Should a serious service disruption occur, there will not be a consistent corporate view of priorities and the organisation will not have the infrastructure, plans and procedures in place to deal with the crisis. In these circumstances, survival of the business will then depend entirely on:

- how management reacts to the crisis – without tested crisis management plans, this may be chaos and panic;
- the time within which key infrastructure services can be resumed in the absence of tested stand-by arrangements – it may take weeks or months to obtain replacement computer or communications hardware if provisions have not been made in advance and recovery rehearsed;
- the adequacy of insurance arrangements – these may exclude certain risks, limit the cover (e.g. to additional cost of working but not consequential loss) and place a time limit on the benefit (e.g. so that long term loss of market share is not covered).

8.6.2 Future Developments

These days, such a reactive approach to dealing with risks to business continuity is not acceptable. Directors have a ‘duty of care’ to their shareholders, customers and creditors and could find themselves personally liable if they have been negligent, whatever their professional discipline.

Over the next few years the greatest challenge facing organisations will be the move to electronic business and many organisations will need to re-invent themselves to prosper in the new digital world. ITSCM will need to react to changes in the business world that it is there to protect. These changes will make business even more dependent on technology than it is already. In order to address these risks, resilience, security and ITSCM will have to be designed into electronic business services from the outset in a way that is only partly done today.

Technological development,the need to optimise efficiency and reduce overheads will force changes to other parts of the business infrastructure. For example, the emergence of a smaller number of huge warehouse and distribution facilities from which goods can be delivered direct to the home, at the expense of the existing large network of retail outlets. Disruption to a major distribution centre could halt the business at a time when timely delivery and customer service levels are a key differentiator between competitors or could result in increased costs of operations to meet customer requirements, negating the efficiencies and cost reductions expected. Again, one would expect part of the investment in this new infrastructure to be in building resilience and security into the infrastructure.

Risk management and business continuity management will be fundamental to the re-design of business for the information age. Technology will be able to support this and, in the future, successful organisations will be those that can not only exploit the opportunities of new technology but can also harness it within their business to minimise risk by building fully resilient and secure systems and business processes. This will require ITSCM to be implemented as an integral part of the development process.

8.6.3 In Conclusion

IT Service Continuity Management is a critical tool in the armoury if today's business environment is to continue to operate in spite of the plethora of risks faced. As organisations strive for greater efficiency, so the use of technology and the consequent dependence upon it increases; thereby the requirement for ITSCM increases in order for the business to be able to sustain its operations. Failure to implement adequate ITSCM measures will impact that organisation's ability and its perception by customers following an interruption.

The business continuity infrastructure developed must be able to respond quickly and efficiently to changes in the organisation and be regularly tested to ensure that the different components of the IT environment will work together.

The commercial world today is realising its dependence upon, and requirement for, technology and the disciplines that have to run in parallel, including ITSCM. There are greater demands being placed on organisations by clients, third parties, shareholders and regulators for Directors of organisations to perform better risk management i.e. the appropriate implementation of BS 7799, Data Protection Act and the Turnbull Report (see Bibliography - Chapter 12). This in turn requires ITSCM to become an integral part of corporate management.

The dependence upon technology and the greater demands being placed upon organisations require IT Service Continuity Management to become an integral part of Corporate Management in order for business objectives and targets to be achieved and maintained.

Annex 8A

ITSCM Manager - Role, Responsibilities & Key Skills

Role

To implement and maintain the IT Service Continuity Management process in accordance with the overall requirements of the organisation's Business Continuity Management process and to represent the IT Services function within the Business Continuity Management process.

Responsibilities

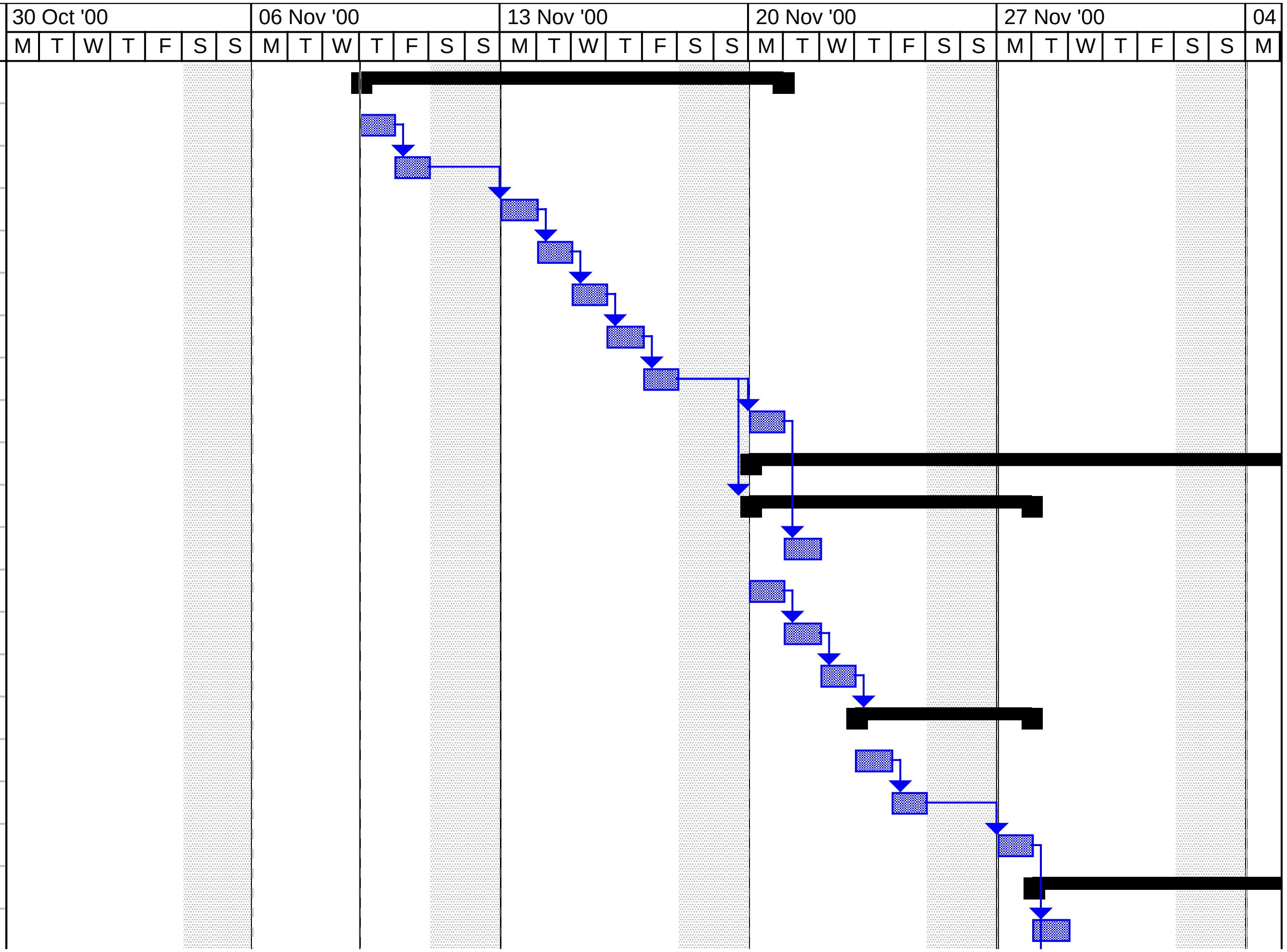
- Develop and manage the IT Service Continuity Management Plan to ensure that, at all times, the recovery objectives of the business can be achieved.
- Ensure that all IT Service areas are prepared and able to respond to an invocation of the continuity plans.
- Maintain a comprehensive IT testing schedule.
- Undertake quality review of all procedures and ensure that these are incorporated into the testing schedule.
- Communicate and maintain awareness of IT Service Continuity Management objectives within the business areas supported and IT Service areas.
- Undertake regular reviews, at least annually, of the continuity plans with the business areas to ensure that they accurately reflect the business processing environment.
- Negotiate and manage contracts with providers of third party recovery services.
- Manage the IT Service delivery during times of crisis including:
    - Co-ordination with the crisis control team;
    - Invocation of the appropriate recovery facilities;
    - Resource management, direction and arbitration;
    - Recovery site management

Skill Set (Specific)

- Senior IT Management level.
- Experience in IT Service Continuity
- Knowledge and experience of Contract Management type roles.
- Ability to translate business recovery requirements into technical requirements and specifications.
- Good IT technical knowledge to enable quality assurance of procedures to be undertaken.
- Ability to communicate at all levels within the organisation.

Annex B - Example Project Plan

ID		Task Name	Duration
1		**Stage 1 – Initiation**	**8 days**
2		Define and set Policy	1 day
3		Define the project organi:	1 day
4		Define Project control str	1 day
5		Develop project plan	1 day
6		Define quality plan	1 day
7		Identify initial costs	1 day
8		Develop Project Initiation	1 day
9		Agree PID	1 day
10		**Stage 2 – Requirements Ana**	**13 days**
11		**Requirements**	**6 days**
12		Identify critical proce	1 day
13		Identify IT service su	1 day
14		Conduct Impact Ana	1 day
15		Extract support requ	1 day
16		**Conduct risk asses**	**3 days**
17		Identify risks	1 day
18		Assess threat a	1 day
19		Assess the leve	1 day
20		**Strategy**	**7 days**
21		determine risk reduc	1 day



Annex 8C - IT Service Continuity Management

Generic Recovery Plan

Contents

1. Document Control 233

1.1. Document Distribution 233

1.2. Document Revision 233

1.3. Document Approval 233

2. Supporting Information 234

2.1. Introduction 234

2.2. Recovery Strategy 234

2.3. Invocation 234

2.4. General Guidance 234

2.5. Dependencies 235

2.6. Recovery Team 235

2.7. Recovery Team Checklist 235

3. Recovery Procedure 237

1. Document Control

This document must be maintained to ensure that the systems, infrastructure and facilities included, appropriately support business recovery requirements.

1.1. Document Distribution

Copy	Issued to	Date	Position
1			
2			
3			
4			

1.2. Document Revision

This document will be reviewed every X months.

Current Revision: *date*

Next Revision: *date*

Revision Date	Version No	Summary of Changes

1.3. Document Approval

This document must be approved by the following personnel.

Name	Title	Signature

2. Supporting Information

2.1. Introduction

This document details the instructions and procedures that are required to be followed to recover or continue the operation of *systems, infrastructure, services or facilities* to maintain service continuity to the level defined or agreed with the business.

2.2. Recovery Strategy

The *systems, infrastructure, services or facilities* will be recovered to *alternative systems, infrastructure, services or facilities*.

It will take approximately *X hours* to recover the *systems, infrastructure, services or facilities*. The system will be recovered to the last known point of stability/data integrity which is *point in day/timing.*

The required recovery time for this *system, infrastructure, service or faciliy* is:

The recovery time and procedures for this *system, infrastructure, service or faciliy* was last tested on:

2.3. Invocation

The following personnel are authorised to invoke this plan

1.

2.

2.4. General Guidance

All requests for information from the media or other sources should be referred to the *Company procedure.*

When notifying personnel of a potential or actual disaster follow the defined operational escalation procedures, and in particular:

- ➢ be calm and avoid lengthy conversation;
- ➢ advise them of the need to refer information requests to *escalation point*;
- ➢ advise them of expectations and actions (avoid giving them details of the incident unless absolutely necessary)
- ➢ If your call is answered by somebody else:
  - ask if the contact is available elsewhere;
  - if they cannot be contacted leave a message to contact you on a given number;
  - do not provide details of the incident.
  - Always document call time details, responses and actions..

All activities and contact/escalation should be clearly and accurately recorded. To facilitate this, actions should be in a checklist format and there should be space to record the date and time the activity was started and completed, and who carried out the activity.

2.5. Dependencies

System, infrastructure, service, facility or interface dependencies should be documented (in priority order) so that related recovery plans or procedures that will need to be invoked in conjunction with this recovery plan can be identified and actioned.

The person responsible for invocation should ensure recovery activities are co-ordinated with these other plans.

System	Document Reference	Contact

2.6. Recovery Team

The following staff/functions are responsible for actioning these procedures or ensuring the procedures are actioned and recording any issues or problems encountered. Contact will be made via the normal escalation procedures.

Name	Title	Contact Details

2.7. Recovery Team Checklist

To facilitate the execution of key activities in a timely manner, a checklist similar to the following should be used.

Task	Target Completion	Actual Completion
Confirm invocation has taken place		

Task	Target Completion	Actual Completion
Initiate call tree and establish Recovery Team		
Identify issues and advise *Crisis Management Team*		
Arrange for backup media, vital records to be shipped from off-site store to recovery site		
Establish recovery team rota		
Confirm progress reporting requirements		
Inform Recovery team of reporting requirements		
Confirm liaison requirements with other recovery teams		
Initiate recovery actions		
Advise the estimate for system recovery and commencement of testing		
Advise estimate for when systems will be ready for user processing.		

3. Recovery Procedure

Enter recovery instructions/procedures here.

Content/format should be in-line with company standards for procedures. If there are none, guidance should be issued by the Manager or Team Leader for the area responsible for the *system, infrastructure*, *services or facility*. The only guideline is that the instructions should be capable of being executed by an experienced professional without undue reliance on local knowledge.

Where necessary, references should be made to supporting documentation (and its location), diagrams and other information sources. This should include the document reference number (if it exists). It is the responsibility of the plan author to ensure that this information is maintained with this plan. If there is only a limited amount of supporting information, it may be easier for this to be included within the plan, providing this plan remains easy to read/follow and does not become too cumbersome.

9 AVAILABILITY MANAGEMENT

9.1 Introduction

The continued advances in IT technology design have resulted in significant availability and reliability improvements of the IT infrastructure. Fault tolerant and error correction design features in hardware and software now reduce the risk of IT component failure to enable improving levels of availability to be delivered.

However, against this background of improving levels of availability and reliability, the need for Availability Management is now greater than ever.

9.1.1 Why Availability Management?

The importance on the availability of IT to the success of the business has never been more apparent. Over the years the inter dependency between the business process and the IT operation has developed to the point where quite simply if the IT stops, then the business stops.

This importance in supporting the business should also be seen in the context of the trends affecting society: trends such as the global economy, 24hr economy, E-Commerce and flexible working. These are the impetus for a growing demand for the availability of IT services that are time and place independent.

The business response to marketplace and consumer demand is to increase their exploitation of IT to create new and innovative ways to provide their customers with flexibility and choice in the way they wish to do business. This is evidenced by the emergence of Internet based online business-to-consumer services. These are now viewed as essential if a business is to not only attract new customers but also retain their existing customers.

In today's competitive marketplace customer satisfaction with the service(s) provided is paramount. Customer loyalty can no longer be relied on and dissatisfaction with the availability and reliability of service can be a key factor in customers taking their business to a competitor.

The role of IT is now pivotal. The availability and reliability of IT can directly influence customer satisfaction and the reputation of the business. This is why today Availability Management is essential in ensuring IT delivers the right levels of availability required by the business to satisfy its business objectives and deliver the quality of service demanded by their customers.

Key message: -

Effective Availability Management will influence customer satisfaction and determine the market place reputation of the business.

9.1.2 The need for a business and end user focus

Given the business dependency on IT availability, it is essential that the deployment of Availability Management has a strong business and end user emphasis. This is to ensure that the IT infrastructure delivers the required levels of availability to support the vital business functions key to the business operation. It will also ensure that availability improvement opportunities are focused to deliver true end user benefits.

This requires Availability Management to understand the business and end user perspective of the IT service provided. This should be evidenced in the key process inputs and outputs.

By having this emphasis, the deployment of Availability Management can make a positive contribution in enhancing the relationship with the business. The IT organisation being seen to recognise and respond to IT availability opportunities and challenges with their needs understood.

9.1.3 Goal and Mission

The goal of the Availability Management process is to optimise the capability of the IT infrastructure and supporting organisation to deliver a cost effective and sustained level of availability that enables the business to satisfy its business objectives.

This is achieved by determining the availability requirements of the business and matching these to the capability of the IT infrastructure and supporting organisation. Where there is a mismatch of the requirement vs capability, Availability Management will ensure the business are provided with available alternatives and associated cost options.

Availability Management should ensure the required level of availability is provided. The measurement and monitoring of IT availability is a key activity to ensure availability levels are being met consistently.

Availability Management should continuously look to optimise the availability of the IT infrastructure and supporting organisation, in order to provide cost effective availability improvements that can deliver evidenced business and end user benefits.

> Hints & Tips: -
>
> A suggested Mission Statement for Availability Management: -
>
> "To understand the availability requirements of the business and to plan, measure, monitor and continuously strive to improve the availability of the IT infrastructure and supporting organisation to ensure these requirements are consistently met"

9.1.4 Objectives

The objectives of the Availability Management process are as follows:-

- To ensure IT services are designed to deliver the levels of availability required by the business.
- To provide a range of IT availability reporting to ensure that agreed levels of Availability, Reliability and Maintainability are measured and monitored on an ongoing basis.
- To optimise the availability of the IT infrastructure to deliver cost effective improvements that deliver tangible benefits to the business and end user.
- To achieve over a period of time a reduction in the frequency and duration of incidents that impact IT availability.

- To ensure shortfalls in IT availability are recognised and appropriate corrective actions are identified and progressed.
- To create and maintain a forward looking Availability Plan aimed at improving the overall availability of IT services and infrastructure components to ensure existing and future business availability requirements can be satisfied.

9.1.5 Scope

Availability Management is concerned with the design, implementation, measurement and management of IT infrastructure availability to ensure the stated business requirements for availability are consistently met.

- Availability Management should be applied to all new IT services and for existing services where Service Level Requirements (SLRs) or Service Level Agreements (SLA) have been established.
- Availability Management can be applied to those IT services deemed to be business critical regardless of whether formal SLA's exist.
- Availability Management will be applied to the suppliers (internal and external) that form the IT support organisation as a precursor to the creation of a formal SLA.
- Availability Management will consider all aspects of the IT infrastructure and supporting organisation which may impact availability, including training, skills, policy, process effectiveness, procedures and tools.
- Availability Management is not responsible for Business Continuity Management and the resumption of business processing after a major disaster. This is the responsibility of IT Service Continuity Management (ITSCM). However, Availability Management provides key inputs to ITSCM and both have a close relationship.

9.2 Basic Concepts

9.2.1 The IT Infrastructure and IT support organisation

The 'Availability' of the IT infrastructure components that deliver IT services to the business and their end users are influenced by: -

- The complexity of the IT infrastructure design.
- The reliability of the IT infrastructure components and environment.
- The IT support organisation capability to maintain and support the IT infrastructure.
- The levels and quality of maintenance provided by suppliers.
- The quality, pattern and extent of deployment of operational process and procedures.

The IT provider will be accountable to the business for the delivery of the IT service(s). The required levels of availability for the IT service(s) will be documented within a formal Service Level Agreement (SLA).

The IT provider will need to formally agree with each infrastructure supplier and maintainer the appropriate conditions and controls necessary for the SLA to be met.

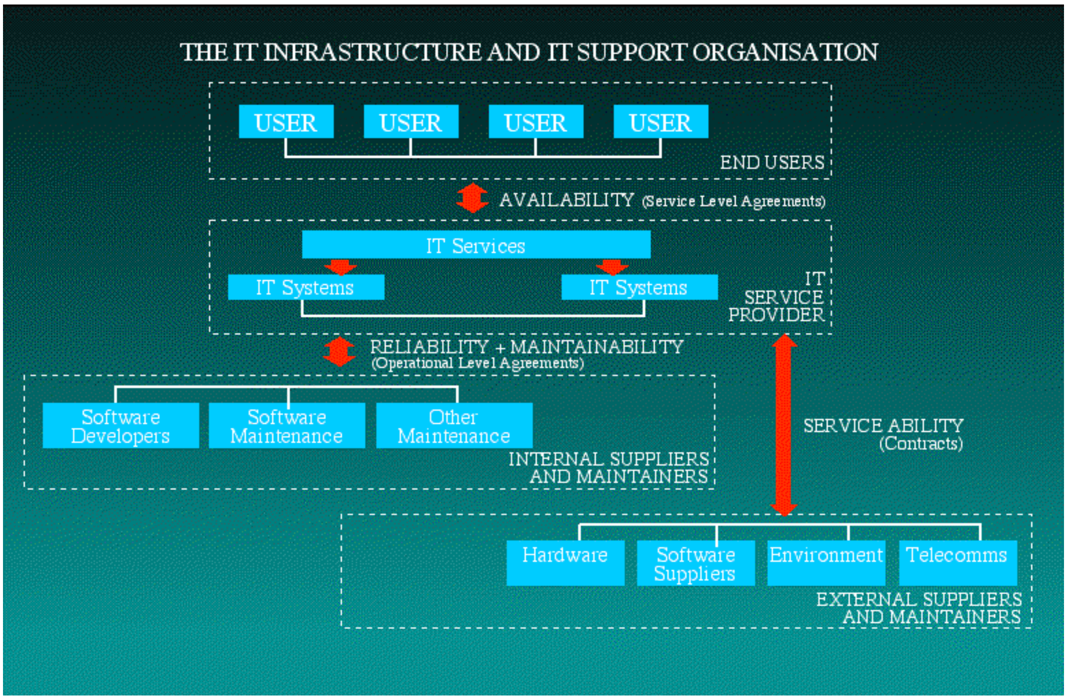


Figure 9.1: - The relationships with suppliers and maintainers of the IT infrastructure

For internal suppliers these requirements would be documented within an Operational Level Agreement (OLA). For external suppliers they must be documented within an underpinning contract.

To ensure that the required level of availability is being delivered to the business will require all formal agreements to be measured, monitored and reviewed on a regular basis.

9.2.2 Guiding Principles

An effective Availability Management process can ***'make a difference'*** and will be recognised as such by the business if the deployment of Availability Management within the IT organisation has a strong emphasis on the needs of the business and end user.

To reinforce this emphasis there are **three guiding principles** that should underpin the Availability Management process and thinking. These are: -

- Availability is at the core of business and end user satisfaction.
- Recognising that when things go wrong, you can still achieve business and end user satisfaction.
- Improving availability can only begin when you understand how the technology supports the business.

Guiding Principle #1

"AVAILABILITY IS AT THE CORE OF BUSINESS AND END USER SATISFACTION"

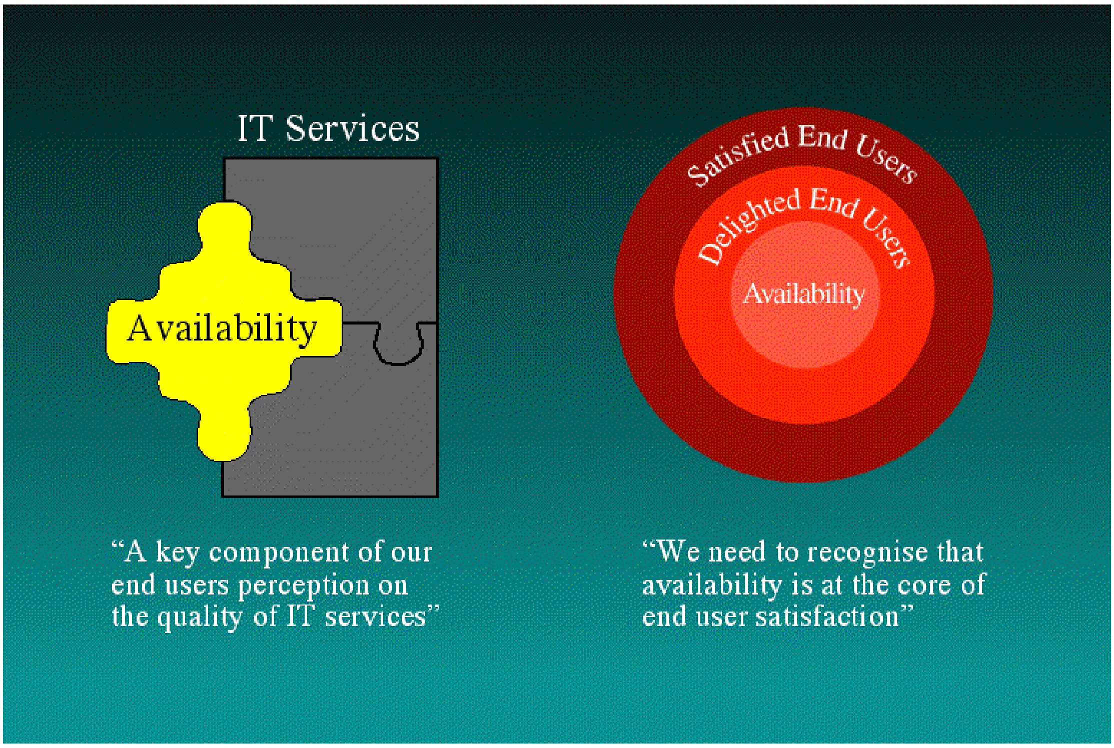


Figure 9.2: - Availability is at the core of business and end user satisfaction

The above are simple messages that make an important point. As a consumer of an IT service, its availability and reliability can directly influence both the perception and satisfaction of the overall IT services provision.

A good analogy for example is to consider the purchasing of a new car. Cost, style, performance and features are all important. However, if the car breaks down frequently it is unlikely you will ever purchase another car from that manufacturer again.

Within the IT organisation it is unlikely that anyone would disagree with the above messages. However, with today's commercial and business pressures it is all too easy for new IT strategies and initiatives to create diversionary activities. The result is that focus on availability drifts with the potential for the IT organisation to become so internally focused that it looses sight of the business and end user(s) it serves. By embracing this principle, Availability Management will ensure that availability is recognised by all (Senior Management to junior staff) as the primary IT deliverable.

Guiding Principle #2

"RECOGNISING THAT WHEN THINGS GO WRONG, YOU CAN STILL ACHIEVE BUSINESS AND END USER SATISFACTION"

Whilst emphasising the importance of availability to the business operation and its influence on end user satisfaction and business reputation, the reality is that on occasions things do go wrong.

For the IT organisation this is their 'moment of truth' in the eyes of the business. How the incident is managed and resolved will play an important role in how the business view the responsiveness and quality of the IT organisation.

To make a point you often need to show the extreme, the following illustration challenges the conventional thinking that assumes the 'quickest' recovery is necessarily viewed by the business as the best if their needs during the failure were not satisfied.

Figure 9.3: - Business and End User satisfaction in failure scenarios

The Availability Management process plays a crucial role in the anticipation of IT failure and the assessment of the business and IT needs that must be satisfied. For the business, these will revolve around informational needs that enable the business to manage the impact of failure on their business and customers. For IT, these will revolve around the provision of process, procedures and tools to enable the technical recovery to be completed in an optimal time.

Industry view: -

“The customer doesn’t always expect everything will go right all the time, the big test is what you do when things go wrong............

Occasional service failure is unavoidable”

Source:- Sir Colin Marshall - Chairman, British Airways

Guiding Principle #3

“IMPROVING AVAILABILITY CAN ONLY BEGIN WHEN YOU UNDERSTAND HOW THE TECHNOLOGY SUPPORTS THE BUSINESS.”

Availability Management should not simply understand the availability of each component. By taking a business and end user perspective it is important to understand how the technology components support the vital business functions upon which the business operation relies. Thus understanding in business terms the contribution of each component to the business operation and the end user.

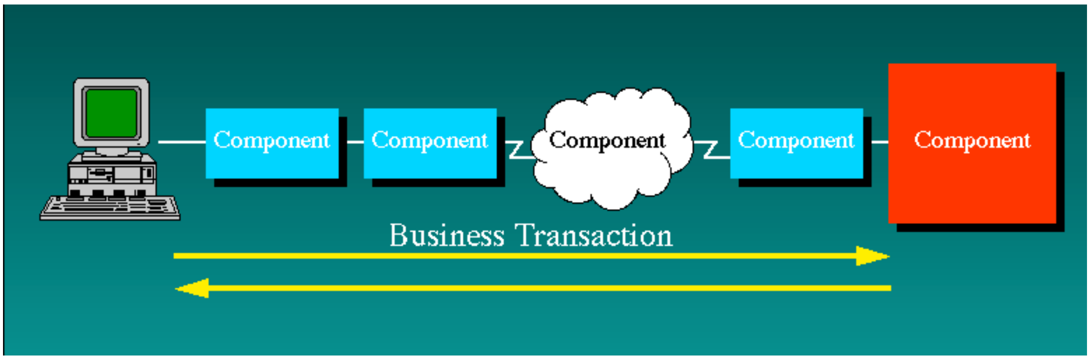


Figure 9.4: - Understanding how the technology supports the business by taking an end – end business and end user perspective.

This is fundamental if Availability Management is to optimise the IT infrastructure and supporting organisation to deliver the levels of availability required by the business and to drive availability improvements that deliver true business and end user benefits.

Key Message

Availability Management will need consider availability from an IT service perspective and from an IT component perspective. These are entirely different aspects. Whilst the underlying concept is similar, the measurement, focus and impact is entirely different.

Validation: -

A Financial services organisation had for years continually looked to improve its Auto Teller Machine (ATM) availability by looking at how individual component availability could be improved. The reasoned thinking being by improving the availability of IT component(s) that this would improve the overall end – end availability. Despite this approach, customer satisfaction levels did not improve and subsequent benchmarking indicated competitors were achieving higher levels of availability.

A new approach was required and a project was defined to establish the underlying reasons that prevented customers from obtaining cash from its ATM's. This required new methods of measurement and reporting which identified a wide range of causes.

The IT organisation was able to identify the range of IT component causes that prevented customers obtaining cash and proceed with a programme of focused improvement activities. This had two effects. Firstly, it improved the overall End - End availability of the ATM service to provide improved availability of the ATM to the customer and secondly, resulted in the number of customers unable to withdraw cash due to IT component non-availability being significantly reduced.

Key observations from this exercise were: -

- The IT improvements were about 'changing the way' things are done, e.g. scheduling
- No technology upgrades were necessary
- Measurement identified improvement opportunities that traditional measures masked
- The IT support organisation now gained valuable insight into the End – End service
- That further improvements to the End to End service were also achieved by business process improvements, e.g. cash replenishment procedures.

9.2.3 Terminology

To aid and provide an ease of understanding, the most important terms used throughout the Chapter 9 - Availability Management - are explained as follows: -

Availability

Availability is the ability of an IT service or component to perform its required function at a stated instant or over a stated period of time.

Availability (or rather unavailability) is the key indicator of service quality perceived by the business and end user. Availability is underpinned by the reliability and maintainability of the IT infrastructure and effectiveness of the IT support organisation. In summary, availability will depend on: -

- Availability of components
- Resilience to failure
- The quality of maintenance and support
- The quality, pattern and extent of deployment of operational process and procedures

An IT service that consistently meets its SLA availability targets will have the characteristics of low frequency of failure and rapid resumption of service after an incident has occurred.

Reliability

The reliability of an IT service can be qualitatively stated as freedom from operational failure. The reliability of an IT service is determined by: -

- The reliability of each component within the IT infrastructure delivering the IT service, i.e. the probability that a component will fail to provide its required functions.
- The level of resilience designed and built into the IT infrastructure, i.e. the ability of an IT component failure to be masked to enable normal business operations to continue.
- The levels of preventive maintenance applied to prevent failures occurring. (Can also be considered an aspect of Maintainability)

Maintainability

Maintainability relates to the ability of an IT infrastructure component to be retained in, or restored to, an operational state. Maintainability of an IT infrastructure component can be divided into 6 separate stages:

- The anticipation of failures
- The detection of failures
- The diagnosing of failures
- The resolving of failures
- The recovery from failures
- The restoration of the data and IT service

Serviceability

Serviceability describes the contractual arrangements made with Third Party IT service providers, e.g. Facilities Management. This is to assure the Availability, Reliability and Maintainability of IT services and components under their care.

It is important to recognise that Serviceability in itself cannot be measured as a specific metric. It is the Availability, Reliability and Maintainability of IT service and components under their care that must be measured.

Vital Business Function

The term Vital Business Function (VBF) is used to reflect the business critical elements of the business process supported by an IT service. An IT service may support a number of business functions that are less critical. E.g. for an Auto Teller Machine (ATM) service the VBF would be the dispensing of cash. However the ability to obtain a mini statement print from an ATM may not be considered as vital. This distinction is important and should influence availability design and associated costs.

End User

The term end user is used to describe the consumer of an IT service. Thus no distinction is made or necessary between describing availability from the internal (business end user) or external (end customer of a business to consumer service) perspective.

IT support organisation

The term IT support organisation is used to describe the IT functions necessary to support, maintain and manage the IT infrastructure to enable an IT service to meet the level of availability defined within the SLA. The IT support organisation can consist of internal and/or external suppliers.

The requirements for internal suppliers will be documented within Operational Level Agreement(s). Underpinning contracts are commissioned for external suppliers

Availability Management (the activity)

The term 'Availability Management' is used throughout each section of this chapter to indicate the execution of activities within the process. A process will however require people for its successful deployment and execution. An organisation may wish to assign this to one individual, one organisational function or have roles and responsibilities associated with the process assigned across multiple organisational areas.

9.3 The Availability Management process

The scope of Availability Management covers the design, implementation, measurement and management of IT infrastructure availability. This is reflected in the process description shown diagrammatically in Figure 9.5 and described in the following paragraphs.

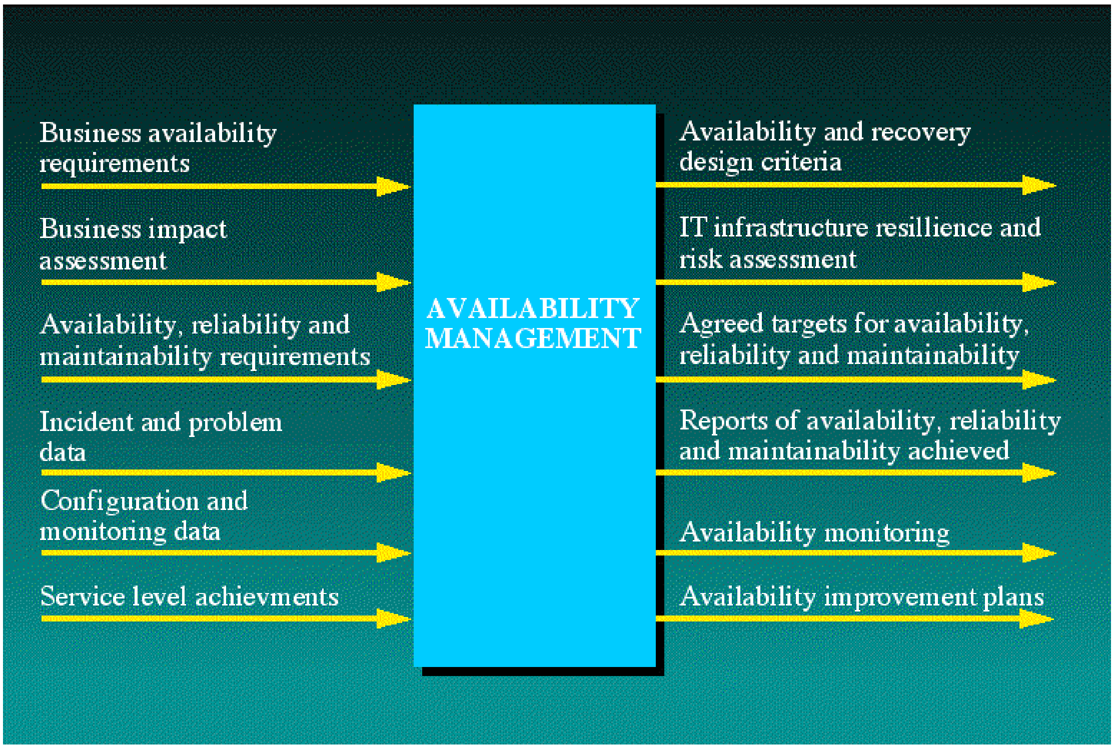


Figure 9.5: - High level Availability Management process diagram

Availability Management commences as soon as the availability requirements for an IT service is clear enough for these to be articulated. It's an ongoing process, finishing only when the IT service is decommissioned.

9.3.1 Key inputs

The key inputs to the Availability Management process are: -

- The availability requirements of the business for a new or enhanced IT service.
- A business impact assessment for each vital business function underpinned by the IT infrastructure.
- The Availability, Reliability and Maintainability requirements for the IT infrastructure components that underpin the IT service(s).
- Information on IT service and component failure(s), usually in the form of Incident and Problem records.
- A wide range of configuration and monitoring data pertaining to each IT service and component..

- Service level achievements against agreed targets for each IT service that has and agreed SLA.

9.3.2 Key outputs

The key outputs from the Availability Management process are: -

- Availability and Recovery design criteria for each new or enhanced IT service.
- Details of the availability techniques that will be deployed to provide additional infrastructure resilience to prevent or minimise the impact of component failure to the IT service.
- Agreed targets of Availability, Reliability and Maintainability for the IT infrastructure components that underpin the IT service(s).
- Availability reporting of Availability, Reliability and Maintainability to reflect the business, end user and IT support organisation perspectives.
- The monitoring requirements for IT components to ensure that deviations in Availability, Reliability and Maintainability are detected and reported.
- Availability Plan for the proactive improvement of the IT infrastructure availability.

9.3.3 Key activities

The key activities of the process are as follows: -

- Determining the availability requirements from the business for a new or enhanced IT service and formulating the availability and recovery design criteria for the IT infrastructure
- In conjunction with ITSCM determining the vital business functions and impact arising from IT component failure. Where appropriate reviewing the availability design criteria to provide additional resilience to prevent or minimise impact to the business.
- Defining the targets for Availability, Reliability and Maintainability for the IT infrastructure components that underpin the IT service to enable these to be documented and agreed within SLAs, OLAs and contracts.
- Establishing measures and reporting of Availability, Reliability and Maintainability that reflects the business, end user and IT support organisation perspectives.
- Monitoring and trend analysis of the Availability, Reliability and Maintainability of IT components.
- Reviewing IT Service and component availability and identifying unacceptable levels of availability.
- Investigation of the underlying reasons for unacceptable availability.
- Production and maintenance of an Availability Plan which prioritises and plans IT availability improvements.

9.3.4 Relationships with other IT Service Management disciplines

All IT Service Management disciplines will have an influence on IT availability therefore Availability Management will by implication interface with all disciplines. These can be considered bi-directional with Availability Management providing inputs whilst also having a dependency on the outputs from the other IT Service Management disciplines.

Example 1: - Service Level Management An input from Availability Management to the Service Level Management process will be an assessment of the availability that can be delivered for a new IT service to enable the SLA to be negotiated and agreed. An output from Service Level Management to the Availability Management process will be details of the agreed SLA that will enable the appropriate availability measurement and reporting to be instigated.

Example 2: - IT Service Continuity Management An output from IT Service Continuity Management will be a business impact assessment detailing the vital business functions dependent on IT infrastructure availability. An Input from Availability Management to IT Service Continuity Management will be the availability and recovery design criteria to maintain ‘business as usual’ by preventing or minimising the impact of failures by use of techniques such as Component Failure Impact Assessment (CFIA).

Example 3: - IT Financial Management An input from Availability Management to IT Financial Management will be the cost of non-availability arising from the loss of an IT service(s) to help cost justify improvements defined within the Availability Plan. An output from IT Financial Management to Availability Management will be the costs associated with proposed upgrades to the IT infrastructure to deliver increased levels of availability.

Example 4: - Capacity Management An input from Availability Management to Capacity Management will be a completed CFIA for a new IT service denoting where availability techniques will be deployed to provide additional infrastructure resilience. An output from Capacity Management to Availability Management will be the capacity plan detailing how the capacity requirements associated with the provision of additional infrastructure resilience will be met.

Example 5: -

Change Management

An input from Availability Management to Change Management will be details of the planned maintenance regime, ie frequency, duration and impact, for components underpinning a new IT service. An output from Change Management to Availability Management will be a schedule of planned maintenance activities for IT components detailing the times and IT services that will be impacted.

Deficiencies in any of the IT Service Management processes may impact Availability. Techniques such as Systems Outage Analysis (SOA) may identify process improvement opportunities for the other IT Service Management processes. In this context, Availability Management can be a driver for process improvements within the overall IT Service Management framework deployed within an organisation. Please refer to Paragraph 9.9.8 for more information on SOA.

9.3.5 Benefits, Costs and Problems

Benefits of Availability Management

The principal benefit of Availability Management is that IT services with an availability requirement are designed, implemented and managed to consistently meet that target. The IT availability requirement being delivered at a known and justified cost and to a predetermined level of quality and security.

Availability Management if deployed with a strong emphasis on the business and end user will ensure the IT organisation recognise availability as the primary IT deliverable to generate and sustain the behaviours of continuous improvement and a service culture.

The benefits of Availability Management can be summarised as follows: -

- A single point of accountability for availability (process owner) is established within the IT organisation.
- IT services are designed to meet the IT availability requirements determined from the business
- The levels of IT availability provided are cost justified
- The levels of availability required are agreed, measured and monitored to fully support Service Level Management.
- Shortfalls in the provision of the required levels of availability are recognised and appropriate corrective actions identified and implemented.
- A business and end user perspective of IT availability is taken to ensure optimal usage and performance of the IT infrastructure is achieved to deliver maximum benefit.
- The frequency and duration of IT failures is reduced over time.
- IT support organisation mindset moves from error correction to service enhancement; from reactive to proactive attitude.

- The IT support organisation is seen to ‘add value’ to the business.

Without an effective deployment of Availability Management the following are examples of the likely difficulties and issues that can arise: -

- It is difficult to define service levels regarding availability that are specific, measurable, achievable and comprehensible to both the business and IT support organisation.
- The management and performance of internal and external suppliers is jeopardised by the lack of agreed and measured availability targets.
- It becomes difficult to assess what level of IT availability is achievable and cost effective
- New IT services are implemented without full consideration of how IT availability will be delivered. The lack of design focus on availability and recovery may lead to costly retrospective changes to improve availability.
- Instability of new IT services results in lost business opportunity and damage to business reputation
- Availability issues are not recognised or owned within the IT support organisation due to a lack of clear accountability.
- The consequences of failing to deliver consistent and agreed levels of availability inevitably leads to dissatisfaction, a lack of trust and conflict between the business and the IT support organisation

Key Message: -

“Where today’s problems are seen as addressed by the latest technology and tools, it is important to reflect that *availability* cannot be purchased....it must be designed, implemented, measured and managed. This is Availability Management”.

Costs

The costs associated with implementing and executing Availability Management will include: -

- The staff costs associated with the process owner role for Availability Management. These would include salary, training, recruitment costs and if needed the cost of initial consultancy.
- Accommodation costs
- Process deployment costs to define and implement the necessary process, procedures and techniques.
- Support tools for monitoring and reporting

Availability Management may identify investment opportunities to improve availability, e.g. new systems management tools. However, these costs should be considered against the availability requirement and the business case justified or rejected as appropriate. These are not the costs of implementing Availability Management.

Possible Problems

The possible problems that may inhibit the establishment and deployment of Availability Management are often of an organisational nature. Typical problems encountered are: -

- The IT Directorate view availability as a responsibility of all senior managers and therefore are reluctant to justify the costs of appointing a single individual as accountable for availability.
- The IT Directorate and supporting organisation have difficulty understanding how 'Availability Management' can make a difference particularly where the existing disciplines of Incident, Problem and Change Management are already deployed.
- The IT Directorate view current levels of availability as good so see no compelling reason for the creation of a new role within the organisation.
- There is resistance to process ownership and the concept of an accountable individual/role who has authority over all the IT support organisation.
- The IT Directorate fail to delegate the appropriate authority and empowerment to enable the process owner for Availability Management to influence all areas of the IT support organisation.

The above are clearly major inhibitors. However, assuming the IT Directorate intends to establish the process and associated role, then other problems can impact implementation and ongoing activity. These can include: -

- Lack of available resources with the required skills and competencies to establish Availability Management.
- The lack of specific Availability Management tools to underpin and support the process
- The need to create bespoke tool and reporting solutions in the absence of appropriate in-house or marketplace offerings.
- The lack of existing service management processes that provide a key input to Availability Management, i.e. Service Level Management, Configuration Management and Problem Management.
- Availability Management can be viewed by some managers as everything or nothing.

None of the problems identified are insurmountable providing there is strong leadership and commitment from the IT Directorate in creating and empowering the role and pragmatism during its implementation to focus on areas of immediate benefit and to accept support tool shortcomings.

VALIDATION: -

Following a period of IT systems instability the senior executive of an IT organisation requested a major IT consultancy group to perform an independent availability review.

A key emphasis placed on the review was not simply to focus on the incidents that had occurred but to assess the capability of the IT organisation to meet the anticipated demand for increasing levels of availability for future IT services.

When the review concluded, two key points within the executive summary presented to the senior executive and sponsor of the review were: -

"There is a lack of focus on availability within the IT support organisation"
"The end user view has been lost"

The major recommendation proposed to address these issues was to implement an Availability Management process and appoint the role of an Availability Manager as the process owner.

The recommendation was accepted and progressed. Two years later, the same consultancy group was asked to undertake an audit of the Availability Management process and its implementation to assess progress. The executive summary of this review concluded: -

"A dramatic change in attitude in the last 2 years"
"Availability now recognised as the key element in providing IT services"
"Decisions now being made on the basis of value and cost of downtime"
"Recognised as important to the success of the business"
"Gets things done"
"Greater sharing of responsibilities between IT and the business"

It is important to reflect that the above achievements were accomplished by appointing a single individual as process owner with sole accountability for the implementation and ongoing execution of the Availability Management process.

9.4 The Cost of (Un)availability

9.4.1 Availability

The level of availability required by the business will influence the overall cost of the IT service provided. In general, the higher the level of availability required by the business the higher the cost. These costs are not just the procurement of the base IT technology and services required to underpin the IT infrastructure. Additional costs will be incurred in providing the appropriate service management processes, systems management tools and high availability solutions required to meet the more stringent availability requirements.

When considering how the availability requirements of the business will be met, it is important to ensure that the level of availability to be provided for an IT service is at the level actually required and is affordable and cost justified by the business.

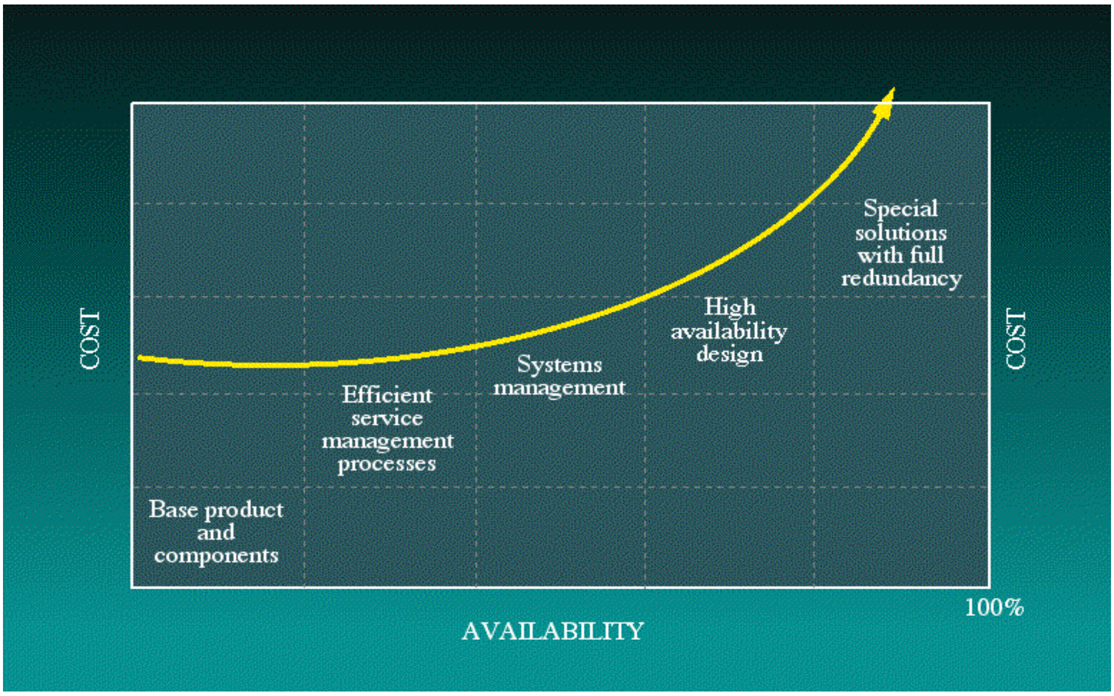


Figure 9.6: - Illustrating how the higher level of availability required the higher the overall cost.

New IT Services

Where new IT services are being developed it is essential that Availability Management takes an early and participating role in determining the availability requirements. This will enable Availability Management to positively influence the IT infrastructure design to ensure that it can deliver the level of availability required.

The importance of this participation early in the design of the IT infrastructure cannot be underestimated. The benefits of this early involvement are: -

- Provide an early indication of the costs required to meet the availability requirement.
- Where costs are seen as too high, enables alternative options with their associated costs and consequences to be presented to the business.
- Higher levels of availability can be achieved when availability is designed in, rather than added on.
- Avoids the costs and delays of late design changes to meet the required levels of availability.
- Ensures the IT infrastructure design will deliver the required levels of availability.

Key Message: -

"The costs of having to retrofit and re-engineer late in the design stage or once in production will incur significant additional costs and in many cases this can be more expensive than building or procuring the initial IT service"

Existing IT Services

Changing business needs and consumer demand may require the levels of availability provided for an IT service to be reviewed. Such reviews should form part of the regular service reviews with the business undertaken by Service Level Management.

Where high levels of availability are already being delivered it may take considerable effort and incur significant cost to achieve a small incremental availability improvement.

A key activity for Availability Management is to continually look at opportunities to optimise the availability of the IT infrastructure. The benefits of this approach being that enhanced levels of Availability may be achievable but with much lower costs.

The optimisation approach is a sensible first step to deliver better value for money. A number of Availability Management techniques can be applied to identify optimisation opportunities. It is recommended that the scope should not be restricted to the IT infrastructure, but also include a review of the business process and other end – end business owned responsibilities.

Validation: -

The previous example of how a financial organisation raised Auto Teller Machine (ATM) availability levels without additional technology investment included a review of their business process.

The approach to identify the reasons for customers being unable to obtain cash from their ATM's highlighted many IT component availability root causes. However, it also identified weaknesses in business process. This lead to a range of improvements that contributed significantly to improved end – end availability of the ATM service to its customers.

Areas included: -

- Improving cash replenishment procedures
- Increased frequency of cash replenishment at high usage ATM locations
- Improved 'cash out' detection and reporting
- Improved ATM hardware fault detection and reporting
- Improved cash demand forecasts to ensure sufficient cash available for peak demand periods

Where the business requirement is to go beyond the optimum level of availability that the IT Infrastructure can deliver; this will incur a significant increase in costs. These increased costs are driven by major redesign of the IT infrastructure and the changing of requirements for the IT support organisation.

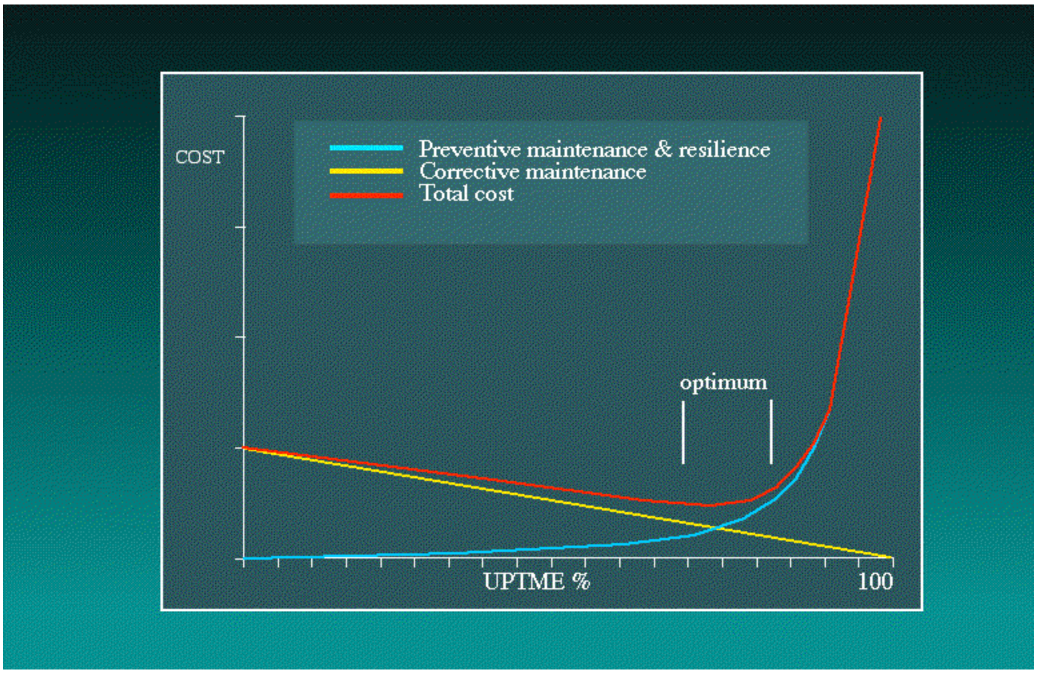


Figure 9.7: - Illustrating how to go beyond optimum availability will significantly impact costs

(See the section *Availability Management Techniques* for guidance on how to apply a range of approaches to further optimise availability for a given service). Information relating to changes in availability requirements and the associated costs should be included in the Availability Plan. If the business have an urgent need to progress this enhanced level of availability an exception report should be produced as the basis for an initial business case.

Key Message: -

"Improving levels of availability can often be achieved by a small step change in thinking rather than a large step change in technology"

9.4.2 Unavailability

The overall costs for an IT service will be influenced by the levels of availability required and the investments required in technology and services provided by the IT support organisation to meet this requirement. Availability certainly does not come for free.

However, it is important to reflect that the unavailability of IT also has a cost .....therefore unavailability isn't for free either. For highly critical business systems it is necessary to consider not only the cost of providing the service but also the costs that are incurred from failure. The optimum balance to strike being the cost of the availability solution weighed against the costs of unavailability.

9.4.3 How much did that failure cost?

Impact on vital business functions

The cost of an IT failure could simply be expressed as the number of business or IT transactions impacted, either as an actual figure (derived from instrumentation) or based on an estimation. When measured against the vital business functions that support the business operation this can provide an obvious indication on the consequence of failure.

The advantage of this approach is the relative ease of obtaining the impact data and the lack of any complex calculations. It also becomes a *'value'* that is understood by both the business and IT organisation. This can be the stimulus for identifying improvement opportunities and can become a key metric in monitoring the availability of the IT service. This approach is recommended as the basis for enhanced SLA reporting. (See the later section *Availability Measurement and Reporting*).

The major disadvantage of this approach is that it offers no obvious monetary value that would be needed to justify any significant financial investment decisions for improving availability.

The monetary impact

Where significant financial investment decisions are required it is better to express the cost of failure arising from system, application or function loss to the business as a monetary *'value'*.

The monetary value can be calculated as a combination of the tangible costs associated with failure, but can also include a number of intangible costs. The monetary value should also reflect the cost impact to the whole organisation, i.e. the business and IT organisation.

Tangible costs

These can include: -

- Lost end user productivity
- Lost IT staff productivity
- Lost revenue
- Overtime payments
- Wasted goods and material
- Imposed Fines or Penalties

These costs are often well understood by the finance area of the business and IT organisation and in relative terms are easier to obtain and aggregate than the intangible costs associated with an IT failure.

Intangible costs

These can include: -

- Loss of customer goodwill (customer dissatisfaction)
- Loss of customers
- Loss of business opportunity (to sell, gain new customers etc)
- Damage to business reputation
- Loss of confidence in IT
- Damage to staff morale

It is important not to simply dismiss the intangible costs (and the potential consequences) on the grounds that they are difficult to measure.

Key Message:-

The Internet and other electronic delivery channels are accelerating the trend that results in changes to the way business customers will now react to IT service failure.

In the emerging world of E-Commerce these intangible costs (business reputation, customer dissatisfaction, loss of existing customers and loss of business opportunity) become far more important.

9.4.4 The benefit of understanding the cost of unavailability

The major benefit provided is that it provides a true financial cost on the consequences of unavailability and therefore provides an objective view of any ‘cost Vs benefit’ assessment. In organisations where decision-makers invariably have to consider the ‘bottom line’ this provides them with information that should facilitate sensible and informed decisions.
The importance of understanding the cost of unavailability should be highlighted as a key performance indicator within the design and operational phases of Service Level Management.

Problems with deriving a cost of unavailability

There are however some problems to consider and overcome with this approach, namely: -

- The wide scope of what can be factored into the cost calculation.
- The difficulties of quantifying the cost impact of many of the intangible consequences.
- Resistance from the business to divulging cost information (concerns about how the IT organisation will use information, i.e. lowered priority if cost is low or continued approaches to invest if costs are high).
- Reluctance from the IT Directorate to disclose and make public the service costs of each of their business customers.

- The time and effort of obtaining the cost impact data.

A pragmatic approach is recommended and the level to which you attempt to quantify the monetary cost of a failure will invariably be influenced with the cost investment to be justified.

An example of a calculation to be used in quantifying the cost of a failure can be found within Section 9.8 Availability Management Methods & Techniques

9.5 Availability Planning

9.5.1 Determining Availability requirements

Before any service level requirement is accepted and ultimately the SLR or SLA is agreed between the business and the IT organisation it is essential that the availability requirements of the business are analysed to assess if/how the IT infrastructure can deliver the required levels of availability.

This applies not only to new IT services that are being introduced but also any requested changes to the availability requirements of existing IT services.

It is important that the business is engaged early in the development lifecycle so that the business needs regarding the availability of a new or enhanced IT service can be obtained. This will provide sufficient time to influence the appropriate IT infrastructure design and set expectation with the business on the associated costs. This is particularly important where stringent availability requirements may require additional investment in service management processes, systems management tools, high availability design and special solutions with full redundancy.

It is likely that the business need for IT availability cannot be expressed in technical terms. Availability Management therefore provides an important role in being able to translate the business and end user requirements into quantifiable availability terms and conditions. This is an important input into the IT Infrastructure design and will provide the basis for assessing the capability of the IT infrastructure design and IT support organisation in meeting the availability requirements of the business.

The business requirements for IT availability should at least contain: -

- A definition of the vital business functions supported by the IT service.
- A definition of IT service downtime, i.e. the conditions under which the business consider the IT service to be unavailable.
- The business impact caused by loss of service.
- Quantitative availability requirements, i.e. the extent to which the business will tolerate IT service downtime or degraded service.
- The required service hours, i.e. when the service is to be provided.
- An assessment of the relative importance of different working periods.
- Specific security requirements.

Hints & Tips

The translation of end user and business availability requirements into quantifiable terms and conditions is crucial. To avoid any confusion or misunderstanding between the Business and IT the first step should be to document and agree a description or definition of the availability terms and conditions that will be used.

What the business understand by downtime, availability, reliability etc may differ from the IT perspective. This will avoid any misunderstandings and enable subsequent design activities to commence with a clear and unambiguous understanding of what is required.

Availability Management having assessed the combined capability of the IT infrastructure design and IT support organisation will be in a position to confirm if the availability requirements can be met. Where shortfalls are identified, dialogue with the business will be required to present the cost options that exist to enhance the proposed design to meet the availability requirements. This will enable the business to reassess if lower or higher levels of availability are required and understand the appropriate impact and costs associated with their decision.

Determining the availability requirements will likely be an iterative process particularly where there is a need to balance the business availability requirement against the associated costs. The necessary steps are: -

- Determine the business impact caused by loss of service.
- From the business requirements specify the availability, reliability and maintainability requirements for the IT components controlled by the IT support organisation.
- For IT services and components provided externally the serviceability requirements should be identified.
- Estimating the costs involved in meeting the availability, reliability, maintainability and serviceability requirements.
- Determine with the business if the costs identified in meeting the availability requirements are justified
- Determine from the business the costs likely to be incurred from loss or degradation of service.
- Where these are seen as cost justified, the availability, reliability, maintainability and serviceability requirements should be defined in agreements and negotiated into contracts.

Hints & Tips

If costs are seen as prohibitive, either:
Reassess the IT infrastructure design and provide options for reducing costs and assess the consequences on availability.

Or reassess the business use and reliance on the IT service and renegotiate the availability targets to be documented in the SLA

The Service Level Management (SLM) function will normally be responsible for communicating with the business on how their availability requirements will be met and ultimately negotiating the SLA for the IT service. Availability Management will therefore provide important support to the SLM function during this period.

Whilst higher levels of availability can often be provided by technology investment there is no justification for providing a higher level of availability than that needed and afforded by the business. The reality is that satisfying availability requirements will always be a balance between cost and quality.

This is where Availability Management can play a key role in optimising availability of the IT infrastructure to meet increasing availability demands whilst deferring an increase in costs. (See sections titled Designing for Recovery and Availability Management Methods and Techniques for additional guidance.

9.5.2 Design activity

Designing for availability is a key activity driven by Availability Management. This ensures that the required level of availability for an IT service can be met.

Availability Management needs to ensure that the design activity for availability looks at this from two related but distinct perspectives: -

DESIGNING FOR AVAILABILITY: - This relates to the technical design of the IT infrastructure and the alignment of the internal and external suppliers required to meet the availability requirements for an IT service.

> **Key Message:-**
>
> **Designing for Availability:- This can be considered the proactive perspective aimed at *avoiding* IT service downtime.**

DESIGNING FOR RECOVERY: - This relates to the design points required to ensure that in the event of an IT service failure, that the service can be re-instated to enable normal business operations to resume as quickly as is possible.

> **Key Message:-**
>
> **Designing for Recovery:- This can be considered the reactive perspective aimed at *minimising* the business and end user impact from an IT service failure.**

By taking this two phased approach to the design activity will ensure new IT services do not suffer unnecessary and extended recovery when the first failure situation occurs.

Additionally, the ability to recover quickly may be a crucial factor. In simple terms it may not be possible or cost justified to build a design that is highly resilient to failure(s). The ability to meet the availability requirements within the cost parameters may rely on the ability to recover in a timely and effective manner, consistently.

9.5.3 Designing for Availability

Availability should be considered in the design process at the earliest possible stage of the development lifecycle. This will avoid the potential for: -

- Increased development costs arising from re-work
- Unplanned expenditure on IT upgrades necessary to meet the availability requirement.
- Unplanned expenditure on additional systems management tools required to operationally control and manage the IT service and supporting IT infrastructure.
- Unplanned expenditure to eliminate single points of failure (SPOF) within the configuration
- The inability of internal or external suppliers to meet the maintainability and serviceability requirements.
- Delays to the planned implementation and the potential for missed business opportunity.

- Incurring avoidable IT service failures

The following figure illustrates a high level outline of how initial availability requirements are progressed by Availability Management to ensure these can be met by the IT infrastructure and IT support organisation. This simple framework can be applied to new IT services or existing IT services where a change of availability requirements has required major redesign.

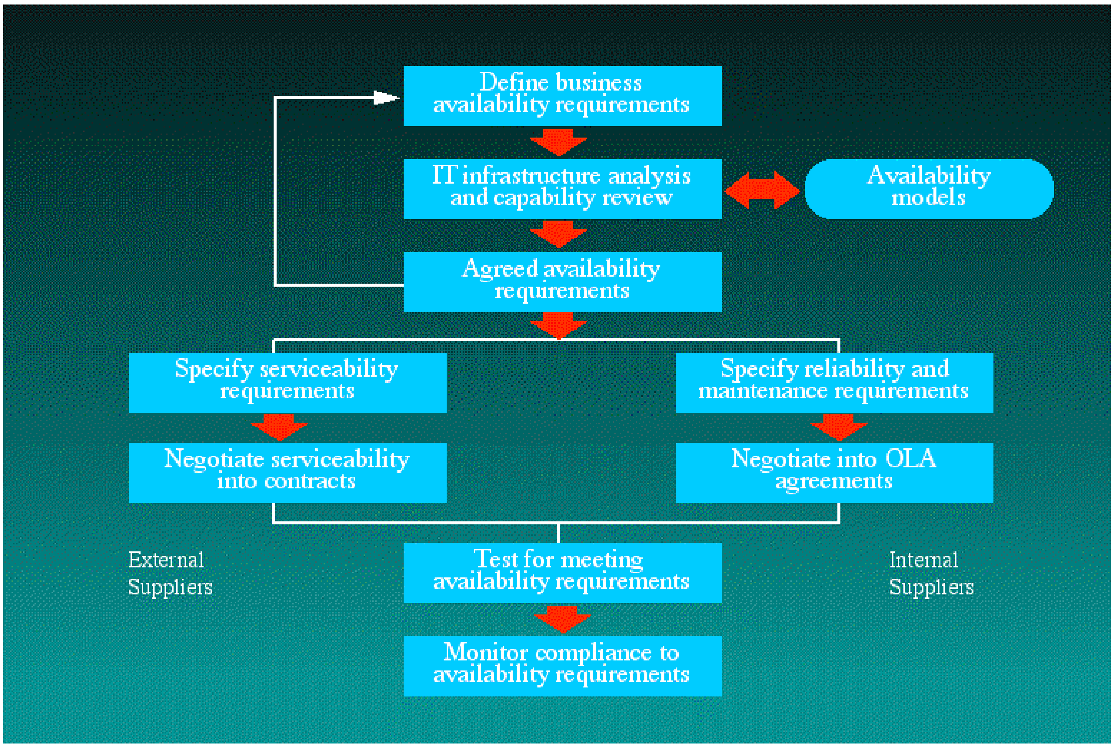


Figure 9.8: - Progressing availability requirements for an IT service

The role of Availability Management within the design activities is to provide: -

- The specification of the availability requirements for hardware and software
- The requirements for availability measurement points (instrumentation)
- The requirements for new/enhanced systems management
- Participation in the IT infrastructure design
- The specification of the reliability, maintainability and serviceability requirements for components supplied by internal and external suppliers
- Validation of the final design to meet the minimum levels of availability required by the business for the IT service

The framework that Availability Management should utilise to determine the appropriateness of a given design to meet the stated availability requirements will consist of the following: -

- The minimum levels of availability required by the business for the IT Service
- The proposed technical design which should cover the full end to end configuration

- Identification of the internal and external suppliers
- The minimum reliability, maintainability and serviceability levels for each component
- The provision of an availability modelling tool
- The ability to test or simulate that new components within the design match the specified requirements

Method of approach

IT Infrastructure analysis – Capability Review

The first stage is to understand the vulnerability to failure of the proposed IT infrastructure design.

- **Single Points of Failure (SPOF)**

A single point of failure is any component within the IT infrastructure that has no back-up capability and can cause impact to the business and end user when it fails.

It is important that no unrecognised single points of failure exist within the IT infrastructure design. The use of Component Failure Impact Assessment (CFIA) as a technique to identify single points of failure is recommended.

Where these are identified, CFIA can be used to identify the business and end user impact and help determine what alternatives can or should be considered to cater for this weakness in the design. Refer to Paragraph 9.9.1 for additional guidance on the use of CFIA.

- **Risk Analysis and Management**

To assess the vulnerability of failure within the configuration and capability of the IT support organisation it is recommended that the proposed IT infrastructure design and supporting organisation (internal and external suppliers) are subject to a formal risk analysis. The CCTA Risk Analysis and Management Method (CRAMM) is a technique that can be used to identify justifiable countermeasures that can protect the Availability of IT Systems. (Refer to the methods and techniques section of this chapter for additional guidance on CRAMM)

- **Testing or Simulation**

To assess if new components within the design can match the stated requirements it is important that the testing regime instigated ensures that the availability expected can be delivered. Simulation tools to generate the expected end user demand for the new IT service should be seriously considered to ensure components continue to operate under volume and stress conditions.

Improving the design

The second stage is to re-evaluate the IT infrastructure design if the availability requirements cannot be met and identify cost justified design changes.

Improvements in design to meet the availability requirements can be achieved by reviewing the capability of the technology to be deployed in the proposed IT Infrastructure design, e.g.

- The exploitation of fault tolerant technology to mask the impact of planned or unplanned component downtime.

- Duplexing or the provision of alternative system components to allow one system component to take over the work of another system component.
- Improving component reliability by enhancing testing regimes.
- Improved release and change management for changes to key components.
- Improved software design and development.
- Improved processes and procedures.
- Systems Management enhancements / exploitation

Chapter 8 – IT Service Continuity Management can provide additional guidance on this aspect of IT infrastructure design.

Hints & Tips

Consider documenting the Availability design requirements and considerations for new IT services and make available to the areas responsible for design and implementation. Longer term seek to mandate these requirements and integrate within the appropriate governance mechanisms that cover the introduction of new IT services.

Considerations for High Availability

Where the business operation has a high dependency on IT availability and the cost of failure or loss of business reputation is considered not acceptable, the business may define stringent availability requirements. These factors may be sufficient for the business to justify the additional costs required to meet these more demanding levels of availability.

Achieving high levels of availability begins with the procurement and/or development of good quality products and components. However, these in isolation are unlikely to deliver the sustained levels of availability required.

To achieve a consistent and sustained level of high availability will require investment and deployment of effective service management processes, systems management tools, high availability design and ultimately special solutions with full redundancy.

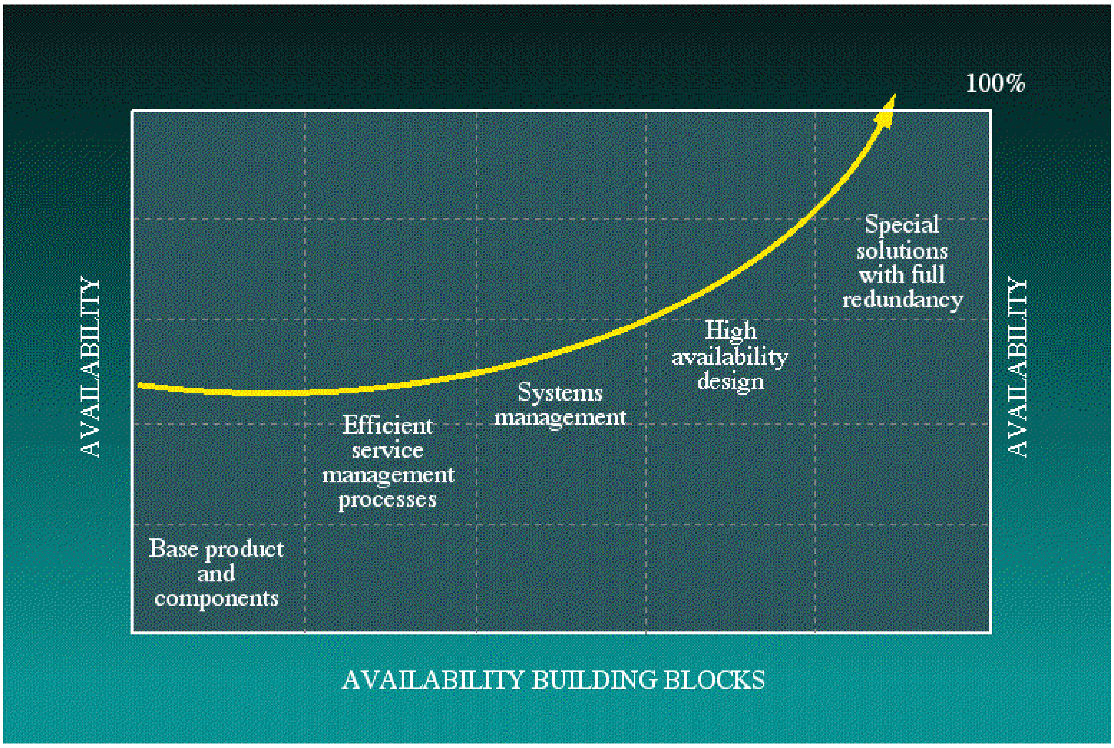


Figure 9.9: - The building blocks to meet the more stringent availability requirements

Figure 9.9 illustrates quite simply, that to achieve higher levels of availability will require investment in more than just the base product and components. This is a similar diagram to that presented in Figure 9.6 which emphasises how the higher the availability deamand the higher the overall costs.

The above can therefore be viewed as a framework for what needs to be considered within the overall design for availability where stringent availability requirements are set.

This suggested framework is described as follows: -

Base product and components

The procurement or development of the base product and components should be based on their capability to meet stringent availability and reliability requirements. These need to be considered the cornerstone of the availability design. The additional investment required to achieve even higher levels of availability will be wasted and availability levels not met if these base products and components are unreliable and prone to failure.

Service Management processes

Efficient service management processes will contribute to higher levels of availability. Processes such as Availability Management, Incident management, Problem Management, Change Management etc play a crucial role in the overall management of the IT service.

Systems Management

Systems Management should be exploited to provide the monitoring, diagnostic and automated error recovery to enable fast detection and resolution of potential and actual IT failure situations.

High Availability Design

The design for high availability will need to consider the elimination of single points of failure and/or the provision of alternative components to provide minimal disruption to the business operation should an IT component failure occur.

The design will also need to eliminate or minimise the effects of planned downtime to the business operation normally required to accommodate maintenance activity, the implementation of changes to the IT infrastructure or business application.

Recovery criteria will define rapid recovery and IT service reinstatement as a key objective within the designing for recovery phase of design.

Special Solutions with full redundancy

To approach continuous availability in the range of 100% will require expensive solutions that incorporate full redundancy. Redundancy is the technique of improving availability by using duplicate components. For stringent availability requirements to be met these need to be working autonomously in parallel. These solutions are not just restricted to the IT components, but also the IT environment, i.e. power supplies, air conditioning, telecommunications.

Helpful additional definitions when defining stringent vailability requirements.

As stated in the earlier section (Determining availability requirements), the translation of the business and end user availability requirements into quantifiable availability terms and conditions is crucial.

Where stringent levels of availability are required additional definitions should be documented and agreed between the business and IT to ensure both parties understand the specific high availability conditions.

The suggested additional definitions are: -

- **HIGH AVAILABILITY**

A characteristic of the IT service that minimises or masks the effects of IT component failure to the end user.

- **CONTINUOUS OPERATION**

A characteristic of the IT service that minimises or mask the effects of planned downtime to the end user.

- **CONTINUOUS AVAILABILITY**

A characteristic of the IT service that minimise or masks the effects of ALL failures and planned downtime to the end user.

These definitions help better define the often used term of 'High availability'. This provides a better structure for determining which areas of availability design are most important to the business.

Industry view:-

Many suppliers will commit to high availability or continuous availability solutions only if stringent environmental standards are used. They will often only agree to such contracts after a site survey has been completed and additional, sometimes-costly improvements have been made.

9.5.4 Designing for Recovery

Designing for availability is a key activity driven by Availability Management. This ensures that the stated availability requirements for an IT service can be met.

However, Availability Management should also ensure that within this design activity there is focus on the design elements required to ensure that when IT services fail, that the service can be reinstated to enable normal business operations to resume as quickly as is possible.

'Designing for recovery' may at first sound negative. However, the reality is that failures will occur. The way the IT organisation manages failure situations can have the following positive outcomes: -

- Normal business operations are resumed quickly to minimise impact to the business and end user.
- The availability requirements are met within the cost parameters set as a result of timely and effective recovery reducing the amount of downtime incurred by the business.
- The IT organisation are seen as responsive and business focused.

A key message for the IT organisation being that every failure is a 'moment of truth' with every failure being an opportunity to make or break your reputation with the business.

By providing focus on the 'designing for recovery' aspects of the overall availability design will ensure that every failure is an opportunity to maintain and even enhance business and end user satisfaction.

Designing for Recovery - needs

To provide an effective 'design for recovery' it is important to recognise that both the Business and the IT organisation have needs that must be satisfied to enable an effective recovery from IT failure.

- **Business needs**

These are informational needs which the business require to help them manage the impact of failure on their business and set expectation within the business, end user community and their business customers.

- **IT needs**

These are the process, procedures and tools required to enable the technical recovery to be completed in an optimal time.

> **Hints & Tips**
>
> Consider documenting the Recovery design requirements and considerations for new IT services and make available to the areas responsible for design and implementation. Longer term seek to mandate these requirements and integrate within the appropriate governance mechanisms that cover the introduction of new IT services.

Key elements in the Design for Recovery

The role of Incident Management and the Service Desk

A key aim is to avoid small incidents becoming major by ensuring the right people are involved early enough to avoid mistakes being made and to ensure the appropriate business and technical recovery procedures are invoked at the earliest opportunity.

This is the responsibility of the Incident management process and role of the Service Desk.

To ensure business needs are met during major IT service failures and to ensure the most optimal recovery, the Incident management process and Service Desk will need to have defined and execute: -

- Stringent escalation procedures
- Clearly defined roles & responsibilities for handling major Incidents
- A communications plan to support the wider informational needs associated with major incidents

Key Message: -

The above are not the responsibilities of Availability Management. However, the effectiveness of the Incident Management process and Service Desk can strongly influence the overall recovery period. The use of Availability Management methods and techniques to further optimise IT recovery may be the stimulus for subsequent continuous improvement activities to the Incident management process and Service Desk.

Understanding the Incident 'lifecycle'

It is important to recognise that every incident passes through a number of stages. These are described as follows: -

- Incident Start
- Incident Detection
- Incident Diagnosis
- Incident Repair
- Incident Recovery
- Service Restoration (& verification)

This 'lifecycle' view provides an important framework in determining amongst others, systems management requirements for incident detection, diagnostic data capture requirements and tools for diagnosis, recovery plans to aid speedy recovery and how to verify that IT service has been restored.

The ITIL Service Support book provides a chapter on Incident Management and the incident 'lifecycle' as viewed from the handling of incidents. To aid the designing for recovery this lifecycle has been expanded to reflect an IT availability perspective of an incident.

Please refer to Paragraph 9.9.9 for additional guidance and illustration on the use of the expanded incident lifecycle.

Systems Management

The provision of Systems Management tools will positively influence the levels of availability that can be delivered. Implementation and exploitation should have strong focus on achieving high availability and enhanced recovery objectives.

I the context of recovery, such tools should be exploited to provide automated failure detection, assist failure diagnosis and support automated error recovery.

Diagnostic data capture procedures

When IT components fail it is important that the required level of diagnostics are captured, to enable problem determination to identify the root cause. For certain failures the capture of diagnostics may extend service downtime. However, the non-capture of the appropriate diagnostics creates and exposes the service to repeat service failures.
Where the time required taking diagnostics is considered excessive; a review should be instigated to identify if techniques and/or procedures can be streamlined to reduce the time required. Equally the scope of the diagnostic data available for capture can be assessed to ensure only the diagnostic data considered essential is taken.

The additional downtime required to capture diagnostics should be included in the recovery metrics documented for each IT component.

Determine backup and recovery requirements

The backup and recovery requirements for the components underpinning a new IT service should be identified as early as possible within the development or selection cycle. These requirements should cover hardware, software and data. The outcome from this activity should be a documented set of recovery requirements that will enable the development of appropriate recovery plans.

Develop and test a backup and recovery strategy and schedule

To anticipate and prepare for performing recovery such that reinstatement of service is effective and efficient will require the development and testing of appropriate recovery plans based on the documented recovery requirements.

The outcome from this activity should be clear, operable and accurate recovery plans that are available to the appropriate parties immediately the new IT Service is introduced.

Wherever possible, the operational activities within the recovery plan should be automated.

The testing of the recovery plans will also deliver approximate timings for recovery. These recovery metrics can be used to support the communication of estimated recovery of service and validate or enhance the CFIA documentation.

Recovery Metrics

The provision of a timely and accurate estimation of when service will be restored is the key informational need of the business. This information will enable the business to make sensible decisions on how they will manage the impact of failure on their business and their customers.
To enable this information to be communicated to the business will require the creation and maintenance of recovery metrics for each IT component and cover a variety of recovery scenarios.

Please refer to the section Availability management techniques for guidance on how Component Failure Impact Analysis (CFIA) can be used to derive recovery metrics to support the communications element by providing estimated recovery timings.

Backup and recovery performance

Availability Management must continuously seek and promote faster methods of recovery for all potential incidents. This can be achieved via a range of methods including automated failure detection, automated recovery, more stringent escalation procedures, exploitation of new and faster recovery tools and techniques.

It is recommended that this aspect of availability measurement is included in the basic set of IT availability measures utilised to measure and report IT availability.

Section 9.9 within Availability Management Methods & Techniques describes how to use the expanded 'Incident lifecycle' as a model for metrics creation. These metrics could be used to drive the 'review backup and recovery performance' process.

Service restoration and verification

An incident can only be considered 'closed' once service has been restored and normal business operation has resumed. It is important that the restored IT service is verified as working correctly as soon as service restoration has completed and before any technical staff involved in the incident are stood down. In the majority of cases this will simply be a case of getting confirmation from the end user. However, the end user for some services may be a customer of the business, i.e. ATM services, Internet based services.

For these types of services it is recommended that IT service verification procedures are developed to enable the IT support organisation to verify that a restored IT service is now working as expected. These could simply be visual checks of transaction throughput or end user simulation scripts that validate the end – end service.

Hints & Tips
There is potential for confusion in distinguishing the aspects of Incident Management appropriate to the Service Desk and to that appropriate for Availability Management.

The goal of the Incident Management process and the aims of Availability Management in designing for recovery are completely complementary, i.e. to restore normal business operation as quickly as possible and to minimise the impact to the business and end user.

The Incident Management process is used by the Service Desk to provide a structured and consistent approach to the handling, tracking and ultimate resolution of all incidents. This is the management perspective best described by the incident lifecycle (ITIL Service Support).

Availability Management is concerned with the methods, tools and techniques employed by the IT support organisation within each stage of the Incident Management 'lifecycle'. This is the technical perspective best described by the expanded incident 'lifecycle' (ITIL Service Delivery).

9.5.5 Security considerations

Availability management is concerned with the ***technical*** availability of IT components. It manages information ***security*** in terms of the availability criterion, namely reliability, maintainability, serviceability and by the resilience designed into the IT infrastructure.

Availability Management is closely connected with Security Management. The importance of availability being recognised as one third of the security 'CIA' tenet:

- **C**onfidentiality
- **I**ntegrity
- **A**vailability

Example: -

An IT service may not be available due to the erroneous deletion of data. This (security) incident resulting in a breach of the Service Level Agreement.

The overall aim of IT security is 'balanced security in depth' with justifiable controls implemented to ensure continued IT service within secure parameters (viz., ***Confidentiality***, ***Integrity*** *and* ***Availability****)*.

During the gathering of availability requirements for new IT services it is important that requirements that cover IT security are defined. These requirements need to be applied within the design phase for the supporting IT infrastructure. The points made in the earlier sections about designing availability into the design at the earliest opportunity equally apply to security controls.

For many organisations the approach taken to IT security will be covered by an IT security policy owned and maintained by Security Management. In the execution of security policy, Availability Management will play an important role in its operation for new IT services.

Hints & Tips

There is potential for confusion between the process owners for Security Management and Availability Management with regard to security requirements for new IT services.

Security Management can be viewed as *accountable* for ensuring compliance to IT security policy for the implementation of new IT services. Availability Management is *responsible* for ensuring security requirements are defined and incorporated within the overall availability design.

Availability Management will gain guidance from the information contained within the organisations IT security policy and associated procedures and methods. However, the following are typical security considerations that must, amongst others be addressed: -

- Products and services must only be available to authorised personnel
- Products and services must be recoverable following failure to ensure confidentiality and integrity are not compromised and availability of service not further compromised.
- Products and services must be recoverable within secure parameters, i.e. must not compromise IT security policy

- Physical access to computer and network equipment should be restricted to authorised personnel only
- Logical access to software should be restricted to authorised personnel only
- Operating System and Systems Management command authority should be commensurate with role and responsibility
- Data must be available to authorised personnel at agreed times as specified in the Service Level Agreement (SLA)
- Operational Level Agreements (OLA) and Service contracts must reflect the adherence to security controls required by the IT support organisation.

Industry view: - <u>The Purpose of information security</u>? "Information Security protects information from a wide range of threats in order to ensure business continuity, minimise business damage and maximise return on investments and business opportunity" Source: BS 7799 – The UK code of practice for Information Security.

For further reference to Information Security refer to ITIL Security Management, CCTA, ISBN 011330014X.

9.5.6 Managing planned downtime

Maintenance

All IT components should be subject to a planned maintenance strategy. The frequency and levels of maintenance required will vary from component to component taking into account the technologies involved, criticality and the potential business benefits that may be introduced.

Planned maintenance activities enable the IT support organisation to provide: -

- Preventative maintenance to avoid failures
- Software or Hardware upgrades to provide new functionality or additional capacity
- Business requested changes to the business applications
- Activation of new infrastructure features for exploitation

The requirement for planned downtime will clearly influence the level of availability that can be delivered for an IT service, particularly those that have stringent availability requirements.

In determining the availability requirements for a new or enhanced IT service the amount of downtime and the resultant loss of income required for planned maintenance may not be acceptable to the business. This is becoming a growing issue in the area of E commerce.
In these instances it is essential that continuous operation is a core design feature to enable maintenance activity to be performed without impacting the full IT service.

Where the required service hours for an IT service are less than 24 hrs per day and/or 7 days per week, it is likely that the majority of planned maintenance can be accommodated without impacting IT service availability.

However, IT services where the business need a level of availability on a 24 hour and 7 day basis will require Availability Management to determine the most effective approach in balancing the requirements for planned maintenance against the loss of service to the business. Unless mechanisms exist to allow continuous operation, scheduled downtime for planned maintenance is essential if high levels of availability are to be achieved and sustained. For all IT services there should logically be a ‘low impact’ period for the implementation of maintenance.

Once the requirements for managing scheduled maintenance have been defined and agreed, these should be documented as a minimum in the following: -

- SLAs
- OLAs
- Underpinning contracts
- Change Management schedules
- Release Management schedules

The areas responsible for implementing and managing change, i.e. Service Desk, Network Management and Computer Operations, will need to be made aware of the maintenance targets and any future revisions.

Key Message: -

Availability Management should ensure that building in a low impact period for preventative maintenance is one of the prime design considerations for a 24x7 IT service.

Minimising Business Impact

Assessing Service Impact

The output from the Component Failure Impact Analysis (CFIA) will indicate for a given component the impact on the end user when the component is not available. The definition of IT service downtime obtained when determining the availability requirements will establish the level of business impact arising from the non-availability of this component.

The CFIA will also indicate if an alternative configuration item can continue to provide the service. Where an alternative configuration item is available service impact is minimal dependent on how quickly the alternate configuration item is activated.

For components that have an alternative configuration item, the maintenance policy agreed with the internal or external supplier should be to ensure planned maintenance to these components are not scheduled concurrently.

Scheduling downtime

The most appropriate time to schedule planned downtime is clearly where the impact to the business and their customers is least. This information should be provided initially by the business when determing the availability requirements.

For an existing IT service or once the new service has been established, monitoring of business and customer transactions will help establish the hours where IT service usage is at its lowest. This should determine the most appropriate timing window for the component(s) to be removed for planned maintenance activity.

Aggregation of maintenance activity

To accommodate the individual component requirements for planned downtime whilst balancing the IT service availability requirements of the business provides an opportunity to consider scheduling planned maintenance to multiple components concurrently,

The benefit of this approach is that the number of service disruptions required to meet the maintenance requirements is reduced.

Whilst this approach has benefits, there are potential risks that need to be assessed, for example: -

- The capability of the IT support organisation to co-ordinate the concurrent implementation of a high number of changes.
- The ability to perform effective problem determination where the IT service is impacted after the completion of multiple changes
- The impact of change dependency across multiple components where backout of a failed change requires multiple changes to be removed.

Service Maintenance Objectives

The effective management of planned downtime will be an important contribution in meeting the required levels of availability for an IT service.

Where planned downtime is required on a cyclic basis to an IT component(s), the time that the component is unavailable to enable the planned maintenance activity to be undertaken should be defined and agreed with the internal or external supplier. This becomes a stated objective that can be formalised, measured and reported.

The Service Outage Objective (SMO) for a given planned maintenance activity should be the total time required for the IT component to be unavailable. The SMO should therefore be an aggreagate of the following timeline events:-

- The time required to prepare the component for the maintenance activity, e.g. closedown of the system software
- The time required to complete the implementation of the maintenance activity
- The time required to reinstate the component to enable IT service to be resumed

The benefits of defining Service Maintenance Objectives for cyclic planned maintenance activity are:

- To agree realistic implementation times for the planned maintenance activity.
- To provide Change Management with advance information on change duration.
- To provide the business with advance information on the duration of service downtime.
- To provide the time necessary to abort implementation and backout defective change(s).
- To enable excess downtime to recorded as a non-compliance with the appropriate supplier.
- To provide a focus on implementation defects.
- To provide a measure on quality of change.
- To provide a focus on planned downtime reduction and/or continuous operation techniques.

In addition they also provide an early warning during the maintenance activity of the time allocated to the planned outage duration being breached. This can enable an early decision to be made on whether the activity is allowed to complete with the potential to further impact service or to abort the activity and instigate the backout plan.

Planned downtime and performance against the stated SMO for each component should be recorded and used in service reporting.

Validation: -

An IT organisation that supported a 24 hr IT service looked at ways that they could reduce the amount of downtime required for scheduled maintenance. A review of the scheduled outages revealed wide ranges in the amount of downtime incurred. These were due to a combination of change quality issues impacting the implementation and operational issues with system closedown and restart procedures. Responsibilities for the successful implementation of the changes were split between the area of the IT support organisation supplying the change and the operational area responsible for implementing the change.

The first step taken was to get both parties to agree on what should be a realistic time to complete the change implementation (closedown, application of change, restart of system).

Once agreed this was documented as a formal agreement between both parties upon which there was a shared ownership for meeting the outage objective.

Formal reporting and regular reviews were held. The reasons for failing to meet the target were investigated and addressed. In addition both parties explored opportunities to reduce the target time allocated. Where improvements were made and the results showed a sustained improvement a new outage objective was set.

This process continued year on year until the implementation time was considered optimal. In this time the duration for planned mainframe maintenance had been reduced from extremes of 3 hrs to consistently less then one hour.

9.6 Availability Improvement

A key output from the Availability Management process is the creation of an Availability Plan.

The Availability plan should be a long-term plan for the proactive improvement of IT availability within the imposed cost constraints.

The impetus to improve availability will be from one or more of the following: -

- The inability for a new IT service to meets its SLA on a consistent basis
- Period(s) of IT service instability resulting in unacceptable levels of availability
- Availability measurement trends indicating a gradual deterioration in availability
- Unacceptable IT service recovery and restoration time
- Requests from the business to increase the level of availability provided
- Increasing impact to the business and their customers from IT service failures as a result of growth and/or increased business functionality
- A request from SLM to improve availability as part of an overall SIP
- Availability Management monitoring and trend analysis

9.6.1 Important considerations

Availability Management monitoring and trend analysis

Availability Management should take a proactive role in identifying and progressing cost justified availability improvement opportunities. The ability to do this will place reliance on having appropriate and meaningful availability measurement and reporting.

To ensure availability improvements deliver benefits to the business and end users it is important that availability measurement and reporting reflects not just IT component availability but availability from a business operation and end user perspective. (See section titled Availability Measurement and Reporting for additional guidance).

Determining (changed) Availability Requirements

Where the business have a requirement to improve availability the process outlined in the Availability Planning section should be followed to reassess the IT infrastructure and IT support organisation capability to meet these enhanced requirements.

An output of this activity will be enhanced availability and recovery design criteria.

The cost of improving availability

To satisfy the business requirement for increased levels of availability may drive additional financial investment to enhance the underpinning IT infrastructure and/or extend the services provided by the IT support organisation.

It is important that any additional investment to improve the levels of availability delivered can be cost justified. Determining the cost of an IT failure(s) can help support any financial investment decision. Section 9.4 provides additional guidance on how to derive the cost of IT failure.

However, a key benefit of Availability Management is the opportunity to optimise the availability of the IT infrastructure to deliver an improved level of availability with reduced cost. This optimisation approach is a sensible first step to deliver better value for money. A range of Availability Management methods and techniques can be applied to identify the potential for improved levels of availability at a much lower cost.

9.6.2 Methods and Techniques

There are a number of methods and techniques that can be utilised to identify availability improvement opportunities. These are described as follows: -

Component Failure Impact Assessment (CFIA)

CFIA can be used to predict and evaluate the impact on IT service arising from component failures within the IT infrastructure. The output from a CFIA can be used to identify where additional infrastructure resilience should be considered to prevent or minimise the impact of component failure to the business operation and end users.

Fault Tree Analysis (FTA)

FTA is a technique that can be used to determine the chain of events that causes a disruption to IT services. FTA in conjunction with calculation methods can offer detailed models of availability. This can be used to assess the availability improvement that can be achieved by individual IT infrastructure design options.

The CCTA Risk Analysis and Management Method (CRAMM)

CRAMM can be used to identify new risks and provide appropriate countermeasures associated with any change to the business availability requirement and revised IT infrastructure design.

Systems Outage Analysis (SOA)

SOA is a technique designed to provide a structured approach to identify the underlying causes of service interruption to the end user. SOA will utilise a range of data sources to assess where and why shortfalls in availability are occurring. SOA will enable an holistic view to be taken to drive not just IT infrastructure improvements but improvements to the IT support organisation process, procedures and tools.

SOA is run as an assignment and may utilise other Availability Management methods and techniques to formulate the recommendations for improvement.

The Expanded Incident 'lifecycle'

An aim of Availability Management is to ensure the duration and impact from incidents impacting IT service are minimised to enable business operations to resume as quickly as is possible.

The expanded incident 'lifecycle' enables you to break down the total IT service downtime for any given incident and map this against the major stages that all incidents progress through. (The lifecycle)

This will enable you to identify where 'time is being lost' and provide the basis for the identification of improvements that can improve recovery and restoration times.

Continuous Improvement

Availability Management can play an important role in helping the IT support organisation recognise where they can add value by exploiting their technical skills and competencies in an availability context. The continuous improvement technique can be used by Availability Management to harness this technical capability. This can be used with either small groups of technical staff or a wider group within a workshop environment.

Technical Observation Post (TOP)

A TOP is a prearranged gathering of specialist technical support staff from within the IT support organisation brought together to focus on specific aspects of IT availability. Its purpose being to monitor events, real time as they occur with the specific aim of identifying improvement opportunities or bottlenecks which exist within the current IT infrastructure.

For more detailed information and guidance on how these methods and techniques can be deployed please refer to Section 9.9 within this chapter.

9.6.3 The Availability Plan

To provide structure and aggregation of the wide range of initiatives that will be undertaken to improve availability these should be formulated within a single Availability Plan.

The Availability Plan should have goals, objectives and deliverables and should consider the wider issues of people, process, tools and techniques as well as having a technology focus. In the initial stages it may be aligned with an implementation plan for Availability Management, but the two are different and should not be confused.

As the Availability Management process matures the plan should evolve to cover the following:

- Actual levels of availability Vs agreed levels of availability for key IT services. Where possible availability measurements should be business focused to report availability as experienced by the business and end user.
- Activities being progressed to address shortfalls in availability for existing IT services. Where investment decisions are required options with associated costs and benefits should be included.
- Details of changing availability requirements for existing IT services. The plan should document the options available to meet these changed requirements. Where investment decisions are required the associated costs of each option should be included.
- Details of the availability requirements for forthcoming new IT services. The plan should document the options available to meet these new requirements. Where investment decisions are required the associated costs of each option should be included.
- A forward looking schedule for the planned Systems Outage Analysis (SOA) assignments.
- Regular reviews of SOA assignments should be completed to ensure that infrastructure availability is being proactively improved.
- A technology futures section to provide an indication of the potential benefits and exploitation opportunities that exist for planned technology upgrades. Anticipated availability benefits should be detailed, were possible based on business focused measures. The effort required to realise these benefits where possible should also be quantified.

During the production of the Availability Plan, it is recommenced that liaison with the following functional areas is undertaken: -

- Service Level Management, concerning changing business and end user requirements for existing IT services.
- IT Service Continuity Management concerning business impact and resilience improvements.
- Business Relationship Management to understand major customer concerns and/or future needs that relate to IT availability.
- Capacity Management, concerning the scenarios for upgrading (or downgrading) the software, hardware and network layers of the IT infrastructure.
- IT Financial Management concerning the cost and budget implications of the various options identified for availability improvement.
- Application Management, concerning the availability requirements for new services.
- Areas responsible for IT supplier management and the managing of relationships and contracts with suppliers.
- Technical support groups responsible for testing and maintenance functions, concerning the reliability and maintainability of existing services.

The Availability Plan should cover a period of one to two years with a more detailed view and information for the first six months. The plan should be reviewed regularly with minor revisions every quarter and major revisions every half year. Where the IT infrastructure is only subject to a low level of change this may be extended as appropriate.

It is recommended that the Availability Plan is considered complementary to the capacity plan and publication aligned with the capacity and business budgeting cycle.

If a demand is foreseen for high levels of availability that cannot be met due to the constraints of the existing IT infrastructure or budget, then exception reports may be required for the attention of both senior IT and Business management.

Hints & Tips

There is potential for confusion on the purpose of an Availability Plan Vs a Service Improvement Programme.

The Availability Plan is a forward looking plan aimed at improving the overall availability of the IT infrastructure to ensure that existing and future levels of availability can be provided on a timely and cost effective basis.

The Availability Plan is a key output and deliverable of the Availability Management process. It is reviewed and revised on an ongoing basis. Improvements to the IT infrastructure may benefit many IT services where common infrastructure is utilised.

Service Level Management are responsible for instigating service improvements that improve the overall quality of the whole IT service provision.
The mechanism used to achieve this is the Service Improvement Programme (SIP). This SIP should be used to co-ordinate all IT service improvement opportunities into an overall programme of improvement activities.

The SIP has a defined start and end and its scope can include all elements of the service provided to the business. This may include availability, but equally may not dependent on the overall service indicators and the areas of desired improvement.

Availability Management can play a key role in supporting a SIP by the appropriate use of availability management techniques, e.g. Systems Outage Analysis.

9.7 Availability Measurement & Reporting

A key output from the Availability Management process is the measurement and reporting of IT availability. This will provide the basis for: -

- Establishing measures of availability and agreeing availability targets with the business.
- Monitoring of the actual availability delivered versus agreed targets.
- Identifying unacceptable levels of availability that impact the business and end user.
- Reviewing availability with the business and end user representatives.
- Reviewing availability with the IT support organisation.
- Continuous improvement activities to optimise availability.

Availability measures should be incorporated into Service Level Agreements (SLA), Operational Level Agreements (OLA) and any underpinning contracts. These should be reviewed regularly at Service Level review meetings.

Famous sayings and truths about measurements: -

"If you don't measure it, you can't manage it"
"If you don't measure it, you can't improve it"
"If you don't measure it, you probably don't care"
"If you can't influence it, then don't measure it"

9.7.1 IT Availability reporting

"What to measure and how to report it" will inevitably depend on which activity is being supported, who the recipients are and how the information will be utilised. It is important to recognise the differing perspectives of availability to ensure measurement and reporting satisfies these varied needs: -

- The ***IT Support Organisation*** perspective will consider IT component availability with regard to Availability, Reliability and Maintainability.
- The ***End User*** perspective will consider IT Service availability as a combination of three factors, namely; the frequency, the duration and the scope of impact, i.e. all users, some users, all business functions or certain business functions. The End User will also consider IT service availability in terms of response times. For many performance centric applications poor response times are considered equal in impact to IT component failures.
- The ***Business*** perspective will consider IT service availability in terms of its contribution or impact on the vital business functions that drive the business operation.

In order to satisfy the differing perspectives of availability, Availability Management will need to consider the spectrum of measures needed to report the 'same' level of availability in different ways. Measurements need to be meaningful and add value if availability measurement and reporting are to ultimately deliver benefit to the IT and business organisations. This will be strongly influenced in the combination of 'what you measure' and 'how you report it'.

Key Message

Availability Management will need consider availability from an IT service perspective and from an IT component perspective. These are entirely different aspects. Whilst the underlying concept is similar, the measurement, focus and impact is entirely different.

9.7.2 IT Availability measures – the traditional view

The IT support organisation have for many years measured and reported on their perspective of availability. Traditionally these measures have concentrated on component availability and have been somewhat divorced from the business and end user views.

Typically these traditional measures are based on a combination of an availability percentage (%), time lost and the frequency of failure. Some examples of these traditional measures are as follows:-

- **% Available**

The truly ‘traditional’ measure which represents availability as a percentage. It is typically used to track and report achievement against a service level target. It tends to emphasise the ‘big number’ such that if the service level target was 98.5% and the achievement was 98.3%, then it does not seem that bad. This can encourage a complacent behaviour within the IT support organisation.

- **% Unavailable**

The inverse of above. This representation however has the benefit of focusing on non-availability. Based on the above example, if the target for non-availability is 1.5% and the achievement was 1.7% then this is a much larger relative difference. This method of reporting tio is more likely to create awareness of the shortfall in delivering the level of availability required.

- **Duration**

Achieved by converting the percentage unavailable into hours and minutes. This provides a more ‘human’ measure that people can relate to. If the weekly downtime target was 2 hrs but one week the actual was 4 hrs; this would represent a trend leading to an additional 4 days of non-availability to the business over a full year. This type of measure and reporting is more likely to encourage focus on service improvement.

- **Frequency of failure**

Used to record the number of interruptions to the IT service. It helps provide a good indication of reliability from an end user perspective. It is best used in combination with ‘duration’ to take a balanced view of the level of service interruptions and the duration of time lost to the business.

The business have for many years accepted as a fait accompli that the IT availability that they experience is represented in this way. However, this is no longer being viewed as acceptable and the business are keen to better represent availability in measure(s) that demonstrate the positive and negative consequences of IT availability on their business and end users.

9.7.3 *The disadvantages of traditional IT Availability measures*

The continued advances in IT technology design and service management focus will result in availability improvements. These improvements will be clearly evidenced by the traditional IT availability measurements.

A positive consequence of technology advancement and IT availability improvements is the creation of new business opportunity i.e. to offer new and enhanced services, to gain more customers and to increase customer value. However, with this business benefit comes an increased business dependency on IT availability and a greater business impact when IT failure occurs.

Here lies an interesting paradox. The traditional IT availability measures show a trend of continuous improvement with non-availability declining. Whereas the business oriented measures used by the business show the impact from non availability increasing!

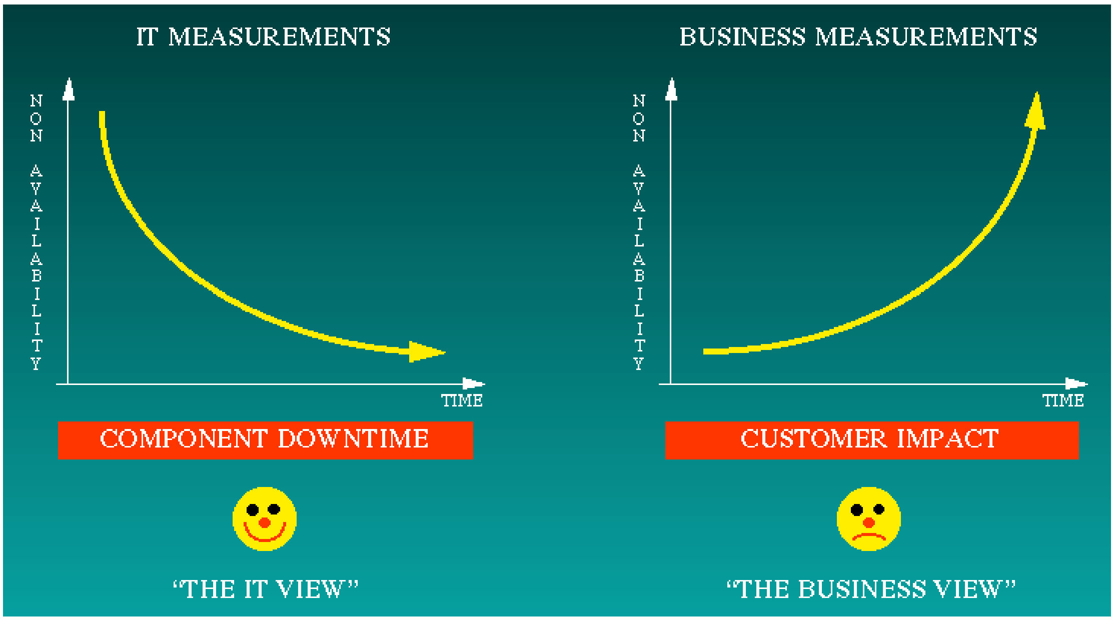


Figure 9.10: - The disadvantages of traditional IT availability measurements

The traditional IT approach to measurement and reporting provides an indicator on IT availability and component reliability which is important for the internal IT support organisation. However, to the business and end user these measures fail to reflect availability from their perspective and are rarely understood. This often fuels mistrust between the business and IT where despite periods of instability the '%' target has been met even though significant business disruption has occurred and customer complaints have been received.

Furthermore, this method of measurement and reporting can often hide the benefits delivered to the business from IT improvements. The traditional IT availability measures can simply mask real IT 'value add' to the business operation.

Whilst the traditional IT availability measurement and reporting can be considered appropriate for internal IT reporting, the disadvantages of this approach can be summarised as follows: -

- They fail to reflect IT availability as experienced by the business and the end user
- They can conceal service 'hot spots' whereby regular reporting shows the SLA 'met', but the business and/or end user is becoming increasingly dissatisfied with the IT service
- They do not easily support continuous improvement opportunities to drive improvements that can benefit the business and the end user
- They can mask IT 'value add' where tangible benefits to the business and end user have been delivered but the method of measurement and reporting does not make this visible.

9.7.4 Measuring end user Availability

Availability when measured and reported to reflect the experience of the end-user provides a more representative view on overall IT service quality.

The end user view of availability is influenced by three factors: -

- The frequency of downtime.
- The duration of downtime
- The scope of impact

Measurements and reporting of end user availability should therefore embrace these factors.

The methodology employed to reflect end user availability could consider two approaches: -

- Impact by end user minutes lost. This is to base calculations on the duration of downtime multiplied by the number of users impacted. This can be the basis to report availability as lost end user productivity or to calculate the availability percentage from an end user perspective.
- Impact by business transaction. This is to base calculations on the number of business transactions that could not be processed during the period of downtime. This provides a better indication of business impact reflecting differing transaction processing profiles across the time of day, week etc. In many instances it may be the case that the end user impact correlates to a vital business function, e.g. if the end user takes customer purchase orders and a vital business function is customer sales. This single measure is the basis to reflect impact to the business operation and end user.

The method to employ should be influenced by the nature of the business operation. A business operation supporting data entry activity is well suited to reporting that reflects end user productivity loss. Business operations that are more customer facing, e.g. ATM services, benefit from reporting transaction impact.

9.7.5 Business driven measurement and reporting

The IT support organisation needs to have a keen awareness of the end user experience of availability. However, the real benefits will come from aggregating the end user view into the overall business view.

A guiding principle of the Availability Management process is that ***'Improving availability can only begin when you understand how the technology supports the business'***. Therefore Availability Management is not just about understanding the availability of each IT component.

From the business perspective, an IT service can only be considered available when the business is able to perform all vital business functions required to drive the business operation. For the IT service to be available it is therefore reliant on all components which the service depends to be available, i.e. system, key components, network, data and application.

The traditional IT approach would be to measure individually the availability of each of these components. However, the true measure of availability has to be based upon the positive and negative impacts on the vital business functions upon which the business operation is dependent.

This approach ensures that SLA's and IT availability reporting are based on measures that are understood by both the business and IT. By measuring the vital business functions that rely upon IT, measurement and reporting becomes business driven with the impact of failure reflecting the consequences to the business.

> Industry view: -
>
> "Effective management requires the ability to view availability from the perspective of the business application"
>
> Source: - International Data Corporation

Establishing the VBF to measure

This clearly needs to be undertaken and agreed with the business. A simple way to establish this is to simply ask the business *'what is the major consequence on your business or customers if there is an IT failure?'*

Whilst for most IT services the major consequence will be understood and obvious, it is still worth probing the business on their initial response to ensure that jointly the true consequence is established.

> Example: - Auto Teller Machine (ATM) service
> Q: - *what is the major consequence on your business or customers if there is an IT failure?*
>
> Is it: -
> A1: - The consequence of IT failure impacting an ATM service is that our ATM machines are unavailable.
> A2: - The consequence of IT failure impacting an ATM service is that our customers cannot obtain cash from our ATM machines.
>
> Whilst both are correct, the vital business function for an ATM service is the ability to dispense cash not provide an ATM. The more meaningful and powerful metric to measure therefore is the obtaining of cash.

Benefits

The benefits of basing availability measurement and reporting upon the vital business functions can be summarised as follows: -

- Provides a ‘common’ measure which both parties to the SLA can understand
- Can visibly demonstrate to the business tangible service improvement enabling the IT organisation to show the ‘value add’.
- Can more easily identify degrading levels of service to enable the IT support organisation to be proactive in responding without formal business escalation.
- Can be used to demonstrate the business and end user impact with suppliers to drive and influence positive supplier behaviour.

Problems

Some of the possible problems with trying to develop this approach are: -

- How to relate business experience to incidents, especially if no end – end monitoring is available.
- An absolute measure, e.g. number of business transaction impacted, on a rapidly expanding IT service can show a downward trend even when the overall % availability improves.
- Who owns the VBF measurements and data?
- Integration and mapping of this data with IT component availability data.

For some example approaches you can undertake to develop business and end user driven measurement please refer to the later section ***Availability Management techniques.***

9.7.6 What to measure

To satisfy the differing reporting needs of the Business and IT support organisation it will be necessary to provide a spectrum of measures that reflect the differing perspectives of IT service and component availability.

The following are examples of the different elements of an IT service that should be measured and reported: -

- **Vital Business Functions**

To provide measurement and reporting that demonstrate the consequences (contribution and impact) of IT availability on the business function(s) key to the business operation, e.g. for a Tele-sales business a vital business function would be customer order fulfilment.

- **Application Services**

To provide measurement and reporting of the application services required to run the business operation and service end user input.

- **Data**

To provide the measurement and reporting of data availability that is essential to support the business operation. E.g., many business operations are dependent on all data being updated by overnight batch processing to provide a new ‘start of day’ position.

- **Key Components**

To provide measurement and reporting that reflects Availability, Reliability and Maintainability of IT infrastructure components supplied and maintained by the IT Support Organisation.

- **Platform**

To provide measurement and reporting of the IT platform that ultimately supports the processing of the business application(s).

Reporting Dimensions

To provide a balanced and meaningful view of the availability of an IT service or component, reporting should consider the following dimensions: -

- **Availability**

To provide measurement and reporting that reflects availability against defined and agreed targets.

- **Reliability**

To provide measurement and reporting that reflects the frequency of failures.

- **Maintainability**

To provide measurement and reporting that reflects the duration of failures.

- **Response Times**

To provide measurement and reporting that reflects the performance as experienced by the end user. Capacity Management may supply (Reports)

9.7.7 IT Availability Metrics Model

The following IT Availability Metrics Model (ITAMM) is suggested to help you consider the range of measures and reporting dimensions that you should consider when establishing availability measurement and reporting.

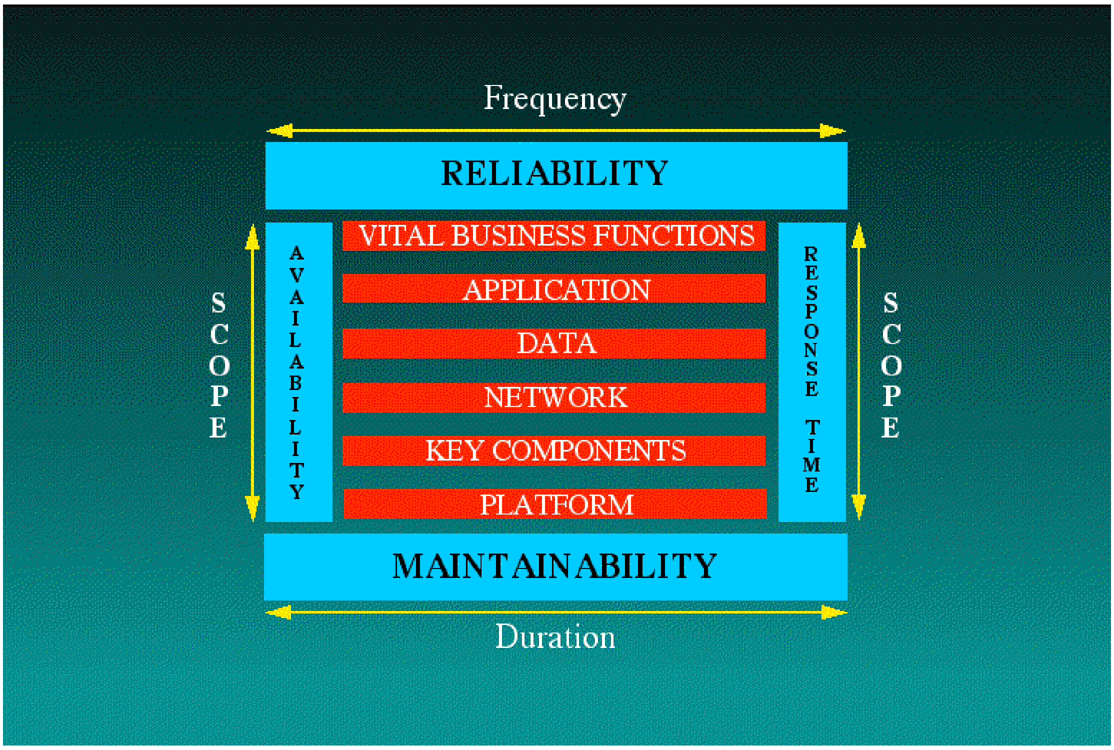


Figure 9.11: - The IT Availability Metrics Model (ITAMM*)*

Key Message: -
Measurements need to be meaningful and add value if availability measurement and reporting are to ultimately deliver benefit to the IT and business organisation. This will be strongly influenced in the combination of 'what you measure' and 'how you report it'.

Gaining additional value from measurement and reporting

Consider where the availability measurement and reporting produced to support the Availability Management process can be used as input to other IT service management processes. Some examples are as follows: -

- Capacity Management: - To highlight availability trends that indicate capacity or performance issues.
- IT Financial Management: - To provide cost of failure information, incorporate availability levels into profit/cost models.
- Service Level Management: - Provide reporting for SLA and OLA activities.
- Incident and Problem Management: - Highlight problem blackspots impacting availability, useful areas for increasing preventative maintenance.
- Change Management: - Availability impact due to poor quality change, % of planned maintenance activities that have overrun their agreed Service Outrage Objectives.

9.8 Availability Management Tools

Availability Management to be effective will have a requirement for a range of tools to support key activities of the process. The initial tool requirements to support the establishment of Availability Management can be summarised as follows: -

- IT component downtime data capture and recording.
- Database repositories for the collection of appropriate availability data and information.
- Report generation.
- Statistical analysis.
- Availability modelling.
- Systems Management

It is important when considering the requirements for tools to initially distinguish between those that are required to support the activities associated with the Availability Management process and those required to improve IT availability, i.e. Systems Management.

However tool selection should be influenced where an integrated solution is available that meets the needs of the process whilst helping achieve the overall goal of improved levels of availability.

9.8.1 IT component downtime data capture and recording

The capture and recording of data pertaining to IT component downtime (planned and unplanned) is a key requirement for the forecasting and reporting of availability. It would be usual to find this information available from the Service Desk, Problem Management and Change Management sections of the IT support organisation.

To improve the accuracy and quality of information it is recommended that component downtime detection and data recording be automated.

9.8.2 Database repositories

Shared database repositories

A number of database repository tools will be required to support other IT infrastructure management processes. Therefore, the implementation and deployment of Availability Management can be simplified where these database tools already exist.

The database support tools that can be utilised by Availability management are as follows: -

Configuration Management Database (CMDB)

The CMDB will provide information regarding the IT infrastructure configuration. Information from the CMDB can provide essential input to assist Availability Management with forecasting the availability of a given IT infrastructure, assessing single points of failure, identification of IT component owners etc.

The CMDB is also an essential source of information relating to incidents, problems and changes affecting the IT infrastructure and IT services.

Capacity Database (CDB)

The CDB will provide information regarding the Capacity Management of the IT infrastructure. Information from the CDB can provide essential input to assist Availability Management identify planned upgrades to hardware, software and network components, workload and application capacity/performance data.

The Availability Management Database (AMDB)

Availability Management may wish to consider having its own database repository. The Availability Management Database (AMDB) can be utilised to record and store selected data and information required to support key activities such as report generation, statistical analysis and availability forecasting.

The AMDB should be the main repository for the recording of IT availability metrics based on the ITAMM model.

The AMDB can also be used to store Availability Management deliverables such as the Availability Plan and SOA assignment report and action plans.

All other information pertaining to availability should reside within the CMDB and CDB to provide reusability of data to multiple IT service management processes and avoid unnecessary duplication and cost.

Hints & Tips: -

Be pragmatic. Define the initial tool requirements and identify what is already deployed that can be used and shared to get started as quickly as possible.

Where basic tools are not already available, work with other IT Service Management disciplines to identify common requirements with the aim of selecting shared tools and minimising costs.

The AMDB should address specific reporting needs of Availability Management not provided/or accommodated by existing repositories.

9.8.3 Report generation & statistical analysis

Reporting and analysis tools are required for the manipulation of data stored in the various databases utilised by Availability Management. These tools can either be platform or PC based. This will be influenced by the database repository technologies selected and the complexity of data processing and reporting required.

Availability Management once implemented and deployed will be required to produce regular reports on an agreed basis, e.g. monthly availability reports, Availability Plan, SOA status reports etc.

However the proactive nature of the role in seeking opportunities to further optimise the availability of the IT infrastructure and the need to respond to emerging availability requirements will require flexibility. When defining requirements the need for interactive database query tools should be stated as an essential requirement.
For reporting purposes organisational reporting standards should be used wherever possible. If these don’t exist then IT standards should be developed so that IT reports can be developed using standard tools and techniques. This will mean that the integration and consolidation of reports will subsequently be much easier to achieve.

9.8.4 *Availability Modelling*

Modelling tools are required to forecast availability and to assess the impact of changes to the IT infrastructure. Inputs to the modelling process will include descriptive data of the component reliability, maintainability and serviceability.

A spreadsheet package to perform calculations is usually sufficient. If more detailed and accurate data is required a more complex modelling tool may need to be developed. The lack of readily available availability modelling tools in the marketplace may require such a tool to be developed and maintained ‘in-house’.

Unless there is a clearly perceived benefit from such a development and the ongoing maintenance costs, the use of existing tools and spreadsheets should be sufficient.

9.8.5 *Systems Management*

An output from the Availability Management process will be the real time monitoring requirements for IT services and components. To achieve the levels of availability required and/or ensure the rapid restoration of service following an IT failure will require investment and exploitation of a systems management toolset.

Systems Management tools are an essential building block for IT services that require a high level of availability (See figure 9.9) and can provide an invaluable role in reducing the amount of downtime incurred.

Availability Management requirements will cover the detection and alerting of IT service and component exceptions, automated escalation and notification of IT failures and the automated recovery and restoration of components from known IT failure situations.

9.9 Availability Management Methods & Techniques

The capability of the Availability Management process will be positively influenced by the range and quality of methods and techniques that are available for deployment and execution within the process.

To provide the reader and prospective process owners with the benefits of established best practice within the field of Availability Management this section documents a number of proven methods and techniques that can be applied to support key activities within the process.

These are as follows: -

1) Component Failure Impact Analysis (CFIA)
2) Fault Tree Analysis (FTA)
3) The CCTA Risk Analysis and Management Method (CRAMM)
4) Calculating Availability
5) Calculating the cost of unavailability
6) Developing basic IT availability measurement and reporting
7) Developing business and end user measurement and reporting
8) Systems Outage Analysis (SOA)
9) The Incident ‘lifecycle’
10) Continuous Improvement methodology

11) Technical Observation Post (TOP)

The above techniques support the three key facets of Availability Management, namely the planning for availability, the improvement of availability and the reporting of availability.

The following table provides guidance on which aspect of Availability Management these techniques can be utilised: -

Facet / Technique	Availability Planning	Availability Improvement	Availability Reporting
CFIA	✓	✓	✓
FTA	✓	✓	
CRAMM	✓	✓	
Calculating Availability	✓	✓	✓
Calculating the Cost of Unavailability	✓	✓	✓
Developing Basic IT Availability Measurement and Reporting		✓	✓
Developing Business and End User Measurement and Reporting		✓	✓
SOA		✓	
The Expanded Incident ‘Lifecycle’	✓	✓	✓
Continuous Improvement		✓	
TOP		✓	

Figure 9.12: - Guidance on the use of Availability Management techniques

9.9.1 Component Failure Impact Analysis (CFIA)

During the 'Design for Availability' activities it will be necessary to predict and evaluate the impact on IT service availability arising from component failures within the proposed IT infrastructure design.

Component Failure Impact Analysis (CFIA) is a relatively simple technique that can be used to provide this information. IBM devised CFIA in the early 1970's with its origins based on hardware design and configuration. However, it is recommended that CFIA be used in a much wider context to reflect the full scope of the IT infrastructure, i.e. Hardware, Network, Software, Application and End Users.

Additionally the technique can also be applied to identify impact and dependencies on IT support organisation skills and competencies amongst staff who will support the new IT service.

This activity is often completed in conjunction with IT Service Continuity Management (ITSCM)

Benefits of CFIA

The output from a CFIA provides vital information to ensure that the Availability and Recovery design criteria for the new IT service is influenced to prevent or minimise the impact of failure to the business operation and end user.

CFIA achieves this by providing: -

- Identification of single points of failure that can impact IT availability.
- The impact of component failure on the business operation and end users.
- Identification of component and people dependencies.
- Provide component recovery timings.
- The need to identify and document recovery options.
- The need to identify and implement risk reduction measures.

The above will also be the stimulus for input to ITSCM to consider the balance between recovery options and risk reduction measures, i.e. where the potential business impact is high there is a need to concentrate on high availability risk reduction measures, i.e. increased resilience or standby systems.

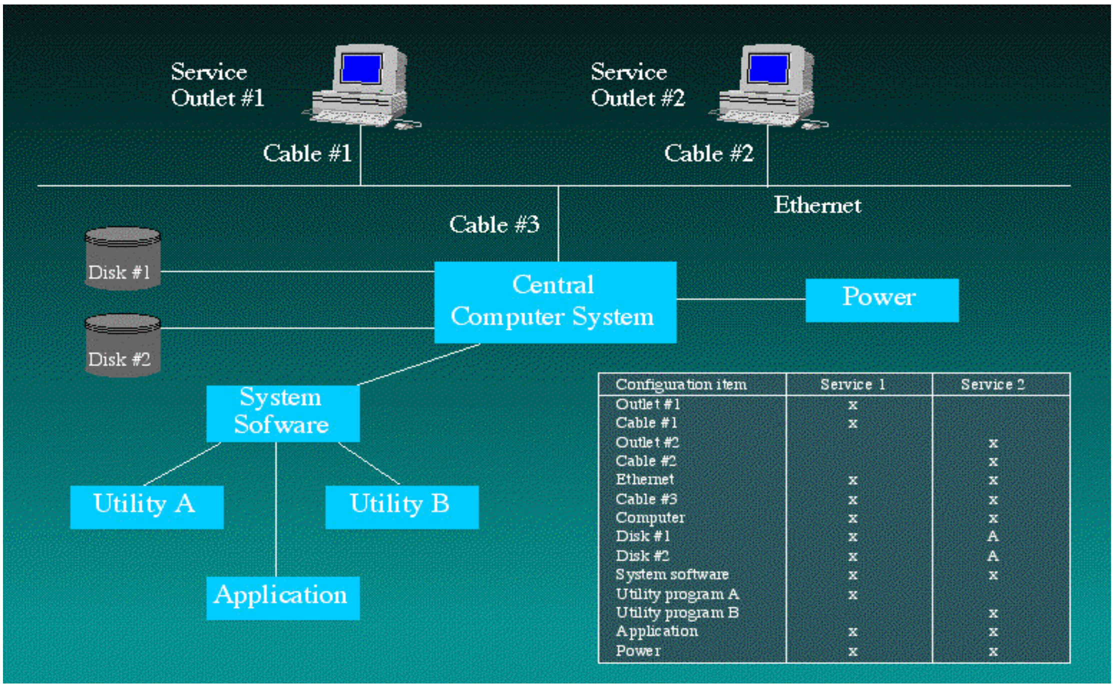
Performing a CFIA

Basic CFIA

Having determined the IT infrastructure configuration to be assessed, the first step is to create a grid with configuration items on one axis and the IT services which have a dependency on the configuration item on the other. This information should be available from the Configuration management Database (CMDB). Alternatively this can be built using documented configuration charts and Service Level Agreements (SLA).

The next step is to perform the following procedure at each intersection point in the grid: -

- Leave a blank when a failure of the configuration item does not impact the service in any way
- Insert an ‘X’ when the failure of the configuration item causes the IT service to be inoperative
- Insert an ‘A’ when there is an alternative configuration item to provide the service
- Insert a ‘B’ when there is an alternative configuration item, but the service has to be recovered first.



Configuration item	Service 1	Service 2
Outlet #1	x	
Cable #1	x	
Outlet #2		x
Cable #2		x
Ethernet	x	x
Cable #3	x	x
Computer	x	x
Disk #1	x	A
Disk #2	x	A
System software	x	x
Utility program A	x	
Utility program B		x
Application	x	x
Power	x	x

Figure 9.13: -Sample configuration and basic CFIA grid.

Having built the grid, Configuration items that have a large number of Xs are critical to many services and can result in high impact should the configuration item fail. Equally, IT services that have high counts of Xs are complex and are vulnerable to failure.

This basic approach to CFIA can provide valuable information in quickly identifying single points of failure, IT services at risk from configuration item failure and what alternatives are available should the configuration item fail.

It should also be used to assess the existence and validity of recovery procedures for the selected configuration items.

The above example assumes common infrastructure supporting multiple IT services. The same approach can be used for a single IT service by mapping the component configuration items against the vital business functions and end users supported by each component. Thus understanding the impact of a component failure on the business and end user.

CFIA MATRIX TO DENOTE VITAL BUSINESS FUNCTIONS
AND END USER IMPACT

Configuration Item	VBF	End users impacted
Power	All	1000
Central computer	All	1000
Application	All	1000
Disk #1	Payments	50
Disk #2	Orders	100
Utility A	Enquiry	25

Figure 9.14: - CFIA matrix reflecting association between components, VBF and end user population.

Advanced CFIA

The above approach can be expanded to provide more detailed information and/or to extend the coverage of the CFIA; e.g. data feeds from 3rd party organisations.

> **Hints and Tips**
>
> **An online IT service can be impacted by failures in batch processing, e.g. the IT service is unable to be started due to dependencies on the completion of overnight database updates.**
>
> **A suggested first step prior to undertaking an advanced CFIA is to identify all the software configuration items that are essential to the 24hr processing cycle. These would include for example transaction processing subsystems, network management subsystems, security management subsystems, systems management subsystems such as tape management, silo management, job scheduling.**
>
> **It can also include software configuration items that support key IT operational processes such as Incident, Problem, Change and Configuration management.**

To undertake an advanced CFIA will require the CFIA matrix to be expanded to provide any additional fields required to for the more detailed analysis.

Some examples of the additional fields that can be included are as follows: -

- **Probability of Failure**

This can be based on the Mean Time Between Failure (MTBF) information if available or on the current availability trends. This can be expressed as a Low/Med/High indicator or as a numeric representation.

- **Recovery Time**

This is the estimated recovery time to recover the configuration item. This can be based on recent recovery timings, recovery information from Disaster Recovery testing or a scheduled test recovery.

- **Recovery procedures**

This is to verify that up to date recovery procedures are available for the configuration item.

- **Device Independence**

Where software configuration items have duplex files to provide resilience this is to ensure that file placements have been verified as being on separate hardware disk configurations.

- **Dependent ID's**

This is to show any dependencies between configuration items. If one configuration item is to fail, there may be an impact on other configuration items, e.g. if the security configuration item was to fail, the operating system may prevent tape processing.

Using CFIA to improve recovery from failure

A detailed CFIA will provide additional information that can be used to improve recoverability. Whilst technology design is continually improving as yet no IT configuration item can ever be considered fail-safe.

A detailed CFIA can provide important information to support the recovery process: -

- **Recovery timings**

To enable IT to provide the business with accurate estimations of when service can be restored.

- **Alternative Available**

To identify what alternative recovery options are available in the event of a configuration item failure.

- **Recovery procedures**

Confidence that valid recovery procedures exist for each configuration item.

- **Dependencies**

What other configuration items may have a dependency? A scenario being that a failed configuration item may have a low impact but if not recovered quickly may cause a dependent configuration item to fail or stop processing causing high impact.

Example: -

A tape management software failure preventing tape usage may not immediately impact the end user. However, the Online Transaction Processing Systems (OLTP) may need to regularly archive system logs to tape. If this cannot be achieved, after a period of time the OLTP may suspend processing directly impacting the end user.

9.9.2 *Fault Tree Analysis (FTA)*

Fault Tree Analysis is a technique that can be used to determine the chain of events that causes a disruption of IT services. This technique, in conjunction with calculation methods can offer detailed models of availability.

The main advantages of FTA are:

- FTA can be used for availability calculations.
- Operations can be performed on the resulting fault tree; these operations correspond with design options.
- The desired level of detail in the analysis can be chosen.

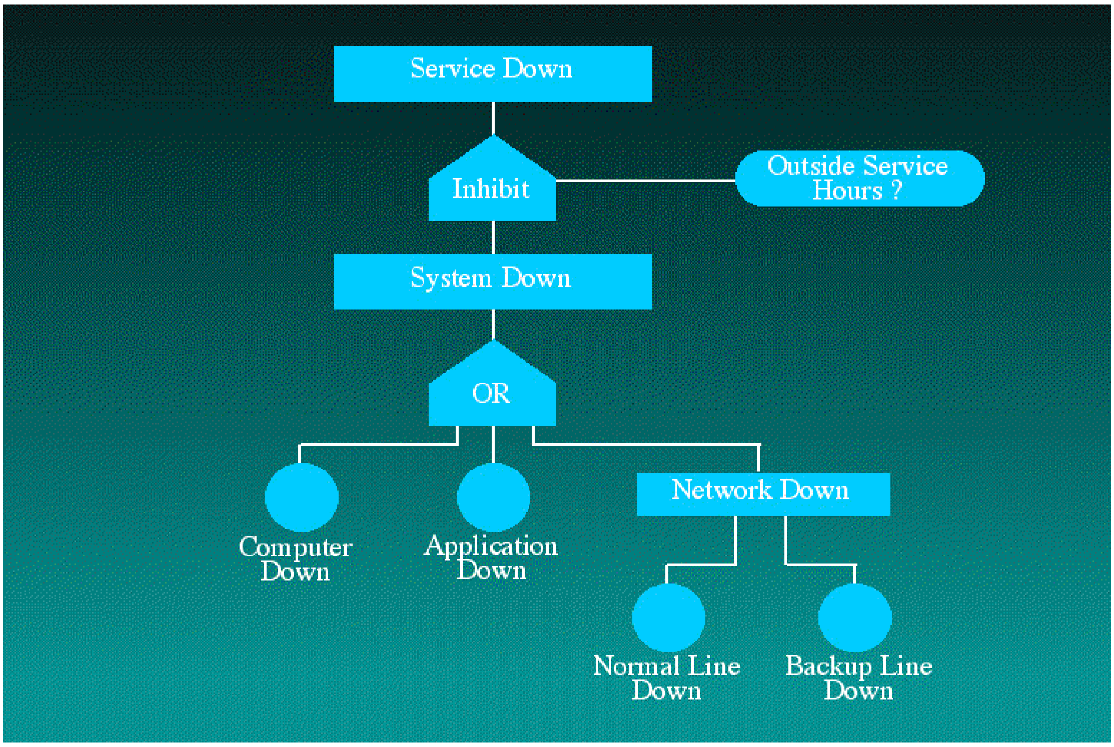


Figure 9.15: - Example Fault Tree

FTA makes a representation of a chain of events using Boolean notation. The above illustration gives an example of a fault tree.

Essentially FTA distinguishes the following events:

Basic events – terminal points for the fault tree, eg, power failure, operator error. Basic events are not investigated in greater depth. If basic events are investigated in further depth, they automatically become resulting events.

Resulting events – intermediate nodes in the fault tree resulting from a combination of events. The top most point in the fault tree is usually a failure of the IT service.

Conditional events – events that only occur under certain conditions, eg, failure of the air-conditioning equipment only affects the IT service if equipment temperature exceeds the serviceable values.

Trigger events – events that trigger other events, eg, power failure detection equipment can trigger automatic shutdown of IT services.

These events can be combined using logic operators, ie:

- **AND-gate** – the resulting event only occurs when all input events occur simultaneously.
- **OR-gate** – the resulting event occurs when one or more of the input events occurs.

Exclusive OR-gate – the resulting event occurs when one and only one of the input events occurs.

Inhibit gate – the resulting event only occurs when the input condition is not met.

The mathematical evaluation of a fault tree is beyond the scope of this chapter.

9.9.3 The CCTA Risk Analysis And Management Method (CRAMM)

The identification of risks and the provision of justified countermeasures to reduce or eliminate the threats posed by such risks can play an important role in achieving the required levels of availability for a new or enhanced IT service.

Risk analysis should be undertaken during the design phase for the IT infrastructure to identify: -

- Risks that may incur non-availability for the IT components within the IT infrastructure design.
- Risks that may incur confidentiality and/or integrity exposures within the IT infrastructure design.

The CCTA Risk Analysis and Management Method (CRAMM) describes a means of identifying justifiable countermeasures to protect Confidentiality, Integrity and Availability of the IT infrastructure.

The general concepts can be represented by a simple diagram that shows risk analysis and risk management as being two related but separate activities as shown below: -

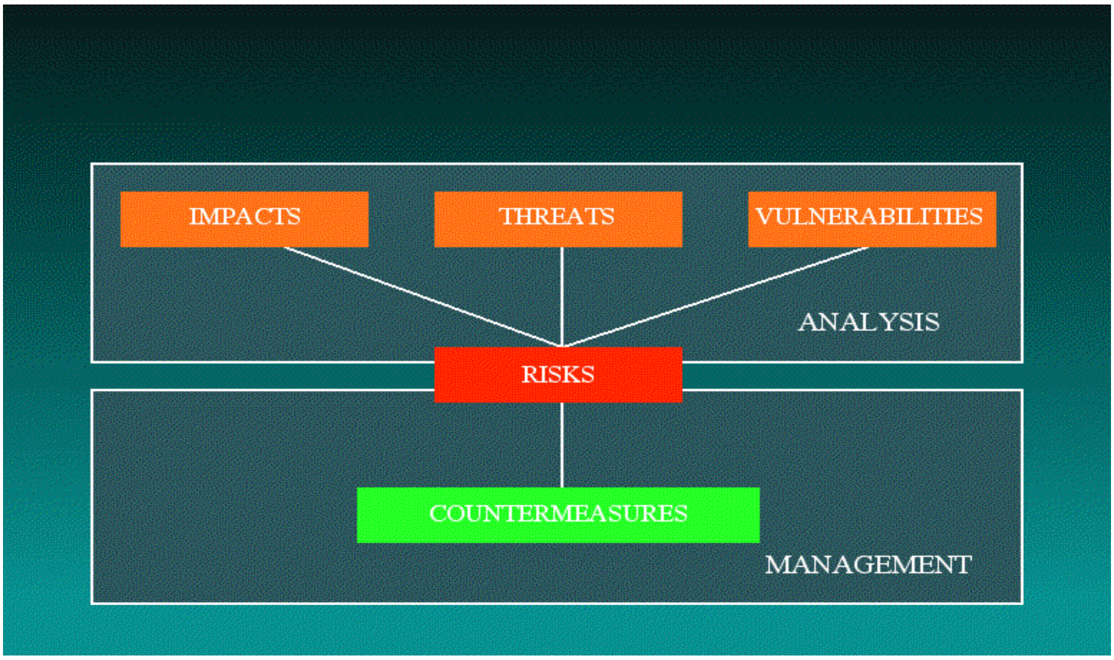


Figure 9.16: - Risk Analysis Management

Risk Analysis involves the identification and assessment of the level (measure) of the risks calculated from the assessed values of assets and the assessed levels of threats to, and vulnerabilities of, those assets.

Risk Management involves the identification, selection and adoption of countermeasures justified by the identified risks to assets, and the reduction of those risks to an acceptable level.

This approach when applied via a formal method will ensure coverage is complete together with sufficient confidence that: -

- All possible risks and countermeasures have been identified

- All vulnerabilities have been identified and their levels accurately assessed
- All threats have been identified and their levels accurately assessed
- All results are consistent across the broad spectrum of the IT infrastructure reviewed
- All expenditure on selected countermeasures can be justified

Formal risk analysis and management methods are now an important element in the overall provision of IT services.

Hints & Tips: -

CRAMM is a methodology that can be utilised by a number of IT service management processes.

CRAMM is used and referenced within Chapter 8 - IT Service Continuity Management (ITSCM).

For more detailed information on CRAMM please refer to the CCTA Management at Risk Library.

9.9.4 *Calculating Availability*

This section describes some of the simple mathematics required to enable component and total infrastructure availability to be calculated. This information will be needed to help formulate availability targets for IT components and IT services. Additionally, these output calculations can also be input to any availability modelling tools that are available.

The examples provided in this section are fairly straightforward with the calculations presented sufficient to provide adequate estimates of availability. Where more detailed estimates of availability are required it may be necessary to research more complex mathematical calculations. The statistical analysis of incident data and the forecasting of availability are a rich study field in many industries outside of IT, i.e. electronics, aviation.

Basic Availability Calculation

To determine the basic availability of a given IT service or component as an availability percentage (%) the following basic formula can be used: -

$$\text{Availability} = \frac{(\text{AST} - \text{DT})}{\text{AST}} \times 100$$

= Service or Component availability (%)

Where: -

AST = Agreed Service Time
DT = Actual downtime during agreed service time

Example: -

A 24x7 IT service requires a weekly 2-hour planned downtime period for application maintenance. Following the completion of the weekly maintenance an application software error occurred which resulted in 3 hours of unplanned downtime.

The weekly availability for the IT service in this reporting period is therefore based on the following: -

The AST should recognise that the planned 2 hr weekly downtime is scheduled
The DT is the 3 hrs of unplanned outage following the application maintenance

The AST value is therefore 24hrs*7days – 2 hrs planned maintenance = 166 hrs/week
The DT value is therefore the 3 hrs unplanned downtime

The availability calculation is: -

$$A = \frac{166 - 3}{166} \times 100 = 98.78\%$$

Total Infrastructure Availability

The availability percentage for each IT component within the total IT infrastructure may be different and as such it will be necessary to provide a calculation that reflects the total infrastructure availability.

The levels of resilience provided will positively influence the availability percentage for the total infrastructure

- *Serial Configuration*

The following illustrates a basic IT infrastructure configuration where no additional components are provided for resilience. The availability percentage for this configuration will be based on the product of all the individual component availability percentages.

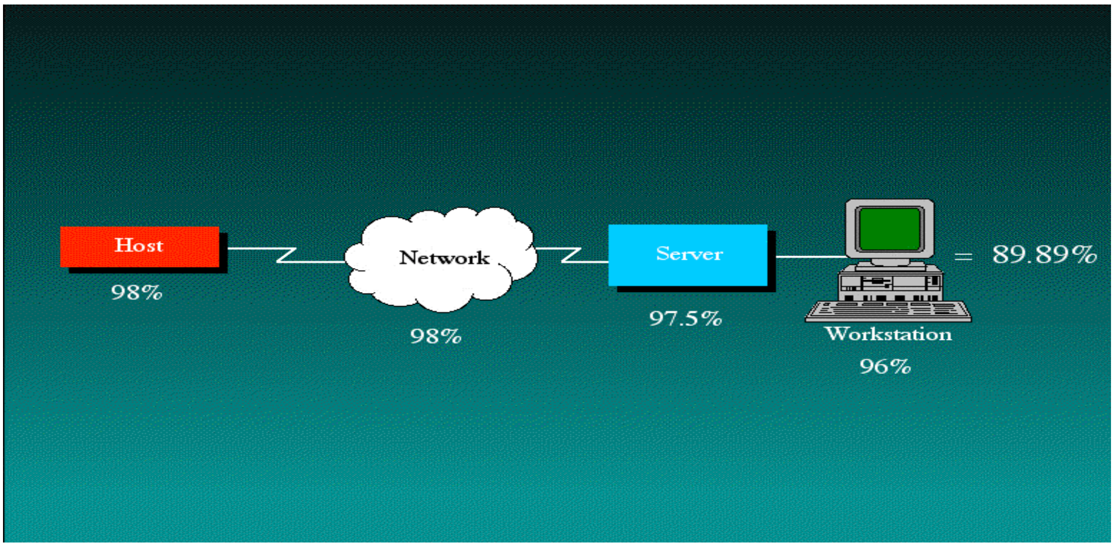


Figure 9.17:- Simple IT infrastructure configuration

Availability as viewed from the end user workstation is therefore calculated as: -

- Availability = Host * Network * Server * Workstation
- Calculation = 0.98 * 0.98 * 0.975 * 0.96 = 0.8989
- Total infrastructure availability = 89.89%

- *Parallel Configuration*

Where additional components are added to provide resilience so that the backup component takes over automatically, then the availability percentage is calculated by multiplying the unavailability (reciprocal of availability) of each component.

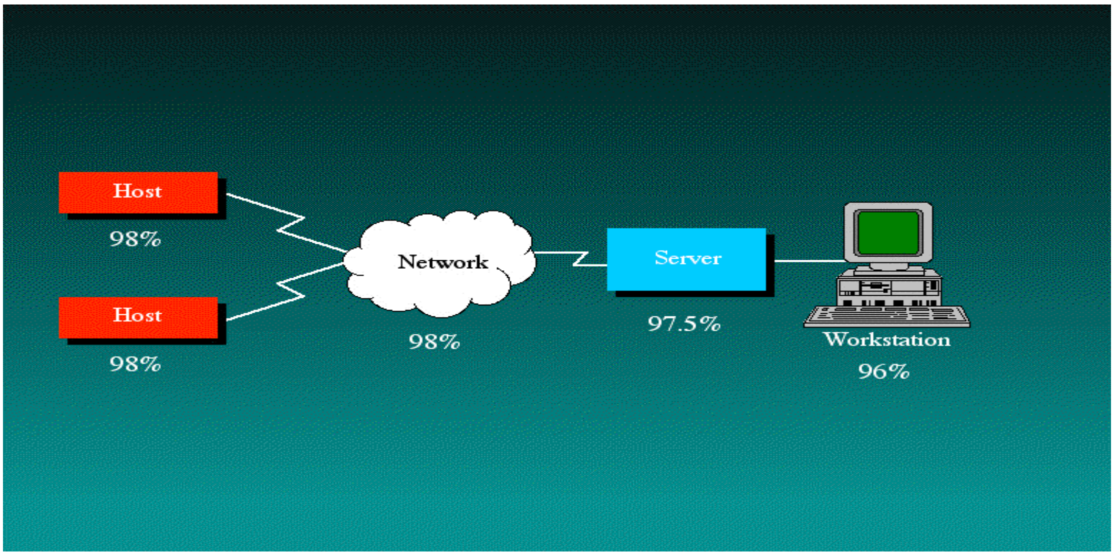


Figure 9.18: - Simple IT configuration with component resilience provided

In the above configuration, the Host component now has a backup component to provide greater resilience. The host component availability percentage is now recalculated as follows: -

- Availability = 1 –((1-0.98)*(1-0.98)) = 0.9996
- Host availability = 99.96%

With the additional resilience provided for the Host component, the total infrastructure availability can now be calculated as follows: -

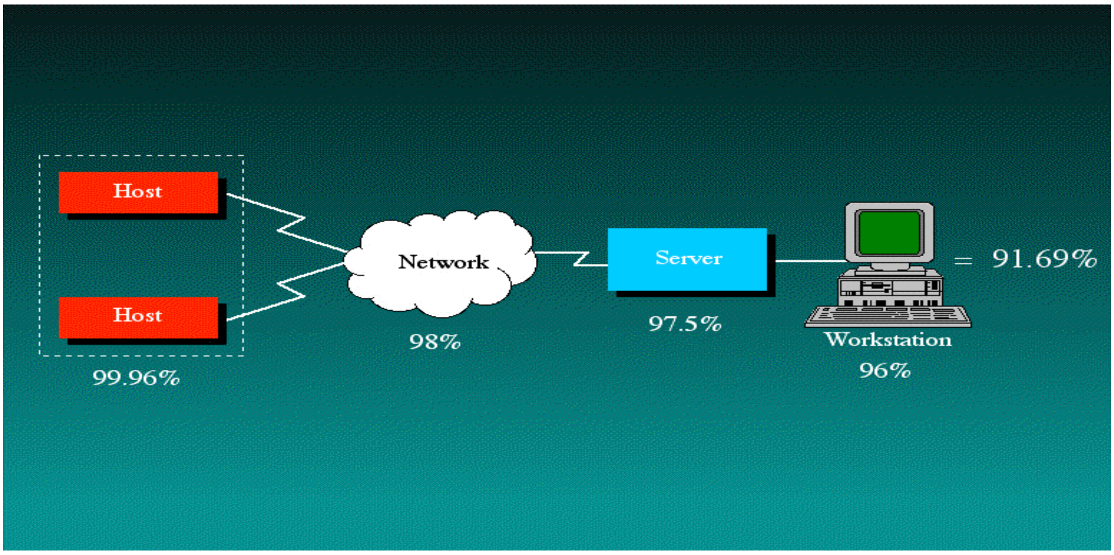


Figure 9.19: - Revised host availability percentage to reflect additional resilience provided

- Availability = Host * Network * Server * Workstation
- Calculation = 0.9996 * 0.98 * 0.975 * 0.96 = 0.9169
- Total infrastructure availability = 91.69%

Availability design implications

To deliver the required levels of availability for an IT service requires focus on all components within the IT infrastructure design that will underpin the IT service. The availability of each individual component will influence the overall availability that can be provided by the total infrastructure.

When viewing availability consider the following points: -

- The total infrastructure availability is limited by the weakest link**.**
- The total service availability cannot exceed the availability percentage of the least available component, unless the weakest components are used in parallel with automatic fail over..
- Improving the availability of one component may have minimal end – end benefit.
- One unreliable component can lead to significant damaging impact to availability.

9.9.5 Calculating the cost of Unavailability

To cost justify improvements to the IT infrastructure that will improve availability it will be necessary to demonstrate how the proposed improvements will deliver tangible business benefits.

Where the proposed improvements required a significant re-investment in the IT infrastructure the benefits will often need to be expressed in financial terms, i.e. the business case.

A good technique to justify IT infrastructure improvements is to quantify the total cost to the organisation of an IT service failure(s). These costs can then be used to support a business case for additional IT infrastructure investment and provide an objective 'cost Vs benefit' assessment.

The following is a sample calculation that can used to quantify the costs associated with IT service failure: -

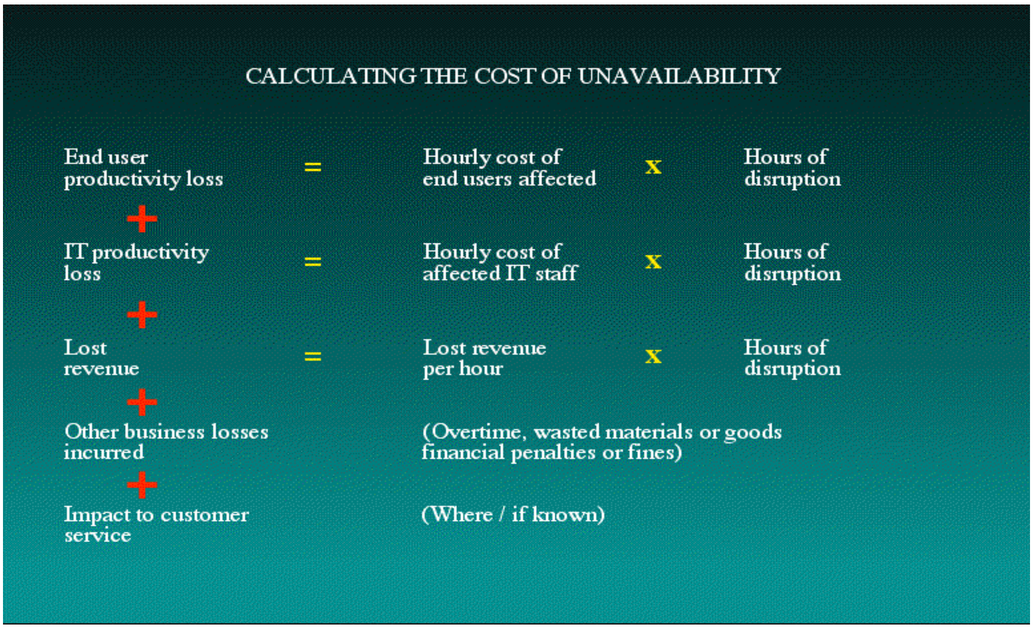


Figure 9.20: - Example calculation for the cost of an outage

Hints & Tips:-

Consider undertaking this exercise 'once only' to deliver what the business and IT organisation agree is an indicative cost of failure for a single or range of time periods, e.g.

The cost of one hours peak SLA outage = £xx
The cost of a full day SLA outage = £yy

This will then enable an indicative cost to be assigned to each IT service failure and included in the regular service reporting.

These figures should then be reviewed at least annually.

Another approach to obtain an indicative cost of a failure is to take the annual cost to the business of taking the service and simply divide by the number of service hours contracted in the SLA for a year. This will provide the IT expenditure cost to the business by hour.

9.9.6 Developing basic IT Availability measurement & reporting

At the component level there are a number of metrics that should be defined, measured and reported to provide a truly holistic view on how the component meets the range of availability criterion, e.g.: -

- Availability
- Reliability
- Maintainability
- Serviceability ♣

♣ *Serviceability is in itself not a metric but is reported in terms of service or component availability, reliability and maintainability*

Within the Availability Management topic a number of sections refer to the creation of metrics that can be utilised to provide this range of component availability reporting. A suggested framework for these metrics is as follows: -

Basic Availability calculation

The simplest form of measurement is to report the proportion of time that a component is actually available for use by the business within the agreed service time. This is usually expressed as the availability percentage.

This is a simple method of providing a measurement of availability for hardware, software application and network components and requires minimal investment in measurement and reporting tools. Consequently many SLA’s are constructed with availability measures based on the availability percentage (%)

Downtime reporting

Further basic forms of measurement can be considered to provide information concerning the availability or non-availability of an IT component, for example the reporting of downtime to reflect the total amount of time a dependent IT service was unavailable. This can be represented as: -

- **Unscheduled**

The amount of downtime incurred during service hours due to IT component failure(s)

- **Planned**

The amount of planned downtime incurred during service hours to apply maintenance

- **Extended**

The amount of excess downtime required to complete the planned maintenance activity

Where Service Maintenance Objectives (SMO) have been agreed for a given planned maintenance activity the total downtime incurred should be recorded as follows: -

- Planned downtime: - The agreed SMO time
- Actual downtime: - The total downtime incurred

- Extended downtime: - The excess downtime (Actual downtime – SMO time)

This method of recording will enable a clear distinction to made between agreed planned downtime and the extended downtime incurred due to deficiencies within the implementation process.

Extended downtime reporting can be used to review change quality issues with internal and external suppliers and be formally reported as non-compliance within OLA and service contract reporting.

Incident based reporting

Measures that reflect the overall reliability and maintainability of an IT service and supporting components can be derived from incident reporting. These can be represented as follows: -

- **Frequency of failure**

The number of occurrences that component failure impacted the IT service

- **Backup and Recovery performance**

The time taken to recover from a component failure (measured against the recovery metrics defined within the 'designing for recovery' and/or CFIA activity).

In addition, incident based reporting can also enable data to be produced which provides an indication of improving or deteriorating trends.

- **MTBF (Mean Time Between Failures)**

The average elapsed time from the time an IT service or supporting component is fully restored until the next occurrence of a failure to the same service or component.

- **MTBSI (Mean Time Between System Incidents)**

The average elapsed time between the occurrence of one failure, and the next failure.

MTTR (Mean Time To Repair)

The average elapsed time from the occurrence of an incident to resolution of the incident.

Please refer to Paragraph 9.9.9 and Figure 9.25 for additional guidance.

9.9.7 Developing business and end user measurement & reporting

The final word on the quality of the IT service provided rests with the business. Whilst traditional IT measures may show the '%' SLA target met, this will do little to change the feeling of dissatisfaction if IT service problems have impacted the business operation.

In the section Availability Measurement and Reporting it is recommended that a wide range of measures be produced to reflect availability from a number of perspectives. The key measures being those that reflect the consequence of IT availability on the business and end user.

This business and end user approach to availability reporting provides a number of benefits: -

- Provides a 'common' measure which both parties to the SLA can understand
- Can visibly demonstrate to the business tangible service improvements enabling the IT organisation to show the 'added value'.
- Can more easily identify degrading levels of service to enable the IT organisation to be proactive in responding without formal business escalation.
- Can be used to demonstrate the end user impact with suppliers to drive and influence positive supplier behaviour.

Approaches to develop business and end user driven metrics

So how can you develop business and end user driven metrics to gain the above benefits? The remainder of this topic provides a number of approaches that can be undertaken.

CFIA

CFIA is used to help predict and evaluate the impact on IT availability arising from component failures within the IT infrastructure design. The CFIA matrix used during this activity can be expanded to include fields that can map the number of end users supported by each component.

Component Description	End Users Affected
Host	1000
OLTP1	750
OLTP2	250
Order application	800
Payments application	50
Order database	800
Payments database	50
Server XYZ	20
Workstation A	1

Figure 9.21: - CFIA matrix denoting the number of end users affected by each component

Thus when a component is unavailable, the number of end users impacted is understood. This can enable availability calculations to be based on the number of end users impacted and/or amount of lost end user processing time.

- Advantages: - An easy to use approach that enables impact to be readily identified at the component level.

- Disadvantages: - The number of end user workstations does not necessarily equate to the number of end users at any one point in time, e.g. lunch hours, staggered shift patterns or weekday vs. weekend periods. In these instances a suggestion would be to base the number of end users within the CFIA matrix upon an average for end users active within selected periods.

Instrumentation

For new IT services the requirement for enhanced availability reporting should be captured in the design phase. It is easier and more cost effective to provide this instrumentation within the application during design than to attempt to retrofit once the IT service is live. Availability Management requirements should be based on the capture of information that relates to impact on the vital business functions arising from IT component failure.

- Advantages: - Requires no external measurement systems or methods and provides immediate availability reporting for new IT services.

- Disadvantages: - If not included in design may be resisted on the grounds of cost to retrofit. For performance centric applications there may be resistance to add any processing overhead to each transaction.

Estimation

A valid technique is to estimate the impact of IT failure against the transaction volumes (related to the vital business functions or end user activity) normally processed during the period of failure. The data to base your estimations against should be captured and maintained by Capacity Management.

- Advantages: - Relatively simple and easy to produce for all IT services using existing data sources.

- Disadvantages: - Provides only the indicative impact. Requires access to up to date volume information to ensure business growth trends etc are reflected, e.g. using November data to assess the ATM customer impact during December Xmas peak periods will not be an accurate reflection of impact.

End user assessment

For organisations unable to justify the costs of more advanced availability measurement techniques, the use of a daily ‘end user assessment’ is a simple technique that provides a business and end user view of IT availability. In its basic form, this would be an agreed set of criteria against which the business will assess IT availability and service quality in support of their business operation. Each day the business user representative should be contacted for their end of day assessment, this could be reported as a GOOD day, ACCEPTABLE day or BAD day. This can be recorded and reported using the 'RAG' (Red, Amber, and Green) method.

This approach can be expanded to enable more specific assessment information to be gathered, e.g. the business can assess a number of categories, e.g. Availability, Performance, Recoverability and be extended to include other IT service management functions, i.e. Service Desk.

- Advantages: - Low cost and effective. Provides the business view of IT availability and service quality.

- Disadvantages: - Requires business commitment. Without a framework of assessment criteria may be open to subjective opinion. May not provide sufficient detailed information to drive continuous improvement activities.

End user simulation tools

Systems Management tools which seek to simulate end user activity can be deployed to provide end – end availability reporting (real-time and offline reports). These tools execute scripts to generate sample transactions and monitor and report on areas such as availability, performance, throughput etc.

- Advantages: - Provide real-time exception reporting and range of offline reports. Additionally as a by-product continually verify the end – end infrastructure and provide an immediate post change implementation healthcheck.

- Disadvantages:- The cost of procurement and need to customise scripts to reflect your application(s)

Customer complaints

The correlation of customer complaints received against specific IT failures can provide an indication on true end customer impact and frustration.

- Advantages: - Low cost and reflects a true sense of customer dissatisfaction.

- Disadvantages: - Not every customer complains so not a true reflection of impact. Time lag between incidents and complaints being received delays and dilutes availability reporting.

Compensation claims & penalty payments

For certain businesses a consequence of IT failure may be claims for financial compensation by impacted customers. An example being for the loss of interest due to delayed or missed payments. This is not restricted to the financial services sector, for example some motorist assistance companies will pay compensation (or a penalty) for failing to meet a request for assistance within a set time period.

- Advantages: - Provides a good indication on the cost of failure.
- Disadvantage: - Not all customers make claims or exercise penalty payment. Again time lag between the incidents and resultant claims (often months) can delay and dilute availability reporting.

Sample end user availability calculations

Where the number of end users impacted by an IT failure is known, this information can be used to report end user availability as follows: -

- End user impact reported as an absolute value per incident or reporting period.
- End user productivity loss as a time based value per incident or reporting period.
- End user availability as an availability percentage (%) for the reporting period.

Calculating end user availability

End user impact (absolute)

This can be derived from the CFIA documentation to associate the end user population impacted by each with a component failure.

WEEKLY IT SERVICE REPORT

Incident No.	Date	Time	Duration (mins)	Incident Description	Failed Component	End user Impact
1	01Oct	09:25	60	Payments database full	Payments database	50
2	04Oct	12:48	25	Server hang-rebooted	Server XYZ	20
3	05Oct	09:56	125	Host operating system failure	Host	1000
4	05Oct	16:40	20	Fuse blown in power supply	Workstation A	1
						1071

Figure 9.22: - Example IT service reporting to denote end user impact per incident

End user impact (productivity)

To enable the reported end user impact to reflect end user productivity loss, requires the amount of downtime (DT) per incident to be expressed as the total amount of end user downtime (EUDT). This EUDT needs to reflect the number of end users affected by the incidents.

This will enable the amount of end user downtime to be derived that can then optionally be used to report as man-hours or man-days lost productivity.

- EUDT is calculated by multiplying the DT by the number of end users impacted.

WEEKLY IT SERVICE REPORT

Incident No.	Date	Time	Duration (mins)	Incident Description	Failed Component	End user Impact	EUDT (mins)
1	01Oct	09:25	60	Payments database full	Payments database	50	3,000
2	04Oct	12:48	25	Server hang -rebooted	Server XYZ	20	500
3	05Oct	09:56	125	Host operating system failure	Host	1000	125,000
4	05Oct	16:40	20	Fuse blown in power supply	Workstation A	1	20
			230				128,520

Figure 9.23: - Example IT service reporting to denote end user downtime

End user impact (availability percentage)

To provide an end user view of availability, the basic availability calculation contained in Paragraph 9.9.4 needs to be developed. The Agreed Service Time (AST) and the Down Time (DT) values need to be replaced with End User Processing Time (EUPT) and End User downtime (EUDT).

- EUPT is calculated by multiplying the AST by the total number of end users.
- EUDT is calculated by multiplying the DT by the number of users impacted and summing all incidents within a period as illustrated in Figure 9.23. (The sum of (Actual downtime during Agreed service time X number of users))

End User Availability (EUA) can therefore be calculated in a reporting period based on the following calculation:

$$\text{Availability} = \frac{(\text{EUPT} - \text{EUDT})}{\text{EUPT}} \times 100 = \text{End user availability (\%)}$$

Example:
Consider a 24 hour x 7 day service having 1,000 end users and a 2 hour planned downtime slot per week. The calculations in 9.9.4 can now be developed using the values obtained from Figure 9.2.3 as follows:

The weekly Agreed Service Time for the service would be:

AST = (24 x 7) - 2 = 168 - 2 = 166 hours

EUPT = AST x number of users = 166 x 1000 = 166,000 hours or (166,000 x 60) = 9,960,000 minutes

EUDT = (60 x 50) + (25 x 20) + (125 x 1000) + (20 x 1) = 128,520 minutes

Therefore End User Availability (EUA) can be calculated as follows:

$$EUA = \frac{(EUPT - EUDT)}{EUPT} \times 100 = \frac{(9960000-128520)}{9960000} \times 100 = .987 \times 100 = 98.7\%$$

Hints & Tips

The approaches outlined in this section should enable you to start moving towards business driven measurement and reporting to compliment the existing traditional IT measures. Cost and effort will influence the extent to which you develop this reporting. As always the key principle with measurement and reporting applies, in that the cost and effort of producing the measures and reports should not outweigh the benefits.

9.9.8 Systems Outage Analysis

The detailed analysis of service interruptions will identify opportunities to enhance levels of availability.

Systems Outage Analysis (SOA) is a technique designed to provide a structured approach to identify end - end availability improvement opportunities that deliver benefits to the end user. Some of the activities involved in SOA are closely aligned with those of Problem Management. In a number of organisations these activities are performed jointly by Problem and Availability Management.

The high level objectives of SOA are: -

- To identify the underlying causes of service interruption to the end user.
- To assess the effectiveness of the IT support organisation and key processes.
- To produce reports detailing the major findings and recommendations.
- To initiate a programme of activities to implement the agreed recommendations.
- That availability improvements derived from SOA driven activities are measured.

The key principles of the SOA approach are: -

- That the underlying reasons for service interruption can be caused by shortfalls in technology, process, procedure or behaviours (culture).
- That wider ranges of data sources are used to support the analysis.
- Business and end user input is fundamental.
- A specifically mobilised cross-functional team undertakes that analysis.
- That SOA assignments have a recognised sponsor(s) (Ideally joint sponsorship from the IT and business).

The reasons for adopting an SOA approach: -

- Traditional IT availability reporting often only provides an IT component perspective.
- Business and end user input provides an ultimate view of availability and the important issues from their perspective.
- Provides a structured, focused and detailed analysis of a selected IT Service or set of infrastructure.components.
- Provides a mechanism to ensure the IT infrastructure delivers optimal availability.

The benefits from taking an SOA approach: -

- Can enable requests for enhanced levels of availability to be met without major cost.
- Provides the business with visible commitment from the IT support organisation.
- Develops in-house skills and competencies to avoid expensive consultancy assignments related to availability improvement.
- Cross-functional team approach is an enabler to ‘think outside of the box’ to challenge traditional thinking and provide innovative and often inexpensive solutions.

- SOA delivers a programme of improvement opportunities that can make a real difference.
- SOA improvement opportunities are focused on delivering benefit to the end user.
- Provides an independent 'healthcheck' of IT service management processes and be the stimulus for process improvements.

A structured approach

To maximise both the time of individuals allocated to the SOA assignment and the quality of the delivered report a structured approach is required. The approach taken is similar to many consultancy models utilised within the industry and in many ways Availability Management can be considered as providing via SOA a form of internal consultancy.

Figure 9.24:- The structured approach for a Systems Outage Analysis assignment

The above high level structure is described briefly as follows: -

Select Opportunity

Prior to scheduling an SOA assignment there needs to be agreement as to which IT service or Infrastructure is to be selected. Within the Availability Plan it is recommended that 4 assignments are scheduled per year and if possible the IT service is selected in advance as part of the proactive approach to Availability Management.

Before commencing with the SOA it is important that the assignment has a recognised sponsor from within the IT organisation and/or the business. This ensures organisational visibility to the SOA and ensures recommendations are endorsed at a senior level within the organisation.

Scope Assignment

This is to state explicitly what areas will and will not be covered within the assignment. This would normally be documented in a Terms of Reference issued prior to the assignment.

Plan Assignment

The assignment needs to be planned a number of weeks in advance of the assignment commencing. The typical areas that require advance planning are: -

- The Start and end dates of the assignment
- Key milestones, e.g. delivery of final report
- The individuals who will form the SOA team
- Role and responsibilities of the individual team members
- The data sources required to provide the data for analysis
- Premises and equipment, i.e. a dedicated room, whiteboards, terminals etc.
- An interview schedule for key IT and business personnel
- A visit to the business operation and the IT operation.

The SOA assignment should be looking at identifying improvement opportunities that benefit the end user. It is therefore important that an end - end view of the data and MIS requirements is taken. A suggested list of data sources is as follows: -

- Incident Management records and MIS
- Problem Management records and MIS
- Change management records and MIS
- SLAs and service level reporting
- Vital Business Function measures that reflect end user impact
- Formal complaints to the business from their customers
- Formal complaints from the business to the IT organisation
- Customer satisfaction survey results
- Process metrics

For practical reasons the coverage period for the above should be limited to approximately 6 months. This will limit the amount of data to analyse but importantly ensure that only current issues are being investigated.

To support the team with analysis, supporting documentation should be available to the team, e.g. operational procedures. Process documentation, IT policies, configuration diagrams, Industry best practice reference material, e.g. ITIL.

Build Hypotheses

This is a useful method of building likely scenarios, which can help the study team draw early conclusions within the analysis period. These hypotheses can be built from discussing the forthcoming assignment with key individuals, e.g. Senior Management, Problem Manager, Change Manager, Service level Manager or by using the planning session to brainstorm the list by the assembled team.

The completed hypotheses list should be documented and input to the analysis period to provide some early focus on data and MIS that match the individual hypotheses.

It should also be noted that this approach also eliminates perceived issues, i.e. no data or MIS substantiates what is perceived to be a service issue.

Example: -

If an SOA was planned to review availability for a Call Centre based IT Service; it is likely that system performance is crucial. Hypotheses to help assess if performance issues are impacting the end user availability could be based on the following: -

"Performance issues are the single largest cause of 'service unavailability' incidents impacting Call Centre operation"

"Existing system and performance monitors do not enable 'service unavilaibility' incidents to be identified, diagnosed and resolved effectively.

Analyse Data

The number of individuals that form the SOA team will dictate how you wish to allocate specific analysis responsibilities.

During this analysis period the hypotheses list should be used to help draw some early conclusions.

Interview key personnel

It is essential that key business representatives and end users are interviewed to ensure the business and end user perspective is captured. It is surprising how this dialogue can identify quick win opportunities as often what the business view as a big issue can often be addressed by a simple IT solution.

The study team should also seek input from key individuals within the IT support organisation to identify additional problem areas and possible solutions which can be fed back to the study team.

The dialogue also helps capture those issues that are not easily visible from the assembled data and MIS reports.

Findings and Conclusions

After analysis of the data and MIS provided, interviews and continual revision of the hypothesis list the study team should be in a position to start documenting initial findings and conclusions.

It is recommended that the team meet immediately after the analysis period to share their individual findings and to then take an aggregate view to form the draft findings and conclusions.

It is important that all findings can be evidenced by facts gathered during the analysis. During this phase of the assignment it may be necessary to validate finding(s) by additional analysis to ensure the SOA team can back up all findings with clear documented evidence.

Recommendations

After all findings and conclusions have been validated the SOA team should be in a position to formulate recommendations. In many cases the recommendations to support a particular finding will be straightforward and obvious.

However, the benefits of bringing a cross functional team together for the SOA assignment is to create an environment for innovative 'think outside of the box' approaches. The SOA assignment leader should facilitate this session with the aim of identifying recommendations that are practical and sustainable once implemented.

Report

The final report should be issued to the Sponsor with a management summary. Reporting styles will normally be determined by the individual organisations.

It is important that the report clearly shows where availability loss is being incurred and how the recommendations will address this. If the report contains many recommendations an attempt should be made to quantify the availability benefit of each recommendation together with the estimated effort to implement.

This will enable informed choices to be made on how to take the recommendations forward and how these should be prioritised and resourced.

Validate

It is recommended that for each SOA, key measures that reflect the business and end user perspectives prior to the assignment are captured and recorded as the 'before' view.

As SOA recommendations are progressed the positive impacts on availability should be captured to provide the 'after' view for comparative purposes. Where anticipated benefits have not be delivered this should be investigated and remedial actions taken.

Hints & Tips

Consider categorising the recommendations under the following headings: -

AVOIDANCE
Recommendations that if implemented will eliminate this particular cause of IT service interruption.
MINIMISE
Recommendations that if implemented will reduce the end user impact from IT service interruption; e.g. recovery and/or restoration can be enhanced to reduce impact duration.
DETECTION
Recommendations that if implemented will provide enhanced reporting of key indicators to ensure underlying IT service issues are detected early to enable a proactive response.

Build Programme

Having invested time and effort in completing the SOA assignment it is important that the recommendations once agreed by the sponsor are then taken forward for implementation.

The best mechanism for achieving this is by incorporating the recommendations as activities to be completed within the Availability Plan or SIP.

It is recommended that these activities are also managed and tracked by Programme Management, Project Management and Change Management processes.

The SOA team

The team should consist of experienced IT practitioners selected from a range of areas within the IT organisation.

For example the SOA team could consist of individuals from the following functions: -

- Availability Management process owner (SOA assignment leader)
- Computer Operations
- Network Management
- Problem Management
- Change Management
- Service Desk
- Service Level Management

The size of the team will be influenced by the size of the IT organisation and the topic selected for the SOA. A team of at least 3 is the recommended minimum.

Dependent on the focus of the SOA assignment you may wish to include or schedule within the assignment plan: -

- A leading technical expert
- Supplier representation

As scheduled events, the Availability Management process owner should have these events defined within the Availability Plan and identified resources committed in advance.

Measure SOA effectiveness

SOA should be viewed as a key element of the Availability Plan that underpins the Availability Management process. Measures should be established to monitor the effectiveness of SOA as an organisational activity and in optimising service availability.

To measure the effectiveness of each SOA the following metrics could be used: -

- Number of recommendations
- Number of recommendations rejected
- Number of recommendations completed
- Number of recommendations in progress

- Number of recommendations with no progress

The above measures provide a clear indication on how progress is being made with each completed SOA assignment. The number of recommendations rejected may reflect the quality of recommendation's made. Conversely a high completion rate would indicate the 'doability' of the recommendations made.

9.9.9 *The expanded incident 'lifecycle'*

A guiding principle of Availability Management is to recognise that you can still gain customer satisfaction even when things go wrong. One approach to help achieve this requires Availability Management to ensure that the duration of any incident is minimised to enable normal business operations to resume as quickly as is possible.

Availability Management should work closely with Incident Management and Problem Management in the analysis of unavailability incidents.

A good technique to help with the technical analysis of incidents affecting the availability of components and IT services is to take an incident 'lifecycle' view.

Every incident passes through several major stages. The time elapsed in these stages may vary considerably. For Availability Management purposes the standard incident lifecycle as described within Incident Management has been expanded to provide additional help and guidance particularly in the area of designing for recovery. The following diagram illustrates the expanded incident 'lifecycle'.

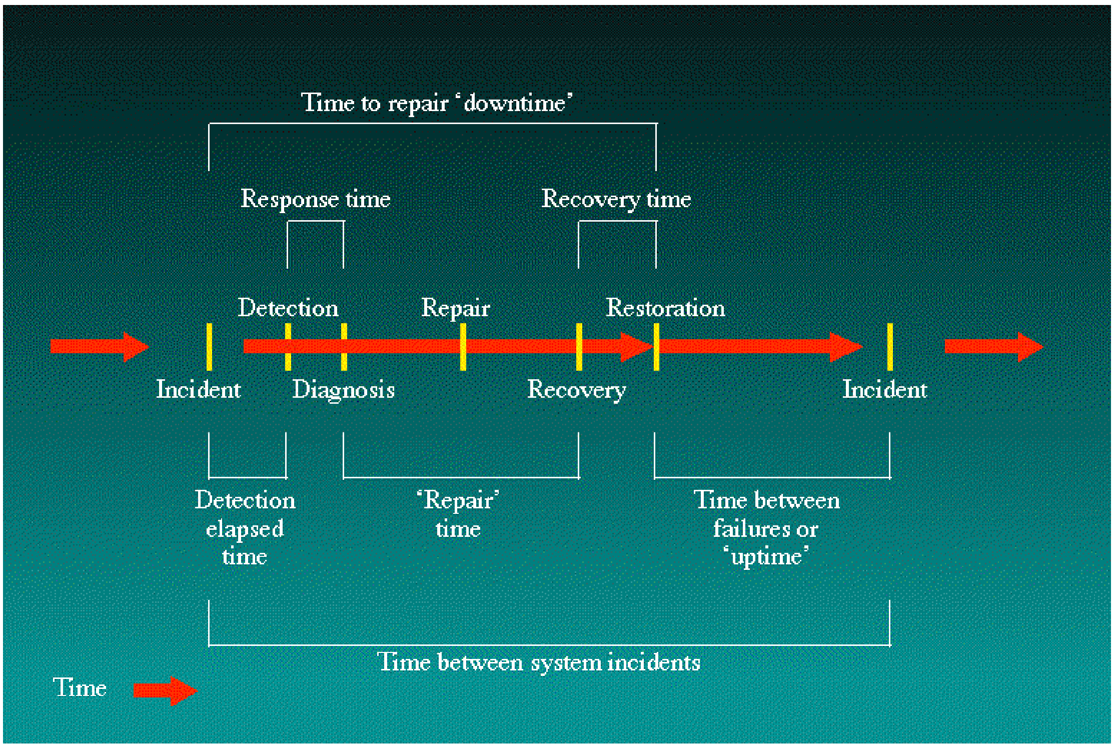


Figure 9.25: - The expanded Incident 'Lifecycle'

From the above you can see that an incident can be broken down into stages which can be timed and measured. These stages are described as follows: -

- Incident Start: - The time at which the customer recognises a loss or deviation of service or the time at which the incident is first reported. Whichever is the earliest..
- Incident Detection:- The time at which the IT organisation is made aware of an incident
- Incident Diagnosis:- The time at which diagnosis to determine the underlying cause has been completed
- Incident repair:- The time at which the failure has been repaired/fixed
- Incident Recovery:- The time at which component recovery has completed
- Incident Restoration: - The time normal business operations resume.

Each stage and the associated time taken influences the total downtime perceived by the user. By taking this approach it is possible to see where time is being 'lost' for the duration of an incident, eg the service was unavailable to the business for 60 mins, yet it only took 5 mins to apply a fix, where did we loose the other 55 minutes?

By using this approach it will identify possible areas of inefficiency that combine to make the loss of service experienced by the business greater than it need necessarily be. These could cover areas such as poor automation (alerts, automated recovery etc.), poor diagnostic tools and scripts, unclear escalation procedures (which delays the escalation to the appropriate technical support group or supplier), lack of comprehensive operational documentation.

Availability Management will need to work in close association with Incident and Problem Management to ensure repeat occurrences are eliminated.

It is recommended that these measures are established and captured for all incidents. This will provide Availability Management with metrics for both specific incidents and trending information. This information can be used as input to SOA assignments, Service Improvement Programmes and regular Availability management reporting and provide an impetus for continuous improvement activity to pursue cost effective improvements.

It can also enable targets to be set for specific stages. Whilst accepting that each incident may have a wide range of technical complexity, a number of stages should be expected to be consistent and reflect consistency in how the IT support organisation responds.

9.9.10 *Continuous improvement*

The primary purpose of the Availability Management process is to ensure that the availability requirements agreed with the business for IT service(s) are consistently met. It is the responsibility of Availability Management to ensure that corrective actions are being progressed to address any shortfalls in meeting the levels of availability required and expected by the business.

Availability Management can also play a key role in further optimisation of the existing IT infrastructure to provide improved levels of availability at a lower cost when availability requirements change.

The Availability Management process should wherever possible contribute activities to support an overall SIP.

To help achieve these aims Availability Management will need to be recognised as a leading influence over the IT support organisation to ensure continued focus on availability and stability of the IT infrastructure.

As the 'champion' for availability in the IT organisation the function should embrace and engender the ethos of 'continuous improvement' within the IT support organisation.

Continuous Improvement is a key element of 'Quality Management' utilised to empower staff to drive improvements that benefit the business and end user. There are a number of Quality Management methodologies available, e.g. Total Quality Management (TQM), however 'continuos improvement' can be embraced without the need for an organisation to adopt a Quality Management methodology.

A suggested technique

The 'continuous improvement' methodology can be employed as a technique by Availability Management to facilitate improvements that can be progressed by the IT support organisation to deliver availability improvements that benefit the business and end user.

Availability Management can provide the IT support organisation with a real business and end user perspective on how deficiencies within the IT infrastructure and the underpinning process and procedures impact the business operation and ultimately their customers.

The use of business driven metrics can demonstrate this impact in real terms and importantly also help quantify the benefits of improvement opportunities.

Availability Management can play an important role in helping the IT support organisation recognise where they can add value by exploiting their technical skills and competencies in an availability context. The continuous improvement technique can be used by Availability Management to harness this technical capability. This can be used with either small groups of technical staff or a wider group within a workshop environment.

The basic steps of the continuous improvement methodology are described as follows: -

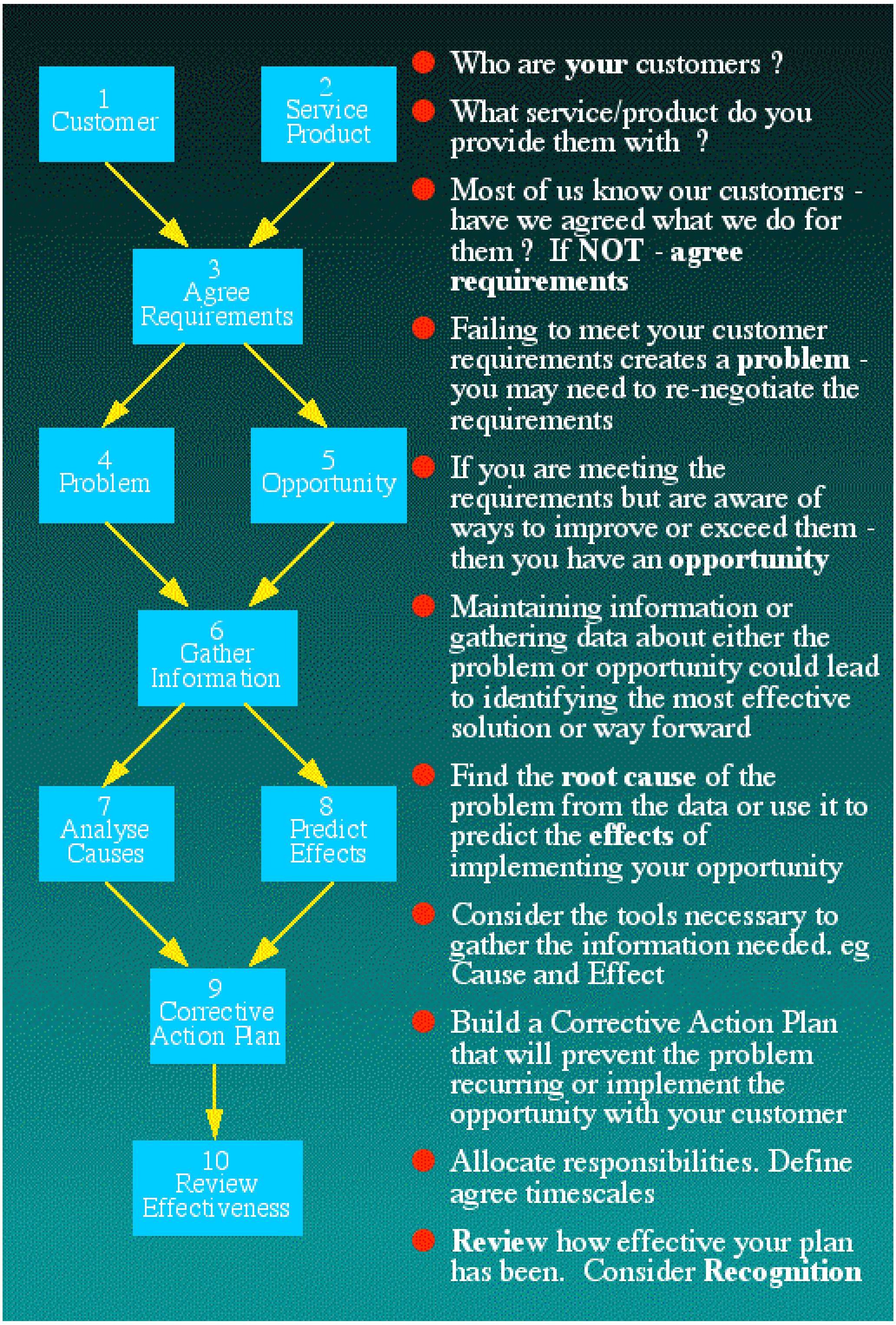


Figure 9.26: - Example of a continuous improvement methodology

The wider benefits of adopting this approach within the IT support organisation are: -

- Provides direction to best exploit skills and competencies
- Creates an understanding of how the business use the technology
- Can identify ‘quick win’ low cost improvements
- Deliver incremental availability improvement
- Provide positive feedback to staff on ‘how they have made a difference’
- Demonstrate to the business the value add of the IT support organisation
- Help establish a ‘service culture’

9.9.11 Technical Observation Post (TOP)

Continuous Improvement is an ethos fundamental to all Service Management processes and Service Management as a whole. An alternative approach for progressing continuous improvement opportunities is the establishment of a Technical Observation Post (TOP).

The TOP is best suited for delivering proactive business and end user benefits from within the real-time IT environement.

What is a TOP?

A TOP is a prearranged gathering of specialist technical support staff from within the IT support organisation brought together to focus on specific aspects of IT availability. Its purpose being to monitor events, real time as they occur with the specific aim of identifying improvement opportunities or bottlenecks which exist within the current IT infrastructure.

Why?

A wide range of systems management products and tools are available to provide real-time and retrospective analysis on specific aspects of components within the IT infrastructure. Some are better than others. However, to acquire a global end - end view of the IT infrastructure or a given IT service using real-time monitors or historical data can often be difficult, time consuming and require significant effort.

Another consideration is the reality that the people who design and support IT systems are not the people who run and operate them. Assumptions and misunderstandings can occur between both parties which unconsciously result in inefficient operational processes, e.g. how many IT support staff actually observe and understand the overnight application processing lifecycle and the key operational events?

By bringing together specialist technical staff to observe specific activities and events within the IT infrastructure and operational processes creates an environment to identify improvement opportunities.

Scope

The scope of a TOP can be wide ranging but must be focused with an overall objective set.

Example: -

A TOP is convened with an objective to ‘Improve the efficiency of the overnight batch window’. The areas of focus to identify contributory improvements to the TOP objective could include: -

- Batch scheduling
- Batch restart and recovery procedures
- Automation
- Application performance
- Infrastructure performance
- Operational processes and procedures

When?

A TOP can be convened at anytime where this approach is considered appropriate and its invocation would be planned and scheduled by Availability Management.

Considerations

The provision of a small area or room with terminals that enables the cross-functional team of specialist technical staff to work together is an important success factor that ensures: -

- The TOP has visibility
- The TOP team remains focused
- The TOP team members are able to interact unhindered by physical or organisational barriers

Availability Management as sponsors of the TOP should play a facilitation role providing guidance and ensure and ensuring the team remains focused on the TOP objective(s). The role should also ensure that all observations and outline recommendations are captured and subsequently create an action plan that will form part of either the Availability Plan or the appropriate SIP.

Benefits

The benefits of using a TOP as an approach to continuous improvements are as follows: -

- Informal structure which technical staff are comfortable with and has limited management overhead
- Cost effective
- Creates an environment that can positively harness the technical capabilities of staff
- Creates a cross functional team that is focused and shares a common sense of purpose
- Creates an environment for the sharing of information to the benefit of all attending
- Enables IT support organisation staff to observe the operational environment
- Will identify areas of improvement masked by inefficient tools, processes and procedures

Industry view: -

An organisation had failed to recognise the early warning signs of gradual erosion of the overnight ‘batch window’.

A steady decline in overnight batch processing completion times resulted in SLA’s for key services being regularly breached at start of day. The factors behind this increase in elapsed time were difficult to identify after the event, so the decision was taken to form a TOP in order to observe the batch real time.

In this particular case, the TOP ran for several consecutive nights, producing recommendations each evening for progression (wherever possible) during the following working day.

A daily report detailing both observations and recommendations was created on a daily basis. By the end of the TOP, over fifty opportunities for improvement had been identified and many implemented.

Continued...

The initial ‘quick win’ actions taken resulted in a reduction in overnight batch elapsed times such that the workload could now be completed within the SLA.

As a result of completing all the main recommendations from the TOP exercise, the overnight batch processing elapsed time was reduced by almost 4 hours.

Annex A - Outline Role Description

AVAILABILITY MANAGER - ROLE, RESPONSIBILITIES & KEY SKILLS

ROLE PURPOSE

In today's competitive market place the business are demanding more and more from IT to enhance their business and deliver cost effective, high levels of service to their customers. As the business needs and customer expectations increase, the availability and reliability of IT services is essential in ensuring the business can meet its business objectives.

The objectives of Availability Management are:-

1. To define and deploy the Availability Management process within the organisation and be accountable for execution and compliance of the process across the IT support organisation.
2. To ensure IT services are designed to deliver the required levels of availability required by the business.
3. To provide a range of IT availability reporting to ensure that agreed levels of Availability, Reliability and Maintainability are measured and monitored on an ongoing basis.
4. To optimise the availability of the IT infrastructure to deliver cost effective improvements that deliver tangible benefits to the business and end user.
5. To achieve over a period of time a reduction in the frequency and duration of incidents that impact IT availability.
6. To ensure shortfalls in IT availability are recognised and appropriate corrective actions are identified and progressed.
7. To create and and maintain a forward looking Availability Plan aimed at improving the overall availability of IT services and infrastructure components to ensure that existing and future business availability requirements can be met.

KEY RESPONSIBILITIES

1) To be accountable for the deployment of the Availability Management process and associated methods and techniques.
2) To be responsible for ensuring the Availability Management process, its associated techniques and methods are regularly reviewed and audited, and that all of these are subjected to continuous improvement and remain fit for purpose.
3) To be responsible for the determining of availability requirements from the business for new or enhanced IT services.
4) To be responsible for the creation of availability and recovery design criteria to be applied to new or enhanced infrastructure design.
5) To be responsible for ensuring the levels of IT availability required are cost justified.

6) To be responsible for defining the targets of availability required for the IT infrastructure and its components that underpin a new or enhanced IT service as the basis for SLA agreement.
7) To be responsible of the establishment of measures and reporting that reflects business, end user and IT support organisation requirements.
8) To be responsible for the monitoring of actual IT availability achieved Vs targets and to ensure shortfalls are addressed.
9) To be responsible for the production and maintenance of an Availability Plan which prioritises and plans IT availability improvements.
10) To promote Availability Management awareness and understanding within the IT support organisation.
11) To maintain an awareness of technology advancements and IT best practice, eg ITIL.

10 SERVICE MANAGEMENT SOFTWARE TOOLS

The first question you should ask yourself on this topic is *'Do I really need software tools*?' If you do, assess the need formally with a well-researched selection process.

If you look at the glossy brochures and listen to the sales talk, Service Management tools are indispensable. However, good people, good process descriptions, and good procedures and working instructions are the basis for successful Service Management**.** The need for, and the sophistication of, the tools required will depend on the business need for IT services and, to some extent, the size of the organisation**.**

In a very small organisation a simple in-house developed database system may be sufficient for logging and controlling Incidents. However, in very large organisation, a very sophisticated distributed integrated Service Management toolset may be required, linking all the processes with event-management systems. While tools can be an important asset in today's IT-dependent organisations, they are a means, not an end in themselves. When implementing Service Management processes, looking at the way your processes work and your need for management information should always be the starting point. This will provide information needed to define the specifications for a tool best suited to assist you.

Why the need? Here are some of the reasons:

- more sophisticated Customer demands
- a shortage of IT skills
- budget constraints
- business dependence on quality IT services
- integration of multi-vendor environments
- increasing complexity of IT infrastructure
- the emergence of international standards
- increased range and frequency of IT Changes.

Automated tools allow:

- the centralisation of key functions
- the automation of core Service Management functions
- the analysis of raw data
- the identification of trends
- the preventive measures to be implemented.

10.1 Types of tools

Software tools range from simple to complex and from inexpensive to very expensive. They generally fall into one of the two following categories:

- CMDB & Service Desk; traditional Service Desk tools without separate databases and modules for the Service Management processes
- integrated Service Management tools comprising modern client server based tools**,** with or without a knowledge database.

10.2 Summary of tool evaluation criteria

The following criteria should generally be used to assess software tools under consideration:

- an 80% fit to operational requirements
- a meeting of *ALL* mandatory requirements
- little (if any) product customisation
- ITIL compliance
- a sound data structure and handling
- business-driven not technology-driven
- administration and maintenance costs within budget.

Software tools should handle processes in conformity with the practices discussed in the IT Infrastructure Library. A set of guidance (the Appraisal and Evaluation Library) is available for the guidance of organisations wishing to select Service Support and Service Delivery tools. The prime areas to consider are:

- functional requirements support, and the level of integration with, for example, service delivery processes and tools
- data structure, data handling and integration, including the capability to support the required functionality
- integration of multi-vendor infrastructure components, and the need to absorb new components in the future – these will place particular demands on the data handling and modelling capabilities of the tool
- conformity to international open standards
- flexibility in implementation, usage and data sharing
- usability: the overall ease of use permitted by the User interface
- service levels: performance and availability
- distributed clients with a centralised shared database (e.g. client server)
- back-up, control and security provisions
- the quality of information provided by the supplier, and its validation by contact with other users.

10.2.1 Service Management tools

Few enterprises have no Service Management tools and many are considering replacing or upgrading those that are in use. The range and sophistication of tools for Service Management automation has grown rapidly in recent years.

Tools for the automation of core processes such as Incident logging and tracking have been supplemented by computer-integrated telephony, software capable of handling complex and multiple Service Level Agreements (with separate targets and business clocks), and remote support technology. Other tools include:

- interactive voice response (IVR) systems

- the Internet, internal electronic mail, voice mail
- self-help knowledge
- case-based reasoning/search systems
- network management tools (including remote support capabilities)
- system monitoring
- Configuration and Change Management systems
- release and distribution systems
- security monitoring and control, including password control, detection of violations and virus protection
- capacity planning
- IT Service Continuity Management (including automatic back ups).

Although some of the tools are not yet commonly used, there are few areas of Service Management that cannot be helped by automation. Some areas of Service Management are too resource intensive to be performed effectively without automation. Each tool for the automation of Service Management has advantages and disadvantages but automation is still recognised as vital.

It is necessary to ensure that the combination of technology, processes and people are integrated and meet the needs of the Customers. Automation should be used to enhance Service Management, not replace it.

Automation is increasingly being treated as part of workflow management, linking each task in the life-cycle from a new service being planned through to disposal. The technology should be used to complement and enhance service delivery, not replace it.

Automation that provides support for distributed computing has revolutionised the ability of an enterprise to diagnose Problems remotely, and in many cases also to fix them remotely (and therefore faster). Remote support technology has also made it possible for an enterprise to make changes by downloading the new versions of software and to monitor the capacity of the infrastructure, identifying capacity problems before they become serious.

Automation has enabled easier contingency planning, with work being switched in the event of a local overload or a serious problem that has taken the service out from a specific area.

Some final considerations:

- *supplier and product credibility and viability* - installed base and degree of support; consider issues such as the financial viability of the vendor (are they likely to be around in a few years when you need them?); also consider large time-zone differences between the supplier and your organisation and language differences
- *costs, including ongoing cost to upgrade and support* - consider which is better :
  - buying a standard package at reasonable initial cost, where the trade-off is that customisation may be very expensive and complex
  - or a more flexible package at higher initial costs where customisation may be relatively easy and cheap
- *adaptability* - will the tool be able to meet organisation specific requirements and constraints in the years to come.

10.3 Product training

To ensure effective use of software tools, product training is required. Therefore budget provision should be made at the planning stage of the implementation project. Furthermore, the training supplier you employ should have a suitable portfolio of training programmes, covering the requirements of practitioners, supervisors and managers.

11 PLANNING FOR THE IMPLEMENTATION OF SERVICE MANAGEMENT

More and more organisations are recognising the importance of Service Management to their business. However, it is common for working practices to be based on historical or political considerations rather than the current needs of the business or best practice. It is therefore essential, before implementing any or all of the components of Service Management, to gain management commitment, to understand the working culture of your organisation, and to assess any existing processes and to compare these to the needs of the business and to best practice.

11.1 A Service Management project

To analyse the needs of the organisation and implement the desired solution requires a temporary organisation to be set up to undertake these activities. Thus this can readily be considered to be a project, or a series of projects, to implement the required Service Management processes.

One of the benefits of adopting a project approach to this activity is that you can undertake the necessary investigations and have designated decision points where you can opt to continue with the project, change direction, or stop.

The project needs to consider your current position and where you would like to be, and to plot the path between these states. For each option identified, you can begin to articulate:

- the business benefits
- the risks, obstacles and potential problems
- the costs of the move plus longer-term running costs
- the costs of continuing with the current structure.

You can then begin to see how the business needs can be supported and see the associated costs. The benefits can then be balanced against costs and risks. Undertaking the investigative work could be considered to be one project that can then be followed by an implementation project.

11.2 Feasibility study

It is essential to investigate and understand the current service levels and costs by baselining all appropriate aspects of the current service before making any major changes. This enables you to measure the impact of the improved service management processes on the baseline service levels and costs. (When baselining involves comparison with other enterprises, it is usually referred to as ‘benchmarking’.)

Some advice is necessary before you start; the IT Infrastructure Library is not a magic wand. Do not expect miracles to happen when you implement the process framework. In the past, many organisations have tried to use process implementations as the basis for company reorganisations, or to assist with company mergers. Too many disparate goals for the project will lead to failure and disappointment. So, the targets and objectives associated with a Service Management project should relate to the objectives of the organisation itself. The target should be to enable the delivery of quality IT services aligned to business need.

11.3 Assessing the current situation

11.3.1 Introduction

Figure 3.1 represents a model that can be used by an organisation as the framework for process improvement. The model is also a framework for benchmarking. It can be used generically for any change-related situation, strategic, tactical or operational. The current situation is compared with best practices; the result is the input for the transition plans, together with the goals related to the change process.

The transition plans describe the way that procedural changes will take place and result in the actual changes. Thanks to continuous measurement, *through measurement of defined processes*, assessment of the changes compared with the goals is possible, and may result in the changing of actions, facilitating the process of continuous improvement.

11.3.2 A 'health check'

A 'health check' based on the current procedures can be used as an objective way of assessing the effectiveness of Service Management processes in an organisation. This assessment should aim to identify those aspects that are functioning well, thus determining which best practices are in current use and should be retained, and also to pinpoint problem areas and constraints. Using the recommendations from a health check, you will be better equipped to define your implementation or improvement priorities.

To summarise, a health check should:

- objectively assess the effectiveness of your Service Management processes
- identify constraints and problem areas
- provide you with advice on how to manage your processes more effectively
- provide you with advice on how to improve your processes.

Non-IT-related issues that can influence your performance in delivering services such as people management and resource management can be assessed by using (self-) assessment provided by Total Quality Management methods. Be aware that in many cases it is these factors that have the major effect upon the actual performance from the Service Management processes. For further reading, see Appendix D, Quality.

Some examples of general topics that should be addressed by questions in a health check include:

- the existence of a strategic business plan
- how the plan supports IT planning
- the extent to which IT supports the needs of the business
- the alignment of IT and business growth plans.

Some examples of process topics, which should be addressed by questions in a health check, include:

- activities for each process
- the organisation of tasks and responsibilities
- communication lines between processes
- the overall control of Service Management
- an IT Infrastructure description
- control over changes to the IT infrastructure
- the level of Customer satisfaction with IT services.

Some improvements will require major change to the current processes, within the organisation, and it may be a considerable time before they can be implemented. Wherever possible, some 'quick wins' should be implemented and communicated so that everyone involved can see that improvements are being achieved prior to the final implementation. Using the health check can assist in the identification of quick-win areas.

Health checks and self-assessments can also help to determine the maturity level of your organisation. This is important if your target is better IT and business alignment. Maturity of process is an important issue, but has to be modified by the knowledge of what the business requires and will pay for in terms of maturity.

11.4 General guidelines on project planning

The CCTA project management method PRINCE2 (Projects IN Controlled Environments, version 2) is widely adopted internationally and is used to describe an approach to projects within the ITIL context. The guidelines below are consistent with a PRINCE2 approach.

11.4.1 Project characteristics

A project can be defined as:

a temporary organisation that is needed to achieve a predefined result at a predefined time using predefined resources.

A ‘temporary organisation’ means that a project has a beginning and a clear ending and is conducted alongside day-to-day activities. By doing this, the project activity can be isolated from ongoing work.

PRINCE2 concentrates on creating an appropriate management environment to achieve the stated aim of the project. To achieve this a PRINCE2 project requires the following to exist:

- a finite and defined lifespan for the project
- a set of defined and measurable business products (to achieve quality requirements)
- a set of activities to achieve the business products (i.e. the ‘doing’ of the project)
- a defined amount of resources
- a project organisation structure, with defined responsibilities, to manage the project.

Before starting a project, an organisation should have a vision about what the results are intended to be. By defining the means necessary to achieve the project result, it is possible to isolate these assets (people, budget, etc) from the day-to-day activities. This increases the success rate of the project.

Before a project actual starts, management should have an overall ‘feel’ for the project and be able to document:

- project definition, explaining what the project needs to achieve - this should give background information, project objectives and scope, and it should outline the desired outcome and state constraints on the project
- the business case, describing how the project outcome will support business needs and justify its existence - including reasons for selection of the approach
- known quality expectations of the business solution
- acceptance criteria for the final outcome
- known risks
- a high-level plan identifying necessary roles and, if possible, assigning them to individuals, as well as identifying the major ‘go/no go’ decision points.

11.4.2 Business case for the project

The business case describes the added value of the project for the organisation: why should this project be carried out? Of course, to establish the answer, the project costs and revenues -- perhaps this is more accurately described as savings - should be compared. The difficulty in doing this, however, is that while the *costs* are relatively easy to quantify (people, budget, etc.), this is not the case with the *revenues*. Some ideas that you might find useful are described in Appendix E with its example cost-benefit analysis.

Particularly with process-oriented projects, assessing and describing the revenues/savings is a hard task. This has to do with the fact that process implementation results in higher quality service provision, higher service levels and a more flexible organisation; these are not always quantifiable financial results. Sometimes you find that you are making investments (or incurring costs, depending on your point of view) without knowing clearly the benefit; it is no use telling the budget holder that you sometimes have to take chances to if you're going to succeed. The business case should enable the reader to understand the value of investing in Service Management process improvement.

11.4.3 Critical success factors and possible problems

Successful Service Management should:

- provide a good understanding of the Customer's requirements, concerns and business activities, and deliver business-driven rather than technology-driven services
- enhance Customer satisfaction
- improve value for money, resource utilisation and service quality
- deliver an infrastructure for the controlled operation of ongoing services by formalised and disciplined processes
- equip staff with goals and an understanding of the Customers' needs.

Problems with Service Management processes that may be encountered include:

- excessively bureaucratic processes, with a high percentage of the total support headcount dedicated to Service Management
- inconsistent staff performance for the same process (often accompanied by noticeable lack of commitment to the process from the responsible staff)
- lack of understanding on what each process should deliver
- no real benefits, service-cost reductions, or quality improvements arising from the implementation of service management processes
- unrealistic expectations, so that service targets are rarely hit
- no discernible improvement.

Some of the major issues concerned with defining and running a successful Service Management function and the project to implement it are discussed in Appendix C.

11.4.4 Project costs

When you build the business case for a project, it is essential that you are clear about what the project costs are and what will be the ongoing running costs of the Service Management processes. Project costs are one-off costs, while the running costs form a commitment for the organisation that may involve long-term contracts with suppliers.

The costs of implementing IT Infrastructure Library processes clearly vary according to the scale of operations. The costs associated with the implementation and running of the processes are roughly categorised as follows:

- project management costs
- project delivery costs (consultancy fees, project team for implementation, process owner)
- equipment and software
- training costs (including awareness, training in specific tools, and training in business awareness)
- documentation costs
- ongoing staff and accommodation costs (for running the processes, including subsequent training needs).

The costs of failing to provide effective processes can be considerable. Some examples are provided in Appendix E.

11.4.5 Organisation

A project needs to be managed, as well as to produce something, in order to achieve the stated end result. Managing a project needs to take account of three viewpoints:

- business – will the outcome support a real business need?
- User – when using the product, will it achieve the objective the User wants?
- supplier/technical – can the product be created (particularly within any given constraints)? Can the product be supported effectively when in operation?

A project needs to balance these three views if it is to achieve a viable result that is 'fit for its business purpose' and be achievable within the other constraints of time and cost.

Typically senior managers provide direction in these areas but wish to leave the day-to-day activity of managing the project to a project manager. PRINCE2 identifies a project board to cover these three interests and provide direction and advice to the project manager without being involved in day-to-day activities.

The project board is responsible for ensuring that the project results in the desired outcome and so should ensure that quality assurance is applied to the project in an appropriate manner. This activity needs to be separated from the project manager to ensure the board gets an objective answer to the question 'Are things really going as well as we are being told?'

In a Service Management process-implementation/improvement project, the three viewpoints are likely to be represented thus:

- the business executive will probably be a director or board member
- the User role could be a (senior) Customer
- the supplier would be the IT manager.

11.4.6 Products

Many contemporary project management methods have a product approach. The advantage is that products - unlike activities - can be described even before the project starts. In this way, a certain outcome of the project can be guaranteed by setting norms for the products to be delivered.

The principle of product planning presupposes good product description and management.
For further information, look at PRINCE2 or refer to your own project management approach.

11.4.7 Planning

After having defined the results in terms of products, the project manager should work out what other products need to be produced on the way to achieving the final outcome. A clear view of the activities can be built up for those products needed in the short-term (typically three to six months) with a high-level view of what is required in the longer term.

The project manager outlines the total project and produces a detailed plan for short-term activity. At this point, resources can be assigned to the activities to build products in the short term and skills requirements for the longer term can also be assessed.

While it is widely accepted that effective project planning and management are essential to project success, many IT projects continue to be poorly planned and badly managed. Often, people won't back out once (large) investments have been made. *"We've already invested heavily in this project, we can't stop now!"* has been used more than once in most organisations. Sometimes the runaway cost is the very reason for dissatisfaction and project cancellation should have taken place much earlier. To resist the temptation of an endless pursuit of a positive outcome of the project, an experienced project manager should assign a number of ‘go/ no go’ decision moments to his project.

This means that before the project is started, it is accepted that it may never be finished successfully. In fact, with each ‘go/ no go’ decision moment the business case of the project should be re-evaluated. In PRINCE2 these decision points come at the end of ‘stages’ where project progress to date is assessed and the ongoing viability of the future of the business case (and so the project) is actively revisited before the next stage is started.

11.4.8 Communication plan

Managing change can only succeed with the correct use of communication. A Service Management project will involve a lot of people but, typically, the outcome will affect the working lives of many more. Implementing or improving Service Management within an organisation requires a change of mindset by IT management and IT employees as well as IT Customers and Users. Communication around this transformation is essential to its success.

In order to ensure that all parties are aware of what is going on and can play a relevant part in the project, it is advisable to clarify how the project will communicate with all interested parties. A well-planned and executed communication plan will have a direct positive contribution to the success of the project.

A good communication plan should be built on a proper conception of what communication is. Communication is more than a one-way information stream. It requires continuous attention to the signals (positive and negative) of the various parties involved. Managing communications effectively involves the following nine steps:

1. describe the communications process in the Change process from the start
2. analyse the communication structure and culture
3. identify the important target groups
4. assess the communication goals for each target group
5. formulate a communication strategy for each target group
6. choose the right communication media for each target group
7. write a communication plan
8. communicate

9 measure and redirect if necessary.

A communication plan describes how target groups, contents and media are connected in a timeframe. Much like a project plan, a communication plan shows how actions, people, means and budget are to be allocated for the communication process.

11.5 Project review and management reporting

When a project is set up it is important to consider the reporting needs. Project management should be used to ensure appropriate decisions can be made. By exercising control over a project it should be possible to show that the project:

- is producing the required results, namely the results which meet predefined quality criteria
- is being carried out to schedule and in accordance with its previously agreed resource and cost plans
- remains viable against its business case (balancing benefits against costs/risks).

To support the decision-making processes, organisations should expect a number of reports throughout the life of a project. At the very least, a project should produce:

- regular progress reports
- post-completion project evaluation (of the way the project was run)
- post-completion project review to assess if the projected benefits have materialised.

As part of the need to evaluate the project is essential to maintain records that enable the project to be audited. Auditing may cover compliance and efficiency as well as looking at improvements that have been achieved or that could still be attempted.

On completion of a project, management will require further, regular, reports to show how well the Service Management processes are supporting the business needs.

11.5.1 Progress reporting

Progress against plans should be assessed on a regular basis, so that problems can be identified early and can be dealt with in a timely manner. The project manager should ensure that progress reports are produced for the project board at regular, agreed, intervals. The reports should include statements regarding:

- achievements in the current period
- achievements expected in the next period
- actual or potential problems and suggestions for their prevention or resolution.

Progress reports should provide a clear picture of the status of the project against plan and the business case so that adequately informed decisions can be made as to whether or not to continue expending resources on the project. It is important to look at the risks, at any changes in these within the current period and, if appropriate, identify their impact on subsequent project activities.

If a problem arises between progress reports that the project manager is not authorised to sort out, then an interim report should be compiled for the project board without waiting for the next progress report to be due. Within this report, the project manager states the nature and scope of the problem that has arisen, identifies options for its resolution, and recommends a course of action.

11.5.2 *Evaluation of the project*

As the project draws to a close, it is important to analyse how the project was managed and to identify lessons that were learned along the way. This information can then be used to benefit the project team as well as the organisation as a whole. An End Project Report typically covers:

- achievement of the project's objectives
- performance against plan (estimated time and costs versus actuals)
- effect on the original plan and business case over the time of the project
- statistics on issues raised and changes made
- total impact of changes approved
- statistics on the quality of the work carried out (in relation to stated expectations)
- lessons learned
- a post-project review plan.

11.5.3 *Post-project review*

The business case will have been built from the premise that the outcome of the project will deliver benefits to the business over a period of time. Thus, delivery of benefits needs to be assessed at a point after the project products have been put into use. The post-project review is used to assess whether the expected benefits have been realised, as well as to investigate whether problems have arisen from use of the products.

Each of the benefits mentioned in the business case should be assessed to see how well, if at all, it has been achieved. Other issues to consider are whether there were additional benefits - or unexpected problems. Both of these can be used to improve future business cases.

If necessary, follow-up actions may be identified to improve the situation that then exists.

11.5.4 *Auditing for compliance using quality parameters*

Process quality parameters can be seen as the 'operational thermometer' of the IT organisation. Using them, you can determine whether the IT organisation is effective and efficient. Quality parameters need to be quantified for your own circumstances. However, this task will be easier once you have determined the required Service Levels and Internal Service Requirements. There are two types of quality parameters, process specific and generic.

Generic quality parameters for IT Service Management

Generic quality parameters that need to be considered include:

- Customer satisfaction
- staff satisfaction
- efficiency
- effectiveness.

Appropriate information should be collected to rate the organisation's performance relative to these parameters. The nature of the information required will vary depending on how you decide to judge each aspect, but what information is required should be clearly thought through from the start of the project so that measurement is possible during the post-project review.

Process-specific parameters

Process-specific metrics for each process are discussed in each of the process-specific chapters of this book.

11.5.5 Auditing for improvement using key performance indicators

Introduction to the 'Balanced Scorecard'

The 'Balanced Scorecard' is an aid to organisational performance management. It helps to focus, not only on the financial targets but also on the internal processes, Customers and learning and growth issues. The balance should be found between these four perspectives[1].

The four perspectives are focused around the following questions:

1 Customers: what do our Customers desire?
2 internal processes: how do we generate the added value for our Customers?
3 learning & growth: how do we guarantee we will keep generating added value in the future?
4 financial: how did we do financially?

As you can see, the first three questions focus on the future, the last question reviews what has gone before. It is worthwhile discussing the Balanced Scorecard further at this point:

- The Balanced Scorecard is not complex but to implement the scorecard successfully *is* complex. In practice, it can take an organisation up to three years to see the benefits of a Balanced Scorecard approach.
- The Balanced Scorecard is not an exclusive IT feature. On the contrary, many organisations use scorecards in other departments - even at board level.
- When implementing the Balanced Scorecard, it pays to start very conservatively. Start with three or four goals for each perspective. To do this, an organisation has to make choices; for many, this is extremely difficult and time consuming to do.
- The most difficult part of using the Balanced Scorecard is not the implementation; it is the consolidation. Usually, consultants are employed to assist in the introduction of the Balanced Scorecard. The challenge is to keep measuring once they have gone.
- The danger is in the temptation to fall back on prior measuring techniques or not measuring at all.

The Balanced Scorecard is complementary to ITIL. It is a way of measuring the effectiveness of the performance of the organisation. Some of the links include the following:

---

1 Kaplan and Norton introduced the Balanced Scorecard in the early 1980's. Contemporary measurements from the time, particularly in US companies, focused on financial targets. Because financial figures are solely concentrated on past events, Kaplan and Norton sought a means to measure and steer using future activities as well. See the appropriate entry in the Bibliography to the book (Chapter 12).

- *client perspective*: this is relevant to most processes and is particularly relevant to Service Level Management where it is documented in Service Level Agreements
- *internal processes*, these of course cover the ITIL processes
- *financial*: Financial Management covers the way costs are allocated to the Customer organisation
- *learning and growth* refers to staffing, training and investments in software.

11.5.6 Management reporting

After implementation of an initial Service Management system, or some improvements to it, a regular system for management reporting has to be set up. The following types of management reports should be considered:

- IT management reports that are used for planning and control of services
- reports matching achieved internal service levels, with service levels as described in the Service Level Agreements
- internal process-management reports, used by the process manager to determine the process' efficiency and effectiveness and for auditing for compliance
- Service Management reports; which the Service Manager uses for higher-level process control.

12 BIBLIOGRAPHY

Note: that the entries in this Bibliography are given throughout in alphabetical order of title within each section.

12.1 References

Acquisition
CCTA IS Management Guide, www.ccta.gov.uk Tel: +44(0) 1603 704567
Published by Format
ISBN 1-90309-1-03-9

Balanced Scorecard (The): Translating strategy into action
Robert S. Kaplan, David P. Norton / Hardcover / Published 1996
Harvard Business School Press
ISBN: 0-87584-651-3

Capability Maturity Model (The): Guidelines for Improving the Software Process
Carnegie Mellon University, Software Engineering Institute, Addison-Wesley
ISBN 0-201-54664-7

Code of Practice for IT Service Management (A), DISC 0005
Extracts are reproduced with the permission of BSI under licence number PD\ 1999 0877. Complete copies of the standard can be obtained by post from BSI Customer Services, 389 Chiswick High Road, London W4 4AL.
www.bsi.ork.uk/disc/products, Tel: 020-8996 9001

Data Protection Act
Available from The Stationery Office, www.the-stationery-office.co.uk
ISBN 010 543 5848

Cultures of work organisations (The)
Trice/Beyer (1993). Prentice Hall, N.J. U.S.A.
ISBN 0-13-191438-3

In Search of Excellence: Lessons From America's Best-Run Companies
Thomas Peters, Robert H. Waterman, Tom Peters
Warner Books
ISBN 0-44-638507-7

Leading change
John P. Kotter/ hardcover/1996
Harvard Business School Press
ISBN: 0-875847-47-1

Managing Change
CCTA IS Management Guide, www.ccta.gov.uk Tel: +44(0) 1603 704567
Published by Format
ISBN 1-90309-1-01-2

Managing Performance
CCTA IS Management Guide, www.ccta.gov.uk Tel: +44(0) 1603 704567
Published by Format
ISBN 1-90309-1-05-5

Managing Services
CCTA IS Management Guide, www.ccta.gov.uk Tel: +44(0) 1603 704567
Published by Format
ISBN 1-90309-1-04-7

Managing Successful Projects with PRINCE 2.
CCTA, , www.ccta.gov.uk Tel: +44(0) 1603 704567 or The Stationery Office, tel 020-7873 9090
ISBN 0-11-330855-8

Organisational transformations
Espejo/Schumann/Schwaninger/Bilello (1996). John Wiley&Sons; Chichester, England
ISBN 0-471-96182-5

PRINCE2 (reference manual)
ISBN 0-11-330685 7

Process Innovation
Davenport (1993). HBS, Boston, Massachusetts U.S.A.
ISBN 0-87584-366-2

Risk Management Library
Various titles
CCTA, www.ccta.gov.uk Tel: +44(0) 1603 704567

Service Quality
Brown/Gummesson/Edvardsson/Gustavsson (1991). Lexington Books; N.Y., U.S.A.
ISBN 0-669-21152-4

Strategic Management of IS
CCTA IS Management Guide, www.ccta.gov.uk Tel: +44(0) 1603 704567
Published by Format
ISBN 1-90309-1-02-0

Structures in fives: designing effective organisations
Minzberg (1993). Prentice Hall, N.J. U.S.A.
ISBN 0-13-855479-X

Turnbull Report on Corporate Governance
Institute of Chartered Accountants of England and Wales
ISBN 1 84152 010 1

12.2 Other Sources

European Foundation for Quality Management
Website at //www.efqm.org. '...the battle for Quality is one of the prerequisites for the success of your companies and for our collective success'.
Jacques Delors

W. Edwards Deming Institute (The)
Website at //www.deming.org. 'We should work on our process, not the outcome of our processes.'
W. Edwards Deming

ISO 9000 Information Forum
Website at //www.iso-9000.co.uk

British Standards Institution
Website at //www.bsi.org.uk

APPENDIX A – LIST OF ACRONYMS AND GLOSSARY

A.1 Acronyms used in this book

AMDB	Availability Management Database
ATM	Auto Teller Machine
BCM	Business Continuity Management
BIA	Business Impact Analysis
BRM	Business Relationship Management
CAB	Change Advisory Board
CCTA	Central Computer and Telecommunications Agency
CDB	Capacity Management Database
CFIA	Component Failure Impact Analysis
CI	Configuration Item
CMDB	Configuration Management Database
CRAMM	CCTA Risk Analysis Management Method
CSS	Customer Satisfaction Survey
DT	Down Time
EFQM	European Foundation for Quality Management
EUDT	End User Down Time
FTA	Fault Tree Analysis
ICT	Information and Communication Technology(ies)
ISEB	Information Systems Examination Board
ITAMM	IT Availability Metrics Model
ITIL	Information Technology Infrastructure Library
ITSC	IT Service Continuity
ITSCM	IT Service Continuity Management
*it*SMF	IT Service Management Forum
KIP	Key Performance Indicator
LAN	Local Area Network
MIM	Major Incident Management
MTBSI	Mean Time Between System Incidents
MTTR	Mean Time to Repair
OLA	Operational Level Agreement
OLTP	On-line Transaction Processing
PKI	Public Key Infrastructure
QA	Quality Assurance
RAG	Red-Amber-Green
RAID	Redundant Array of Inexpensive Disks
ROCE	Return on Capital Employed
ROI	Return on Investment
RWO	Real World Object
SIP	Service Improvement Programme
SLA	Service Level Agreement
SLAM	SLA Monitoring
SLM	Service Level Management
SLR	Service Level Requirement
SMO	Service Maintenance Objectives
SOA	System Outage Analysis
SOO	Service Outage Objective
SPOF	Single Point of Failure
TCO	Total Cost of Ownership

TOP	Technical Observation Post
TQM	Total Quality Management
UPS	Uninterruptible Power Supply
VBF	Vital Business Function
WAN	Wide Area Network

A.2 Glossary of CCTA terms

Absorbed overhead	Overhead which, by means of absorption rates, is included in costs of specific products or saleable services, in a given period of time. Under or over-absorbed overhead. The difference between overhead cost incurred and overhead cost absorbed: it may be split into its two constituent parts for control purposes
Absorption costing	A principle whereby fixed as well as variable costs are allotted to cost units and total overheads are absorbed according to activity level. The term may be applied where production costs only, or costs of all functions are so allotted.
Action lists	Defined actions, allocated to recovery teams and individuals, within a phase of a plan These are supported by reference data.
Alert phase	The first phase of a business continuity plan in which initial emergency procedures and damage assessments are activated.
Allocated cost	A cost that can be directly identified with a business unit
Apportioned cost	A cost that is shared by a number of business units (an indirect cost). This cost must be shared out between these units on an equitable basis.
Asset	Component of a business process. Assets can include people, accommodation, computer systems, networks, paper records, fax machines, etc.
Assurance	The processes by which an organisation can verify the accuracy and completeness of its BCM.
Asynchronous /synchronous	Asynchronous in a communications sense is the ability to transmit each character as a self-contained unit of information, without additional timing information. This method of transmitting data is sometimes called start/stop. Synchronous working involves the use of timing information to allow transmission of data, which is normally done in blocks. Synchronous transmission is usually more efficient than the asynchronous method.
Availability	Ability of a component or service to perform its required function at a stated instant or over a stated period of time. It is usually expressed as the availability ratio, i.e. the proportion of time that the service is actually available for use by the Customers within the agreed service hours.
BCM activity	An action or series of actions as part of a BCM process.
BCM Lifecycle	The complete set of activities and processes necessary to manage business continuity - divided into four stages.
BCM process	A set of activities with defined deliverables forming a discrete part of the BCM lifecycle.
Bridge	A bridge is equipment and techniques used to match circuits to each other ensuring minimum transmission impairment.
Build	The final stage in producing a usable configuration. The process involves taking one of more input Configuration Items and processing them (building them) to create one or more output Configuration Items e.g. software compile and load.
Business function	A business unit within an organisation, eg a department, division, branch.
Business process	A group of business activities undertaken by an organisation in pursuit of a common goal. Typical business processes include receiving orders, marketing services, selling products, delivering services, distributing products, invoicing for services, accounting for money received. A business process will usually depend upon several business functions for support, eg IT, personnel, accommodation. A business process will rarely operate in isolation, ie, other business processes will depend on it and it will depend on other processes.
Business recovery objective	The desired time within which business processes should be recovered, and the minimum staff, assets and services required within this time.
Business recovery plan framework	A template business recovery plan (or set of plans) produced to allow the structure and proposed contents to be agreed before the detailed business recovery plan is produced.
Business recovery plans	Documents describing the roles, responsibilities and actions necessary to resume business processes following a business disruption.

Business recovery team	A defined group of personnel with a defined role and subordinate range of actions to facilitate recovery of a business function or process
Business unit	A segment of the business entity by which both revenues are received and expenditure are caused or controlled, such revenues and expenditure being used to evaluate segmental performance.
Capital Costs	Typically those applying to the physical (substantial) assets of the organisation. Traditionally this was the accommodation and machinery necessary to produce the enterprise's product. Capital Costs are the purchase or major enhancement of fixed assets, for example computer equipment (building and plant and are often also referred to as 'one-off' costs.
Capital investment appraisal	The process of evaluating proposed investment in specific fixed assets and the benefits to be obtained from their acquisition. The techniques used in the evaluation can be summarised as non-discounting methods (ie simple pay-back), return on capital employed and discounted cash flow methods (ie yield, net present value and discounted pay-back).
Capitalisation	Many organisations choose to identify major expenditure as Capital, whether there is a substantial asset or not, to reduce the impact on the current financial year of such expenditure and this is referred to as 'Capitalisation'. The most common item for this to be applied to is software, whether developed in-house or purchased.
Category	Classification of a group of Configuration Items, Change documents or problems.
Change	The addition, modification or removal of approved, supported or baselined hardware, network, software, application, environment, system, desktop build or associated documentation.
Change Advisory Board	A group of people who can give expert advice to Change Management on the implementation of Changes. This board is likely to be made up of representatives from all areas within IT and representatives from business units.
Change authority	A group that is given the authority to approve Change, e.g. by the project board. Sometimes referred to as the Configuration Board.
Change control	The procedure to ensure that all Changes are controlled, including the submission, analysis, decision making, approval, implementation and post implementation of the Change.
Change document	Request for Change, Change control form, Change order, Change record.
Change history	Auditable information that records, for example, what was done, when it was done, by whom and why.
Change log	A log of Requests for Change raised during the project, showing information on each Change, its evaluation what decisions have been made and its current status, e.g. Raised, Reviewed, Approved, Implemented, Closed.
Change Management	Process of controlling Changes to the infrastructure or any aspect of services, in a controlled manner, enabling approved Changes with minimum disruption.
Change record	A record containing details of which CIs are affected by an authorised Change (planned or implemented) and how.
Channel	Channel is the physical connection from CPU to an I/O device, usually a controller, or indeed another CPU.
Charging	The process of establishing charges in respect of business units, and raising the relevant invoices for recovery from customers.
Classification	Process of formally grouping Configuration Items by type, e.g. software, hardware, documentation, environment, application. Process of formally identifying Changes by type e.g. project scope change request, validation change request, infrastructure change request. Process of formally identifying incidents, problems and known errors by origin, symptoms and cause.
Closure	When the Customer is satisfied that an incident has been resolved.
Cold stand-by	See 'Gradual Recovery'

Computer Aided Systems Engineering	A software tool for programmers. It provides help in the planning, analysis, design and documentation of computer software.
Command, control and communications	The processes by which an organisation retains overall co-ordination of its recovery effort during invocation of business recovery plans.
Configuration Baseline	Configuration of a product or system established at a specific point in time, which captures both the structure and details of the product or system, and enables that product or system to be rebuilt at a later date. A snapshot or a position which is recorded. Although the position may be updated later, the baseline remains unchanged and available as a reference of the original state and as a comparison against the current position (PRINCE 2).
Configuration control	Activities comprising the control of Changes to Configuration Items after formally establishing its configuration documents. It includes the evaluation, co-ordination, approval or rejection of Changes. The implementation of Changes includes changes, deviations and waivers that impact on the configuration.
Configuration documentation	Documents that define requirements, system design, build, production, and verification for a configuration item.
Configuration identification	Activities that determine the product structure, the selection of Configuration Items, and the documentation of the Configuration Item's physical and functional characteristics including interfaces and subsequent Changes. It includes the allocation if identification characters or numbers to the Configuration Items and their documents. It also includes the unique numbering of configuration control forms associated with Changes and Problems.
Configuration item (CI)	Component of an infrastructure - or an item, such as a Request for Change, associated with an infrastructure - which is (or is to be) under the control of Configuration Management. CIs may vary widely in complexity, size and type - from an entire system (including all hardware, software and documentation) to a single module or a minor hardware component.
Configuration Management	The process of identifying and defining the Configuration Items in a system, recording and reporting the status of Configuration Items and Requests for Change, and verifying the completeness and correctness of configuration items.
Configuration Management Tool (CM Tool)	A software product providing automatic support for Change, Configuration or version control.
Configuration Management Database	A database which contains all relevant details of each CI and details of the important relationships between CIs.
Configuration Management plan	Document setting out the organisation and procedures for the Configuration Management of a specific product, project, system, support group or service.
Configuration Structure	A hierarchy of all the CIs that comprise a configuration.
Contingency Planning	Planning to address unwanted occurrences that may happen at a later time. Traditionally, the term has been used to refer to planning for the recovery of IT systems rather than entire business processes.
Cost	The amount of expenditure (actual or notional) incurred on, or attributable to, a specific activity or business unit.
Cost effectiveness	Ensuring that there is a proper balance between the quality of service on the one side and expenditure on the other. Any investment that increases the costs of providing IT services should always result in enhancement to service quality or quantity.
Cost management	The term used in this module to describe a!! the procedures, tasks and deliverables that are needed to fulfil an organisation's costing and charging requirements.
Cost unit	In the context of CSBC the cost unit is a functional cost unit which establishes standard cost per workload element of activity, based on calculated activity ratios converted to cost ratios,
Costing	The process of identifying the costs of the business and of breaking them down and relating them to the various activities of the organisation.

CRAMM	CCTA Risk Analysis and Management Method.
Crisis management	The processes by which an organisation manages the wider impact of a disaster, such as adverse media coverage.
Customer	Recipient of the service; usually the Customer management has responsibility for the cost of the service, either directly through charging or indirectly in terms of demonstrable business need.
Data transfer time	Data transfer time is the length of time taken for a block or sector of data to be read from or written to an I/O device, such as a disk or tape.
Definitive Software Library (DSL)	The library in which the definitive authorised versions of all software CIs are stored and protected. It is a physical library or storage repository where master copies of software versions are placed. This one logical storage area may in reality consist of one or more physical software libraries or filestores. They should be separate from development and test filestore areas. The DSL may also include a physical store to hold master copies of bought-in software, e.g. fire-proof safe. Only authorised software should be accepted into the DSL, strictly controlled by Change and Release Management. The DSL exists not directly because of the needs of the Configuration Management process, but as a common base for the Release Management and Configuration Management processes.
Delta Release	A Delta, or partial, Release is one that includes only those CIs within the Release unit that have actually changed or are new since the last full or Delta Release. For example, if the Release unit is the program, a Delta Release contains only those modules that have changed, or are new, since the last full release of the program or the last Delta Release of the modules - see also ‘Full Release’.
Dependency	The reliance, either direct or indirect, of one process or activity upon another.
Depreciation	Depreciation is the loss in value of an asset due to its use and/or the passage of time, The annual depreciation charge in accounts represents the amount of capital assets need up in the accounting period. It is charged in the cost accounts to ensure that the cost of capital equipment is reflected in the unit costs of the services provided using the equipment. There are various methods of calculating depreciation for the period, but the Treasury usually recommend the use of current cost asset validation as the basis for the depreciation charge.
Differential charging	Charging business customers different rates for the same work, typically to dampen demand or to generate revenue for spare capacity. This can also be used to encourage off-peak or night time running.
Direct cost	A cost which is incurred for, and can he traced in full to a product, service, cost centre or department. This is an allocated cost. Direct costs are direct materials, direct wages and direct expenses.
Disaster recovery planning	A series of processes that focus only upon the recovery processes, principally in response to physical disasters, that are contained within BCM.
Discounted cash flow	An evaluation of the future net cash flows generated by a capital project by discounting them to their present-day value. The two methods most commonly used are: a) yield method, for which the calculation determines the internal rate of return (IRR) in the form of a percentage b) net present value (NPV) method, in which the discount rate is chosen and the answer is a sum of money.
Discounting	Discounting is the offering to business customers of reduced rates for the use of off-peak resources (see also Surcharging).
Disk cache controller	Disk cache controllers have memory which is used to store blocks of data which have been read from the disk devices connected to them. If a subsequent I/O requires a record which is still resident in the cache memory, it will be picked up from there, thus saving another physical I/O.
Duplex (full and half)	Full duplex line/channel allows simultaneous transmission in both directions. flail duplex line/channel is capable of transmitting in both directions, but only in one direction at a time.
Echoing	Echoing is a reflection of the transmitted signal from the receiving end, a visual method of error detection in which the signal from the originating device is looped back to that device so that it can be displayed..
Elements of cost	The constituent parts of costs according to the factors upon which expenditure is

	incurred viz, materials, labour and expenses.
End-User	see ‘User’.
Environment	A collections of hardware, software, network communications and procedures that work together to provide a discrete type of computer service. There may be one or more environments on a physical platform e.g. test, production. An environment has unique features and characteristics that dictate how they are administered in similar, yet diverse manners.
Expert User	See ‘Super User’.
External Target	One of the measures against which a delivered IT service is compared, expressed in terms of the customer's business.
Financial year	The financial year is an accounting period covering 12 consecutive months. In the public sector this financial year will generally coincide with the fiscal year which runs from 1 April to 31 March.
Forward Schedule of Changes	Contains details of all the Changes approved for implementation and their proposed implementation dates. It should be agreed with the Customers and the business, Service Level Management, the Service Desk and Availability Management. Once agreed, the Service Desk should communicate to the User community at large any planned additional downtime arising from implementing the Changes, using the most effective methods available.
Full Cost	Calculated as a total cost of ownership, including depreciation / planned renewal)
Full cost	Full cost is the total cost of all the resources used in supplying a service ie the sum of the direct costs of producing the output a proportional share of overhead costs and any selling and distribution expenses. Both cash costs and notional (non-cash) costs should be included, including the cost of capital.
Full Release	All components of the Release unit are built, tested, distributed and implemented together - see also ‘Delta Release’.
Gateway	A gateway is equipment which is used to interface networks so that a terminal on one network can communicate with services or a terminal on another.
Gradual Recovery	Previously called 'Cold stand-by', this is applicable to organisations that do not need immediate restoration of business processes and can function for a period of up to 72 hours, or longer, without a re-establishment of full IT facilities. This may include the provision of empty accommodation fully equipped with power, environmental controls and local network cabling infrastructure, telecommunications connections, and available in a disaster situation for an organisation to install its own computer equipment.
Hard charging	Descriptive of a situation where, within an organisation, actual funds are transferred from the customer to the IT directorate in payment for the delivery of IT services.
Hard fault	Hard faults describe the situation in a virtual memory system when the required page of code or data, which a program was using, has been redeployed by the operating system for some other purpose. This means that another piece of memory must be found to accommodate the code or data, and will involve physical reading/writing of pages to the page file.
Help desk	The single point of contact within the IT directorate for users of IT services.
Host	A host computer comprises the central hardware and software resources of a computer complex, eg CPU, memory, channels, disk and magnetic tape I/O subsystems plus operating and applications software. The term is used to denote all non-network items.
Hot stand-by	See 'Immediate Recovery'
ICT	The convergence of Information Technology, Telecommunications and Data Networking Technologies into a single technology
Immediate Recovery	Previously called 'Hot stand-by', provides for the immediate restoration of services following any irrecoverable incident. It is important to distinguish between the previous definition of ‘hot standby’ and ‘immediate recovery’. Hot standby typically referred to availability of services within a short timescale such as 2 or 4 hours whereas immediate recovery implies the instant availability of services.
Impact	Measure of the business criticality of an Incident. Often equal to the extend to

	which an Incident leads to distortion of agreed or expected service levels.
Impact analysis	The identification of critical business processes, and the potential damage or loss that may be caused to the organisation resulting from a disruption to those processes. Business impact analysis identifies the form the loss or damage will take; how that degree of damage or loss is likely to escalate with time following an incident; the minimum staffing, facilities and services needed to enable business processes to continue to operate at a minimum acceptable level; and the time within which they should be recovered. The time within which full recovery of the business processes is to be achieved is also identified.
Impact scenario	Description of the type of impact on the business that could follow a business disruption. Will usually be related to a business process and will always refer to a period of time, eg customer services will be unable to operate for two days.
Incident	Any event which is not part of the standard operation of a service and which causes, or may cause, an interruption to, or a reduction in, the quality of that service.
Indirect cost	An indirect cost is a cost incurred in the course of making a product providing a service or running a cost centre or department, but which cannot be traced directly and in full to the product, service or department, because it has been incurred for a number of cost centres or cost units. These costs are apportioned to cost centres/cost units. Indirect costs are also referred to as overheads.
Intelligent customer	The purchaser (as distinct from the provider) of services. The term is often used in relation to the outsourcing of IT/IS.
Interface	Physical or functional interaction at the boundary between Configuration Items.
Intermediate Recovery	Previously called 'Warm stand-by', will typically involve the reestablishment of the critical systems and services within a 24 to 72 hour period, and will be used by organisations that need to recover IT facilities within a predetermined time to prevent impacts to the business process.
Internal target	One of the measures against which supporting processes for the IT service are compared. Usually expressed in technical terms relating directly to the underpinning service being measured.
Invocation (of business recovery plans)	Putting business recovery plans into operation after a business disruption.
Invocation (of stand by arrangements)	Putting stand-by arrangements into operation as part of business recovery activities.
Invocation and recovery phase	The second phase of a business recovery plan.
ISO9001	The internationally accepted set of standards concerning quality management Systems.
IT directorate	That part of an organisation charged with developing and delivering the IT services.
ITIL	The CCTA IT Infrastructure Library - a set of guides on the management and provision of operational IT services
Known Error	An Incident or Problem for which the root cause is known and for which a temporary Work-around or a permanent alternative has been identified. If a business case exists, an RFC will be raised, but, in any event, it remains a Known Error unless it is permanently fixed by a Change.
Latency	Latency describes the elapsed time from the moment when a seek was completed on a disk device to the point when the required data is positioned under the read/write heads. Latency is normally defined by manufacturers as being half the disk rotation time.
Lifecycle	A series of states, connected by allowable transitions. The lifecycle represents an approval process for Configuration Items, Problem Reports and Change documents.
Logical I/O	Logical I/O is a read or write request by a program. That request may, or may not, necessitate a physical I/O. For example, on a read request the required record may already be in a memory buffer and therefore a physical I/O wilt not be necessary.
Marginal Cost	The cost of providing the service now, based upon the investment already made

Marginal cost	The variable cost of producing one extra unit of product or service. That is, the cost which would have been avoided if the unit/service was not produced/provided.
Maturity level/Milestone	The degree to which BCM activities and processes have become standard business practice within an organisation.
Multiplexer	Multiplexers divide data channels into two or more independent fixed data channels of lower speed
Operational Costs	Those resulting from the day-to-day running of the IT Services section, e.g. staff costs, hardware maintenance and electricity, and relate to repeating payments whose effects can be measured within a short timeframe, usually the less than the 12-month financial year.
Operational level agreement	An internal agreement covering the delivery of services which support the IT directorate in their delivery of services.
Opportunity cost (or true cost)	The value of a benefit sacrificed in favour of an alternative course of action. That is the cost of using resources in a particular operation expressed in terms of foregoing the benefit that could be derived from the best alternative use of those resources.
Outsourcing	The process by which functions performed by the organisation are contracted out for operation, on the organisation's behalf, by third parties.
Overheads	The total of indirect materials, wages and expenses.
Package assembly /disassembly device	A package assembly/disassembly device permits terminals which do not have an interface suitable for direct connection to a packet switched network to access such a network. A PAD converts data to/from packets and handles call set-up and addressing.
Page fault	A program interruption which occurs when a page that is marked 'not in real memory' is referred to by an active page.
Paging	Paging is the I/O necessary to read and write to and from the paging disks: real (not virtual) memory is needed to process data. With insufficient real memory, the operating system writes old pages to disk, and reads new pages from disk, so that the required data and instructions are in real memory.
PD0005	Alternative title for the BSI publication 'A Code of Practice for IT Service Management'.
Percentage utilisation	Percentage utilisation describes the amount of time that a hardware device is busy over a given period of time. For example, lf the CPU is busy for 1800 seconds in one hour period, its utilisation is said to be 50%.
Phantom line error	A phantom line error is a communications error reported by a computer system which is not detected by network monitoring equipment. It is often caused by changes to the circuits and network equipment (eg re-routing circuits at the physical level on a backbone network) while data communications is in progress.
Physical I/O	Physical I/O means that a read or write request from a program has necessitated a physical read or write operation on an I/O device.
Prime cost	The total cost of direct materials, direct labour and direct expenses. The term prime cost is commonly restricted to direct production costs only and so does not customarily include direct costs of marketing or research and development.
PRINCE2	The standard UK government method for project management.
Priority	Sequence in which an Incident or Problem needs to be resolved, based on impact and urgency.
Problem	Unknown underlying cause of one or more Incidents.
Process	A connected series of actions, activities, Changes etc, performed by agents with the intent of satisfying a purpose or achieving a goal.
Process Control	The process of planning and regulating, with the objective of performing the process in an effective and efficient way.
Programme	A collection of activities and projects that collectively implement a new corporate requirement or function.
Queuing time	Queuing time is incurred when the device, which a program wishes to use, is already busy. The program will therefore have to wait in a queue to obtain service from that device.
Reference data	Information that supports the plans and action lists, such as names and addresses or inventories, which is indexed within the plan.

Release	A collection of new and/or changed CIs which are tested and introduced into the live environment together.
Request for Change (RFC)	Form, or screen, used to record details of a request for a change to any CI within an infrastructure or to procedures and items associated with the infrastructure.
Resolution	Action which will resolve an Incident. This may be a Work-around.
Resource cost	This term is used to describe the amount of machine resource that a given task will consume. This resource is usually expressed in seconds for the CPU or the number of I/Os for a disk or tape device.
Resource profile	Resource profile describes the total resource costs which are consumed by an individual online transaction, batch job or program. It is usually expressed in terms of CPU seconds, number of I/Os and memory usage.
Resource unit costs	Resource unit may be calculated on a standard cost basis to identify the expected (standard) cost for using a particular resource. Because computer resources come in many shapes and forms, units have to be established by logical groupings. Examples are; a) CPU Time or instructions, b) disk I/Os, c) print lines, d) communication transactions.
Resources	The term resources refers to the means the IT Services section needs to provide the customers with the required services. The resources are typically computer and related equipment, software, facilities or organisational (people).
Return to normal phase	The phase within a business recovery plan which re-establishes normal operations.
Risk	A measure of the exposure to which an organisation may be subjected. This is a combination of the likelihood of a business disruption occurring and the possible loss that may result from such business disruption.
Risk reduction measure	Measures taken to reduce the likelihood or consequences of a business disruption occurring (as opposed to planning to recover after a disruption).
Role	A set of responsibilities, activities and authorisations.
Roll in roll out (RIRO)	RIRO is a term which is used on some systems to describe swapping.
Rotational Position Sensing	RPS is a facility which is employed on most mainframes and some minicomputers. When a seek has been initiated the system can free the path from a disc drive to a controller for use by another disc drive, while it is waiting for the required data to come under the read/write heads (latency). This facility usually improves the overall performance of the I/O subsystem.
Seek time	Seek time occurs when the disc read/write heads are not positioned on the required track. It describes the elapsed time taken to move heads to the right track.
Self-insurance	A decision to bear the losses that could result from a disruption to the business as opposed to taking insurance cover on the risk.
Service achievement	The actual service levels delivered by the IT directorate to a customer within a defined lime-span.
Service catalogue	Written statement of IT services, default levels and options.
Service improvement programme	A formal project undertaken within an organisation to identify and introduce measurable improvements within a specified work area or, work process.
Service Level Agreement	Written agreement between a service provider and the Customer(s) that documents agreed service levels for a service.
Service level management	The process of defining, agreeing, documenting and managing the levels of customer IT service, that are required and cost justified.
Service provider	Third-party organisation supplying services or products to customers.
Service quality plan	The written plan and specification of internal targets designed to guarantee the agreed service levels.
Service Request	Every Incident not being a failure in the IT Infrastructure.
Services	Services are the deliverables of the IT Services section as perceived by the customers; the services do not consist merely of making computer resources available for customers to use.

Simulation modelling	Simulation modelling, as the name implies, employs a program which simulates computer processing by describing in detail the path of a job or transaction. It can give extremely accurate results. Unfortunately, it demands a great deal of time and effort from the modeller. It is most beneficial in extremely large or time critical systems where the margin for error is very small.
Soft fault	A soft fault describes the situation in a virtual memory system when the operating system has detected that a page of code or data was due to be reused, ie it is on a list of "free" pages, but it is still actually in memory. It is now rescued and put back into service.
Software Configuration Item (SCI)	As 'Configuration Item', excluding hardware and services.
Software Environment	Software used to support the application such as operating system, database management system, development tools, compilers, and application software.
Software Library	A controlled collection of SCIs designated to keep those with like status and type together and segregated from unlike, to aid in development, operation and maintenance.
Software work unit	Software work is a generic term devised to represent a common base on which all calculations for workload usage and IT resource capacity are then based. A unit of software work for I/O type equipment equals the number of bytes transferred; and for central processors it is based on the product of power and cpu-time.
Solid state devices	Solid state disks are memory devices which are made to appear as if they are disk devices. The advantages of such devices are that the service times are much faster than real disks since there is no seek time or latency. The main disadvantage is that they are much more expensive.
Specsheet	Specifies in detail what the customer wants (external) and what consequences this has for the service provider (internal) such as required resources and skills.
Standard cost	A pre-determined calculation of how much costs should be under specified working conditions. It is built up from an assessment of the value of cost elements and correlates technical specifications and the quantification of materials, labour and other costs to the prices and/or wages expected to apply during the period in which the standard cost is intended to be used. Its main purposes are to provide bases for control through variance accounting, for the valuation of work in progress and for fixing selling prices.
Standard costing	A technique which uses standards for costs and revenues for the purposes of control through variance analysis.
Stand-by arrangements	Arrangements to have available assets which have been identified as replacements should primary assets be unavailable following a business disruption. Typically, these include accommodation, IT systems and networks, telecommunications and sometimes people.
Storage occupancy	Storage occupancy is a defined measurement unit that is used for storage type equipment to measure usage. the unit value equals the number of bytes stored.
Super User	In some organisations it is common to use 'expert' Users (commonly known as Super or Expert Users) to deal with first-line support problems and queries. This is typically in specific application areas, or geographical locations, where there is not the requirement for full-time support staff. This valuable resource however needs to be carefully co-ordinated and utilised.
Surcharging	Surcharging is charging business users a premium rate for using resources at peak times.
Swapping	The reaction of the operating system to insufficient real memory: swapping occurs when too many tasks are perceived to be competing for limited resources. It is the physical movement of an entire task (eg all real memory pages of an address space may be moved at one time from main storage to auxiliary storage).
System	An integrated composite that consists of one or more of the processes, hardware, software, facilities and people, that provides a capability to satisfy a stated need or objective.

Terminal emulation	Terminal emulation is achieved by software running on an intelligent device, typically a PC or workstation, which allows that device to function as an interactive terminal connected to a host system. Examples of such emulation software includes IBM 3270 BSC or SNA, ICL C03, or Digital VT100.
Terminal I/O	Terminal I/O is a read from, or a write to, an online device such as a VDU or remote printer.
Thrashing	A condition in a virtual storage system where an excessive proportion of CPU time is spent between moving data between main and auxiliary storage.
Tree structures	In data structures, a series of connected nodes without cycles. One node is termed the root and is the starting point of all paths, other nodes termed leaves terminate the paths.
Underpinning contract	A contract with an external supplier covering delivery of services that support the IT directorate in their delivery of services.
Unit costs	Unit costs are costs distributed over individual component usage to establish the unit cost. For example, it can be assumed, that if a box of paper with 1000 sheets costs £10, then obviously one sheet costs 1p. Similarly if a CPU costs £lm a year and it is used to process 1,000 jobs that year, each job costs on average £1,000.
Urgency	Measure of the business criticality of an Incident or Problem based on the impact and on the business needs of the Customer.
User	The person who uses the service on a day-to-day basis.
Utility cost centre (UCC)	A cost centre for the provision of support services to other cost centres.
Variance analysis	A variance is the difference between planned, budgeted, or standard cost and actual cost (or revenues). Variance analysis is an analysis of the factors which have caused the difference between the pre-determined standards and the actual results. Variances can be developed specifically related to the operations carried out in addition to those mentioned above.
Version	An identified instance of a Configuration Item within a product breakdown structure or configuration structure for the purpose of tracking and auditing change history. Also used for software Configuration Items to define a specific identification released in development for drafting, review or modification, test or production.
Version Identifier	A version number; version date; or version date and time stamp.
Virtual memory system	Virtual memory systems were developed to increase the size of memory by adding an auxiliary storage layer which resides on disk.
VSI	VSI (virtual storage interrupt) is an ICL VME term for a page fault.
Warm stand-by	See 'Intermediate Recovery'
Waterline	The lowest level of detail relevant to the customer.
Work-around	Method of avoiding an Incident or Problem, either from a temporary fix or from a technique that means the Customer is not reliant on a particular aspect of the service that is known to have a problem.
Workloads	Workloads in the context of Capacity Management Modelling, are a set of forecasts which detail the estimated resource usage over an agreed planning horizons. Workloads generally represent discrete business applications and can be further sub-divided into types of work (interactive, timesharing, batch)
WORM	WORM or CD-WORM is the term which is frequently used to describe optical read only disks, standing for write once read many.

APPENDIX B PROCESS THEORY AND PRACTICE

B.1 Process theory

This Appendix provides a general introduction to process theory and practice, which is the basis for the ITIL process models. We become aware of 'process' through process models that define workflows and provide guidance on performing it. A process model enables understanding and helps to articulate the distinctive features of a process.

A Process can be defined as

*a connected series of actions, activities, changes etc, performed by agents with the intent of satisfying a purpose or achieving a goal***.**

Process control can similarly be defined as

the process of planning and regulating, with the objective of performing a process in an effective and efficient way.

Processes, once defined, should be under control; once under control, they can be repeated and become manageable. Degrees of control over processes can be defined, and then metrics can be built in to manage the control process.

The output produced by a process has to conform to operational norms that are derived from business objectives. If products conform to the set norm, the process can be considered *effective* (because it can be repeated, measured and managed). If the activities are carried out with a minimum effort, the process can also be considered *efficient*.
Process results metrics should be incorporated in regular management reports.

B.1.1 The product-oriented organisation

Process activities exist in many organisations. However, they are often carried out throughout an organisation, but without any process-oriented co-ordination. This results in problems, which have to be addressed during process implementation. Some examples include:

- processes lacking a clear purpose and focus on business results
- similar processes with inconsistent approaches
- actions or processes performed many times instead of once
- activities that are missing
- no focus on existing business oriented results.

B.1.2 Moving towards a process-oriented organisation

Since processes and their activities run through an organisation, they should be mapped and co-ordinated by process managers. Figure B.1 shows how process activities may be assigned to people in several different organisational units. The simple box diagram indicates the apparent consecutive flow of processes in a linear sequence. Reality is better reflected in the organisational view, where the flow is clearly non-linear and where it is possible to think of delays and interactions that might take place.

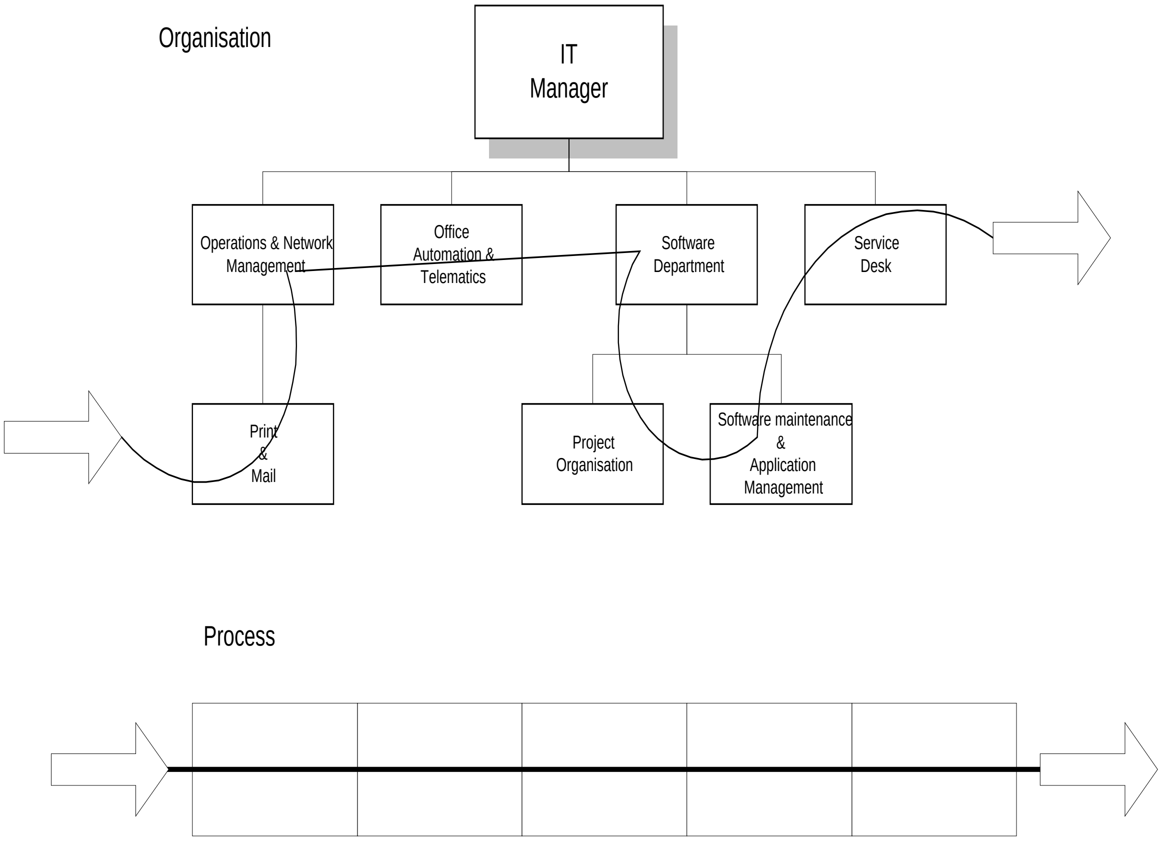


Figure B.1 - Process mapped to organisational unit

In a product-oriented organisation, the flow of activities and processes in Figure B.1 is not generally recognised *at all*; the focus is on the product, and management and control is often lacking. The evidence is in the lack of any useful metrics related to the production process, because the process activities are not clear or even not identified.

B.1.3 The process approach

The model shown in Figure B.2 is a generic process model. Data enters the process, is processed, data comes out, the outcome is measured and reviewed. This very basic description underpins any process description. A process is always organised around a goal. The main output of that process is the result of that goal.

Working with *defined* processes is a novelty for many organisations. By *defining* what your activities are, which inputs are necessary and which outputs will result from the process, you will be able to work in a more efficient and effective manner. Measuring and steering your activities increases this efficacy. Finally, by adding norms to the process, you can add quality measures to your output.

The approach underpins the ‘plan-do-check-act’ cycle of any quality-management system. Plan the purpose of your process in such a way that that the process action can be audited for successful achievement and, if necessary, improved.

The output produced by a process has to conform to operational norms that are derived from business objectives. If the products conform to the set norm, the process can be considered effective. If the activities are also carried out with a minimum effort, the process can also be considered efficient. Process-measurement results should be incorporated in regular management reports.

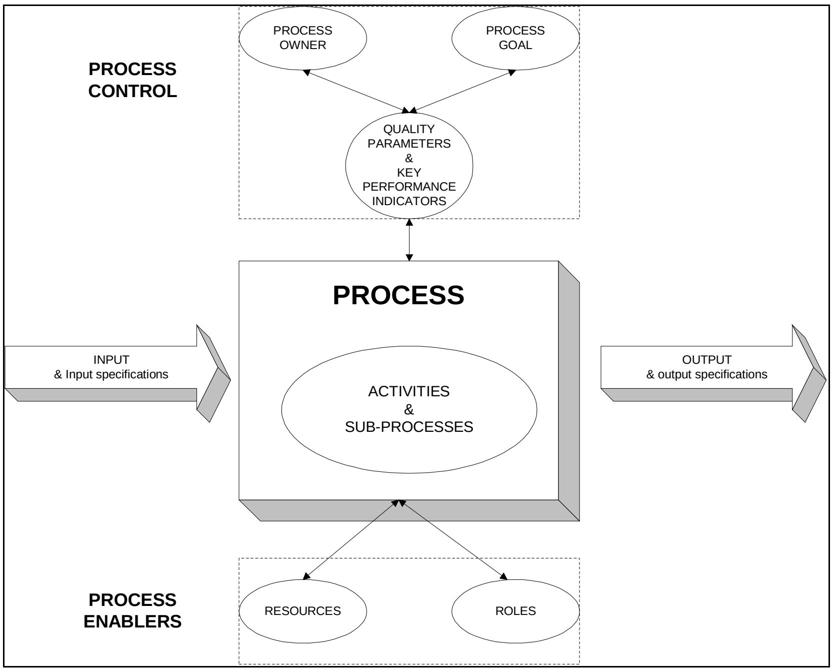


Figure B.2 - The generic process model

'Norms' define certain conditions that the result should meet. By defining norms, you introduce quality aspects to the process. Even before you start, you can think about what the outcome should look like. This enables you to:

consider inputs and activities beforehand because you know what to do
measure and steer effectively because you know what to measure
assess if the result fulfilled your expectations because you know what to expect.

Defining objective norms is a tedious task and also often very complex since objectivity can often be subjective (to slightly misquote Woody Allen).

To discover whether or not your activities are contributing optimally to the business goal of the process, you should measure their effectiveness on a regular basis. Measuring allows you to compare what you have actually done to what you set out to do and to consider the improvement that may be needed.

APPENDIX C. IMPLEMENTING SERVICE MANAGEMENT PROCESSES – ISSUES TO CONSIDER.

When implementing, or improving, any aspects of Service Management processes, there are a wide range of issues to take into consideration. This Appendix looks at some of the issues that need to be taken into account when deciding to implement, or change, Service Management processes. The topics covered also look at planning the project for such an implementation.

C.1 Process implementation

A practical implementation of Service Management should include:

- earning and communicating 'quick wins' to demonstrate the benefits of Service Management
- starting with something simple and adopting a phased approach
- involving Customers, especially those that have been critical of the service
- explaining the differences that will be seen by Customers and Users
- including third-party service suppliers
- managing the changes, and explaining what is being done (and why) to everyone involved or affected - support staff are often cautious about changes, and it is particularly important that they understand the benefits in order to overcome their resistance
- educating staff and managers to become *service* managers.

Vital elements to consider are:

- the extent of the organisation
- the resources at your disposal, including staff numbers
- the level of maturity of staff, of the processes and of the organisation
- the impact of IT on the business
- the culture of the organisation
- continuous communication with the User population.

C.2 Applicability / scalability

The size of an organisation is an important factor when implementing ITIL processes. In a small organisation, many of the roles defined may well be the responsibility of one person.

- Generally it is not a good idea to link planning roles with reactive support roles as the reactive element will always take priority.

- If the person responsible for setting up SLAs also has responsibility for the Incident Management process then there is a danger that the SLA element will not be done.

- The skills required of the person responsible for the Financial Management process are not the same as for IT Service Continuity Management.

- There is a tension between Incident Management and Problem Management, because of their different goals. Incident Management is responsible for minimising the effect of Incidents for Users. Problem Management's task is to find the underlying Problem and is less interested in the continuity of the Users' activities. When combining these two roles, this tension should be acknowledged.

- There is a similar tension between Problem Management and Change Management. When combining these roles, there is the danger of Changes being implemented too quickly by Problem Management. One would do well to separate these functions to ensure that the proper checks and balances exist.

- Configuration Management and Release Management are roles that are quite commonly shared. Both roles have an administrative component and are concerned with maintaining an up-to-date database.

- Configuration Management and Change Management can also be easily shared, because both roles are centered on CMDB information, and no direct conflict of interest exists between the roles.

- Combining the roles of Service Level Management with Financial Management would not present any conflicting interests or issues.

C.2.1 Large and small IT units

In small IT units, one group (or individual) has responsibility for a wide variety of processes. Typically, such a person or group is much more effective in performing one role better than the others. The range of personalities and skills in the group determines which of the processes is done most effectively.

Conversely, a large organisation is able to allocate individual processes to specialist groups composed of people with specialist skills who also have a personality that is a good match for the process. However, over-specialisation has its disadvantages, as specialisation may be perceived as tedious and demotivating if an individual is simply left in place without looking after their needs and aspirations.

C.3 Process implementation projects: a checklist

Most organisations planning to implement ITIL will already have their own 'best practices' in place. ITIL is a best-practice framework, so a totally greenfield situation will hardly ever be found. Some other methods advocate discarding your own best practices; the ITIL view is that those elements that are working for you should not be discarded unless they will not be able to fit within your vision for the future.

IT Infrastructure Library process assessment services are commercially available, and can help you to determine the way in which processes, activities and communication lines are already in place in your organisation. Furthermore, they can help you to determine the maturity level of your process framework. Simple self-assessment could also be provided.

The following checklist is of a general nature, since these topics are discussed in detail in the process-specific chapters. The checklist can function as a guideline for Service Managers controlling the overall implementation of processes.

C.3.1 Procedures

- establish the procedure framework
- implement re-active procedures
- implement pro-active procedures
- implement supporting tools
- establish a managed documentation system
- establish control over procedures used.

C.3.2 Dependencies

- establish a dependency and relationship framework
- describe inter-dependencies with all other processes within the model, both operational and tactical
- establish process interfaces with the IT directorate - these interfaces will be crucial at the outset of process implementations, although the role of the IT directorate relative to tactical matters should diminish over time as tactical processes are put in place
- include vendor relationships
- establish a Customer liaison function on an operational level to organise publicity campaigns.

C.3.3 People

- implement the staff training plan and make this an ongoing activity - focus on both social and technical skills
- assign roles within the ITIL model to people, and make this part of their function description
- delegate tasks and authorisations as low as possible in the organisation.

C.3.4 Timing

- control the project timescale because other stages or projects may depend on it
- consider the timing of ‘going live’, including the timing and communication of the ‘go live’ event, as well as any special considerations for the ‘go live’ day and the period of days or weeks immediately after the ‘going live’.

C.4 Impact on the organisation

An often-asked question in this regard is: “Do I have to change my organisational structure?” The question often crops up because the ITIL process approach means that processes have to be managed over more than one department within traditional hierarchical company structures. Some organisations have tried the ‘matrix’ organisation approach, but whatever structure you choose, there will always be benefits and disadvantages connected to each approach. Consider the following examples:

C.4.1 Hierarchical structure

+ the traditional role model
+ clear lines of communication
+ clear function and task descriptions within each department
– may result in a bureaucracy if you describe procedures in too much detail
– difficult to place process roles in this model
– process approach will require a complex communication structure.

C.4.2 Matrix organisation

+ process oriented structure
+ flexible
+ clear communication model
– no (or less) clear responsibilities
– no (or less) clear leadership roles (informal leadership).

C.4.3 Self learning teams (coaching management)

+ continuous quality improvement from within
+ equality within the different teams
– requires quality awareness
– no control over performance
– possible role confusion.

C.5 Benchmarking

In some circumstances, it may be possible to compare a service with that provided by other organisations. This comparison is only useful to the degree that the compared organisations are either the same or very similar. In the latter case, the differences must be understood and quantified before the comparison can provide useful information. Benchmarking is used to find out if a service is cost-effective, responsive to the Customer’s needs and effective in comparison with outside. Some Customers use benchmarking to decide if they should change their service provider.

A number of organisations provide benchmarking services. These generally fall into four categories:

1. a baseline set at a certain point in time for the same system or department (service targets are a form of benchmark)
2. comparison to industry norms provided by an external organisations
3. direct comparisons with similar organisations
4. comparison with other systems or departments within the same company.

Differences in benchmarks between organisations are normal. All organisations and service-provider infrastructures are unique, and most are continually changing. There are also intangible but influential factors that cannot be measured, such as growth, goodwill, image and culture.

Of the four types of benchmark listed above, the first is usual for Service Management.
The second and third involve comparisons with other organisations. Comparison against industry norms provides a common frame of reference but may be misleading if the comparisons are used without an understanding of the differences that exist across a wide variety of organisations. The differences between organisations may be greater than the similarities, and comparison with a ‘typical’ result may not be useful as a consequence.

Direct comparison with *similar* organisations are most effective if there is a sufficiently large group of organisations with similar characteristics. It is important to understand the size and nature of the business area, including the geographical distribution and the extent to which the service is used for business, or time-critical, activities.

The culture of the Customer population also has an influence. Many support services are influenced by the extent to which Customers will or will not accept restrictions on what they may do with the technology provided. For example, it is difficult to have good security standards with Customers who will not keep their passwords secure, or who load unlicensed or untested software. Finally, comparison with other groups in the same organisation normally allows a detailed examination of the features being compared, so that it can be established whether or not the comparison is of ‘like with like’.

Most benchmarks include some financial measures, such as ‘cost per unit’, and an assessment of cost-effectiveness is a common reason for benchmarking against other organisations. This is particularly so for organisations that have only limited historical information and that are therefore unable to use service or financial trends to measure objectively whether the service is getting better or worse. Financial benchmarking is very difficult. Establishing genuine baselines is nearly impossible, and organisational factors involved in arriving at the cost of similar processes make it hard to make true comparisons.

C.6 A sample implementation strategy

In general, the impact of current weaknesses on IT service quality should determine priorities. For example, if User services are less affected by 'real' errors than those that arise from poor implementation of Changes, Change Management must have priority. While each organisation must therefore set its own priorities, the following phased approach may be used as a starting point:

C.6.1 Phase 1:

- determine the base line
- start with an assessment to determine priorities.

C.6.2 Phase 2:

- survey the services/system(s) currently used by the organisation for providing day-to-day User support and for handling Incidents, Problems and Known Errors
- review the support tools used, and the interfaces to Change Management and Configuration Management, including inventory management, and the operational use of the current system within the IT provider; identify strengths to be retained, and weaknesses to be eliminated
- identify and review the agreements in place between service providers and Customers.

C.6.3 Phase 3:

- determine and document service level requirements

- plan and implement the service desk using tools designed for this function that support Incident control - these tools should either support, or be capable of integration with, tools for Problem Management, Configuration Management and Change Management
- implement at least the inventory elements of Configuration Management that are required for Incident Management and Change Management.

C.6.4 Phase 4:

- extend the Incident control system to allow other domains, such as Computer Operations and Network Control staff, to log Incidents directly
- negotiate and set up SLAs.

C.6.5 Phase 5:

- develop the management reporting system.

C.6.6 Phase 6:

- implement the balance of 'reactive' Problem Management (Problem control, error control and management reporting) and Configuration Management
- realise the proactive parts of Problem Management as staff are released from reactive duties by gradually improving service quality
- establish the Release Management process.

This six-phased approach reduces the development overhead experienced at any given time for the four IT infrastructure management systems under consideration (Incident Management, Problem Management, Change Management and Configuration Management). It should, nevertheless, be noted that, although busy sites will appreciate this smoothing of the development bulge, the approach will increase the overall timescale for implementation.

C.7 Process improvement

Regrettably, even a high standard of Service Management may not be adequate for rapid and major Changes. This can be an issue, for example, when two organisations merge and two sets of Service Management processes, functional groups and support technology have to be rationalised. The most common reason for normal service tuning not being adequate are when one or more of the components of Service Management are missing or deficient, so that the service has degraded and Customers are dissatisfied with its quality or cost.

Under these circumstances, management is faced with a potential or real crisis and should react by initiating a project or series of projects to address the situation. These are required to make much faster improvements to the service, costs or the processes than are possible within the scope of normal ongoing Service Management.

Projects may be directly related to the activities of Service Level Management. For example they may be part of a Service Improvement Programme (SIP) or may be focused on improving Customer and/or staff satisfaction. However, apart from this reactive reason to start a process improvement project, many proactive ones should also be considered, such as:

- providing the operational processes with a tactical planning horizon
- aiming for a higher IT maturity level
- improving alignment with, and support for, business objectives
- reducing costs or improving business profits
- implementing a planning and control system
- increasing quality awareness amongst staff
- making the “feedback loop work”.

In all cases, management needs to have a clear view of the Service Catalogue and the internal service requirements, and also needs to have a so-called ‘Helicopter View’ of the IT organisation.

More guidance can be obtained from the CCTA ITIL volume *Planning and Control for IT Services*, which covers information flows and the development of an appropriate planning-and-control system to meet the requirements of the organisation.

APPENDIX D QUALITY

D.1 Quality Management

Quality management for IT Services is a systematic way of ensuring that all the activities necessary to design, develop and implement IT services which satisfy the requirements of the organisation and of Users take place as planned and that the activities are carried out cost effectively.

The way that an organisation plans to manage its operations so that it delivers quality services, is specified by its quality management system. The quality management system defines the organisational structure, responsibilities, policies, procedures, processes, standards and resources required to deliver quality IT services. However, a quality management system will only function as intended if management and staff are committed to achieving its objectives.

D.1.1 Ongoing quality improvement: The Deming Cycle

> "We have learned to live in a world of mistakes and defective products as if they were necessary to life. It is time to adopt a new philosophy..."
> (W. Edwards Deming, 1900-93)

W. Edwards Deming is best known for his management philosophy establishing quality, productivity, and competitive position. As part of this philosophy he formulated 14 points of attention for managers. Some of these points are more appropriate to Service Management than others, for example:

Excerpts from Deming's fourteen points relevant to Service Management

- break down barriers between departments (improves communications and management)
- management must learn their responsibilities, and take on leadership (process improvement requires commitment from the top; good leaders motivate people to improve themselves and therefore the image of the organisation)
- improve constantly (a central theme for Service Managers is continuous improvement; this is also a theme for quality management. A process-led approach is key to achieving this target)
- institute a programme of education and self-improvement (learning and improving skills has been the focus of Service Management for many years)
- training on the job (linked to continuous improvement)
- transformation is everyone's job (the emphasis being on teamwork and understanding).

For quality improvement Deming proposed the Deming Cycle (or Circle). The four key stages are plan, do, check and act after which a phase of consolidation prevents the 'Circle' from 'rolling down the hill' as illustrated in Figure D.1. The consolidation phase enables the organisation to take stock of what has been taking place and to ensure that improvements are embedded. Often, a series of improvements have been made to processes that require documentation (both to allow processes to be repeatable and to facilitate recognition of the achievement of some form of quality standard).

As discussed in the chapter on process theory, the cycle is underpinned by a process-led approach to management, where defined processes are in place, the activities are measured for compliance with expected values, and outputs are audited to validate and improve the process.

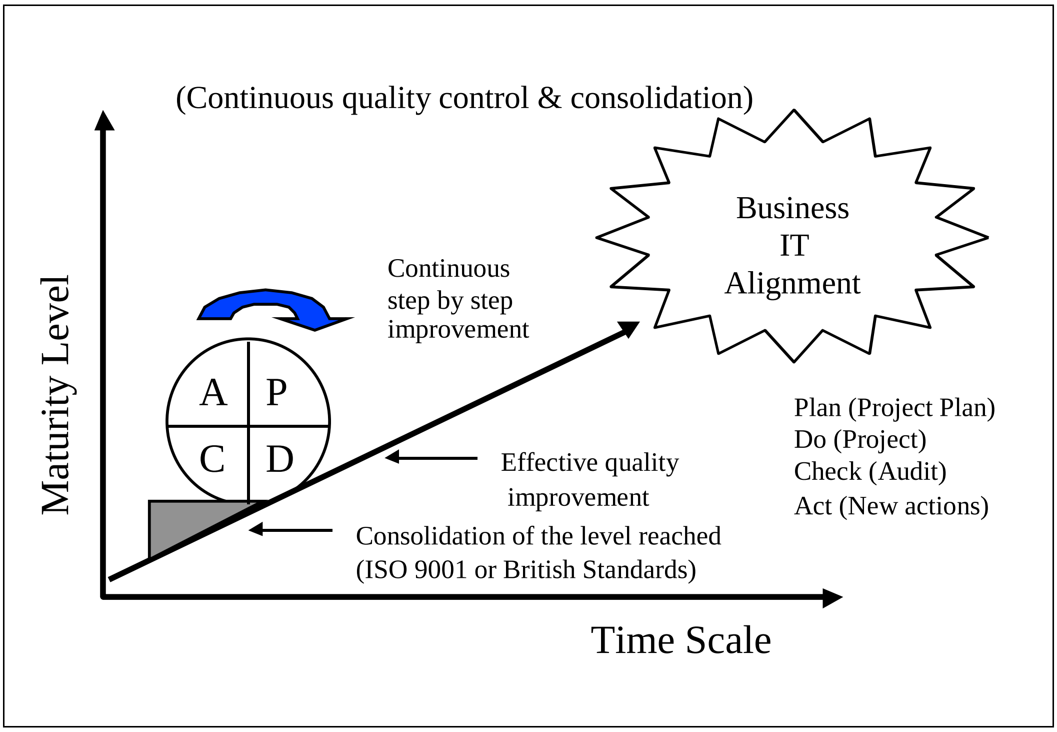


Figure D.1 – The Deming Cycle

D.1.2 Quality standards

<u>International Standards Organisation ISO 9000</u>

An important set of International Standards for Quality Assurance is the ISO 9000 range, a set of five universal standards for a Quality Assurance system that is accepted around the world. At the turn of the millennium, 90 or so countries have adopted ISO 9000 as the cornerstone of their national standards. When you purchase a product or service from a company that is registered to the appropriate ISO 9000 standard, you have important assurances that the quality of what you will receive will be as you expect. For reference purposes, *ISO Standards* can be ordered directly from the International Standards Institute (see Chapter 12, Bibliography, for details).

The most comprehensive of the standards is ISO 9001. It applies to industries involved in the design, development, manufacturing, installation and servicing of products or services. The standards apply uniformly to companies in any industry and of any size.

More detailed guidance of the IT Infrastructure Library (including comprehensive mapping of ISO clauses to IT Infrastructure Library processes) can be obtained from the ITIL book *Quality Management for IT Services* – ISBN 0-11-330555-9. This book is concerned primarily with putting in place in IT services an ISO-9001-conformant quality management system based on Service Management processes; it also touches upon organisational and planning-and-control aspects.

The BSI Code of Practice for IT Service Management is a modern update of the original document PD0005, which was published in 1995. The new code has been written to serve as an introduction to Service Management, and in fact can be used as an introduction to the IT Infrastructure Library. However, the guide is not an official BSI *standard;* indeed, a *standard* for Service Management does not yet exist. The IT Infrastructure Library is in many countries the *de facto* standard, and with the help of BSI and ISO it is hoped that a true international standard based on ITIL will soon be in place. The BSI Code of Practice covers the established Service Support and Service Delivery processes, as well as some additional topics such as implementing the processes.

D.1.3 Total Quality Systems: EFQM

> *"...the battle for Quality is one of the prerequisites for the success of your companies and for our collective success".2*
>
> *(Jacques Delors)*

The EFQM excellence model

The European Foundation for Quality Management (EFQM) was founded in 1988 by the Presidents of 14 major European companies, with the endorsement of the European Commission. The present membership is in excess of 600 very well respected organisations, ranging from major multinationals and important national companies to research institutes in prominent European universities.

EFQM provides an excellent model for those wishing to achieve business excellence in a programme of continuous improvement.

EFQM mission statement

The mission statement is:

> To stimulate and assist organisations throughout Europe to participate in improvement activities leading ultimately to excellence in Customer satisfaction, employee satisfaction, impact on society and business results; and to support the managers of European organisations in accelerating the process of making Total Quality Management a decisive factor for achieving global competitive advantage.

Depiction of the EFQM excellence model

The EFQM Excellence Model consists of 9 criteria and 32 sub-criteria; it is illustrated in Figure D2:

2 Jacques Delors, President of the European Commission, at the signing of the letter of intent in Brussels to establish EFQM on 15 September 1988.

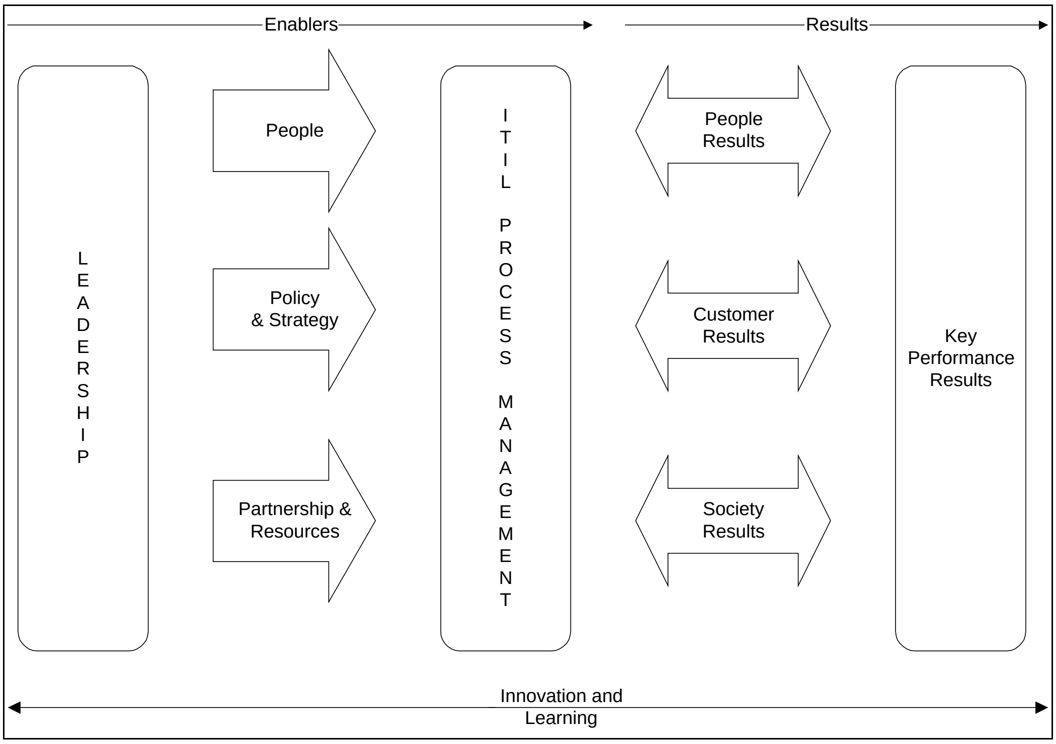


Figure D.2 - The EFQM excellence model

In the model there is explicit focus on the value to Users of the 'Plan, Do, Check, Act' cycle to business operations[3], and the need to relate everything that is done, and the measurements taken, to the goals of business policy and strategy.

<u>Self-assessment and maturity; the EFQM maturity scale</u>

One of the tools provided by EFQM is the self-assessment questionnaire. The self-assessment process allows the organisation to discern clearly its strengths and also any areas where improvements can be made. The questionnaire process and culminates in planned improvement actions, which are then monitored for progress.

In this assessment progress can be checked against a five-point maturity scale:

1. product orientation
2. process orientation (the maturity stage aimed for by the original ITIL)
3. system orientation (the maturity target for ITIL-compliant organisations in the new millennium)
4. chain orientation
5. total quality.

<u>Quality awards</u>

To demonstrate a successful adaptation of the EFQM model, some companies aim for the European Quality Award, a process that allows Europe to recognise its most successful organisations and promote them as role models of excellence for others to copy.

The US equivalent to this award is the Malcolm Baldridge Quality Award for Quality management. The Malcolm Baldrige National Quality Improvement Act of 1987, established an annual US National Quality Award. The purpose of the Award was (and still is) to promote awareness of quality excellence, to recognise quality achievements of US companies, and to publicise successful quality strategies.

3 See the paragraph on W. Edwards Deming.

For the Malcolm Baldridge Award, there are three categories:

- manufacturing companies or subunits
- service companies or subunits
- small businesses.

The criteria against which firms are judged are:

1. leadership
2. strategic planning
3. Customer and market focus
4. information and analysis
5. human resource development and management
6. process management
7. business results.

For the European Quality Award, there are four possible categories:

- companies
- operational units of companies
- public-sector organisations
- small and medium enterprises.

The criteria against which candidate organisations are measured are:

1. leadership
2. people
3. policy and strategy
4. partnerships and resources
5. processes
6. people results
7. customer results
8. society results
9. key performance results.

In the EFQM excellence model the first four criteria are defined as enablers. Best practice in IT Infrastructure Library process implementations show that placing proper emphasis on these topics will increase the chances for success. The key points for the four enablers are listed below.

<u>Leadership</u>

- organise a kick off session involving all of your people
- be a role model
- encourage and support your staff.

<u>People management</u>

- create awareness
- recruit new staff and/or hire temporary staff to prevent service levels being affected during implementation stages
- develop people through training and experience
- align human resource plans with policy and strategy

- adopt a coaching style of management
- align performance with salaries.

Policy and strategy

- communicate mission, vision and values
- align communication plans with the implementation stages.

Partnerships and resources

- establish partnerships with subcontractors and customers
- use financial resources in support of policy and strategy
- utilise existing assets.

Readers interested in either of the quality programmes mentioned above should refer to Chapter 12, Bibliography, for details of contacts.

APPENDIX E EXAMPLE COST-BENEFIT ANALYSIS FOR SERVICE MANAGEMENT PROCESSES

This Appendix is intended as an example of how to quantify the costs and benefits of implementing the processes described in the IT Infrastructure Library. It is not intended to be comprehensive. Please be sure to substitute your own organisation's specific assumptions, purposes, costs, and benefits to get an example that is more suitable to your own circumstances.

In this example, the following assumptions are made:

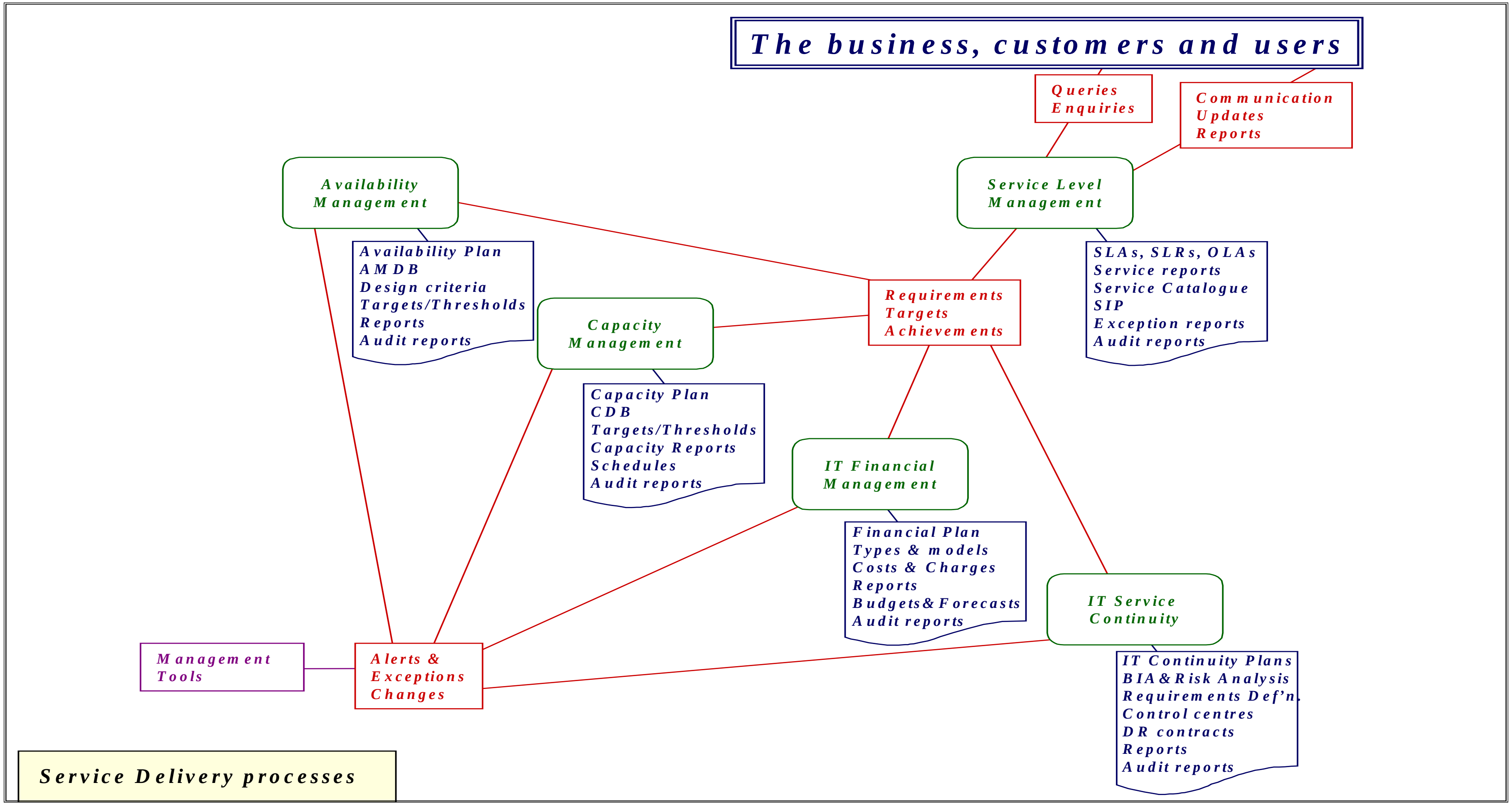
- all employees cost $50, an hour
- your organisation comprises 500 users
- the total number of Incidents is 5,000 per year
- the average time to fix an Incident is 10 minutes
- a working year has 200 days.

Example costs and benefits are set out below.

Process:	**Purpose:**	**Cost/Benefit Examples:**
Configuration Management	Controlling the IT infrastructure Ensuring that only authorised hardware and software is in use	Following the implementation of Configuration Management, the Service Desk has a much greater insight into the relationship between Users, CIs and Incidents. The 3 people assigned to Incident matching can be reduced to 2, resulting in a benefit of 200*8*$50 = $80,000 a year.
Incident Management	Continuity of the service levels Underpin Service Desk function	The implementation of Incident Management has resulted in a decrease in down time per User; this is defined as the amount of time a User is on the phone to the Service Desk or cannot work because of a failure. If the downtime per User has gone down by 1 minute per person per day, this would save the organisation 1/60*500*200*$50 = $83,300 per year.
Problem Management	Minimise disruption of the service level	Suppose that the implementation of Problem Management decreases the amount of recurring Incidents by 500 (10% of total) per year. This means a revenue of 500*10/60*$50 = > $4,000 per year.
Change Management	Efficient handling of Changes	Two Changes are implemented simultaneously, resulting in a major problem. The Customer support system fails, resulting in the loss of 50 Customers with an average purchasing power of $500. This has just cost your company $25,000.

Release Management	Ensuring authorised software modules are used Provide means to build Change Releases Automating release of software	Suppose that a new software module is released containing a bug. The previous version should be reinstalled, but due to poor version management, the wrong version is used, resulting in a system shutdown that lasts for 3 hours and affects two thirds of all employees. This would cost the organisation 2/3*500*$50*3 = $50,000
Service Level Management	Agree on and control the service levels Understand business needs	Thanks to a clear set of agreements, the Service Desk is less troubled with calls that are not part of the services offered. This way the 4 Service Desk employees work 5% more efficiently, resulting in a gain of 4*5%*$500*8*200 = $ 16,000 a year.
Availability Management	Ensure high availability of services	Due to a physical error on a hard disk, a server supporting 100 people crashes. It takes 3 hours to have a new disk delivered and installed before starting up the system again. Costs: 100*3*$50 = $15,000 On a critical system, Availability Management processes would have highlighted the need for a mirror disk, which could automatically take over.
Capacity Management	Ensure the optimal use of IT	There is an overcapacity of 20%. Assuming your IT infrastructure cost you $5 million, you could gain up to $1 ml by implementing Capacity Management and frequently reassessing the necessary capacity.
IT Service Continuity Management	Ensure quick recovery after a disaster	A water pipe breaks, flooding the server room. It takes 2 days to be fully operational. The average User has missed 10 hours of work. Total costs (apart from the pumping): 500*10*$50 = $250,000. Please note that a good contingency plan doesn't come cheap; however the recovery costs (as in this example) could be dramatic – that is, if your organisation is still in business!
Financial Management	Provide insight , control and charge the costs of IT services	Imagine that the costs of IT services are charged to the departments that take them. A 10% reduction in the requests for new services, would directly result in a 10% reduction of IT expenditure. The insight into the real costs in IT services proves to be surprising in practice; most Users don't have a clue about the costs.

APPENDIX F – SERVICE DELIVERY PROCESSES



INDEX

1 INTRODUCTION TO THIS BOOK 5

1.1 The IT Infrastructure Library 5
1.1.1 Public domain framework 5
1.1.2 Best practice framework 5
1.1.3 De facto standard 5
1.1.4 Quality approach 6
1.1.5 itSMF 6

1.2 Restructuring the IT Infrastructure Library 7

1.3 Target audience 7

1.4 Navigating the IT Infrastructure Library 7

1.5 Why choose a jigsaw concept? 9

1.6 The Service Delivery book 9

1.7 Service Management 10

1.8 Customers and Users 11

1.9 A Code of Practice for IT Service Management - PD0005 11

1.10 Service Management: a process approach 11

1.11 Recommended reading 12

2 RELATIONSHIP BETWEEN PROCESSES 14

2.1 Service Level Management 14

2.2 Financial Management for IT Services 14

2.3 Capacity Management 14

2.4 IT Service Continuity Management 14

2.5 Availability Management 15

2.6 Configuration Management 15

2.7 Change Management 16

2.8 Release Management 17

2.9 Incident Management 17

2.10 Problem Management 17

2.11 Service Desk 17

2.12 ICT Infrastructure Management 17

2.13 Application Management 18

2.14 Security Management 18

2.15 Environmental infrastructure processes 18

2.16 Project Management 18

3 GETTING STARTED 19

3.1 Service Management benefits 19

3.2 A process led approach 19

3.3 Management commitment 21
3.3.1 Aspects of management commitment 21
3.3.2 Management commitment in the planning stage 21

3.4 Cultural aspects 22
3.4.1 What is culture? 23
3.4.2 Responsibilities 23
3.4.3 What is meant by 'service culture'? 23
3.4.4 How is this relevant to IT service provision? 24
3.4.5 What do Customers want? 24
3.4.6 Common excuses for conducting 'business as usual' 25
3.4.7 How much will all this cost? 26
3.4.8 What are the potential benefits of Customer care? 26
3.4.9 Service Management training 26

4 BUSINESS COMMUNICATION 28

4.1 Introduction 28

4.2 Why is Communication with the Business so Important 28

4.3 Who is involved? 30

4.4 Benefits and Problems 31
4.4.1 Benefits 31
4.4.2 Problems 32

4.5 Planning for Communication 33
4.5.1 The Communication plan 33
4.5.2 Awareness Campaign 33

4.6 Techniques 34
4.6.1 The importance of Listening 35

4.7 Performance Measurement 36

Annex A Customer survey questionnaires 37
Al Purpose 37
A2 Design 37
A3 Questions 37
A4 Check the draft 38
A5 Produce the questionnaire 38
A6 Covering letter 39
A7 Pilot the questionnaire 39
A8 Useful checklist 39
A9 Example customer satisfaction questionnaire 39

5 SERVICE LEVEL MANAGEMENT 48

5.1 Introduction 48
5.1.1 Why Service Level Management SLM? 48
5.1.2 Mission for SLM 48
5.1.3 Scope for SLM 48
5.1.4 Basic Concept of Service Level Management 48

5.2 The Service Level Management Process 49
5.2.1 Benefits of Service Level Management 50
5.2.2 Costs 51
5.2.3 Possible Problems 52

5.3 Planning the Process 53
5.3.1 Initial Planning Activities 53
5.3.2 Plan Monitoring Capabilities 54
5.3.3 Establish Initial Perception of the Services. 54
5.3.4 Underpinning Contracts and Operational Level Agreements 54

5.4 Implementing the Process 54
5.4.1 Produce a Service Catalogue 54
5.4.2 Expectation Management 55
5.4.3 Plan the SLA Structure 55
5.4.4 Establish Service Level Requirements and Draft SLA 57
5.4.5 Wording of SLAs 58
5.4.6 Seek Agreement 58
5.4.7 Establish Monitoring Capabilities 60
5.4.8 Review Underpinning Contracts and Operational Level Agreements 62
5.4.9 Define Reporting and Review Processes 64
5.4.10 Publicise the existence of SLAs 64

5.5 Executing the Process 64
5.5.1 Monitoring and Reporting 65
5.5.2 Service Review Meetings 65
5.5.3 Service Improvement Programme (SIP) 66
5.5.4 Maintenance of SLAs, Contracts and OLAs 66

5.6 SLA Contents and Key Targets 67
5.6.1 New Services 69

5.7 KPIs and Metrics for SLM Efficiency and Effectiveness 69

Annex A - Example of a simple Service Catalogue 71

Annex B - Example SLAM Chart 72

Annex C - Service Level Manager - Role, Responsibilities & Key Skills 73

6 FINANCIAL MANAGEMENT FOR IT SERVICES 75

6.1 Introduction 75
6.1.1 Why Introduce Formal Financial Management for IT Services ? 75
6.1.2 Basic Concepts of Financial Management for IT Services 75
6.1.3 Scope of IT Financial Management 77
6.1.4 Mission for Financial Management for IT Services 78
6.1.5 Relationship with other IT Service Disciplines 78
6.1.6 Impact on the Enterprise 79
6.1.7 Benefits of Financial Management for IT Services 79
6.1.8 Costs 82
6.1.9 Possible problems 82
6.1.10 Accountancy 83

6.2 Budgeting 84
6.2.1 Introduction 84
6.2.2 Estimating the cost of budget items 85
6.2.3 Estimating the cost of workload dependent budget items 86

6.3 Developing the IT Accounting System 87

6.3.1	Scope of IT Accounting	87
6.3.2	Business Perspective	87
6.3.3	Building the Cost Model	88
6.3.4	Depreciation	94
6.3.5	Apportioning the IT Services Costs	95
6.3.6	Level of detail required in cost calculations	97
6.3.7	Variable Costs or Indirect Costs requiring apportionment	97
6.3.8	Calculating the Cost-by-service	98
6.3.9	Calculating the costs of Cost Units	102
6.3.10	Changes Affecting Costs	103
6.3.11	Investment Appraisal	105
6.3.12	Total Cost of Ownership (TCO)	107
6.3.13	Budgeting, IT Accounting and Charging cycles	108
6.3.14	Charging in Profit Centres	108
6.4	Developing the Charging System	110
6.4.1	Scope of Charging	110
6.4.2	Charging Policies	110
6.4.3	Leading practice	112
6.4.4	Deciding Chargeable Items	113
6.4.5	Variable costs and charges	114
6.4.6	Pricing	114
6.4.7	The internal market	116
6.4.8	Differential Charging	117
6.4.9	Pricing flexibility	117
6.4.10	Billing	118
6.4.11	Conditions for Charging to be effective	119
6.4.12	Case Studies	121
6.5	Planning for IT Accounting & Charging	126
6.5.1	The project	126
6.5.2	Feasibility study	126
6.5.3	The project plan	127
6.5.4	Designing the Cost Model and related processes	128
6.5.5	Performance indicators	129
6.5.6	Dependencies	130
6.5.7	Organisation	130
6.5.8	Tools	131
6.5.9	Training	131
6.6	Implementation	132
6.6.1	Documentation	132
6.6.2	Preparation	133
6.6.3	Awareness	134
6.6.4	Pilot the system	136
6.6.5	Monitor the system	136
6.6.6	Plan IT Accounting and Charging Continuity	136
6.7	Ongoing management and operation	137
6.7.1	Ongoing operation	137
6.7.2	Change management	138
6.7.3	Reasons for changing the IT Accounting & Charging system	139
6.7.4	Dealing with variances	139
6.7.5	Workload variance	140
6.7.6	Cost variance	140
6.7.7	Revenue variance	141
6.7.8	Service Level issues	141

6.7.9 Ongoing planning 142
6.7.10 Management reporting 142
6.7.11 Auditing the systems 143

Annex A - IT Finance Manager - Role, Responsibilities & Key Skills 145

7 CAPACITY MANAGEMENT 147

7.1 Introduction 147
7.1.1 Why Capacity Management? 148
7.1.2 Mission for Capacity Management 149
7.1.3 Scope of Capacity Management 149

7.2 The Capacity Management process 151
7.2.1 Business Capacity Management 154
7.2.2 Service Capacity Management 156
7.2.3 Resource Capacity Management 156

7.3 Activities in Capacity Management 158
7.3.1 Monitoring 159
7.3.2 Analysis 161
7.3.3 Tuning 162
7.3.4 Implementation of tuning activities 163
7.3.5 Storage of Capacity Management Data - the CDB 163
7.3.6 Demand Management 166
7.3.7 Modelling 167
7.3.8 Application sizing 168
7.3.9 Production of the Capacity Plan 169

7.4 Costs, benefits and possible problems 170
7.4.1 Costs 170
7.4.2 Benefits 170
7.4.3 Possible problems 171

7.5 Planning and implementation 173
7.5.1 Review what exists already 173
7.5.2 Planning the process 173
7.5.3 Implementation of the process 176

7.6 Review of the Capacity Management process 177

7.7 Interfaces with other SM processes 179
7.7.1 Incident Management 179
7.7.2 Problem Management 179
7.7.3 Change Management 179
7.7.4 Release Management 179
7.7.5 Configuration Management 180
7.7.6 Service Level Management 180
7.7.7 Financial Management for IT Services 180
7.7.8 IT Service Continuity Management 180
7.7.9 Availability Management 180
7.7.10 Application Design and Development 180

Annex A – Contents of a Capacity Plan 182

Annex B- Capacity Management's role 184

8 IT Service Continuity Management 185

8.1 Introduction 185
8.1.1 Why IT Service Continuity Management? 185
8.1.2 Mission for IT Service Continuity Management 185
8.1.3 Scope for IT Service Continuity Management 185
8.1.4 Basic Concepts of IT Service Continuity Management 185
8.1.5 Benefits of IT Service Continuity Management 186
8.1.6 Business Continuity Management and IT Service Continuity Management 188
8.1.7 Management Commitment 188
8.1.8 Relationship with other IT Service Disciplines 189

8.2 Scope of IT Service Continuity Management 190
8.2.1 Risks 'in scope' 190
8.2.2 Risks 'out of scope' 191
8.2.3 Roles and Responsibilities 192

8.3 The Business Continuity Lifecycle 193
8.3.1 Stage 1 – Initiation 194
8.3.2 Stage 2 – Requirements Analysis and Strategy Definition 195
8.3.3 Stage 3 – Implementation 208
8.3.4 Stage 4 – Operational Management 213
8.3.5 Invocation 214

8.4 Management Structure 217
8.4.1 Management Roles 218
8.4.2 Responsibilities 220

8.5 Generating Awareness 221
8.5.1 Awareness of the Need for ITSCM 222
8.5.2 Initial Awareness 222
8.5.3 Ongoing Awareness 223

8.6 The Future of IT Service Continuity Management 225
8.6.1 The Corporate Challenge 225
8.6.2 Future Developments 226
8.6.3 In Conclusion 227

Annex A 229
ITSCM Manager - Role, Responsibilities & Key Skills 229

Annex B - Example Project Plan 231

Annex C - An Example ITSCM Plan **Error! Bookmark not defined.**

9 AVAILABILITY MANAGEMENT 238

9.1 Introduction 238
9.1.1 Why Availability Management? 238
9.1.2 The need for a business and end user focus 238
9.1.3 Goal and Mission 239
9.1.4 Objectives 239
9.1.5 Scope 240

9.2 Basic Concepts 241
9.2.1 The IT Infrastructure and IT support organisation 241
9.2.2 Guiding Principles 242
9.2.3 Terminology 246

9.3 The Availability Management process 248
9.3.1 Key inputs 248
9.3.2 Key outputs 249

9.3.3 Key activities 249
9.3.4 Relationships with other IT Service Management disciplines 250
9.3.5 Benefits, Costs and Problems 251

9.4 The Cost of (Un)availability 255
9.4.1 Availability 255
9.4.2 Unavailability 258
9.4.3 How much did that failure cost? 258
9.4.4 The benefit of understanding the cost of unavailability 259

9.5 Availability Planning 261
9.5.1 Determining Availability requirements 261
9.5.2 Design activity 263
9.5.3 Designing for Availability 263
9.5.4 Designing for Recovery 270
9.5.5 Security considerations 274
9.5.6 Managing planned downtime 276

9.6 Availability Improvement 280
9.6.1 Important considerations 280
9.6.2 Methods and Techniques 281
9.6.3 The Availability Plan 282

9.7 Availability Measurement & Reporting 285
9.7.1 IT Availability reporting 285
9.7.2 IT Availability measures – the traditional view 286
9.7.3 The disadvantages of traditional IT Availability measures 287
9.7.4 Measuring end user Availability 288
9.7.5 Business driven measurement and reporting 289
9.7.6 What to measure 290
9.7.7 IT Availability Metrics Model 291

9.8 Availability Management Tools 293
9.8.1 IT component downtime data capture and recording 293
9.8.2 Database repositories 293
9.8.3 Report generation & statistical analysis 294
9.8.4 Availability Modelling 295
9.8.5 Systems Management 295

9.9 Availability Management Methods & Techniques 295
9.9.1 Component Failure Impact Analysis (CFIA) 297
9.9.2 Fault Tree Analysis (FTA) 302
9.9.3 The CCTA Risk Analysis And Management Method (CRAMM) 304
9.9.4 Calculating Availability 306
9.9.5 Calculating the cost of Unavailability 310
9.9.6 Developing basic IT Availability measurement & reporting 311
9.9.7 Developing business and end user measurement & reporting 313
9.9.8 Systems Outage Analysis 320
9.9.9 The expanded incident ‘lifecycle’ 326
9.9.10 Continuous improvement 328
9.9.11 Technical Observation Post (TOP) 331

Annex A - Outline Role Description 334

10 SERVICE MANAGEMENT SOFTWARE TOOLS 336

10.1 Types of tools 336

10.2 Summary of tool evaluation criteria 337

10.2.1 Service Management tools 337

10.3 Product training 339

11 PLANNING FOR THE IMPLEMENTATION OF SERVICE MANAGEMENT 340

11.1 A Service Management project 340

11.2 Feasibility study 340

11.3 Assessing the current situation 340
11.3.1 Introduction 340
11.3.2 A ‘health check’ 341

11.4 General guidelines on project planning 342
11.4.1 Project characteristics 342
11.4.2 Business case for the project 343
11.4.3 Critical success factors and possible problems 343
11.4.4 Project costs 343
11.4.5 Organisation 344
11.4.6 Products 344
11.4.7 Planning 345
11.4.8 Communication plan 345

11.5 Project review and management reporting 346
11.5.1 Progress reporting 346
11.5.2 Evaluation of the project 347
11.5.3 Post-project review 347
11.5.4 Auditing for compliance using quality parameters 347
11.5.5 Auditing for improvement using key performance indicators 348
11.5.6 Management reporting 349

12 BIBLIOGRAPHY 350

12.1 References 350

12.2 Other Sources 351

APPENDIX A – LIST OF ACRONYMS AND GLOSSARY 353

A.1 Acronyms 353

A.2 Glossary 355

APPENDIX B PROCESS THEORY AND PRACTICE 365

B.1 Process theory 365
B.1.1 The product-oriented organisation 365
B.1.2 Moving towards a process-oriented organisation 365
B.1.3 The process approach 366

APPENDIX C. IMPLEMENTING SERVICE MANAGEMENT PROCESSES – ISSUES TO CONSIDER. 368

C.1 Process implementation 368

C.2 Applicability / scalability 368
C.2.1 Large and small IT units 369

C.3 Process implementation projects: a checklist 369
C.3.1 Procedures 370
C.3.2 Dependencies 370
C.3.3 People 370
C.3.4 Timing 370

C.4 Impact on the organisation 371
C.4.1 Hierarchical structure 371
C.4.2 Matrix organisation 371
C.4.3 Self learning teams (coaching management) 371

C.5 Benchmarking 371

C.6 A sample implementation strategy 372
C.6.1 Phase 1: 372
C.6.2 Phase 2: 372
C.6.3 Phase 3: 372
C.6.4 Phase 4: 373
C.6.5 Phase 5: 373
C.6.6 Phase 6: 373

C.7 Process improvement 373

APPENDIX D QUALITY 375

D.1 Quality Management 375

D.1.1 Ongoing quality improvement: The Deming Cycle 375

D.1.2 Quality standards 376

D.1.3 Total Quality Systems: EFQM 377

APPENDIX E EXAMPLE COST-BENEFIT ANALYSIS FOR SERVICE MANAGEMENT PROCESSES 381

APPENDIX F – SERVICE DELIVERY PROCESSES 383

INDEX 384