establishing realistic service targets. When targets and timeframes have been confirmed, the SLAs must be signed.

Once the initial SLA has been completed, and any early difficulties overcome, then move on and gradually introduce SLAs for other services/customers. If it is decided from the outset to go for a multi-level structure, it is likely that the corporate-level issues have to be covered for all customers at the time of the initial SLA. It is also worth trialling the corporate issues during this initial phase.

Hints and tips

Don't go for easy targets at the corporate level. They may be easy to achieve, but have no value in improving service quality or credibility. Also, if the targets are set at a sufficiently high level, the corporate SLA can be used as the standard that all new services should reach.

4.3.9.2 Risks

Some of the risks associated with service level management are:

- A lack of accurate input, involvement and commitment from the business and customers
- Lack of appropriate tools and resources required to agree, document, monitor, report and review agreements and service levels
- The process becomes a bureaucratic, administrative process, rather than an active and proactive process delivering measurable benefit to the business
- Access to and support of appropriate and up-to-date CMS and SKMS
- Bypassing the use of the SLM processes
- Business and customer measurements are too difficult to measure and improve, so are not recorded
- Inappropriate business and customer contacts and relationships are developed
- High customer expectations and low perception
- Poor and inappropriate communication is achieved with the business and customers.

4.4 AVAILABILITY MANAGEMENT

Availability is one of the most critical parts of the warranty of a service. If a service does not deliver the levels of availability required, then the business will not experience the value that has been promised. Without availability the utility of the service cannot be accessed. Availability management process activity extends across the service lifecycle.

4.4.1 Purpose and objectives

The purpose of the availability management process is to ensure that the level of availability delivered in all IT services meets the agreed availability needs and/or service level targets in a cost-effective and timely manner. Availability management is concerned with meeting both the current and future availability needs of the business.

Availability management defines, analyses, plans, measures and improves all aspects of the availability of IT services, ensuring that all IT infrastructure, processes, tools, roles etc. are appropriate for the agreed availability service level targets. It provides a point of focus and management for all availability-related issues, relating to both services and resources, ensuring that availability targets in all areas are measured and achieved.

The objectives of availability management are to:

- Produce and maintain an appropriate and up-to-date availability plan that reflects the current and future needs of the business
- Provide advice and guidance to all other areas of the business and IT on all availability-related issues
- Ensure that service availability achievements meet all their agreed targets by managing services and resources-related availability performance
- Assist with the diagnosis and resolution of availability-related incidents and problems
- Assess the impact of all changes on the availability plan and the availability of all services and resources
- Ensure that proactive measures to improve the availability of services are implemented wherever it is cost-justifiable to do so.

Availability management should ensure the agreed 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 look to continually optimize and proactively improve the availability of the IT infrastructure, the services and the supporting organization, in order to provide cost-effective availability improvements that can deliver business and customer benefits.

4.4.2 Scope

The scope of the availability management process covers the design, implementation, measurement, management and improvement of IT service and component availability. Availability management commences as soon as the availability requirements for an IT service are clear enough to be articulated. It is an ongoing process, finishing only when the IT service is decommissioned or retired.

The availability management process includes two key elements:

- **Reactive activities** These involve the monitoring, measuring, analysis and management of all events, incidents and problems involving unavailability. These activities are principally performed as part of the operational roles.
- **Proactive activities** These involve the proactive planning, design and improvement of availability. These activities are principally performed as part of the design and planning roles.

These activities are described in detail in section 4.4.5.

Availability management needs to understand the service and component availability requirements from the business perspective in terms of the:

- Current business processes, their operation and requirements
- Future business plans and requirements
- Service targets and the current IT service operation and delivery
- IT infrastructure, data, applications and environment and their performance
- Business impacts and priorities in relation to the services and their usage.

Understanding all of this will enable availability management to ensure that all the services and components are designed and delivered to meet their targets in terms of agreed business needs. The availability management process:

- Should be applied to all operational services and technology, particularly those covered by SLAs. It can also be applied to those IT services deemed to be business-critical, regardless of whether formal SLAs exist
- Should be applied to all new IT services and for existing services where SLRs or SLAs have been established
- Should be applied to all supporting services and the partners and suppliers (both internal and external) that form the IT support organization as a precursor to the creation of formal agreements
- Should consider all aspects of the IT services and components and supporting organizations that may impact availability, including training, skills, process effectiveness, procedures and tools.

The availability management process should include:

- Monitoring of all aspects of availability, reliability and maintainability of IT services and the supporting components, with appropriate events, alarms and escalation, with automated scripts for recovery
- Maintaining a set of methods, techniques and calculations for all availability measurements, metrics and reporting
- Actively participating in risk assessment and management activities
- Collecting measurements and the analysis and production of regular and ad hoc reports on service and component availability
- Understanding the agreed current and future demands of the business for IT services and their availability
- Influencing the design of services and components to align with business availability needs
- Producing an availability plan that enables the service provider to continue to provide and improve services in line with availability targets defined in SLAs, and to plan and forecast future availability levels required, as defined in SLRs
- Maintaining a schedule of tests for all resilience and fail-over components and mechanisms
- Assisting with the identification and resolution of any incidents and problems associated with service or component unavailability

- Proactively improving service or component availability wherever it is cost-justifiable and meets the needs of the business.

The availability management process does not include business continuity management (BCM) and the resumption of business processing after a major disaster. The support of BCM is included within ITSCM. However, availability management does provide key inputs to ITSCM, and the two processes have a close relationship, particularly in the assessment and management of risks and in the implementation of risk reduction and resilience measures.

4.4.3 Value to the business

The availability management process ensures that the availability of systems and services matches the evolving agreed needs of the business. The role of IT within the business is now pivotal. The availability and reliability of IT services can directly influence customer satisfaction and the reputation of the business. This is why availability management is essential in ensuring IT delivers the levels of service availability required by the business to satisfy its business objectives and deliver the quality of service demanded by its customers.

In today's competitive marketplace, customer satisfaction with service(s) provided is paramount. Customer loyalty can no longer be relied on, and dissatisfaction with the availability and reliability of IT service can be a key factor in customers taking their business to a competitor. Availability can also improve the ability of the business to follow an environmentally responsible strategy by using green technologies and techniques in availability management.

4.4.4 Policies, principles and basic concepts

The availability management process is continually trying to ensure that all operational services meet their agreed availability targets, and that new or changed services are designed appropriately to meet their intended targets, without compromising the performance of existing services.

4.4.4.1 Policies

As a matter of policy, the availability management process, just like capacity management, must be involved in all stages of the service lifecycle, from strategy and design, through transition and operation to improvement. The appropriate availability and resilience should be designed into services and components from the initial design stages. This will ensure not only that the availability of any new or changed service meets its expected targets, but also that all existing services and components continue to meet all of their targets. This is the basis of stable service provision.

The service provider organization should establish policies defining when and how availability management must be engaged throughout each lifecycle stage. Policies should also be established regarding the criteria to be used to define availability and unavailability of a service or component and how each will be measured.

4.4.4.2 Guiding principles of availability

An effective availability management process, consisting of both the reactive and proactive activities, can 'make a big difference' and will be recognized as such by the business, if the deployment of availability management within an IT organization has a strong emphasis on the needs of the business and customers. To reinforce this emphasis, there are several guiding principles that should underpin the availability management process and its focus:

- Service availability is at the core of customer satisfaction and business success: there is a direct correlation in most organizations between service availability and customer and user satisfaction, where poor service performance is defined as being unavailable.
- Recognizing that when services fail, it is still possible to achieve business, customer and user satisfaction and recognition: the way a service provider reacts in a failure situation has a major influence on customer and user perception and expectation.
- Improving availability can only begin after understanding how the IT services support the operation of the business.
- Service availability is only as good as the weakest link in the chain: it can be greatly increased by the elimination of single points of failure or an unreliable or weak component.
- Availability is not just a reactive process. The more proactive the process, the better service availability will be. Availability should not purely react to service and component failure. The

more often events and failures are predicted, pre-empted and prevented, the higher the level of service availability.

- It is cheaper to design the right level of service availability into a service from the start, rather than try and 'bolt it on' subsequently. Adding resilience into a service or component is invariably more expensive than designing it in from the start. Also, once a service gets a bad name for unreliability, it becomes very difficult to change the image. Resilience is also a key consideration of ITSCM, and this should be considered at the same time.

Availability management is completed at two interconnected levels:

- **Service availability** This involves all aspects of service availability and unavailability and the impact of component availability, or the potential impact of component unavailability on service availability.
- **Component availability** This involves all aspects of component availability and unavailability.

4.4.4.3 Aspects of availability

Availability management relies on the monitoring, measurement, analysis and reporting of the following aspects.

Availability

Availability is the ability of a service, component or CI to perform its agreed function when required. It is often measured and reported as a percentage. Note that downtime should only be included in the following calculation when it occurs within the agreed service time (AST). However, total down time should also be recorded and reported.

$$\text{Availability (\%)} = \frac{\text{Agreed service time (AST)} - \text{downtime}}{\text{AST}} \times 100$$

Reliability

Reliability is a measure of how long a service, component or CI can perform its agreed function without interruption. The reliability of the service can be improved by increasing the reliability of individual components or by increasing the resilience of the service to individual component failure (i.e. increasing the component redundancy, for example by using load-balancing techniques). It is often measured and reported as the mean time between service incidents (MTBSI) or mean time between failures (MTBF):

$$\text{Reliability (MTBSI in hours)} = \frac{\text{Available time in hours}}{\text{Number of breaks}}$$

$$\text{Reliability (MTBF in hours)} = \frac{\text{Available time in hours} - \text{Total downtime in hours}}{\text{Number of breaks}}$$

Maintainability

Maintainability is a measure of how quickly and effectively a service, component or CI can be restored to normal working after a failure. It is measured and reported as the mean time to restore service (MTRS) and should be calculated using the following formula:

$$\text{Maintainability (MTRS in hours)} = \frac{\text{Total downtime in hours}}{\text{Number of service breaks}}$$

MTRS should be used to avoid the ambiguity of the more common industry term mean time to repair (MTTR), which in some definitions includes only repair time, but in others includes recovery time. The downtime in MTRS covers all the contributory factors that make the service, component or CI unavailable:

Example: measuring availability, reliability and maintainability

A situation where a 24 × 7 service has been running for a period of 5,020 hours with only two breaks, one of six hours and one of 14 hours, would give the following figures:

Availability	= (5,020–(6+14))/5,020 × 100	= 99.60%
Reliability (MTBSI)	= 5,020/2	= 2,510 hours
Reliability (MTBF)	= 5,020–(6+14)/2	= 2,500 hours
Maintainability (MTRS)	= (6+14)/2	= 10 hours

- Time to record
- Time to respond
- Time to resolve
- Time to physically repair or replace
- Time to recover.

Serviceability

Serviceability is the ability of a third-party supplier to meet the terms of its contract. This contract will include agreed levels of availability, reliability and/or maintainability for a supporting service or component. These aspects and their interrelationships are illustrated in Figure 4.8.

Although the principal service target contained within SLAs for customers and the business is availability, as illustrated in Figure 4.8, some customers also require reliability and maintainability targets to be included in the SLA as well. Where these are included they should relate to end-to-end service reliability and maintainability, whereas the reliability and maintainability targets contained in OLAs and contracts relate to component and supporting service targets and can often include availability targets relating to the relevant components or supporting services.

Vital business function

The term vital business function (VBF) is used to reflect the part of a business process that is critical to the success of the business. An IT service may support a number of business functions that are less critical. For example, an automated teller machine (ATM) or cash dispenser service VBF would be the dispensing of cash. However, the ability to obtain a statement from an ATM may not be considered as vital. This distinction is important and should influence availability design and associated costs. The more vital the business function generally, the greater the level of resilience and availability that needs to be incorporated into the design required in the supporting IT services. For all services, whether VBFs or not, the availability requirements should be determined by the business and not by IT. The initial availability targets are often set at too high a level, and this leads to either over-priced services or an iterative discussion between the service provider and the business to agree an appropriate compromise between the service availability and the cost of the service and its supporting technology.

Certain VBFs may need special designs, which are now being used as a matter of course within service design plans, incorporating:

- **High availability** A characteristic of the IT service that minimizes or masks the effects of IT component failure to the users of a service.
- **Fault tolerance** The ability of an IT service, component or CI to continue to operate correctly after failure of a component part.
- **Continuous operation** An approach or design to eliminate planned downtime of an IT service. Note that individual components or CIs may be down even though the IT service remains available.
- **Continuous availability** An approach or design to achieve 100% availability. A continuously available IT service has no planned or unplanned downtime.

Industry view

Many suppliers commit to high availability or continuous availability solutions only if stringent environmental standards and resilient processes are used. They often agree to such contracts only after a site survey has been completed and additional, sometimes costly, improvements have been made.

4.4.4.4 Role of measurement

The availability management process depends heavily on the measurement of service and component achievements with regard to availability.

Key messages

'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 or control it, then don't measure it.'

What to measure and how to report it inevitably depends on which activity is being supported, who the recipients are and how the information is to be utilized. It is important to recognize the differing perspectives of availability to ensure measurement and reporting satisfies these varied needs:

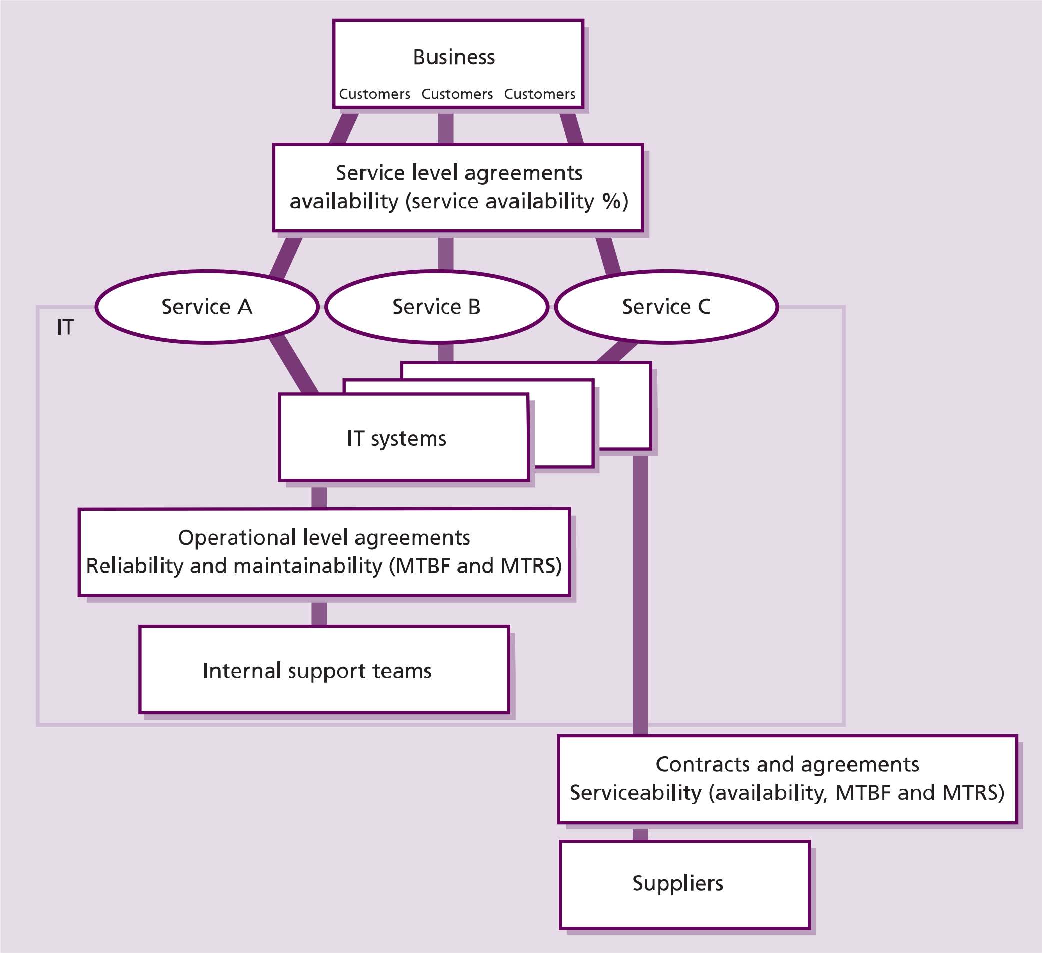

Business
Customers
Customers
Customers
Service level agreements
availability (service availability %)
Service A
Service B
Service C
IT
IT systems
Operational level agreements
Reliability and maintainability (MTBF and MTRS)
Internal support teams
Contracts and agreements
Serviceability (availability, MTBF and MTRS)
Suppliers

Figure 4.8 Availability terms and measurements

- The business perspective considers IT service availability in terms of its contribution or impact on the VBFs that drive the business operation.
- The user perspective considers 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 user also considers IT service availability in terms of response times. For many performance-centric applications, poor response times are considered equal in impact to failures of technology.
- The IT service provider perspective considers IT service and component availability with regard to availability, reliability and maintainability.

In order to satisfy the differing perspectives of availability, availability management needs 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 ultimately to deliver benefit to the IT and business organizations. This is influenced strongly by the combination of 'what you measure' and 'how you report it'.

4.4.5 Process activities, methods and techniques

Availability management should perform the reactive and proactive activities illustrated in Figure 4.9.

4.4.5.1 Reactive activities

The reactive aspect of availability management involves work to ensure that current operational services and components deliver the agreed levels of availability and to respond appropriately when they do not. The reactive activities include:

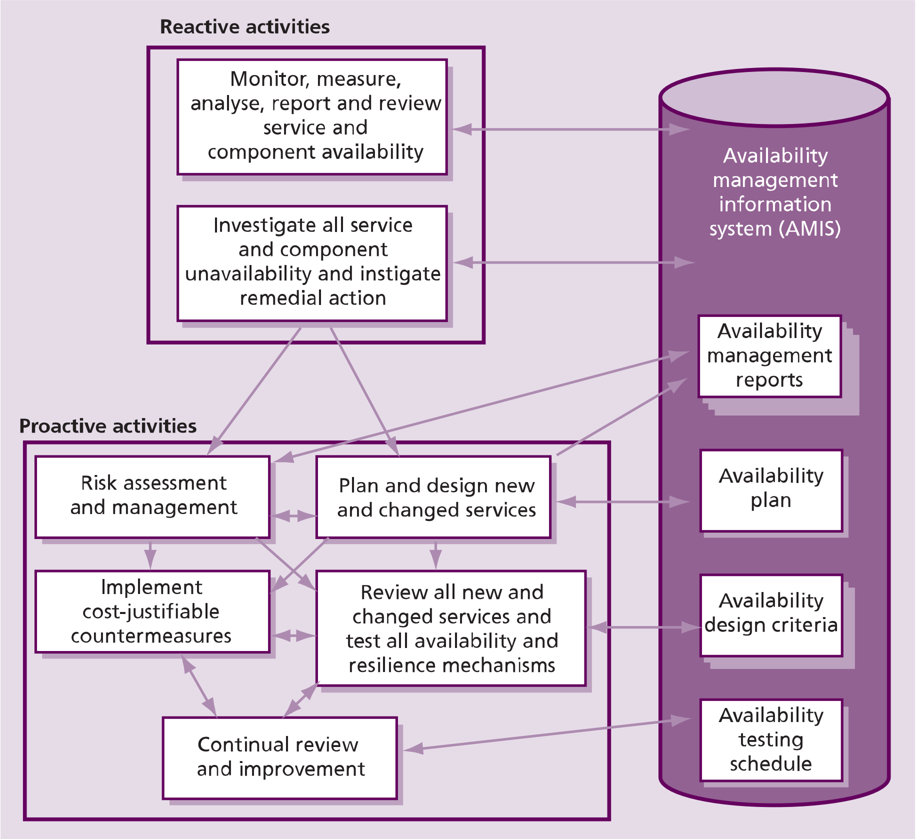

Reactive activities
Monitor, measure, analyse, report and review service and component availability
Investigate all service and component unavailability and instigate remedial action
Proactive activities
Risk assessment and management
Plan and design new and changed services
Implement cost-justifiable countermeasures
Review all new and changed services and test all availability and resilience mechanisms
Continual review and improvement
Availability management information system (AMIS)
Availability management reports
Availability plan
Availability design criteria
Availability testing schedule

Figure 4.9 The availability management process

- Monitoring, measuring, analysing, reporting and reviewing service and component availability
- Investigating all service and component unavailability and instigating remedial action. This includes looking at events, incidents and problems involving unavailability.

These activities are primarily conducted within the service operation stage of the service lifecycle and are linked into the monitoring and control activities and incident management processes (see *ITIL Service Operation*).

4.4.5.2 Proactive activities

The proactive activities of availability management involve the work necessary to ensure that new or changed services can and will deliver the agreed levels of availability and that appropriate measurements are in place to support this work. They include producing recommendations, plans and documents on design guidelines and criteria for new and changed services, and the continual improvement of service and reduction of risk in existing services wherever it can be cost-justified. These are key aspects to be considered within service design activities.

Proactive activities include:

- Planning and designing new or changed services:
  - Determining the VBFs, in conjunction with the business and ITSCM
  - Determining the availability requirements from the business for a new or enhanced IT service and formulating the availability and recovery design criteria for the supporting IT components
  - 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
  - Performing risk assessment and management activities to ensure the prevention and/or recovery from service and component unavailability

  - Designing the IT services to meet the availability and recovery design criteria and associated agreed service levels
  - Establishing measures and reporting of availability, reliability and maintainability that reflect the business, user and IT support organization perspectives
- Risk assessment and management:
  - Determining the impact arising from IT service and component failure in conjunction with ITSCM and, where appropriate, reviewing the availability design criteria to provide additional resilience to prevent or minimize impact to the business
- Implementing cost-justifiable counter-measures, including risk reduction and recovery mechanisms
- Reviewing all new and changed services and testing all availability and resilience mechanisms
- Continual reviewing and improvement:
  - Producing and maintaining an availability plan that prioritizes and plans IT availability improvements.

4.4.5.3 Reactive availability management

Monitoring, measuring, analysing and reporting service and component availability

A key output from the availability management process is the measurement and reporting of IT availability. Availability measures should be incorporated into SLAs, OLAs and any underpinning contracts. These should be reviewed regularly at service level review meetings. Measurement and reporting provide the basis for:

- Monitoring the actual availability delivered versus agreed targets
- Establishing measures of availability and agreeing availability targets with the business
- Identifying unacceptable levels of availability that impact the business and users
- Reviewing availability with the IT support organization
- Continual improvement activities to optimize availability.

IT service provider organizations 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 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:

- **Per cent available** The truly 'traditional' measure that represents availability as a percentage and, as such, much more useful as a component availability measure than a service availability measure. It is typically used to track and report achievement against a service level target. It tends to emphasize 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 complacent behaviour within the IT support organization.
- **Per cent unavailable** The inverse of the 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 is more likely to create awareness of the shortfall in delivering the level of availability required.
- **Duration** This is 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 is two hours, but in one week the actual downtime was four hours, this would represent a trend leading to an additional four 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** This is used to record the number of interruptions to the IT service. It helps provide a good indication of reliability from a 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.
- **Impact of failure** This is the true measure of service unavailability. It depends on mature incident recording where the inability of users to perform their business tasks is the most important piece of information captured. All other measures suffer from a potential to mask the real effects of service failure and are often converted to a financial impact.

A business may have, for many years, accepted that the IT availability that it experiences is represented in terms of component availability rather than overall service or business availability. However, this is no longer being viewed as acceptable and businesses are keen to better represent availability in measure(s) that demonstrate the positive and negative consequences of IT availability on their business and users.

Key messages

The most important availability measurements are those that reflect and measure availability from the business and user perspective.

Availability management needs to consider availability from both a business/IT service provider perspective and from an IT component perspective. These are entirely different aspects, and while the underlying concept is similar, the measurement, focus and impact are entirely different.

The sole purpose of producing these availability measurements and reports, including those from the business perspective, is to improve the quality and availability of IT service provided to the business and users. All measures, reports and activities should reflect this purpose.

Availability, when measured and reported to reflect the experience of the user, provides a more representative view on overall IT service quality. The user view of availability is influenced by three factors:

- Frequency of downtime
- Duration of downtime
- Scope of impact.

Measurements and reporting of user availability should therefore embrace these factors. The methodology employed to reflect user availability could consider two approaches:

- **Impact by 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 user productivity, or to calculate the availability percentage from a user perspective, and can also include the costs of recovery for lost productivity (e.g. increased overtime payments).
- **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 user impact correlates to a VBF – for example, if the user takes customer purchase orders and a VBF is customer sales. This single measure is the basis to reflect impact to the business operation and user.

The method employed should be influenced by the nature of the business operation. A business operation supporting data entry activity is well suited to reporting that reflects user productivity loss. Business operations that are more customer-facing, for example ATM services, benefit from reporting transaction impact. It should also be noted that not all business impact is user-related. With increasing automation and electronic processing, the ability to process automated transactions or meet market cut-off times can also have a large financial impact that may be greater than the ability of users to work.

The IT support organization needs to have a keen awareness of the user experience of availability. However, the real benefits come from aggregating the user view into the overall business view. A guiding principle of the availability management process is that: ‘Improving availability can only begin when the way technology supports the business is understood’. Therefore availability management is not just about understanding the availability of each IT component, but is also about understanding the impact of component failure on service and user availability. 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 therefore relies on all components on which the service depends being available, i.e. systems, key components, network, data and applications.

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 on the positive and negative impacts on the VBFs on which the business operation is dependent. This approach ensures that SLAs and IT availability reporting are based on measures that are understood by both the business and IT. By measuring the VBFs that rely on IT services, measurement and reporting becomes business-driven, with the impact of failure reflecting the consequences to the business. It is also important that the availability of the services is defined and agreed with the business and reflected within SLAs. This definition of availability should include:

- What is the minimum available level of functionality of the service?
- At what level of service response is the service considered unavailable?
- Where will this level of functionality and response be measured?
- What are the relative weightings for partial service unavailability?
- If one location or office is impacted, is the whole service considered unavailable, or is this considered to be 'partial unavailability'? This needs to be agreed with the customers.

Hints and tips

When defining 'available' and 'unavailable' for a particular service it is not only important that the agreed levels are measurable, but also that any acceptable 'partial availability' is clearly defined. If it is possible for the scope of service unavailability to be limited, the IT service provider and customer may wish to identify the most likely of these situations and agree to terms for associated service levels. For example, if a vital business function remains available even though less critical features are down, this might be defined as 'partial unavailability' and the agreed service restoration targets might be more lenient. Or if the primary work location remains available, and only satellite offices are affected, this might be considered to be 'partial unavailability'. The key is to be sure that the customer's true business needs are reflected and that the balance of cost and availability are appropriately managed.

Reporting and analysis tools are required for the manipulation of data stored in the various databases utilized by availability management. These tools can either be platform- or PC-based and are often a combination of the two. 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 – for example, monthly availability reports, availability plan and service failure analysis (SFA) status reports. These reporting activities can require much manual effort and the only solution is to automate as much of the report generation activity as possible. For reporting purposes, organizational reporting standards should be used wherever possible. If these do not exist, IT standards should be developed so that IT reports can be developed using standard tools and techniques. This means that the integration and consolidation of reports will subsequently be much easier to achieve.

Investigating all service and component unavailability and instigating remedial action

When a service or component becomes unavailable based on the agreed terms, the situation must be investigated and the appropriate corrective action undertaken.

Unavailability analysis

All events and incidents causing unavailability of services and components should be investigated, with remedial actions being implemented within either the availability plan or the overall SIP. Trends should be produced from this analysis to direct and focus activities such as SFA to those areas causing the most impact or disruption to the business and the users.

The overall costs of an IT service are influenced by the levels of availability required and the investments required in technology and services provided by the IT support organization 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, and therefore unavailability is not free either. For highly critical business processes and VBFs, 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 is the cost of the availability solution weighed against the costs of unavailability.

Before any SLR is accepted, and ultimately the SLR or SLA negotiated and agreed between the business and the IT organization, it is essential that the availability requirements of the business are analysed to assess if/how the IT service can deliver the required levels of availability. This applies not only to new IT services that are being introduced, but also to any requested changes to the availability requirements of existing IT services.

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 VBFs that support the business operation, this can provide an obvious indication of 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 organization. This can be the stimulus for identifying improvement opportunities and can become a key metric in monitoring the availability of IT services.

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. Where significant financial investment decisions are required, it is better to express the cost of failure arising from service, 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 organization, i.e. the business and IT organization.

Tangible costs can include:

- Lost user productivity
- Lost IT staff productivity
- Lost revenue
- Overtime payments
- Wasted goods and material
- Litigation, imposed fines or penalty payments.

These costs are often well understood by the finance area of the business and IT organization, and in relative terms are easier to obtain and aggregate than the intangible costs associated with an IT failure.

Intangible costs can include:

- Loss of customers
- Loss of customer goodwill (customer dissatisfaction)
- Loss of business opportunity (to sell, gain new customers or revenue etc.)
- Damage to business reputation
- Loss of confidence in IT service provider
- Damage to staff morale.

It is important not simply to dismiss the intangible costs (and the potential consequences) on the grounds that they may be difficult to measure. The overall unavailability of service, the total tangible cost and the total intangible costs arising from service unavailability are all key metrics in the measurement of the effectiveness of the availability management process.

The expanded incident lifecycle

A guiding principle of availability management is to recognize that it is still possible to 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 minimized to enable normal business operations to resume as quickly as possible. An aim of availability management is to ensure the duration and impact from incidents impacting IT services are minimized, to enable business operations to resume as quickly as possible. The analysis of the 'expanded incident lifecycle' enables the total IT service downtime for any given incident to be broken down and mapped against the major stages through which all incidents progress (the lifecycle). Availability management should work closely with incident management and problem management in the analysis of all incidents causing unavailability.

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'. Figure 4.10 illustrates the expanded incident lifecycle.

From the above it can be seen that an incident can be broken down into individual stages within

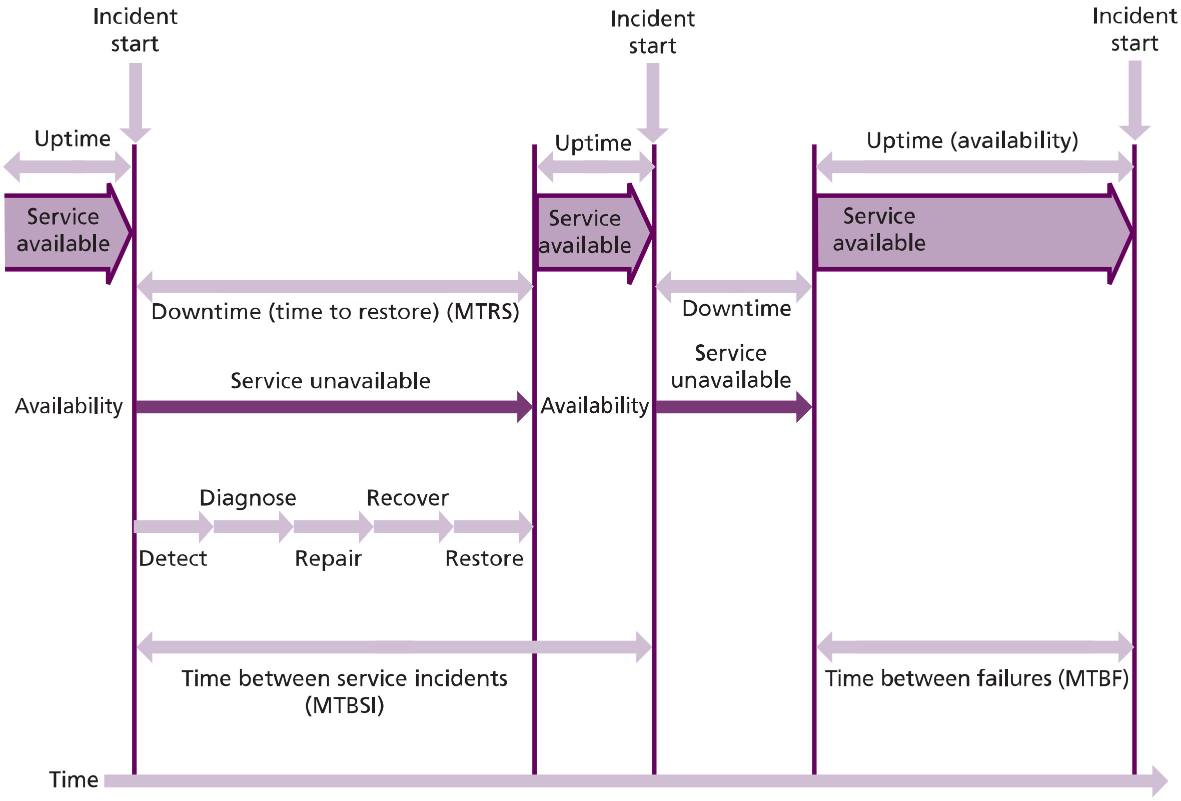


Figure 4.10 The expanded incident lifecycle

a lifecycle that can be timed and measured. This lifecycle view provides an important framework in determining, among others, systems management requirements for event and 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 individual stages of the lifecycle are considered in more detail as follows:

1 **Incident detection** This is the time at which the IT service provider organization is made aware of an incident. Systems management tools positively influence the ability to detect events and incidents and therefore to improve levels of availability that can be delivered. Implementation and exploitation should have a strong focus on achieving high availability and enhanced recovery objectives. In the context of recovery, such tools should be exploited to provide automated failure detection, assist failure diagnosis and support automated error recovery, with scripted responses. Tools are very important in reducing all stages of the incident lifecycle, but principally the detection of events and incidents. Ideally the event is automatically detected and resolved, before the users have noticed it or have been impacted in any way.

2 **Incident diagnosis** This is the time at which diagnosis to determine the underlying cause has been completed. When IT components fail, it is important that the required level of diagnostics is captured, to enable problem determination to identify the root cause and resolve the issue. The use and capability of diagnostic tools and skills is critical to the speedy resolution of service issues. 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 failures. Where the time required to take diagnostics is considered excessive, or varies from the target, 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.

3 **Incident repair** This is the time at which the failure has been repaired/fixed. Repair times for incidents should be continuously monitored and compared against the targets agreed within OLAs, underpinning contracts and other agreements. This is particularly important with respect to externally provided services and supplier performance. Wherever breaches are observed, techniques should be used to reduce or remove the breaches from similar incidents in the future.

4 **Incident recovery** This is the time at which component recovery has been completed. The backup and recovery requirements for the components underpinning a new IT service should be identified as early as possible within the design cycle. These requirements should cover hardware, software and data and recovery targets. The outcome from this activity should be a documented set of recovery requirements that enables the development of appropriate recovery plans. To anticipate and prepare for performing recovery such that reinstatement of service is effective and efficient requires the development and testing of appropriate recovery plans based on the documented recovery requirements. Wherever possible, the operational activities within the recovery plan should be automated. The testing of the recovery plans also delivers approximate timings for recovery. These recovery metrics can be used to support the communication of estimated recovery of service and validate or enhance the component failure impact analysis documentation. 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, and exploitation of new and faster recovery tools and techniques. Availability requirements should also contribute to determining what spare parts are kept within the Definitive Spares to facilitate quick and effective repairs, as described within *ITIL Service Transition*.

5 **Incident restoration** This is the time at which normal business service is resumed. 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 is completed and before any technical staff involved in the incident are stood down. In the majority of cases, this is simply a case of getting confirmation from the affected users. However, the users for some services may be customers 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 service provider organization to verify that a restored IT service is now working as expected. These could simply be visual checks of transaction throughput or user simulation scripts that validate the end-to-end service.

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. For example, the service was unavailable to the business for 60 minutes, yet it only took five minutes to apply a fix – where did the other 55 minutes go?

Using this approach identifies possible areas of inefficiency that combine to make the loss of service experienced by the business greater than it need be. These could cover areas such as poor automation (alerts, automated recovery etc.), poor diagnostic tools and scripts, unclear escalation procedures (which delay the escalation to the appropriate technical support group or supplier), or lack of comprehensive operational documentation. Availability management needs 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 availability incidents. This provides availability management with metrics for both specific incidents and trending information. This information can be used as input to SFA assignments, SIP activities and regular availability management reporting and to provide an impetus for continual improvement activity to pursue cost-effective improvements. It can also enable targets to be set for specific stages of the incident lifecycle. While accepting that each incident may have a wide range of technical complexity, the targets can be used to reflect the consistency of how the IT service provider organization responds to incidents.

An output from the availability management process is 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 requires 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 and can provide an invaluable role in reducing the amount of downtime incurred. Availability management requirements 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. This makes it possible to identify where 'time is being lost' and provides the basis for the identification of factors that can improve recovery and restoration times. These activities are performed on a regular basis within service operation.

Service failure analysis

SFA is a technique designed to provide a structured approach to identifying the underlying causes of service interruptions to the user. SFA utilizes a range of data sources to assess where and why shortfalls in availability are occurring. SFA enables a holistic view to be taken to drive not just technology improvements, but also improvements to the IT support organization, processes, procedures and tools. SFA is run as an assignment or project, and may utilize other availability management methods and techniques to formulate the recommendations for improvement. The detailed analysis of service interruptions can identify opportunities to enhance levels of availability. SFA is a structured technique to identify improvement opportunities in end-to-end service availability that can deliver benefits to the user. Many of the activities involved in SFA are closely aligned with those of problem management, and in a number of organizations these activities are performed jointly by problem and availability management.

The high-level objectives of SFA are to:

- Improve the overall availability of IT services by producing a set of improvements for implementation or input to the availability plan
- Identify the underlying causes of service interruption to users
- Assess the effectiveness of the IT support organization and key processes
- Produce reports detailing the major findings and recommendations
- Ensure that availability improvements derived from SFA-driven activities are measured.

SFA initiatives should use input from all areas and all processes including, most importantly, the business and users. Each SFA assignment should have a recognized sponsor(s) (ideally, joint sponsorship from the IT and business) and involve resources from many technical and process areas. The use of the SFA approach:

- Provides the ability to deliver enhanced levels of availability without major cost
- Provides the business with visible commitment from the IT support organization
- Develops in-house skills and competencies to avoid expensive consultancy assignments related to availability improvement
- Encourages cross-functional team work and breaks barriers between teams, and is an enabler to lateral thinking, challenging traditional thoughts and providing innovative, and often inexpensive, solutions
- Provides a programme of improvement opportunities that can make a real difference to service quality and user perception
- Provides opportunities that are focused on delivering benefit to the user
- Provides an independent 'health check' of IT service management processes and is the stimulus for process improvements.

To maximize both the time of individuals allocated to the SFA assignment and the quality of the delivered report, a structured approach is required. This structure is illustrated in Figure 4.11. This approach is similar to many consultancy models utilized within the industry, and in many ways availability management can be considered as providing via SFA a form of internal consultancy.

The high-level structure in Figure 4.11 is described briefly as follows.

1 **Select opportunity** Prior to scheduling an SFA assignment, there needs to be agreement as to which IT service or technology is to be selected. It is recommended that an agreed number of assignments are scheduled per year within the availability plan and, if possible, the IT services

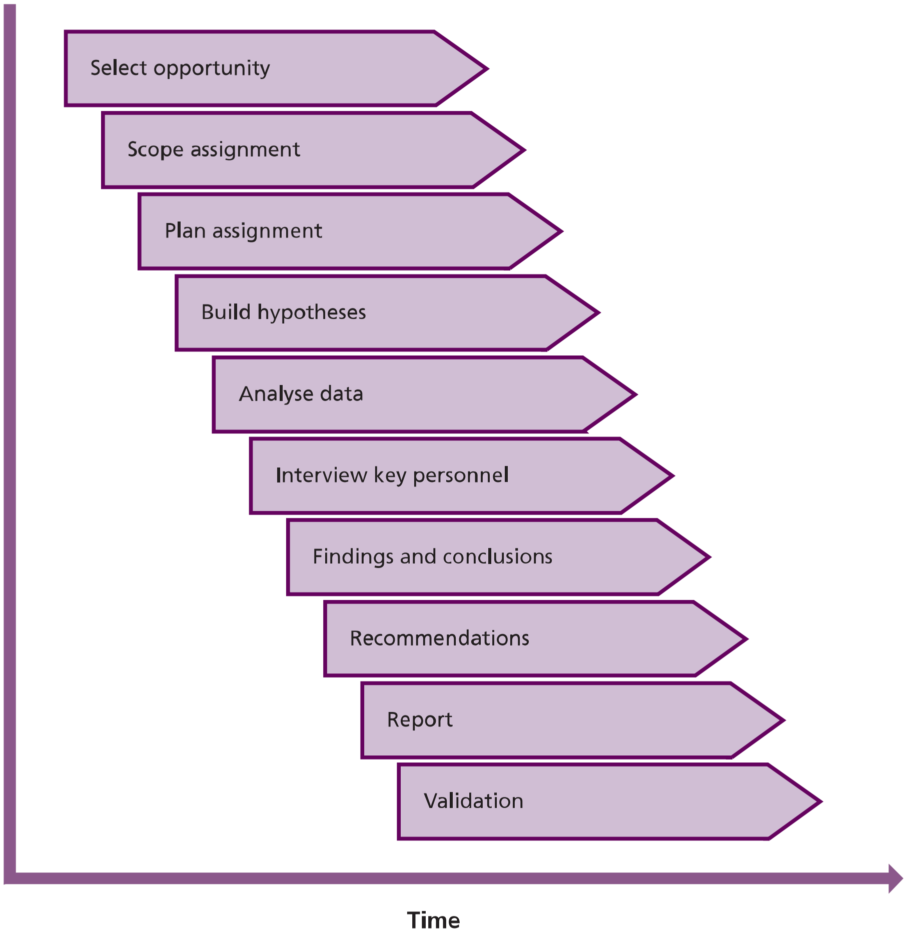


Figure 4.11 The structured approach to SFA

are selected in advance as part of the proactive approach to availability management. Before commencing with the SFA, it is important that the assignment has a recognized sponsor from within the IT organization and/or the business and that they are involved and regularly updated with progress of the SFA activity. This ensures organizational visibility to the SFA and ensures recommendations are endorsed at a senior level within the organization.

2 **Scope assignment** This is to state explicitly what areas are and are not covered within the assignment. This is normally documented in terms of reference issued prior to the assignment.

3 **Plan assignment** The SFA assignment needs to be planned a number of weeks in advance of the assignment commencing, with an agreed project plan and a committed set of resources. The project should look at identifying improvement opportunities that benefit the user. It is therefore important that an end-to-end view of the data and management information system (MIS) requirements is taken. The data and documentation should be collected from all areas and analysed from the user and business perspective. A 'virtual' SFA team should be formed from all relevant areas to ensure that all aspects and perspectives are considered. The size of the team should reflect the scope and complexity of the SFA assignment.

4 **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 scenarios can be built from discussing the forthcoming assignment with key roles (for example, senior management and users) or by using the planning session to brainstorm the list from the assembled team. The completed hypotheses list should be documented and input to the analysis period to provide some early focus on the data and MIS that match the individual

scenarios. It should be noted that this approach also eliminates perceived issues, i.e. no data or MIS substantiates what is perceived to be a service issue.

5 **Analyse data** The number of individuals who form the SFA team dictates how to allocate specific analysis responsibilities. During this analysis period the hypotheses list should be used to help draw some early conclusions.

6 **Interview key personnel** It is essential that key business representatives and users are interviewed to ensure the business and user perspectives are captured. It is surprising how this dialogue can identify quick win opportunities, as often what the business views as a big issue can be addressed by a simple IT solution. Therefore these interviews should be initiated as soon as possible within the SFA assignment. The study team should also seek input from key individuals within the IT service provider organization to identify additional problem areas and possible solutions that 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.

7 **Findings and conclusions** After analysis of the data and MIS provided, interviews and continual revision of the hypotheses 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 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 SFA team can back up all findings with clear documented evidence.

8 **Recommendations** After all findings and conclusions have been validated, the SFA team should be in a position to formulate recommendations. In many cases, the recommendations to support a particular finding are straightforward and obvious. However, the benefit of bringing a cross-functional team together for the SFA assignment is to create an environment for innovative lateral-thinking approaches. The SFA assignment leader should facilitate this session with the aim of identifying recommendations that are practical and sustainable once implemented.

9 **Report** The final report should be issued to the sponsor with a management summary. Reporting styles are normally determined by the individual organizations. It is important that the report clearly shows where loss of availability is being incurred and how the recommendations 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 enables informed choices to be made on how to take the recommendations forward and how these should be prioritized and resourced.

10 **Validation** It is recommended that for each SFA, key measures that reflect the business and user perspectives prior to the assignment are captured and recorded as the 'before' view. As SFA recommendations are progressed, the positive impacts on availability should be captured to provide the 'after' view for comparative purposes. Where anticipated benefits have not been delivered, this should be investigated and remedial action taken. Having invested time and effort in completing the SFA 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 the overall SIP. The success of the SFA assignment as a whole should be monitored and measured to ensure its continued effectiveness.

4.4.5.4 Proactive availability management

The capability of the availability management process is positively influenced by the range and quality of proactive techniques utilized by the process. This activity and those that follow include descriptions of the proactive techniques of the availability management process. All of the proactive activities described interact with each other and the boundaries between them are rarely sharp. Organizations should seek to ensure that all proactive activities consider both the component and the end-to-end service perspectives.

Hints and tips

Consider categorizing the recommendations under the following headings:

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.

Reduction: Recommendations that, if implemented, will reduce or minimize the user impact from IT service interruption – for example, recovery and/or restoration can be enhanced to reduce impact duration.

Avoidance: Recommendations that, if implemented, will eliminate this particular cause of IT service interruption.

Planning and designing new or changed services

The availability management process ensures that new or changed services are designed appropriately to meet the customer's availability-related requirements, defined in service level targets. The design must be developed not only to ensure that the new or changed service will meet its availability specifications, but also to ensure that performance of existing services is not negatively impacted. The work involves producing recommendations, plans and documents on design guidelines and criteria for new and changed services. The availability requirements of the business must be clearly defined and understood so that appropriate availability and recovery design criteria can be developed.

Requirements definition

Where new IT services are being developed, it is essential that availability management takes an early and participative design role in determining the availability requirements. This enables availability management to influence positively 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. There needs to be a dialogue between IT and the business to determine the balance between the business perception of the cost of unavailability and the exponential cost of delivering higher levels of availability.

It is important that the level of availability designed into the service is appropriate to the business needs, the criticality of the business processes being supported and the available budget. The business should be consulted early in the service design lifecycle so that the business availability needs of a new or enhanced IT service can be costed and agreed. This is particularly important where stringent availability requirements may require additional investment in service management processes, IT service and system 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 user requirements into quantifiable availability targets and conditions. This is an important input into the IT service design and provides the basis for assessing the capability of the IT design and IT support organization in meeting the availability requirements of the business.

The business requirements for IT availability should contain at least:

- A definition of the VBFs supported by the IT service
- A definition of IT service downtime, i.e. the conditions under which the business considers the IT service to be unavailable
- The business impact caused by loss of service, together with the associated risk
- Quantitative availability requirements, i.e. the extent to which the business tolerates 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
- The service backup and recovery capability.

Once the IT technology design and IT support organization are determined, the service provider organization is then in a position to confirm if the availability requirements can be met. Where shortfalls are identified, dialogue with the business is required to present the cost options that exist to enhance the proposed design to meet the availability requirements. This enables the business to reassess if lower or higher levels of availability are required, and to understand the appropriate impact and costs associated with its decision.

Determining the availability requirements is likely to be an iterative process, particularly where there is a need to balance the business availability requirement against the associated costs. The necessary steps are:

- Determining the business impact caused by loss of service
- From the business requirements, specifying the availability, reliability and maintainability requirements for the IT service and components supported by the IT support organization
- For IT services and components provided externally, identifying the serviceability requirements
- Estimating the costs involved in meeting the availability, reliability, maintainability and serviceability requirements
- Determining, with the business, if the costs identified in meeting the availability requirements are justified
- Determining, from the business, the costs likely to be incurred from loss or degradation of service
- Where these are seen as cost-justified, defining the availability, reliability, maintainability and serviceability requirements in agreements and negotiating them into contracts.

Identifying vital business functions

VBF refers to the part of a business process that is critical to the success of the business. An IT service may also support less critical business functions and processes, and it is important that the VBFs are recognized and documented to provide the appropriate business alignment and focus. It is the business which determines and validates the VBFs.

Designing for availability

The level of availability required by the business influences 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 are incurred in providing the appropriate service management processes, systems management tools and high-availability solutions required to meet the more stringent availability requirements. The greatest level of availability should be included in

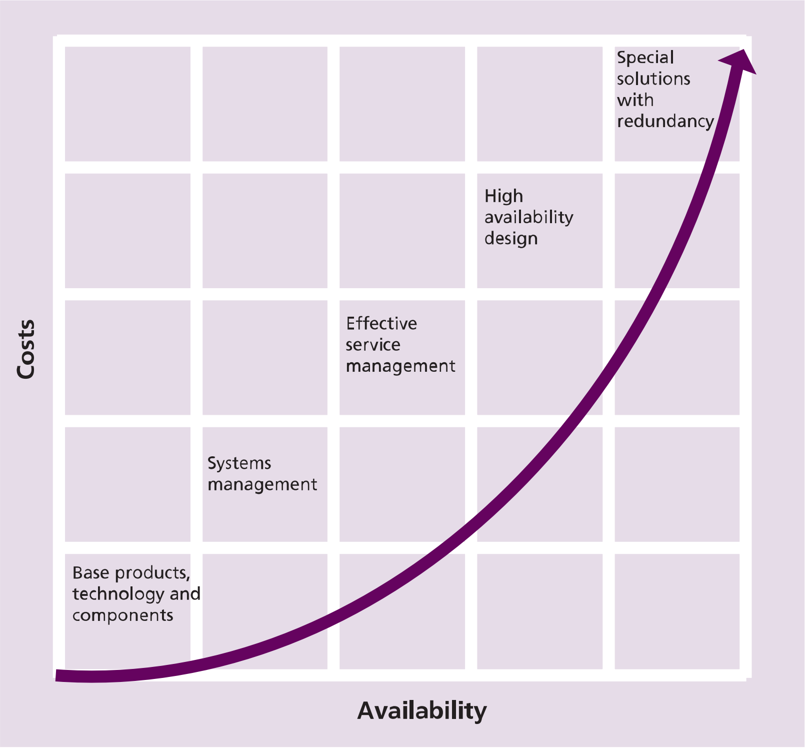

Special solutions with redundancy
High availability design
Effective service management
Systems management
Base products, technology and components
Costs
Availability

Figure 4.12 Relationship between levels of availability and overall costs

the design of those services supporting the most critical of the VBFs.

When considering how the availability requirements of the business are to 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-justifiable to the business. Figure 4.12 indicates the products and processes required to provide varying levels of availability and the cost implications.

Base products and components

The procurement or development of the base products, technology and components should be based on their capability to meet stringent availability and reliability requirements. These should be considered as 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.

Systems management

Systems management should provide monitoring, diagnostic and automated error recovery to enable fast detection and speedy resolution of potential and actual IT failure.

Service management processes

Effective service management processes contribute to higher levels of availability. Processes such as availability management, incident management, problem management, change management, service asset and configuration management play a crucial role in the overall management of the IT service.

High-availability design

The design for high availability needs to consider the elimination of SPOFs and/or the provision of alternative components to provide minimal disruption to the business operation should an IT component failure occur. The design also needs to eliminate or minimize 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 should 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% requires expensive solutions that incorporate full mirroring or 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 to the IT environments, i.e. data centres, power supplies, air conditioning and telecommunications.

As illustrated in Figure 4.12, there is a significant increase in costs when the business requirement is higher than the optimum level of availability that the IT infrastructure can deliver. These increased costs are driven by major redesign of the technology and the changing of requirements for the IT support organization.

Hints and 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 within the SLA.

Service availability design

The SLM process is normally responsible for communicating with the business on how its availability requirements for IT services are to be met and negotiating the SLR/SLA for the IT service design stage. Availability management therefore provides important support and input to both SLM and design processes during this period. While higher levels of availability can often be provided by investment in tools and technology, 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 is always a balance between cost and quality. This is where availability management can play a key role in optimizing availability of the IT service design to meet increasing availability demands while deferring an increase in costs.

Designing service 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 the task from two related, but distinct, perspectives:

- **Designing for availability** This activity relates to the technical design of the IT service and the alignment of the internal and external suppliers required to meet the availability requirements of the business. It needs to cover all aspects of technology, including infrastructure, environment, data and applications.
- **Designing for recovery** This activity relates to the design points required to ensure that in the event of an IT service failure, the service and its supporting components can be reinstated to enable normal business operations to resume as quickly as possible. This again needs to cover all aspects of technology.

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 consistently to recover in a timely and effective manner. Design for recovery should take into consideration the business needs identified through the BIA, such as specific recovery time objectives and recovery point objectives. All aspects of availability should be considered in the service design activity and should consider all stages within the service lifecycle.

The contribution of availability management within the design activities is to provide:

- The specification of the availability requirements for all components of the service
- The requirements for availability measurement points (instrumentation)
- The requirements for new/enhanced systems and service management
- Assistance with the IT infrastructure design
- The specification of the reliability, maintainability and serviceability requirements for components supplied by internal support teams and external suppliers
- Validation of the final design to meet the minimum levels of availability required by the business for the IT service.

If the availability requirements cannot be met, the next task is to re-evaluate the service design 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 design. For example:

- The exploitation of fault-tolerant technology to mask the impact of planned or unplanned component downtime
- Duplexing, or the provision of alternative IT infrastructure components to allow one component to take over the work of another component
- Improving component reliability by enhancing testing regimes
- Improved software design and development
- Improved processes and procedures
- Systems management enhancements/ exploitation
- Improved externally supplied services, contracts or agreements
- Developing the capability of people with more training.

Hints and tips

Consider documenting the availability design requirements and considerations for new IT services and making them available to the design and implementation functions. Longer term seek to mandate these requirements and integrate within the appropriate governance mechanisms that cover the introduction of new IT services.

Part of the activity of designing for availability must ensure that all business, data and information security requirements are incorporated within the service design. The overall aim of IT security is 'balanced security in depth', with justifiable controls implemented to ensure that the information security policy is enforced and that continued IT services within secure parameters (i.e. confidentiality, integrity and availability) continue to operate. 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 stage for the supporting technology. For many organizations, the approach taken to IT security is covered by an information security policy owned and maintained by information security management. In the execution of the security policy, availability management plays an important role in its operation for new IT services.

Where the business operation has a high dependency on IT service availability, and the cost of failure or loss of business reputation is considered to be unacceptable, 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 agreed levels of availability begins with the design, procurement and/or development of good-quality products and components. However, in isolation, these are unlikely to deliver the sustained levels of availability required. Achieving a consistent and sustained level of availability will require investment in and deployment of effective service management processes, systems management tools, high-availability design and ultimately special solutions with full mirroring or redundancy.

Designing for availability is a key activity, driven by availability management, which 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, 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. Clearly good availability design is about avoiding failures and delivering, where possible, a fault-tolerant IT infrastructure. However, with this focus is too much reliance placed on technology, and has as much emphasis been placed on the fault-tolerance aspects of the IT infrastructure? The reality is that failures will occur. The way the IT organization manages failure situations can have a positive effect on the perception of the business, customers and users of the IT services.

Key message

Every failure is an important 'moment of truth' – 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 can ensure that every failure is an opportunity to maintain and even enhance business and user satisfaction. To provide an effective 'design for recovery' and to enable an effective recovery from IT failure, it is important to recognize that both the business and the IT organization have needs that must be satisfied. The service provider has informational needs that will help it manage the impact of failure on its business and set expectations within the business, user community and its business customers. The needs include the skills, knowledge, processes, procedures and tools required to enable the technical recovery to be completed in an optimal time.

A key aim is to prevent minor incidents from becoming major incidents 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. The instigation of these activities is the responsibility of the incident management process and is typically initiated by 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 need to have defined procedures for assessing and managing all incidents and to execute them effectively.

Key message

Ensuring the right people are involved in a recovery early enough is not the responsibility 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 optimize IT recovery may be the stimulus for subsequent continual improvement activities to the incident management process and the service desk.

In order to remain effective, the maintainability of IT services and components should be monitored, and their impact on the 'expanded incident lifecycle' understood, managed and improved.

Component failure impact analysis

Component failure impact analysis (CFIA) can be used to predict and evaluate the impact on IT services arising from component failures within the technology. The output from a CFIA can be used to identify where additional resilience should be considered to prevent or minimize the impact of component failure to the business operation and users. This is particularly important during the service design stage, where it is necessary to predict and evaluate the impact on IT service availability arising from component failures within the proposed IT service design. However, the technique can also be applied to existing services and infrastructure.

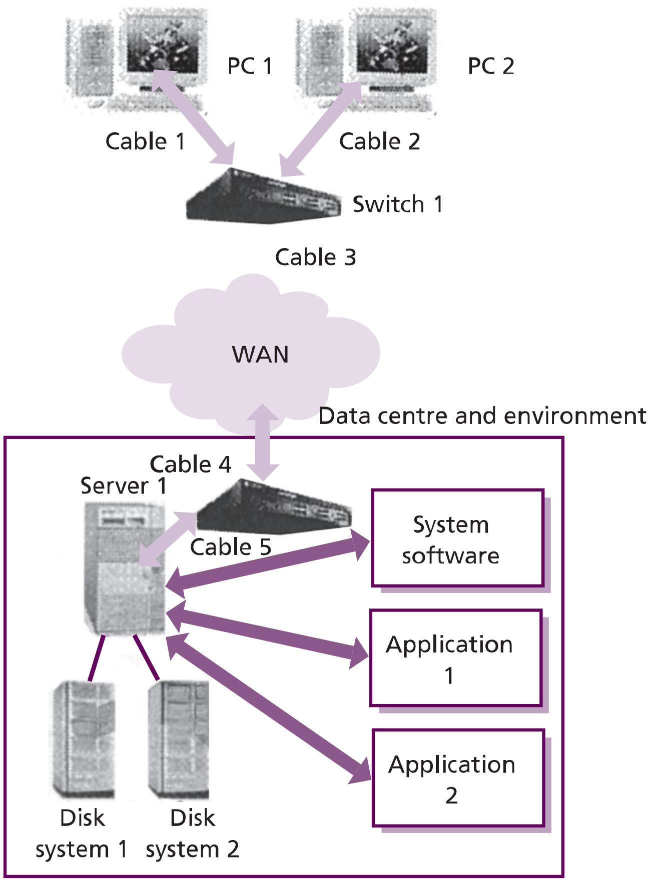
CFIA is a relatively simple technique that can be used to provide this information. IBM devised CFIA in the early 1970s, with its origins based on hardware design and configuration. However, it is recommended that CFIA be used in a much wider context to reflect the full scope of the IT infrastructure and applications, i.e. hardware, network, software, applications, data centres and support staff. Additionally, the technique can also be applied to identify impact and dependencies on IT support organization skills and competencies among staff supporting the new IT service. This activity is often completed in conjunction with ITSCM and possibly capacity management.

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 minimize the impact of failure to the business operation and users. CFIA achieves this by providing and indicating:

- SPOFs that can impact availability
- The impact of component failure on the business operation and users
- Component and people dependencies
- Component recovery timings
- The need to identify and document recovery options
- The need to identify and implement risk reduction measures.

The above can also provide 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.

Having determined the IT infrastructure configuration to be assessed, the first step is to create a grid with CIs on one axis and the IT



CI	Service 1	Service 2
PC 1	M	M
PC 2	M	M
Cable 1	M	M
Cable 2	M	M
Switch 1	X	X
Cable 3	X	X
WAN	X	X
Cable 4	X	X
Switch 2	X	X
Cable 5	X	X
Data centre	X	X
Server 1	X	X
Disk 1	A	A
Disk 2	A	A
System S/W	X	X
Application 1	X	
Application 2		X

Figure 4.13 Component failure impact analysis

services that have a dependency on the CI on the other, as illustrated in Figure 4.13. This information should be available from the CMS, or alternatively it can be built using documented configuration charts and SLAs.

The next step is to perform the CFIA and populate the grid as follows:

- Leave a blank when a failure of the CI does not impact the service in any way
- Insert an 'X' when the failure of the CI causes the IT service to be inoperative
- Insert an 'A' when there is an alternative CI to provide the service
- Insert an 'M' when there is an alternative CI, but the service requires manual intervention to be recovered.

Having built the grid, CIs that have a large number of Xs are critical to many services and can result in high impact should they fail. Equally, IT services having high counts of Xs are complex and vulnerable to failure. This basic approach to CFIA can provide valuable information in quickly identifying SPOFs, IT services at risk from CI failure and what alternatives are available should CIs fail. It should also be used to assess the existence and validity of recovery procedures for the selected CIs. 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 CIs against the VBFs and users supported by each component, thus understanding the impact of a component failure on the business and user. The approach can also be further refined and developed to include and develop 'component availability weighting' factors that can be used to assess and calculate the overall effect of the component failure on the total service availability.

To undertake an advanced CFIA requires the CFIA matrix to be expanded to provide additional fields required for the more detailed analysis. This could include fields such as:

- **Component availability weighting** A weighting factor appropriate to the impact of failure of the component on the total service availability. For example, if the failure of a switch can cause 2,000 users to lose service out of a total service user base of 10,000, then the weighting factor should be 0.2, or 20%.
- **Probability of failure** This can be based on the reliability of the component as measured by the MTBF information if available or on the current trends. This can be expressed as a low/medium/high indicator or as a numeric representation.
- **Recovery time** This is the estimated recovery time to recover the CI. 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 CI.
- **Device independence** Where software CIs have duplex files to provide resilience, this is to ensure that file placements have been verified as being on separate hardware disk configurations. This also applies to power supplies – it should be verified that alternative power supplies are connected correctly.
- **Dependency** This is to show any dependencies between CIs. If one CI failed, there could be an impact on other CIs – for example, if the security CI failed, the operating system might prevent backup processing.

Single point of failure analysis

A SPOF is any configuration item that can cause an incident when it fails, and for which a countermeasure has not been implemented. A single point of failure may be a person or a step in a process or activity, as well as a component of the IT infrastructure. It is important that no unrecognized SPOFs exist within the IT infrastructure design or the actual technology, and that they are avoided wherever possible.

The use of SPOF analysis or CFIA as techniques to identify SPOFs is recommended. SPOF and CFIA analysis exercises should be conducted on a regular basis, and wherever SPOFs are identified, CFIA can be used to identify the potential business, customer or user impact and help determine what alternatives can or should be considered to cater for this weakness in the design or the actual infrastructure. Countermeasures should then be implemented wherever they are cost-justifiable. The impact and disruption caused by the potential failure of the SPOF should be used to cost-justify its implementation.

Fault tree analysis

Fault tree analysis (FTA) is a technique that can be used to determine a chain of events that has caused an incident or may cause an incident in the future. 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 technology component design options. Using FTA:

- Information can be provided that 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.

FTA makes a representation of a chain of events using Boolean notation. Figure 4.14 gives an example of a fault tree.

Essentially FTA distinguishes the following events:

- **Basic events** These are terminal points for the fault tree – for example, power failure, operator error. Basic events are not investigated in great depth. If basic events are investigated in further depth, they automatically become resulting events.
- **Resulting events** These are intermediate nodes in the fault tree, resulting from a combination of events. The highest point in the fault tree is usually a failure of the IT service.
- **Conditional events** These are events that only occur under certain conditions – for example, failure of the air-conditioning equipment only affects the IT service if equipment temperature exceeds the serviceable values.
- **Trigger events** These are events that trigger other events – for example, power failure detection equipment can trigger automatic shutdown of IT services.

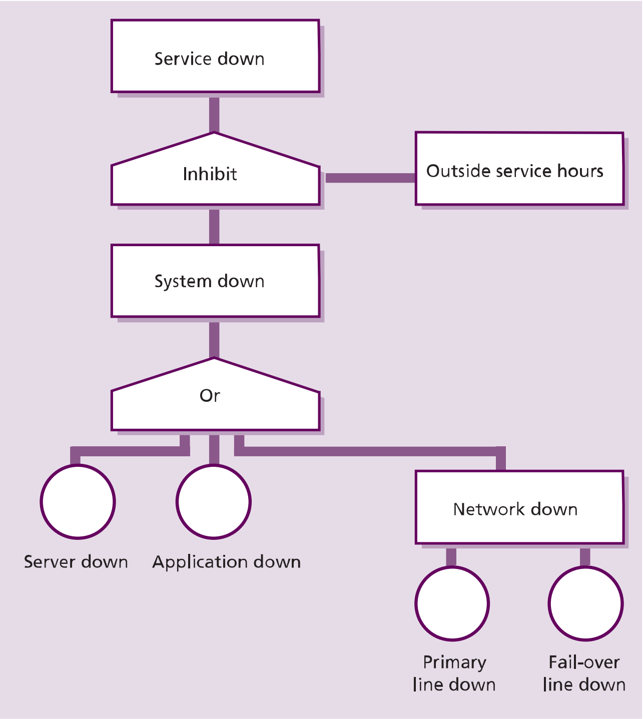


Figure 4.14 Fault tree analysis – example

These events can be combined using logic operators, i.e.:

- **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.

Based on these definitions, Figure 4.14 depicts an analysis in which there is one 'AND' gate on the lower right side. Both the primary and fail-over lines must be down for the network to be down. There is also an 'OR' gate on the left of the diagram; if any one of the three listed events occurs, the system will be down. Finally, the 'Inhibit' gate shows that only if the failure occurs during service hours will the service itself be considered to be down.

This is the basic FTA technique. This technique can also be refined, but complex FTA and the mathematical evaluation of fault trees are beyond the scope of this publication.

Modelling

To assess if new components within a design can match the stated requirements, it is important that the testing regime instigated ensures that the availability expected can be delivered. Simulation, modelling or load testing tools to generate the expected user demand for the new IT service should be seriously considered to ensure components continue to operate under anticipated volume and stress conditions.

Modelling tools are also required to forecast availability and to assess the impact of changes to the IT infrastructure. Inputs to the modelling process 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 or acquired. The lack of readily available availability modelling tools in the marketplace may require such a tool to be developed and maintained 'in-house', but this is a very expensive and time-consuming activity that should only be considered where the investment can be justified. 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. However, some system management tools do provide modelling capability and can provide useful information on trending and forecasting availability needs.

Risk assessment and management

To assess the vulnerability of failure within the configuration and capability of the IT service and support organization it is recommended that existing or proposed IT infrastructure, service configurations, service design and supporting organization (internal and external suppliers) are subject to formal risk assessment and management exercises. Risk assessment and management is a technique that can be used to identify and quantify risks and justifiable countermeasures that can be implemented to protect the availability of IT systems.

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 assessment should be undertaken during the design stage for the IT technology and service to identify:

- Risks that may incur unavailability for IT components within the technology and service design
- Risks that may incur confidentiality and/or integrity exposures within the IT technology and service design.

Most risk assessment and management methodologies involve the use of a formal approach to the assessment of risk and the subsequent mitigation of risk with the implementation of cost-justifiable countermeasures (Figure 4.15). Appendix M describes several broadly known and used approaches to the assessment and management of risk.

As illustrated in Figure 4.15, risk assessment 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 is also determined to a certain extent by its acceptance. Some organizations and businesses may be more willing to accept risk whereas others are not.

Risk management involves the identification, selection and adoption of countermeasures justified by the identified risks to assets in terms of

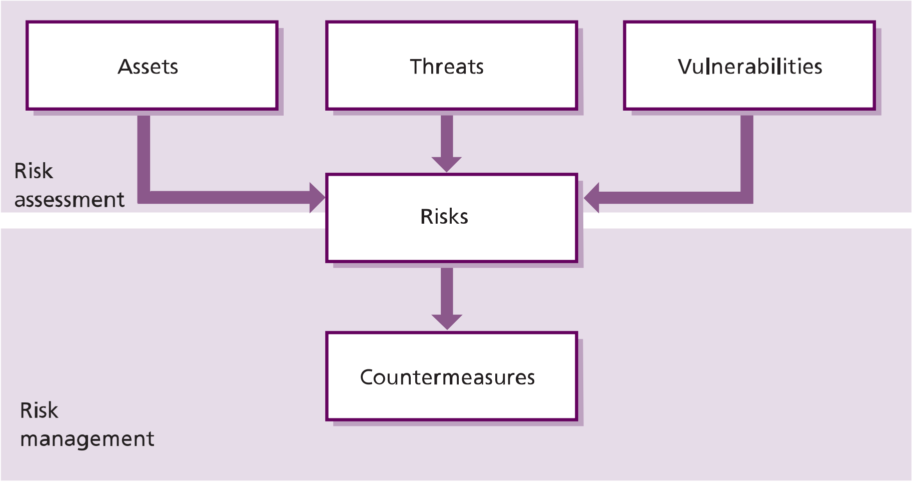


Figure 4.15 Risk assessment and management

their potential impact on services if failure occurs, and the reduction of those risks to an acceptable level. Risk management is an activity that is associated with many other activities, especially the IT service continuity management and information security management processes and the service transition lifecycle stage. All of these risk assessment exercises should be coordinated rather than being separate activities.

This approach, when applied via a formal method, ensures 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 technology reviewed
- All expenditure on selected countermeasures can be justified.

Formal risk assessment and management methods are now an important element in the overall design and provision of IT services. The assessment of risk is often based on the probability and potential impact of an event occurring. Countermeasures are implemented wherever they are cost-justifiable, to reduce the impact of an event, or the probability of an event occurring, or both.

It should be noted that the risk assessment and management described here aligns in its essentials with an asset-focused approach required in ISO/IEC 27001. Management of risk (M_o_R) provides an alternative generic framework for the management of risk across all parts of an organization – strategic, programme, project and operational (see Appendix N for more details).

Implementing cost-justifiable countermeasures

The risks identified to service and component availability should be addressed through appropriate risk reduction measures and the development of effective recovery mechanisms. These countermeasures may be implemented as part of the overall design of the new or changed service, as well as through the implementation of best practice in the areas of maintenance and continual review and improvement.

Planned and preventive maintenance

All IT components should be subject to a planned maintenance strategy. This can be considered part of the risk reduction strategy for services and components. The frequency and levels of maintenance required varies 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 organization to provide:

- Preventive maintenance to avoid failures
- Planned software or hardware upgrades to provide new functionality or additional capacity
- Business requested changes to the business applications
- Implementation of new technology and functionality for exploitation by the business.

The requirement for planned downtime clearly influences 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 24 × 7 service operation. In these instances, it is essential that continuous operation is a core design feature to enable maintenance activity to be performed without impacting the availability of IT services.

Where the required service hours for IT services are less than 24 hours per day and/or seven days per week, it is likely that the majority of planned maintenance can be accommodated without impacting IT service availability. However, where the business needs IT services available on a 24-hour and seven-day basis, availability management needs 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:

- SLAs
- OLAs
- Underpinning contracts
- Change management schedules
- Release and deployment management schedules.

Hints and tips

Availability management should ensure that building in preventive maintenance is one of the prime design considerations for a '24 × 7' IT service.

The most appropriate time to schedule planned downtime is clearly when the impact on the business and its customers is least. This information should be provided initially by the business when determining the availability requirements. For an existing IT service, or once the new service has been established, monitoring of business and customer transactions helps establish the hours when IT service usage is at its lowest. This should determine the most appropriate time for the component(s) to be removed for planned maintenance activity.

To accommodate the individual component requirements for planned downtime while 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. While this approach has benefits, there are potential risks that need to be assessed. For example:

- The capability of the IT support organization to coordinate 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 back-out of a failed change requires multiple changes to be removed.

The effective management of planned downtime is 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 formalized, measured and reported. All planned maintenance should be scheduled, managed and controlled to ensure that the individual objectives are achieved, time slots are not exceeded, and that activities are coordinated with all other schedules of activity to minimize clashes and conflict (e.g. change and release schedules, testing schedules). In addition they 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 back-out or other remediation plan. Planned downtime and performance against the stated objectives for each component should be recorded and used in service reporting.

Reviewing all new and changed services and testing all availability and resilience mechanisms

During the service transition stage all the elements designed to contribute to service and component availability need to be reviewed and tested. Availability review and testing procedures and policies should be embedded into overall transition methods, processes and practices to ensure that the promised levels of availability will be delivered.

In addition to the reviews and tests that occur during service transition, regularly scheduled reviews and tests are required for the most comprehensive approach.

Availability testing schedule

A key deliverable from the availability management process is the 'availability testing schedule'. This is a schedule for the regular testing of all availability mechanisms. Some availability mechanisms, such as 'load balancing', 'mirroring' and 'grid computing', are used in the provision of normal service on a day-by-day basis; others are used on a fail-over or manual reconfiguration basis. It is essential, therefore, that all availability mechanisms are tested in a regular and scheduled manner to ensure that when they are actually needed for real they work. This schedule needs to be maintained and widely circulated so that all areas are aware of its content and so that all other proposed activities can be synchronized with its content, such as:

- The change schedule
- Release plans and the release schedule
- All transition plans, projects and programmes
- Planned and preventive maintenance schedules
- The schedule for testing IT service continuity plans and recovery mechanisms
- Business plans and schedules
- Capacity plans.

Continual review and improvement

Changing business needs and customer 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 SLM. Other inputs should also be considered on a regular basis from ITSCM, particularly from the updated BIA and risk assessment exercises. The criticality of services will often change and it is important that the design and the technology supporting such services is regularly reviewed and improved by availability management to ensure that the change of importance in the service is reflected within a revised design and supporting technology and documentation. Where the required levels of availability are already being delivered, it may take considerable effort and incur significant cost to achieve a small incremental improvement within the level of availability.

A key activity for availability management is to look continually at opportunities to optimize the availability of the IT infrastructure in conjunction with overall CSI activities. The benefits of this regular review approach are that, sometimes, enhanced levels of availability may be achievable, but with much lower costs. The optimization approach is a sensible first step to delivering better value for money. A number of availability management techniques can be applied to identify optimization opportunities. It is recommended that the scope should not be restricted to the technology, but also include a review of both the business process and other end-to-end business-owned responsibilities. To help achieve these aims, availability management needs to be recognized as a leading influence over the IT service provider organization to ensure continued focus on availability and stability of the technology.

Availability management can provide the IT support organization with a real business and user perspective on how deficiencies within the technology and the underpinning process and procedure impact on 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 service provider organization recognize where it can add value by exploiting its technical skills and competencies in an availability context. The continual 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 or SFA environment.

The impetus to improve availability comes from one or more of the following:

- The inability of existing or new IT services to meet SLA targets 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 times
- Requests from the business to increase the level of availability provided

- Increasing impact on the business and its customers of IT service failures as a result of growth and/or increased business priorities or functionality
- A request from SLM to improve availability as part of an overall SIP
- Availability management monitoring and trend analysis.

Availability management should take a proactive role in identifying and progressing cost-justified availability improvement opportunities within the availability plan. The ability to do this places reliance on having appropriate and meaningful availability measurement and reporting. To ensure availability improvements deliver benefits to the business and users, it is important that measurement and reporting reflect not just IT component availability but also availability from a business operation and user perspective – that is, end-to-end service availability.

Where the business has a requirement to improve availability, the process and techniques to reassess the technology and IT service provider organization capability to meet these enhanced requirements should be followed. An output of this activity is enhanced availability and recovery design criteria. To satisfy the business requirement for increased levels of availability, additional financial investment may be required to enhance the underpinning technology and/or extend the services provided by the IT service provider organization. It is important that any additional investment to improve the levels of availability delivered can be cost-justified. Determining the cost of unavailability as a result of IT failure(s) can help support any financial investment decision in improving availability.

Contribute to production of the projected service outage document

Availability management should work closely with change management, the process which produces and maintains the projected service outage (PSO) document. This document describes any variations from the service availability agreed within SLAs. This should be produced based on input from:

- The change schedule
- The release schedules
- Planned and preventive maintenance schedules
- Availability testing schedules
- ITSCM and BCM testing schedules.

The PSO contains details of all scheduled and planned service downtime within the agreed service hours for all services. These documents should be agreed with all the appropriate areas and representatives of both the business and IT. Once the PSO has been agreed, the service desk should ensure that it is communicated to all relevant parties so that everyone is made aware of any additional, planned service downtime.

4.4.6 Triggers, inputs, outputs and interfaces

4.4.6.1 Triggers

Many events may trigger availability management activity. These include:

- New or changed business needs or new or changed services
- New or changed targets within agreements, such as SLRs, SLAs, OLAs or contracts
- Service or component breaches, availability events and alerts, including threshold events, exception reports
- Periodic activities such as reviewing, revising or reporting
- Review of availability management forecasts, reports and plans
- Review and revision of business and IT plans and strategies
- Review and revision of designs and strategies
- Recognition or notification of a change of risk or impact of a business process or VBF, an IT service or component
- Request from SLM for assistance with availability targets and explanation of achievements.

4.4.6.2 Inputs

A number of sources of information are relevant to the availability management process. Some of these are as follows:

- **Business information** From the organization's business strategy, plans and financial plans, and information on their current and future requirements, including the availability requirements for new or enhanced IT services
- **Business impact information** From BIAs and assessment of VBFs underpinned by IT services

- **Reports and registers** Previous risk assessment reports and a risk register
- **Service information** From the service portfolio and the service catalogue
- **Service information** From the SLM process, with details of the services from the service portfolio and the service catalogue, service level targets within SLAs and SLRs, and possibly from the monitoring of SLAs, service reviews and breaches of the SLAs
- **Financial information** From financial management for IT services, the cost of service provision, the cost of resources and components
- **Change and release information** From the change management process with a change schedule, the release schedule from release and deployment management and a need to assess all changes for their impact on service availability
- **Service asset and configuration management** Containing information on the relationships between the business, the services, the supporting services and the technology
- **Service targets** From SLAs, SLRs, OLAs and contracts
- **Component information** On the availability, reliability and maintainability requirements for the technology components that underpin IT service(s)
- **Technology information** From the CMS on the topology and the relationships between the components and the assessment of the capabilities of new technology
- **Past performance** From previous measurements, achievements and reports and the availability management information system (AMIS)
- **Unavailability and failure information** From incidents and problems
- **Planning information** From other processes such as the capacity plan from capacity management.

4.4.6.3 Outputs

The outputs produced by availability management should include:

- The availability MIS (AMIS)
- The availability plan for the proactive improvement of IT services and technology
- Availability and recovery design criteria and proposed service targets for new or changed services
- Service availability, reliability and maintainability reports of achievements against targets, including input for all service reports
- Component availability, reliability and maintainability reports of achievements against targets
- Revised risk assessment reviews and reports and an updated risk register
- Monitoring, management and reporting requirements for IT services and components to ensure that deviations in availability, reliability and maintainability are detected, actioned, recorded and reported
- An availability management test schedule for testing all availability, resilience and recovery mechanisms
- The planned and preventive maintenance schedules
- Contributions for the PSO to be created by change in collaboration with release and deployment management
- Details of the proactive availability techniques and measures that will be deployed to provide additional resilience to prevent or minimize the impact of component failures on the IT service availability
- Improvement actions for inclusion within the SIP.

4.4.6.4 Interfaces

The key interfaces that availability management has with other processes are:

- **SLM** This process relies on availability management to determine and validate availability targets and to investigate and resolve service and component breaches.
- **Incident and problem management** These are assisted by availability management in the resolution and subsequent justification and correction of availability incidents and problems.
- **Capacity management** This provides appropriate capacity to support resilience and overall service availability. The process also uses information from demand management about patterns of business activity and user profiles to understand business demand for IT services and provides this information to availability management for business-aligned availability planning.

- **Change management** This leads to the creation of the PSO with contributions from availability management. When changes are proposed to a service, availability must assess the change for availability-related issues including any potential impact on achievement of availability service levels.
- **IT service continuity management (ITSCM)** Availability management works collaboratively with this process on the assessment of business impact and risk and the provision of resilience, fail-over and recovery mechanisms. Availability focuses on normal business operation and ITSCM focuses on the extraordinary interruption of service.
- **Information security management (ISM)** If the data becomes unavailable, the service becomes unavailable. ISM defines the security measures and policies that must be included in the service design for availability and design for recovery.
- **Access management** Availability management provides the methods for appropriately granting and revoking access to services as needed.

4.4.7 Information management

The availability management process should maintain an AMIS that contains all of the measurements and information required to complete the availability management process and provide the appropriate information to the business on the level of IT service provided. This information, covering services, components and supporting services, provides the basis for regular, ad hoc and exception availability reporting and the identification of trends within the data for the instigation of improvement activities. These activities and the information contained within the AMIS provide the basis for developing the content of the availability plan.

In order to provide structure and focus to a wide range of initiatives that may need to be undertaken to improve availability, an availability plan should be formulated and maintained. The availability plan should have aims, objectives and deliverables and should consider the wider issues of people, processes, 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 versus agreed levels of availability for key IT services. Availability measurements should always be business- and customer-focused and report availability as experienced by the business and users.
- 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 SFA assignments.
- Regular reviews of SFA assignments should be completed to ensure that the availability of technology is being proactively improved in conjunction with the SIP.
- 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, where possible based on business-focused measures, in conjunction with capacity management. The effort required to realize these benefits where possible should also be quantified.

During the production of the availability plan, it is recommended that liaison with all functional, technical and process areas is undertaken. 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 technology 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 financial plan, and that publication is 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.

In order to facilitate the production of the availability plan, availability management may wish to consider having its own database repository. The AMIS can be utilized to record and store selected data and information required to support key activities such as report generation, statistical analysis and availability forecasting and planning. The AMIS should be the main repository for the recording of IT availability metrics, measurements, targets and documents, including the availability plan, availability measurements, achievement reports, SFA assignment reports, design criteria, action plans and testing schedules.

Hints and 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 the other IT service and systems management processes to identify common requirements with the aim of selecting shared tools and minimizing costs. The AMIS should address the specific reporting needs of availability management not currently provided by existing repositories and integrate with them and their contents.

4.4.8 Critical success factors and key performance indicators

The following list includes some sample CSFs for availability management. Each organization should identify appropriate CSFs based on its objectives for the process. Each sample CSF is followed by a small number of typical KPIs that support the CSF. These KPIs should not be adopted without careful consideration. Each organization should develop KPIs that are appropriate for its level of maturity, its CSFs and its particular circumstances. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

- **CSF** Manage availability and reliability of IT service
  - **KPI** Percentage reduction in the unavailability of services and components
  - **KPI** Percentage increase in the reliability of services and components
  - **KPI** Effective review and follow-up of all SLA, OLA and underpinning contract breaches relating to availability and reliability
  - **KPI** Percentage improvement in overall end-to-end availability of service
  - **KPI** Percentage reduction in the number and impact of service breaks
  - **KPI** Improvement in the MTBF
  - **KPI** Improvement in the MTBSI
  - **KPI** Reduction in the MTRS
- **CSF** Satisfy business needs for access to IT services
  - **KPI** Percentage reduction in the unavailability of services
  - **KPI** Percentage reduction of the cost of business overtime due to unavailable IT
  - **KPI** Percentage reduction in critical time failures – for example, specific business peak and priority availability needs are planned for
  - **KPI** Percentage improvement in business and users satisfied with service (by customer satisfaction survey results)
- **CSF** Availability of IT infrastructure and applications, as documented in SLAs, provided at optimum costs:
  - **KPI** Percentage reduction in the cost of unavailability
  - **KPI** Percentage improvement in the service delivery costs
  - **KPI** Timely completion of regular risk assessment and system review
  - **KPI** Timely completion of regular cost-benefit analysis established for infrastructure CFIA
  - **KPI** Percentage reduction in failures of third-party performance on MTRS/MTBF against contract targets
  - **KPI** Reduced time taken to complete (or update) a risk assessment

- **KPI** Reduced time taken to review system resilience
- **KPI** Reduced time taken to complete an availability plan
- **KPI** Timely production of management reports
- **KPI** Percentage reduction in the incidence of operational reviews uncovering security and reliability exposures in application designs.

4.4.9 Challenges and risks

4.4.9.1 Challenges

Availability management faces many challenges, but the main challenge is to actually meet and manage the expectations of the customers, the business and senior management. These expectations are frequently that services will always be available not just during their agreed service hours, but that all services will be available on a 24-hour, 365-day basis. When they are not, it is assumed that they will be recovered within minutes. This is only the case when the appropriate level of investment and design has been applied to the service, and this should only be made where the business impact justifies that level of investment. However, the message needs to be publicized to all customers and areas of the business, so that when services do fail they have the right level of expectation on their recovery. It also means that availability management must have access to the right level of quality information on the current business need for IT services and its plans for the future. This is another challenge faced by many availability management processes.

Another challenge facing availability management is the integration of all of the availability data into an integrated set of information (AMIS) that can be analysed in a consistent manner to provide details on the availability of all services and components. This is particularly challenging when the information from the different technologies is often provided by different tools in differing formats.

Yet another challenge facing availability management is convincing the business and senior management of the investment needed in proactive availability measures. Investment is always recognized once failures have occurred, but by then it is really too late. Persuading businesses and customers to invest in resilience to avoid the possibility of failures that may happen is a difficult challenge. Availability management should work closely with ITSCM, information security management and capacity management in producing the justifications necessary to secure the appropriate investment.

4.4.9.2 Risks

Some of the major risks associated with availability management include:

- A lack of commitment from the business to the availability management process
- A lack of commitment from the business and a lack of appropriate information on future plans and strategies
- A lack of senior management commitment or a lack of resources and/or budget to the availability management process
- Labour-intensive reporting processes
- The processes focus too much on the technology and not enough on the services and the needs of the business
- The AMIS is maintained in isolation and is not shared or consistent with other process areas, especially ITSCM, information security management and capacity management. This investment is particularly important when considering the necessary service and component backup and recovery tools, technology and processes to meet the agreed needs.

4.5 CAPACITY MANAGEMENT

Capacity management is a process that extends across the service lifecycle. A key success factor in managing capacity is ensuring it is considered during the design stage. It is for this reason that the capacity management process is included here. Capacity management is supported initially in service strategy where the decisions and analysis of business requirements and customer outcomes influence the development of patterns of business activity, lines of service (LOS) and service options. This provides the predictive and ongoing capacity indicators needed to align capacity to demand. Capacity management provides a point of focus and management for all capacity- and performance-related issues, relating to both services and resources.

Like availability, capacity is an important part of the warranty of a service. If a service does not deliver the levels of capacity and performance required, then the business will not experience the value that has been promised. Without capacity and performance the utility of the service cannot be accessed.

4.5.1 Purpose and objectives

The purpose of the capacity management process is to ensure that the capacity of IT services and the IT infrastructure meets the agreed capacity- and performance-related requirements in a cost-effective and timely manner. Capacity management is concerned with meeting both the current and future capacity and performance needs of the business.

The objectives of capacity management are to:

- Produce and maintain an appropriate and up-to-date capacity plan, which reflects the current and future needs of the business
- Provide advice and guidance to all other areas of the business and IT on all capacity- and performance-related issues
- Ensure that service performance achievements meet all of their agreed targets by managing the performance and capacity of both services and resources
- Assist with the diagnosis and resolution of performance- and capacity-related incidents and problems
- Assess the impact of all changes on the capacity plan, and the performance and capacity of all services and resources
- Ensure that proactive measures to improve the performance of services are implemented wherever it is cost-justifiable to do so.

4.5.2 Scope

The capacity management process should be the focal point for all IT performance and capacity issues. Capacity management considers all resources required to deliver the IT service, and plans for short-, medium- and long-term business requirements.

The process should encompass all areas of technology, both hardware and software, for all IT technology components and environments. Capacity management should also consider space planning and environmental systems capacity. Capacity management could consider human resource capacity where a lack of human resources could result in a breach of SLA or OLA targets, a delay in the end-to-end performance or service response time, or an inability to meet future commitments and plans (e.g. overnight data backups not completed in time because no operators were present to load required media).

In general, human resource management is a line management responsibility, although the staffing of a service desk should use identical capacity management techniques. The scheduling of human resources, staffing levels, skill levels and capability levels should therefore be included within the scope of capacity management.

The capacity management process should include:

- Monitoring patterns of business activity through performance, utilization and throughput of IT services and the supporting infrastructure, environmental, data and applications components and the production of regular and ad hoc reports on service and component capacity and performance
- Undertaking tuning activities to make the most efficient use of existing IT resources
- Understanding the agreed current and future demands being made by the customer for IT resources, and producing forecasts for future requirements
- Influencing demand in conjunction with the financial management for IT services and demand management processes
- Producing a capacity plan that enables the service provider to continue to provide services of the quality defined in SLAs and that covers a sufficient planning timeframe to meet future service levels required as defined in the service portfolio and SLRs
- Assisting with the identification and resolution of any incidents and problems associated with service or component capacity or performance
- The proactive improvement of service or component performance, wherever it is cost-justifiable and meets the needs of the business.

Capacity management also includes understanding the potential for the delivery of new services. New technology needs to be understood and, if appropriate, used to innovate and deliver the services required by the customer. Capacity management should recognize that the rate of technological change will probably increase and that new technology should be harnessed to ensure that the IT services continue to satisfy changing business expectations. A direct link to the service strategy and service portfolio is needed to ensure that emerging technologies are considered in future service planning.

Capacity management has a close, two-way relationship with the service strategy and planning processes within an organization. On a regular basis, the long-term strategy of an organization is encapsulated in an update of the business plans. The service strategy will reflect the business plans and strategy, which are developed from the organization'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. Often a shorter-term tactical plan or business change plan is developed to implement the changes necessary in the short to medium term to progress the overall business plan and service strategy. Capacity management needs to understand the short-, medium- and long-term plans of the business and IT while providing information on the latest ideas, trends and technologies being developed by the suppliers of computing hardware and software.

The organization's business plans drive the specific IT service strategy, the contents of which capacity management needs to be familiar with, and to which capacity management needs to have had significant and ongoing input. The right level of capacity at the right time is critical. Service strategy plans will be helpful to capacity planning by identifying the timing for acquiring and implementing new technologies, hardware and software.

Capacity management is responsible for ensuring that IT resources are planned and scheduled to provide a consistent level of service that is matched to the current and future needs of the business, as agreed and documented within SLAs and OLAs. In conjunction with the business and its plans, capacity management provides a capacity plan that outlines the IT resources and funding needed to support the business plan, together with a cost justification of that expenditure.

4.5.3 Value to the business

A well-executed capacity management process will benefit the business by:

- Improving the performance and availability of IT services the business needs by helping to reduce capacity- and performance-related incidents and problems
- Ensuring required capacity and performance are provided in the most cost-effective manner
- Contributing to improved customer satisfaction and user productivity by ensuring that all capacity- and performance- related service levels are met
- Supporting the efficient and effective design and transition of new or changed services through proactive capacity management activities
- Improving the reliability of capacity-related budgeting through the use of a forward-looking capacity plan based on a sound understanding of business needs and plans
- Improving the ability of the business to follow an environmentally responsible strategy by using green technologies and techniques in capacity management.

4.5.4 Policies, principles and basic concepts

Capacity management ensures that the capacity and performance of the IT services and systems match the evolving agreed demands of the business in the most cost-effective and timely manner. Capacity management is essentially a balancing act:

- **Balancing costs against resources needed** The need to ensure that processing capacity that is purchased is cost-justifiable in terms of business need, and the need to make the most efficient use of those resources.
- **Balancing supply against demand** The need to ensure that the available supply of IT 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.

4.5.4.1 Policies

The driving force for capacity management should be the business requirements of the organization and to plan the resources needed to provide service levels in line with SLAs and OLAs. Policies should be established defining the required points of interface between the capacity management and SLM processes to ensure this connection to business requirements is appropriately established and maintained. Capacity management needs to understand the total IT and business environments, including:

- The current business operation and its requirements, through the patterns of business activity (as provided by the demand management process)
- The future business plans and requirements via the service portfolio
- The service targets and the current IT service operation though SLAs and standard operating procedures
- All areas of IT technology and its capacity and performance, including infrastructure, data, environment and applications.

Understanding all of this will enable capacity management to ensure that all the current and future capacity and performance aspects of services are provided cost-effectively.

It should be the service provider's policy that capacity management processes and planning must be involved in all stages of the service lifecycle from strategy and design, through transition and operation to improvement. From a strategic perspective, the service portfolio contains the IT resources and capabilities. The advent of service-oriented architecture, virtualization and the use of value networks in IT service provision are important factors in the management of capacity. The appropriate capacity and performance should be designed into services and components from the initial design stages. This will ensure not only that the performance of any new or changed service meets its expected targets, but also that all existing services continue to meet all of their targets. This is the basis of stable service provision.

4.5.4.2 Planning and managing complexity

Managing the capacity of large distributed IT infrastructures is a complex and demanding task, especially when the IT capacity and the financial investment required is ever-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 is usually 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. Capacity management should have input to the service portfolio and procurement process to ensure that the best deals with suppliers are negotiated.

Capacity management provides the necessary information on current and planned resource utilization of individual components to enable organizations to decide, with confidence:

- Which components to upgrade (i.e. more memory, faster storage devices, faster processors, greater bandwidth)
- When to upgrade – ideally this is not too early, resulting in expensive over-capacity, nor too late, failing to take advantage of advances in new technology, 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 processes are 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 SLM process be assured that the SLRs of the new service are achievable, and that the SLAs of existing services will not be affected?
- Can the problem management process properly diagnose the underlying cause of incidents caused by poor performance?
- Can the IT service continuity process accurately determine the capacity requirements of the key business processes?

Capacity management is one of the forward-looking processes that, when properly carried out, can forecast business events and impacts often before they happen. Good capacity management ensures that there are no surprises with regard to service and component design and performance.

The overall capacity management process is continually trying to match IT resources and capacity cost-effectively to the ever-changing needs and requirements of the business. This requires the tuning (or 'optimization') of the current resources and the effective estimation and planning of the future resources.

One of the key activities of capacity management is to produce a plan that documents the current levels of resource utilization and service performance and, after consideration of the service strategy and plans, forecasts the future requirements for new IT resources to support the IT services that underpin the business activities. The plan should indicate clearly any assumptions made. It should also include any recommendations quantified in terms of resource required, cost, benefits, impact etc.

The service provider should establish pre-defined intervals for the production and maintenance of a capacity plan. It is, essentially, an investment plan and should therefore be published annually, in line with the business or budget lifecycle, and completed before the start of negotiations on future budgets. A quarterly reissue of the updated plan may be necessary to take into account changes in service plans, to report on the accuracy of forecasts and to make or refine recommendations. This takes extra effort but, if it is regularly updated, the capacity plan is more likely to be accurate and to reflect the changing business need.

The typical contents of a capacity plan are described in Appendix J.

4.5.4.3 Capacity management sub-processes

Capacity management is an extremely technical, complex and demanding process, and in order to achieve results, it requires three supporting sub-processes: business capacity management, service capacity management and component capacity management. These sub-processes are described briefly in this section and in more detail in section 4.5.5.

Business capacity management

The business capacity management sub-process translates business needs and plans into requirements for service and IT infrastructure, ensuring that the future business requirements for IT services are quantified, designed, planned and implemented in a timely fashion. This can be achieved by using the existing data on the current resource utilization by the various services and resources to trend, forecast, model or predict future requirements. These future requirements come from the service strategy and service portfolio detailing new processes and service requirements, changes, improvements, and also the growth in the existing services.

Service capacity management

The service capacity management sub-process focuses on the management, control and prediction of the end-to-end performance and capacity of the live, operational IT services usage and workloads. It ensures that the performance of all services, as detailed in service targets within SLAs and SLRs, is monitored and measured, and that the collected data is recorded, analysed and reported. Wherever necessary, proactive and reactive action should be instigated, to ensure that the performance of all services meets their agreed business targets. This is performed by staff with knowledge of all the areas of technology used in the delivery of end-to-end service, and often involves seeking advice from the specialists involved in component capacity management. Wherever possible, automated thresholds should be used to manage all operational services, to ensure that situations where service targets are breached or threatened are rapidly identified and cost-effective actions to reduce or avoid their potential impact are implemented.

Component capacity management

The component capacity management sub-process focuses on the management, control and prediction of the performance, utilization and capacity of individual IT technology components. It ensures 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. Again, wherever possible, automated thresholds should be implemented to manage all components, to ensure that situations where service targets are breached or threatened by component usage or performance are rapidly identified, and cost-effective actions to reduce or avoid their potential impact are implemented.

There are many similar activities that are performed by each of the above sub-processes,

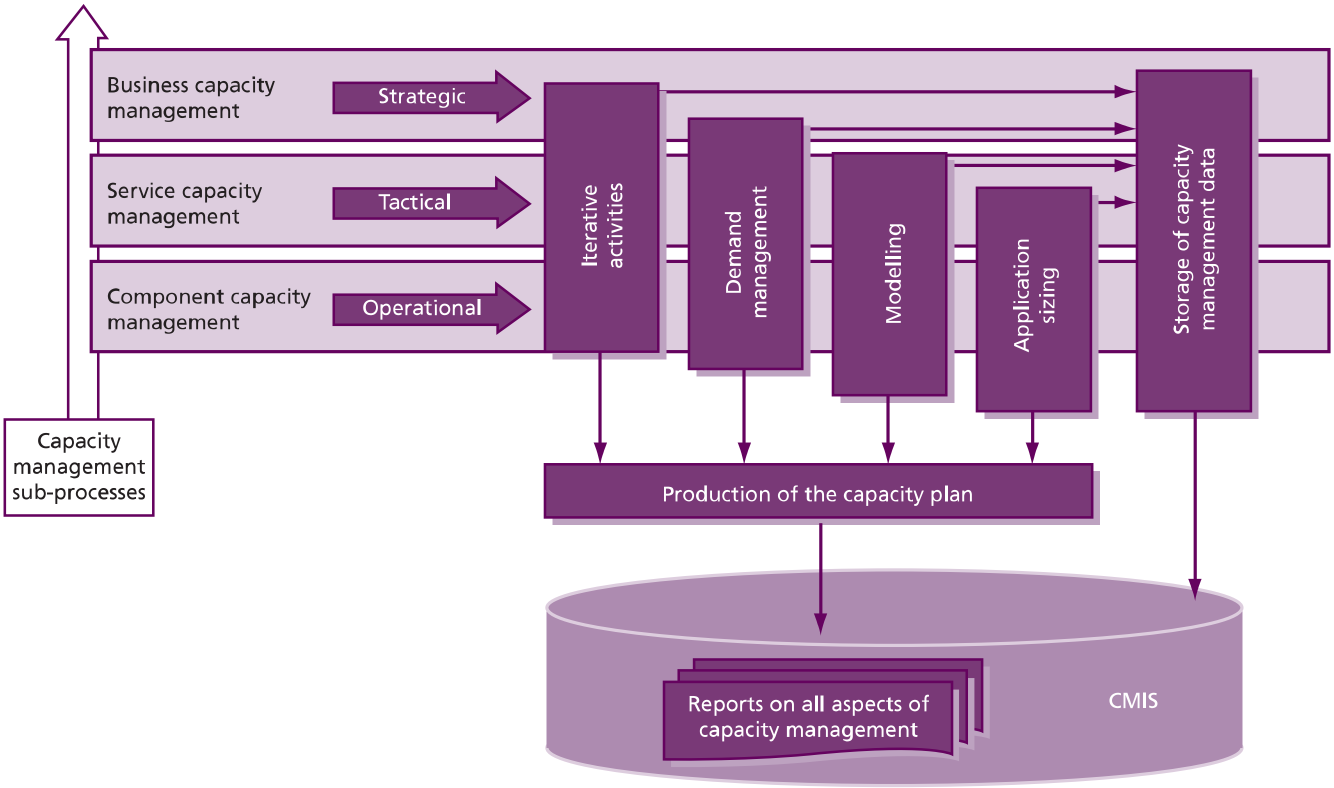

Business capacity management
Strategic
Service capacity management
Tactical
Component capacity management
Operational
Iterative activities
Demand management
Modelling
Application sizing
Storage of capacity management data
Capacity management sub-processes
Production of the capacity plan
Reports on all aspects of capacity management
CMIS

Figure 4.16 Capacity management sub-processes

but each sub-process has a very different focus. Business capacity management is focused on the current and future business requirements, while service capacity management is focused on the delivery of the existing services that support the business, and component capacity management is focused on the IT infrastructure that underpins service provision. The role that each of these sub-processes plays in the overall process, the production of the capacity plan and the storage of capacity-related data is illustrated in Figure 4.16.

The tools used by capacity management need to conform to the organization's management architecture and integrate with other tools used for the management of IT systems and automating IT processes. The monitoring and control activities within service operation will provide a good basis for the tools to support and analyse information for capacity management. The IT operations management function and the technical management departments such as network management and server management may carry out the bulk of the day-to-day operational duties, participating in the capacity management process by providing performance information to the process.

4.5.5 Process activities, methods and techniques

Some activities in the capacity management process are reactive, while others are proactive. The proactive activities of capacity management should include:

- Pre-empting performance issues by taking the necessary actions before they occur
- Producing trends of the current component utilization and estimating the future requirements, using trends and thresholds for planning upgrades and enhancements
- Modelling and trending the predicted changes in IT services (including service retirements), and identifying the changes that need to be made to services and components of the IT infrastructure and applications to ensure that appropriate resource is available
- Ensuring that upgrades are budgeted, planned and implemented before SLAs and service targets are breached or performance issues occur
- Actively seeking to improve service performance wherever it is cost-justifiable

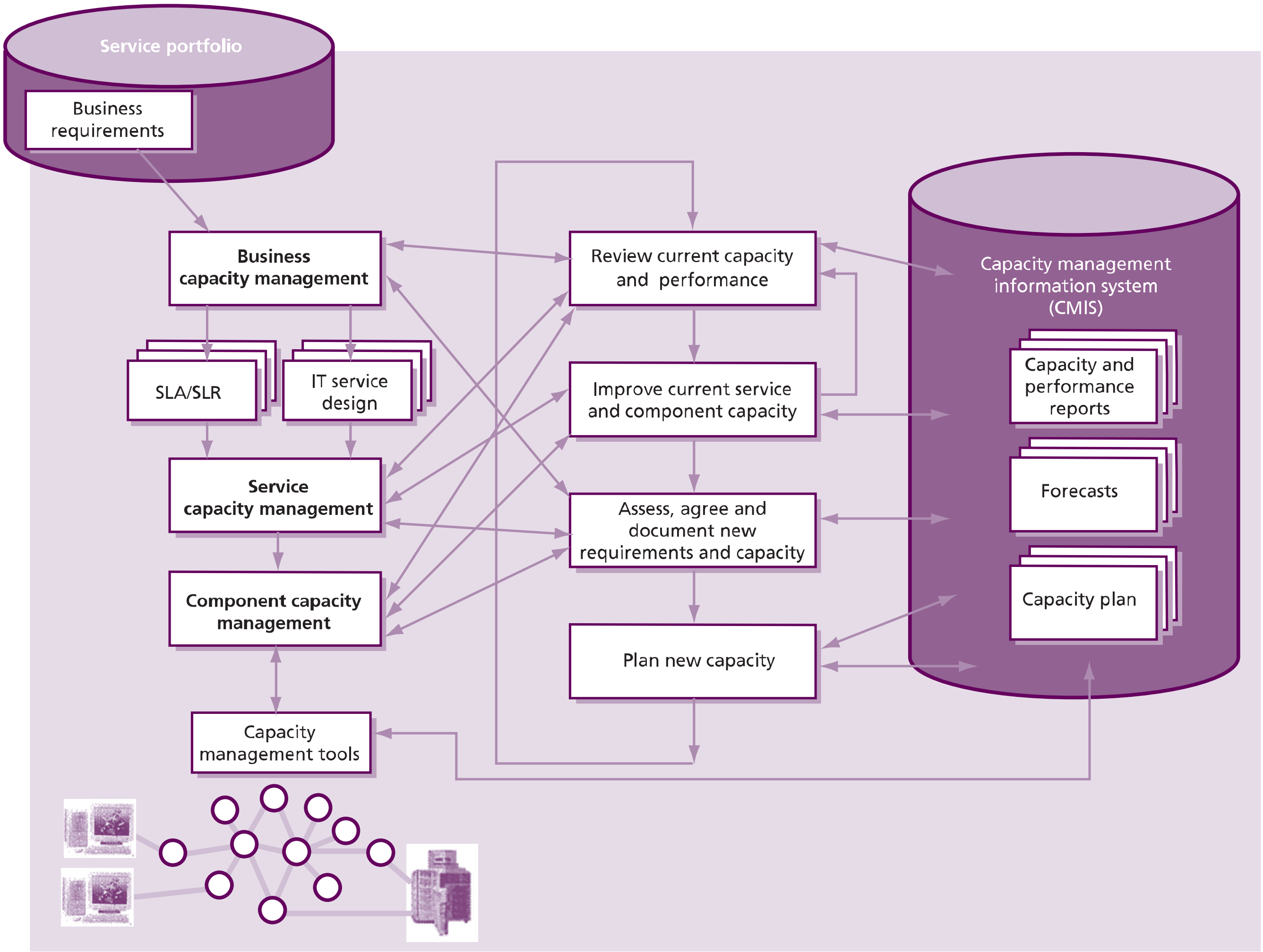

Service portfolio
Business requirements
Business capacity management
SLA/SLR
IT service design
Service capacity management
Component capacity management
Capacity management tools
Review current capacity and performance
Improve current service and component capacity
Assess, agree and document new requirements and capacity
Plan new capacity
Capacity management information system (CMIS)
Capacity and performance reports
Forecasts
Capacity plan

Figure 4.17 Capacity management overview with sub-processes

- Producing and maintaining a capacity plan that reflects all trends, predicted changes, future requirements and plans for meeting them
- Tuning (optimizing) the performance of services and components.

The reactive activities of capacity management should include:

- Monitoring, measuring, reporting and reviewing the current performance of both services and components
- Responding to all capacity-related 'threshold' events and instigating corrective action
- Reacting to and assisting with specific performance issues. For example, the service desk may refer incidents of poor performance to technology management, which will employ capacity management techniques to resolve them.

These individual activities together allow an organization to:

- Assess, agree and document new requirements and capacity
- Plan new capacity
- Review current capacity and performance
- Improve current service and component capacity.

Figure 4.17 provides a high-level overview of the capacity management process with its sub-processes, related documents and data, as well as the relationships to capacity management tools.

Key message

The more successful the proactive and predictive activities of capacity management, the less need there will be for the reactive activities of capacity management.

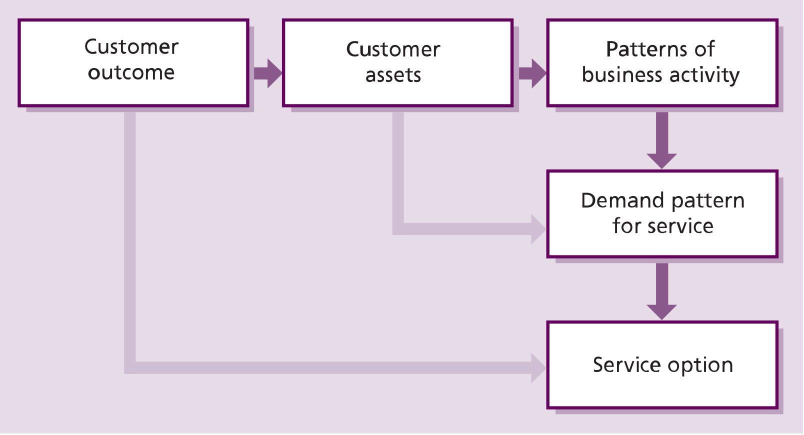


Figure 4.18 Capacity must support business requirements

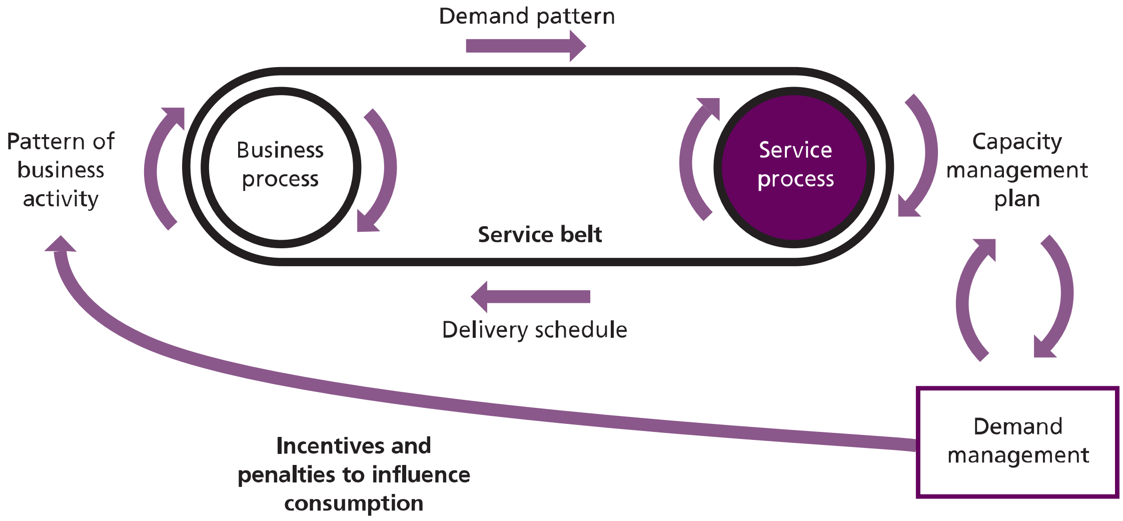


Figure 4.19 Capacity management takes particular note of demand pattern

4.5.5.1 Business capacity management

The main objective of the business capacity management sub-process is to ensure that the future business requirements (customer outcomes) for IT services are considered and understood, and that sufficient IT capacity to support any new or changed services is planned and implemented within an appropriate timescale. Figure 4.18 illustrates that business capacity management is influenced by the patterns of business activity and how services are used.

The capacity management process must be responsive to changing requirements for capacity demand. New services or changed 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' SLRs and SLAs will be affected. It is the responsibility of capacity management to predict the demand for capacity for such changes and manage the demand at a tactical level.

These new requirements may come to the attention of capacity management from many different sources and for many different reasons, but the principal sources of supply should be the patterns of business activity from demand management. The demand management process will analyse patterns of business activity to find out how these patterns generate demand patterns for IT service and, together with service portfolio management, will create service packages and service options at the right level of utility and warranty to efficiently support these patterns (see section 4.4 in *ITIL Service Strategy* for more information on demand management). This information will inform all the work of capacity management and allow for more successful and more proactive capacity management. Examples of resulting actions could be a recommendation to upgrade to take advantage of new technology, or the implementation of a tuning activity to resolve a performance problem. Figure 4.19 shows the cycle of demand management.

Capacity management needs to be included in all strategic, planning and design activities, being involved as early as possible within each activity, such as:

- Assisting and supporting the development of service strategy
- Involvement in the review and improvement of IT strategies and policies
- Involvement in the review and improvement of technology architectures.

Key message

Capacity management should not be a last-minute 'tick in the box' just prior to customer acceptance and operational acceptance.

If early involvement can be achieved from capacity management within these activities, then the planning and design of IT capacity can be closely aligned with business requirements and can ensure that service targets can be achieved and maintained.

Assist with agreeing service level requirements

Capacity management should assist SLM in understanding the customers' capacity and performance requirements, in terms of required service/system response times, expected throughput, patterns of usage and volume of users. Capacity management should help in the negotiation process by providing possible solutions to a number of scenarios. For example, if the number of users is fewer than 2,000, then response times can be guaranteed to be less than two seconds. If more than 2,000 users connect concurrently, then extra network bandwidth is needed to guarantee the required response time, or a slower response time will have to be accepted. Modelling, trending or application sizing techniques are often employed here to ensure that predictions accurately reflect the real situation.

Design, procure or amend service configuration

Capacity management should be involved in the design of new or changing services and make recommendations for the procurement of hardware and software, where performance and/or capacity are factors. In some instances capacity management instigates the implementation of the new requirement through change management, where it is also represented on the change advisory board. In the interest of balancing cost and capacity, the capacity management process obtains the costs of alternative proposed solutions and recommends the most appropriate cost-effective solution.

Verify service level agreements

The SLA should include details of the anticipated service throughputs and the performance requirements. Capacity management advises SLM on achievable targets that can be monitored and on which the service design has been based. Confidence that the service design will meet the SLRs and provide the ability for future growth can be gained by using modelling, trending or sizing techniques.

Support service level agreement negotiation

The results of the predictive techniques provide the verification of service performance capabilities. There may be a need for SLM to renegotiate SLAs based on these findings. Capacity management provides support to SLM should renegotiations be necessary, by recommending potential solutions and associated cost information. Once assured that the requirements are achievable, it is the responsibility of SLM to agree the service levels and sign the SLA.

Control and implementation

All changes to service and resource capacity must follow all IT processes such as change, release, configuration and project management to ensure that the right degree of control and coordination is in place for all changes and that any new or change components are recorded and tracked through their lifecycle.

4.5.5.2 Service capacity management

The main 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 meet their SLA targets, i.e. 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 agreed SLAs or SLRs.

The service capacity management sub-process ensures that the services meet the agreed capacity service targets. The monitored service provides 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 on. Therefore capacity management informs SLM of any service breaches or near misses.

There will be occasions when incidents and problems are referred to capacity management from other processes, or it is 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 component capacity management. For example, when the failure is analysed it may be found that there is no lack of capacity, or no individual component is over-utilized. However, if the design or coding of an application is inefficient, then the service performance may need to be managed, as well as individual hardware or software resources. Service capacity management should also be monitoring service workloads and transactions to ensure that they remain within agreed limitations and thresholds.

The key to successful service capacity management is to forecast issues, wherever possible, by monitoring changes in performance and monitoring the impact of changes. So this is another sub-process that, whenever possible, has to be proactive and predictive, even pre-emptive, rather than reactive. However, there are 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 used, 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.

4.5.5.3 Component capacity management

The main objective of component capacity management is to identify and understand the performance, capacity and utilization of each of the individual components within the technology used to support the IT services, including the infrastructure, environment, data and applications. 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 that, when approached or exceeded, has the potential to cause performance problems.

This sub-process is concerned with components such as processors, memory, disks, network bandwidth, network connections etc. So information on resource utilization needs to be collected on a continuous basis. Monitors should be installed on the individual hardware and software components, and then configured to collect the necessary data, which is accumulated and stored over a period of time. This is an activity generally carried out through monitoring and control within service operation. A direct feedback to component capacity management should be applied within this sub-process.

As in service capacity management, the key to successful component capacity management is to forecast issues, wherever possible, and it therefore has to be proactive and predictive as well. However, there are times when component capacity management has to react to specific problems that are caused by a lack of capacity, or the inefficient use of the component. 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 and hardware or software upgrades can be budgeted and planned. Alternatively, services can be balanced across the existing resources to make most effective use of the current resources.

4.5.5.4 Design-related activities

The three sub-processes of capacity management can all benefit from attention to the exploitation of new technology and to designing resilience into our services and infrastructure. While both of these activities may be seen as associating most directly with component capacity management, they may also be applied to service capacity management and business capacity management. Organizations should aspire to be as proactive as possible in the performance of these activities.

Exploitation of new technology

This involves understanding new techniques and new technology and how they can be used to support the business and innovate improvements. It may be appropriate to introduce new technology to improve the provision and support of the IT services on which the organization is dependent. This information can be gathered by studying professional literature (magazine and press articles) and 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 provides sources of information relating to potential techniques, technology, hardware and software, which might be advantageous for IT to implement to realize business benefits. However, at all times capacity management should recognize that the introduction and use of this new technology must be cost-justified and deliver real benefit to the business. It is not just the new technology itself that is important, but capacity management should also keep aware of the advantages to be obtained from the use of new technologies, using techniques such as 'grid computing', 'virtualization' and 'on-demand computing'.

Designing resilience

Capacity management assists with the identification and improvement of the resilience within the IT infrastructure or any subset of it, wherever it is cost-justified. In conjunction with availability management, capacity management should use techniques such as CFIA (as described in section 4.4 on availability management) to identify how susceptible the current configuration is to the failure or overload of individual components and make recommendations on any cost-effective solutions.

Capacity management should 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. So the provision of spare capacity can act as resilience or fail-over in failure situations.

The requirements for resilience in the IT infrastructure should always be considered at the time of the service or system design. However, for many services, the resilience of the service is only considered after it is in live operational use. Incorporating resilience into service design is much more effective and efficient than trying to add it at a later date, once a service has become operational.

4.5.5.5 The ongoing iterative activities of capacity management

The activities described in this section are necessary to support the sub-processes of capacity management, and these activities can be done both reactively and proactively.

The major differences between the sub-processes are in the data that is being monitored and collected, and the perspective from which the data is analysed. For example, the level of utilization of individual components in the infrastructure – such as processors, disks, and network links – is of interest in component capacity management, while the transaction throughput rates and response times of the entire service are of interest in service capacity management. For business capacity management, the transaction throughput rates for the online service need to be translated into business volumes – for example, in terms of sales invoices raised or orders taken. The biggest challenge facing capacity management is to understand the relationship between the demands and requirements of the business and the business workload, and to be able to translate these in terms of the impact and effect of these on the service and resource workloads and utilizations, so that appropriate thresholds can be set at each level.

A number of the activities need to be carried out iteratively and form a natural cycle, as illustrated in Figure 4.20.

These activities provide the basic historical information and triggers necessary for all of the other activities and processes within 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

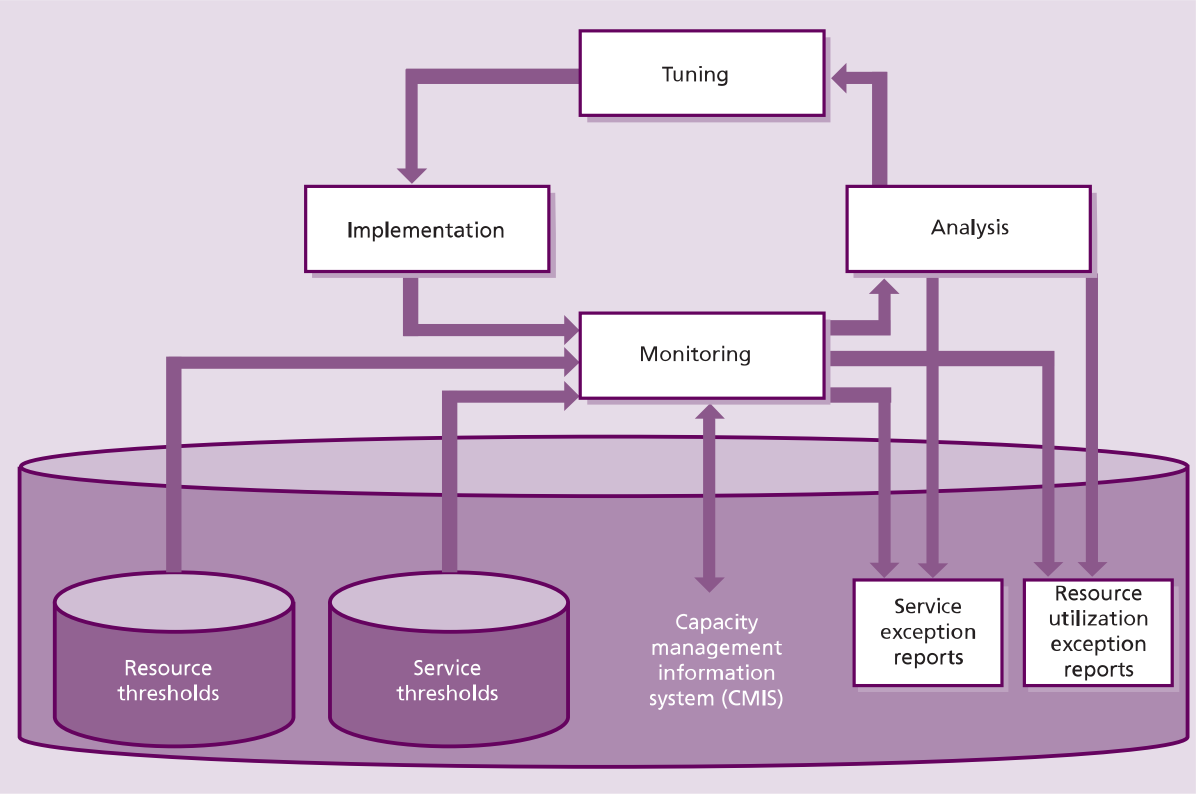

Tuning
Implementation
Analysis
Monitoring
Resource thresholds
Service thresholds
Capacity management information system (CMIS)
Service exception reports
Resource utilization exception reports

Figure 4.20 Ongoing iterative activities of capacity management

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, balancing workloads, changing concurrency levels and adding or removing resources. All of the information accumulated during these activities should be stored in the capacity management information system (CMIS) and the cycle then begins again, monitoring any changes made to ensure they have had a beneficial effect and collecting more data for future actions. These iterative activities are primarily performed as part of the service operation stage of the service lifecycle.

Monitoring

The monitors should be specific to particular operating systems, hardware configurations, applications etc. 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 includes:

- Processor utilization
- Memory utilization
- Per cent processor per transaction type
- I/O rates (physical and buffer) and device utilization
- Queue lengths
- Disk utilization
- Transaction rates
- Response times
- Batch duration
- Database usage
- Index usage
- Hit rates
- Concurrent user numbers
- Network traffic rates.

In considering the data that needs to be included, a distinction needs to be drawn between the data collected to monitor capacity (e.g. throughput) and the data to monitor performance (e.g. response times). Data of both types is required by the service and component capacity management sub-processes. This monitoring and collection needs to incorporate all components in the service, thus monitoring the 'end-to-end' customer experience. The data should be gathered at total resource utilization level and at a more detailed profile for the load that each service places on each particular component. This needs to be carried out across the whole technology, host or server, the network, local server and client or workstation. Similarly the data needs to be collected for each service.

Threshold management and control

Part of the monitoring activity should be of thresholds and baselines or profiles of the normal operating levels. If these are exceeded, alarms should be raised and exception reports produced. These thresholds and baselines should have been determined from the analysis of previously recorded data, and can be set at both the component and service levels. The technical limits and constraints on the individual services and components can be used by the monitoring activities to set the thresholds at which warnings and 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.

All thresholds should be set below the level at which the component or service is over-utilized, or below the targets in the SLAs. When the threshold is reached or threatened, there is still an opportunity to take corrective action before the SLA has been breached, or the resource has become over-utilized and there has been a period of poor performance. The monitoring and management of these events, thresholds and alarms is covered in detail in *ITIL Service Operation*.

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 component capacity management sub-processes.

The management and control of service and component thresholds is fundamental to the effective delivery of services to meet their agreed service levels. It ensures that all service and component thresholds are maintained at the appropriate levels and are continuously, automatically monitored, and alerts and warnings generated when breaches occur. Once defined, thresholds and how they are to be implemented and used should be documented as part of the service design package.

Whenever monitored thresholds are breached or threatened, alarms are raised and breaches, warnings and exception reports are produced. Analysis of the situation should then be completed and remedial action taken whenever justified, ensuring that the situation does not recur. The same data items can be used to identify when SLAs are breached or likely to be breached or when component performance degrades or is likely to be degraded. By setting thresholds below or above the actual targets, action can be taken and a breach of the SLA targets avoided.

Threshold monitoring should not only alarm on exceeding a threshold, but should also monitor the rate of change and predict when the threshold will be reached. For example, a disk-space monitor should monitor the rate of growth and raise an alarm when the current rate will cause the disk to be full within the next N days. If a 1GB disk has reached 90% capacity, and is growing at 100KB per day, it will be 1,000 days before it is full. If it is growing at 10MB per day, it will only be 10 days before it is full. The monitoring and management of these events and alarms is covered in detail in *ITIL Service Operation*.

There may be occasions when optimization of infrastructure components and resources is needed to maintain or improve performance or throughput. This can often be done through workload management, which is a generic term to cover such actions as:

- Rescheduling a particular service or workload to run at a different time of day or day of the week etc. (usually away from peak times to off-peak windows) – which will often mean having to make adjustments to job-scheduling software
- Moving a service or workload from one location or set of CIs to another – often to balance utilization or traffic
- Technical 'virtualization': setting up and using virtualization techniques and systems to allow the movement of processing around the infrastructure to give better performance/ resilience in a dynamic fashion
- Limiting or moving demand for components or resources through demand management techniques, in conjunction with financial management for IT services (see section 4.5.5.6).

It will only be possible to manage workloads effectively if a good understanding exists of which workloads will run at what time and how much resource utilization each workload places on the IT infrastructure. Diligent monitoring and analysis of workloads, together with a comprehensive CMIS, are therefore needed on an ongoing operational basis.

Response time monitoring

Many SLAs have user response times as one of the targets to be measured, but equally many organizations have great difficulty in supporting this requirement. User response times of IT and network services can be monitored and measured in the following ways:

- **Incorporating specific code within client and server applications software** This 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 give the actual response times as perceived by the users of a service.
- **Using 'robotic scripted systems' with terminal emulation software** These systems consist of client systems with terminal emulation software (e.g. browser or Telnet systems) and specialized scripted software for generating and measuring transactions and responses. These systems generally provide sample 'end-to-end' service response times and are useful for providing representative response times, particularly for multi-phase transactions or complex interactions. These only give sample response times, not the actual response times as perceived by the real users of the service.
- **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 (e.g. 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 website. However, again the times received are only indications of the response times and are not the real user response times.
- **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', which are 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 systems can only be used to give an approximation to the actual user response times. These times are often very close to the real-world situation, although this depends on 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 it is an in-house service running on a private network. If this is an external internet service, the process is much more complex because of the sheer number of different organizations and technologies involved.

Anecdote

A private company with a major website implemented a website monitoring service from an external supplier that would provide automatic alarms on the availability and responsiveness of its website. The availability and speed of the monitoring points were lower than those of the website being monitored. Therefore the figures produced by the service were of the availability and responsiveness of the monitoring service itself, rather than those of the monitored website.

Hints and tips

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.

Analysis

The data collected from the monitoring should be analysed to identify trends from which the normal utilization and service levels, or baselines, can be established. By regular monitoring and comparison with this baseline, exception conditions in the utilization of individual components or service thresholds can be defined, and breaches or near misses in the SLAs can be reported and actioned.

Also the data can be used to predict future resource usage, or to monitor actual business growth against predicted growth.

Analysis of the data may identify issues such as:

- 'Bottlenecks' or 'hot spots' within the infrastructure
- Inappropriate distribution of workload across available resources
- Inappropriate database indexing
- Inefficiencies in the application design
- Unexpected increase in workloads or transaction rates
- Inefficient scheduling or memory usage.

The use of each component and service needs to be considered over the short, medium and long term, and the minimum, maximum and average utilization for these periods recorded. Typically, the short-term pattern covers the utilization over a 24-hour period, while the medium term may cover a one- 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. The usefulness of this information is further enhanced by recording any observed contributing factors to peaks or valleys in utilization – for example, if a change of business process or staffing coincides with any deviations from normal utilization.

It is important to understand the utilization in each of these periods, so that changes in the use of any service can be related to predicted changes in the level of utilization of individual components. The ability to identify the specific hardware or software components on which a particular IT service depends is improved greatly by an accurate, up-to-date and comprehensive CMS.

When the utilization of a particular resource is considered, it is important to understand both the total level of utilization and the utilization by individual services of the resource.

Understanding both the individual pieces and the whole

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 service 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.

Tuning

The analysis of the monitored data may identify areas of the configuration that could be tuned to better utilize the service, system and component resources or improve the performance of the particular service.

Tuning techniques that are of assistance include:

- **Balancing workloads and traffic** Transactions may arrive at the host or server at a particular gateway, depending on where the transaction was initiated; balancing the ratio of initiation points to gateways can provide tuning benefits.
- **Balancing disk traffic** Storing data on disk efficiently and strategically – for example, striping data across many spindles may reduce data contention.
- **Definition of an accepted locking strategy** This specifies when locks are necessary and the appropriate level – for example, database, page, file, record and row. Delaying the lock until an update is necessary may provide benefits.
- **Efficient use of memory** This may include looking to utilize more or less memory, depending on the circumstances.

Before implementing any of the recommendations arising from the tuning techniques, it may be appropriate to consider testing the validity of the recommendation. For example, 'Can 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?'

Implementation

The objective of this activity is to introduce to the live operation services any changes that have been identified by the monitoring, analysis and tuning activities. 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 can have major implications on the customers of the service. The impact and risk associated with these types of change are likely to be greater than that of other types of change.

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.

4.5.5.6 Demand management in capacity management

The prime objective of demand management at the tactical level is to influence user and customer demand for IT services and manage the impact on IT resources. Information provided by the strategic demand management process (see *ITIL Service Strategy*) is an important input to the type of demand management occurring as part of the capacity management process.

This tactical demand management 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 a processor within a multi-processor server, 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 processor power required to run the service) and then be able to identify which services can be run while there is a limited amount of processor power available.

Long-term demand management may be required when it is difficult to cost-justify an expensive upgrade. For example, many processors are heavily utilized for only a few hours each day, typically 10.00–12.00 and 14.00–16.00. Within these periods, the processor may be overloaded for only one or two hours. For the hours between 18.00 and 08.00, these processors are only very lightly loaded and the components are under-utilized. Is it possible to justify the cost of an upgrade to provide additional capacity for only a few hours in 24 hours? Or is it possible to influence the demand and spread the requirement for resource across 24 hours, thereby delaying or avoiding altogether the need for a costly upgrade?

Demand management needs to understand which services are utilizing the resource, to what level, and the schedule for its use. Then a decision can be made 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 being available at certain times, or to limit the number of customers who can use a particular service by limiting the number of concurrent users; or 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 in place, reduced rates could be offered for running work at times of the day when there is currently less demand for the resource. This is known as differential charging.

4.5.5.7 Modelling and trending

A prime objective of capacity management is to predict the behaviour of IT services under a given volume and variety of work. Modelling is an activity that can be used to accomplish this to beneficial effect in any of the sub-processes of capacity management.

The different types of modelling range from making estimates based on experience and current resource utilization information, to pilot studies, prototypes and full-scale benchmarks. The former is a cheap and reasonable approach for day-to-day small decisions, while the latter is expensive, but may be advisable when implementing a large new project or service. With all types of modelling, similar levels of accuracy can be obtained, but all are totally dependent on the skill of the person constructing the model and the information used to create it.

Baselining

The first stage in modelling is to create a baseline model that reflects accurately the performance that is currently being achieved. When this baseline model has been created, predictive modelling can be done, i.e. ask the 'What if?' questions that reflect failures, planned changes to the hardware and/or the volume/variety of workloads. If the baseline model is accurate, then the accuracy of the result of the potential failures and changes can be trusted.

Effective capacity management, together with modelling techniques, enables 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 server onto a new server – what will be the effect on the response times of the two services?

Trend analysis

Trend analysis can be done on the resource utilization and service performance information that has been collected by the capacity management process. The data can be analysed in a spreadsheet, and the graphical and trending and forecasting facilities used to show the utilization of a particular resource over a previous period of time, and how it can be expected to change in the future.

Typically, trend analysis only provides estimates of future resource utilization information. Trend analysis is less effective in producing an accurate estimate of response times, in which case either analytical or simulation modelling should be used. Trend analysis is most effective when there is a linear relationship between a small number of variables, and less effective when there are non-linear relationships between variables or when there are many variables.

Analytical modelling

Analytical models are representations of the behaviour of computer systems using mathematical techniques – for example, 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 need to be modelled, and the utilization of the components – for example, processor, memory and disks – 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 utilization, and 15–20% for online application response times, the results are usually satisfactory.

Simulation modelling

Simulation involves the modelling of discrete events (for example, 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.

When simulating transaction arrival rates, have a number of staff enter a series of transactions from prepared scripts, or use software to input the same scripted transactions with a random arrival rate. Either of these approaches takes time and effort to prepare and run. However, it can be cost-justified for organizations with very large services and systems where the major cost and the associated performance implications assume great importance.

4.5.5.8 Application sizing

The primary objective of application sizing is to estimate the resource requirements to support a proposed change to an existing service or the implementation of a new service, to ensure that it meets its required service levels. To achieve this, application sizing has to be an integral part of the service lifecycle.

Application sizing has a finite lifespan. It is initiated at the design stage for a new service, or when there is a major change to an existing service, and is completed when the application is accepted into the live operational environment. Sizing activities should include all areas of technology related to the applications, and not just the applications themselves. This should include the infrastructure, environment and data, and will often use modelling and trending techniques.

During the initial requirements and design, the required service levels must be specified in an SLR. This enables the service design and development to employ the pertinent technologies and products 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 service design considers the required service levels at the very beginning of the service lifecycle, rather than at some later stage.

Other considerations in application sizing are the resilience aspects that it may be necessary to build into the design of new services. Capacity management is able to provide advice and guidance to the availability management process on the resources required to provide the required level of performance and resilience.

The sizing of the application should be refined as design and development progress. Modelling can be used during application sizing.

The SLRs of the planned application developments should not be considered in isolation. The resources to be utilized by the application are likely to be shared with other services, and potential threats to existing SLA targets must be recognized and managed.

When purchasing software packages from external suppliers, it is just as important to understand the resource requirements needed to support the service. Often it can be difficult to obtain this information from the suppliers and it may vary, depending on throughput. Therefore, it is beneficial to identify similar customers of the product and to gain an understanding of the resource implications from them. It may be pertinent to benchmark, evaluate or 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 cost more to add additional capacity after service implementation rather than as part of the original project.

4.5.6 Triggers, inputs, outputs and interfaces

4.5.6.1 Triggers

There are many triggers that will initiate capacity management activities. These include:

- New and changed services requiring additional capacity
- Service breaches, capacity or performance events and alerts, including threshold events
- Exception reports
- Periodic revision of current capacity and performance and the review of forecasts, reports and plans
- Periodic trending and modelling
- Review and revision of business and IT plans and strategies
- Review and revision of designs and strategies
- Review and revision of SLAs, OLAs, contracts or any other agreements
- Request from SLM for assistance with capacity and/or performance targets and explanation of achievements.

4.5.6.2 Inputs

A number of sources of information are relevant to the capacity management process. Some of these are as follows.

- **Business information** From the organization's business strategy, plans and financial plans, and information on their current and future requirements
- **Service and IT information** From service strategy, the IT strategy and plans and current budgets, covering all areas of technology and technology plans, including the infrastructure,

environment, data and applications, and the way in which they relate to business strategy and plans

- **Component performance and capacity information** Of both existing and new technology, from manufacturers and suppliers
- **Service performance issue information** The incident and problem management processes, with incidents and problems relating to poor performance
- **Service information** From the SLM process, with details of the services from the service portfolio and the service catalogue and service level targets within SLAs and SLRs, and possibly from the monitoring of SLAs, service reviews and breaches of the SLAs
- **Financial information** From financial management for IT services, the cost of service provision, the cost of resources, components and upgrades, the resultant business benefit and the financial plans and budgets, together with the costs associated with service and component failure. Some of the costs of components and upgrades to components will be obtained from procurement, suppliers and manufacturers
- **Change information** From the change management process, with a change schedule and a need to assess all changes for their impact on the capacity of the technology
- **Performance information** From the CMIS on the current performance of both all existing services and IT infrastructure components
- **CMS** Containing information on the relationships between the business, the services, the supporting services and the technology
- **Workload information** From the IT operations team, with schedules of all the work that needs to be run, and information on the dependencies between different services and information, and the interdependencies within a service.

4.5.6.3 Outputs

The outputs of capacity management are used within all other parts of the process, by many other processes and by other parts of the organization. Often this information is supplied as electronic reports or displays on shared areas, or as pages on intranet servers, to ensure the most up-to-date information is always used. The information provided is as follows:

- **CMIS** This holds the information needed by all sub-processes within capacity management. For example, the data monitored and collected as part of component and service capacity management is used in business capacity management to determine what infrastructure components or upgrades to components are needed, and when.
- **Capacity plan** This is used by all areas of the business and IT management, and is acted on by the IT service provider and senior management of the organization to plan the capacity of the IT infrastructure. It also provides planning input to many other areas of IT and the business. It contains information on the current usage of service and components, and plans for the development of IT capacity to meet the needs in the growth of both existing service and any agreed new services. The capacity plan should be actively used as a basis for decision-making. Too often, capacity plans are created and never referred to or used.
- **Service performance information and reports** This is used by many other processes. For example, the capacity management process assists SLM with the reporting and reviewing of service performance and the development of new SLRs or changes to existing SLAs. It also assists the financial management for IT services process by identifying when money needs to be budgeted for IT infrastructure upgrades, or the purchase of new components.
- **Workload analysis and reports** This is used by IT operations to assess and implement changes in conjunction with capacity management to schedule or reschedule when services or workloads are run, to ensure that the most effective and efficient use is made of the available resources.
- **Ad hoc capacity and performance reports** These are used by all areas of capacity management, IT and the business to analyse and resolve service and performance issues.
- **Forecasts and predictive reports** These are used by all areas to analyse, predict and forecast particular business and IT scenarios and their potential solutions.
- **Thresholds, alerts and events**
- **Improvement actions** For inclusion in a SIP.

4.5.6.4 Interfaces

The key interfaces that capacity management has with other processes are:

- **Availability management** This process works with capacity management to determine the resources needed to ensure the required availability of services and components.
- **Service level management** This process provides assistance with the determining capacity targets and the investigation and resolution of service and component capacity-related breaches.
- **ITSCM** Capacity management assists with the assessment of business impact and risk and determining the capacity needed to support risk reduction measures and recovery options.
- **Incident and problem management** Capacity management provides assistance with the resolution and subsequent justification and correction of capacity-related incidents and problems.
- **Demand management** By anticipating the demand for services based on user profiles and patterns of business activity, and identifying the means to influence that demand, this process provides strategic decision-making and critical related data on which capacity management can act.

4.5.7 Information management

The aim of the CMIS is to provide the relevant capacity and performance information to produce reports and support the capacity management process. These reports provide valuable information to many IT and service management processes. These reports should include the following:

- **Component-based reports** For each component there should be a team of technical staff responsible for its control and management. Reports must be produced to illustrate how components are performing and how much of their maximum capacity is being used.
- **Service-based reports** Reports and information must also be produced to illustrate how the service and its constituent components are performing with respect to their overall service targets and constraints. These reports will provide the basis of SLM and customer service reports.
- **Exception reports** Reports that show management and technical staff when the capacity and performance of a particular component or service becomes unacceptable are required from analysis of capacity data. Thresholds can be set for any component, service or measurement within the CMIS. An example threshold may be that processor percentage utilization for a particular server has breached 70% for three consecutive hours, or that the concurrent number of logged-in users exceeds the agreed limit.

  In particular, exception reports are of interest to the SLM 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.

  While the focus of exception reporting usually focuses on indications of insufficient capacity, excess capacity should also be identified. Unused capacity may represent an opportunity for cost savings.
- **Predictive and forecast reports** To ensure the IT service provider continues to provide the required service levels, the capacity management process must predict future workloads and growth. To do this, future component and service capacity and performance must be forecast. This can be done in a variety of ways, depending on the techniques and the technology used. Changes to workloads by the development and implementation of new functionality and services must be considered alongside growth in the current functionality and services driven by business growth. A simple example of a capacity forecast is a correlation between a business driver and component utilization – for example, processor utilization against the number of customer accounts. This data can be correlated to find the effect that an increase in the number of customer accounts will have on processor utilization. If the forecasts on future capacity requirements identify a requirement for increased resource, this requirement needs to be input into the capacity plan and included within the IT budget cycle.

Often capacity reports are consolidated together and stored on an intranet site so that anyone can access and refer to them.

4.5.7.1 Capacity management information system data

Often capacity data is stored in technology-specific tools and databases, and full value of the data, the information and its analysis is not obtained. The true value of the data can only be obtained when the data is combined into a single set of integrated, information repositories or set of databases.

The CMIS is a set of tools, data and information that is used to support capacity management and is the cornerstone of a successful capacity management process. Information contained within the CMIS is stored and analysed by all the sub-processes of capacity management because it is a repository that holds a number of different types of data, including business, service, resource or utilization and financial data, from all areas of technology.

However, the CMIS is unlikely to be a single database, and probably exists in several physical locations. Data from all areas of technology, and all components that make up the IT services, can then be combined for analysis and provision of technical and management reporting. Only when all of the information is integrated can 'end-to-end' service reports be produced. The integrity and accuracy of the data within the CMIS need to be carefully managed. If the CMIS is not part of an overall CMS or SKMS, then links between these systems need to be implemented to ensure consistency and accuracy of the information recorded within them.

The information in the CMIS is used to form the basis of performance and capacity management reports and views that are to be delivered to customers, IT management and technical personnel. Also, the data is utilized to generate future capacity forecasts and allow capacity management to plan for future capacity requirements. Often a web interface is provided to the CMIS to provide the different access and views required outside of the capacity management process itself.

The full range of data types stored within the CMIS is as follows:

- **Business data** It is essential to have quality information on the current and future needs of the business. The future business plans of the organization 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 IT infrastructure. Business data should include business transactions or measurements such as the number of accounts, the number of invoices generated, the number of product lines.
- **Service data** To achieve a service-orientated approach to capacity management, service data should be stored within the CMIS. Typical service data are transaction response times, transaction rates, workload volumes etc. In general, the SLAs and SLRs provide the service targets for which the capacity management process needs to record and monitor data. To ensure that the targets in the SLAs are achieved, SLM thresholds should be included, so that the monitoring activity can measure against these service thresholds and raise exception warnings and reports before service targets are breached.
- **Component data** The CMIS also needs to record resource data consisting of utilization, threshold and limit information on all of the technological components supporting the services. Most of the IT components have limitations on the level to which they should be utilized. Beyond this level of utilization, the resource will be over-utilized and the performance of the services using the resource will be impaired. For example, the maximum recommended level of utilization on a processor could be 80%, or the utilization of a shared Ethernet LAN segment should not exceed 40%.

  Also, components have various physical limitations beyond which greater connectivity or use is impossible. For example, the maximum number of connections through an application or a network gateway is 100, or a particular type of disk has a physical capacity of 15 Gb. The CMIS should therefore contain, for each component and the maximum performance and capacity limits, current and past utilization rates and the associated component thresholds. Over time this can mean accumulation of vast amounts of data, so there need to be good techniques for analysing, aggregating and archiving this data.
- **Financial data** The capacity management process requires financial data. For evaluating alternative upgrade options, when proposing various scenarios in the capacity plan,

the financial cost of the upgrades to the components of the IT infrastructure, together with information about the current IT hardware budget, must be known and included in the considerations. 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.

4.5.8 Critical success factors and key performance indicators

The following list includes some sample CSFs for capacity management. Each organization should identify appropriate CSFs based on its objectives for the process. Each sample CSF is followed by a small number of typical KPIs that support the CSF. These KPIs should not be adopted without careful consideration. Each organization should develop KPIs that are appropriate for its level of maturity, its CSFs and its particular circumstances. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

- **CSF** Accurate business forecasts
  - **KPI** Production of workload forecasts on time
  - **KPI** Percentage accuracy of forecasts of business trends
  - **KPI** Timely incorporation of business plans into the capacity plan
  - **KPI** Reduction in the number of variances from the business plans and capacity plans
- **CSF** Knowledge of current and future technologies
  - **KPI** Increased ability to monitor performance and throughput of all services and components
  - **KPI** Timely justification and implementation of new technology in line with business requirements (time, cost and functionality)
  - **KPI** Reduction in the use of old technology, causing breached SLAs due to problems with support or performance
- **CSF** Ability to demonstrate cost effectiveness
  - **KPI** Reduction in last-minute buying to address urgent performance issues
  - **KPI** Reduction in the over-capacity of IT
  - **KPI** Accurate forecasts of planned expenditure
  - **KPI** Reduction in the business disruption caused by a lack of adequate IT capacity
  - **KPI** Relative reduction in the cost of production of the capacity plan
- **CSF** Ability to plan and implement the appropriate IT capacity to match business need
  - **KPI** Percentage reduction in the number of incidents due to poor performance
  - **KPI** Percentage reduction in lost business due to inadequate capacity
  - **KPI** All new services implemented match SLRs
  - **KPI** Increased percentage of recommendations made by capacity management are acted on
  - **KPI** Reduction in the number of SLA breaches due to either poor service performance or poor component performance.

4.5.9 Challenges and risks

4.5.9.1 Challenges

One of the major challenges facing capacity management is persuading the business to provide information on its strategic business plans, to enable the IT service provider organization to provide effective business capacity management. This is particularly true in outsourced situations where there may be commercial or confidential reasons why this data cannot be shared. Even if the data on the strategic business is available there may be issues with regard to the quality or accuracy of the data contained within the business plans with regard to business capacity management.

Another challenge is the combination of all of the component capacity management data into an integrated set of information that can be analysed in a consistent manner to provide details of the usage of all components of the services. This is particularly challenging when the information from the different technologies is provided by different tools in differing formats. Often the quality of component information on the performance of the technology is variable in both its quality and accuracy.

The amounts of information produced by business capacity management, and especially service capacity management and component capacity management, are huge and the analysis of this information is difficult to achieve. The people and the processes need to focus on the key resources and their usage, while not ignoring other areas. In order to do this, appropriate thresholds must be used, and reliance placed on the tools and technology to automatically manage the technology and provide warnings and alerts when things deviate significantly from the 'norm'.

4.5.9.2 Risks

Some of the major risks associated with capacity management include:

- A lack of commitment from the business to the capacity management process
- A lack of appropriate information from the business on future plans and strategies
- A lack of senior management commitment or a lack of resources and/or budget for the capacity management process
- Service capacity management and component capacity management performed in isolation because business capacity management is difficult, or there is a lack of appropriate and accurate business information
- The processes become too bureaucratic or manually intensive
- The processes focus too much on the technology (component capacity management) and not enough on the services (service capacity management) and the business (business capacity management)
- The reports and information provided are too bulky or too technical and do not give the information required or appropriate to the customers and the business.

4.6 IT SERVICE CONTINUITY MANAGEMENT

As technology is a core component of most business processes, continued or high availability of IT is critical to the survival of the business as a whole. This is achieved by introducing risk reduction measures and recovery options. Like all elements of IT service management, successful implementation of the ITSCM process can only be achieved with senior management commitment and the support of all members of the organization. Ongoing maintenance of the recovery capability is essential if it is to remain effective.

Service continuity is an essential part of the warranty of a service. If a service's continuity cannot be maintained and/or restored in accordance with the requirements of the business, then the business will not experience the value that has been promised. Without continuity the utility of the service cannot be accessed.

Recovery is understood

Since effecting recovery when required is a fundamental element of ITSCM, the concept of continuity in this context encompasses risk reduction and recovery. When the word 'continuity' is used in this section, recovery should be understood to be included.

4.6.1 Purpose and objectives

The purpose of the IT service continuity management process is to support the overall business continuity management (BCM) process by ensuring that, by managing the risks that could seriously affect IT services, the IT service provider can always provide minimum agreed business continuity-related service levels.

In support of and alignment with the BCM process, ITSCM uses formal risk assessment and management techniques to:

- Reduce risks to IT services to agreed acceptable levels
- Plan and prepare for the recovery of IT services.

For a definition of BCM, please see the glossary at the end of this publication.

The objectives of ITSCM are to:

- Produce and maintain a set of IT service continuity plans that support the overall business continuity plans of the organization
- Complete regular BIA exercises to ensure that all continuity plans are maintained in line with changing business impacts and requirements
- Conduct regular risk assessment and management exercises to manage IT services within an agreed level of business risk in conjunction with the business and the availability management and information security management processes

- Provide advice and guidance to all other areas of the business and IT on all continuity-related issues
- Ensure that appropriate continuity mechanisms are put in place to meet or exceed the agreed business continuity targets
- Assess the impact of all changes on the IT service continuity plans and supporting methods and procedures
- Ensure that proactive measures to improve the availability of services are implemented wherever it is cost-justifiable to do so
- Negotiate and agree contracts with suppliers for the provision of the necessary recovery capability to support all continuity plans in conjunction with the supplier management process.

4.6.2 Scope

ITSCM focuses on those events that the business considers significant enough to be treated as a 'disaster'. Less significant events will be dealt with as part of the incident management process. What constitutes a disaster will vary from organization to organization. The impact of a loss of a business process, such as financial loss, damage to reputation or regulatory breach, is measured through a BIA exercise, which determines the minimum critical requirements. The specific IT technical and service requirements are supported by ITSCM. The scope of ITSCM within an organization is determined by the organizational structure, culture and strategic direction (both business and technology) in terms of the services provided and how these develop and change over time.

ITSCM primarily considers the IT assets and configurations that support the business processes. If (following a disaster) it is 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 scope will need to take into account the number and location of the organization's offices and the services performed in each.

ITSCM does not usually directly cover longer-term risks such as those from changes in business direction, diversification, restructuring, major competitor failure, and so on. While these risks can have a significant impact on IT service elements and their continuity mechanisms, there is usually 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 overall business and IT change management programme.

Similarly, ITSCM does not usually cover minor technical faults (for example, non-critical disk failure), unless there is a possibility that the impact could have a major impact on the business. These risks would be expected to be covered mainly through the service desk and the incident management process, or resolved through the planning associated with the processes of availability management, problem management, change management, service asset and configuration management and 'business as usual' operational management.

The ITSCM process includes:

- The agreement of the scope of the ITSCM process and the policies adopted
- BIA to quantify the impact loss of IT service would have on the business
- Risk assessment and management – the risk identification and risk assessment to identify potential threats to continuity and the likelihood of the threats becoming reality. This also includes taking measures to manage the identified threats where this can be cost-justified. The approach to managing these threats will form the core of the ITSCM strategy and plans
- Production of an overall ITSCM strategy that must be integrated into the BCM strategy. This can be produced following the BIA and the development of the risk assessment, and is likely to include elements of risk reduction as well as selection of appropriate and comprehensive recovery options
- Production of an ITSCM plan, which again must be integrated with the overall BCM plans
- Testing of the plans
- Ongoing operation and maintenance of the plans.

4.6.3 Value to the business

ITSCM provides an invaluable role in supporting the BCM process. In many organizations, ITSCM is

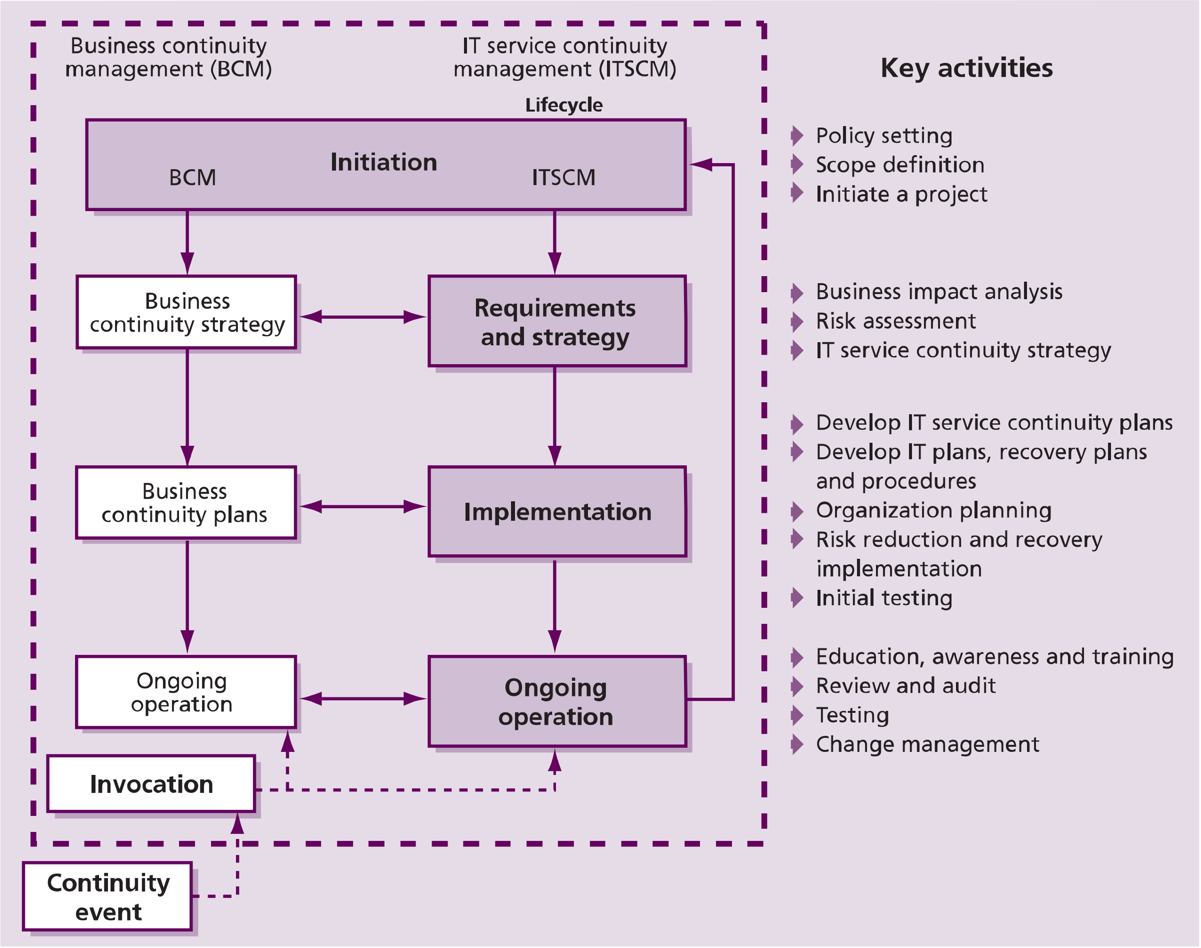

Business continuity management (BCM)
IT service continuity management (ITSCM)
Lifecycle
Initiation
BCM
ITSCM
Business continuity strategy
Requirements and strategy
Business continuity plans
Implementation
Ongoing operation
Ongoing operation
Invocation
Continuity event
Key activities
Policy setting
Scope definition
Initiate a project
Business impact analysis
Risk assessment
IT service continuity strategy
Develop IT service continuity plans
Develop IT plans, recovery plans and procedures
Organization planning
Risk reduction and recovery implementation
Initial testing
Education, awareness and training
Review and audit
Testing
Change management

Figure 4.21 Lifecycle of IT service continuity management

used to raise awareness of continuity requirements and is often used to justify and implement a BCM process and business continuity plans. ITSCM should be driven by business risk as identified by BCM, and ensure that the recovery arrangements for IT services are aligned to identified business impacts, risks and needs.

4.6.4 Policies, principles and basic concepts

A lifecycle approach should be adopted to the setting up and operation of an ITSCM process. Figure 4.21 shows the lifecycle of ITSCM, from initiation through to continual assurance that the protection provided by the plan is current and reflects all changes to services and service levels. ITSCM is a cyclic process through the lifecycle to ensure that once service continuity plans have been developed they are kept aligned with business continuity plans and business priorities. Figure 4.21 also shows the role played within the ITSCM process of BCM.

Initiation and, to a significant extent, the requirements stages are principally BCM activities. ITSCM should only be involved in these stages to support the BCM activities and to understand the relationship between the business processes and the impacts caused on them by loss of IT service. As a result of these initial BIA and risk assessment activities, BCM should produce a business continuity strategy, and the first real ITSCM task is to produce an ITSCM strategy that underpins the BCM strategy and its needs.

The business continuity strategy should principally focus on business processes and associated issues (e.g. business process continuity, staff continuity, buildings continuity). Once the business continuity strategy has been produced, and the role that IT services has to fulfil within the strategy has been determined, an ITSCM strategy can be produced that supports and enables the business continuity strategy. This ensures that cost-effective decisions can be made, considering all the 'resources' to deliver a business process. Failure to do this tends to encourage ITSCM options that are faster, more elaborate and more expensive than actually needed.

The activities to be considered during initiation depend on the extent to which continuity facilities have been applied within the organization. Some parts of the business may have established individual business continuity plans based around manual workarounds, and IT may have developed continuity plans for systems perceived to be critical. This is good input to the process. However, effective ITSCM depends on supporting vital business functions. The only way of implementing effective ITSCM is through the identification of critical business processes and the analysis and coordination of the required technology and supporting IT services.

This situation may be even more complicated in outsourcing situations where an ITSCM process within an external service provider or outsourcer organization has to meet the needs not only of the customer BCM process and strategy, but also of the outsourcer's own BCM process and strategy. These needs may be in conflict with one another, or may conflict with the BCM needs of one of the other outsourcing organization's customers.

However, in many organizations BCM is absent or has very little focus, and often ITSCM is required to fulfil many of the requirements and activities of BCM. The rest of this section has assumed that ITSCM has had to perform many of the activities required by BCM. Where a BCM process is established with business continuity strategies and plans in place, these documents should provide the focus and drive for establishing ITSCM.

4.6.5 Process activities, methods and techniques

The following sections contain details of each of the stages within the ITSCM lifecycle.

4.6.5.1 Stage 1 – Initiation

The initiation process covers the whole of the organization and consists of the following activities.

Policy setting

This should be established and communicated as soon as possible so that all members of the organization 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.

Define scope and specify terms of reference

This includes defining the scope and responsibilities of all staff in the organization. It covers 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 organization stipulations, and compliance with standards such as ISO/IEC 27001, the standard on information security management, which also addresses service continuity requirements.

Initiate a project

The initiation of formal IT service continuity management is best organized into a project. The project can be used to bring ITSCM to the 'ongoing operation' stage. Setting up the project includes:

- **Allocating resources** The establishment of an effective business continuity environment requires considerable resource in terms of both money and personnel. Depending on the maturity of the organization with respect to ITSCM, there may be a requirement to familiarize and/or train staff to accomplish stage 2 tasks. Alternatively, the use of experienced external consultants may assist in completing the analysis more quickly. However, it is important that the organization can then maintain the process going forward without the need to rely totally on external support.
- **Defining the project organization and control structure** ITSCM and BCM projects are potentially complex and need to be well organized and controlled. It is strongly advisable to use a recognized standard project planning methodology such as PRojects IN Controlled Environments (PRINCE2) or Project Management Body of Knowledge (PMBOK).
- **Agreeing 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.

4.6.5.2 Stage 2 – Requirements and strategy

Ascertaining the business requirements for IT service continuity is a critical component in order to determine how well an organization 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 BIA and risk assessment
- **Strategy** Following the requirements analysis, the strategy should document how the risks will be managed through risk reduction measures and recovery options required to support the business.

Requirements – business impact analysis

The purpose of a BIA is to quantify the impact to the business that loss of service would have. This impact could be a 'hard' impact that can be precisely identified – such as financial loss – or 'soft' impact – such as public relations, moral, health and safety or loss of competitive advantage. The BIA will identify the most important services to the organization and will therefore be a key input to the strategy.

The BIA identifies:

- The form that the damage or loss may take – for example:
  - Lost income
  - Additional costs
  - Damaged reputation
  - Loss of goodwill
  - Loss of competitive advantage
  - Breach of law, health and safety regulations
  - Risk to personal safety
  - Immediate and long-term loss of market share
  - Political, corporate or personal embarrassment
  - Loss of operational capability, for example, in a command and control environment
- How the degree of damage or loss is likely to escalate after a service disruption, and the times of the day, week, month or year when disruption will be most severe
- 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 relative business recovery priority for each of the IT services.

One of the key outputs from a BIA exercise is a graph of the anticipated business impact caused by the loss of a business process or the loss of an IT service over time, as illustrated in Figure 4.22.

This graph can then be used to drive the business and IT continuity strategies and plans. More preventive measures need to be adopted with regard to those processes and services with earlier and higher impacts, whereas greater emphasis should be placed on continuity and recovery measures for those where the impact is lower and takes longer to develop. A balanced approach of both measures should be adopted to those in between.

These items 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 on a cost-benefit analysis, or it may be that comprehensive disaster prevention measures will need to be implemented.

These assessments enable the mapping of critical service, application and technology components to critical business processes, thus helping to identify the ITSCM elements that need to be provided. The business requirements are ranked and the associated ITSCM elements confirmed and prioritized in terms of risk reduction and recovery planning. The results of the BIA, discussed earlier, are invaluable input to several areas of process design including SLM to understand the required service levels.

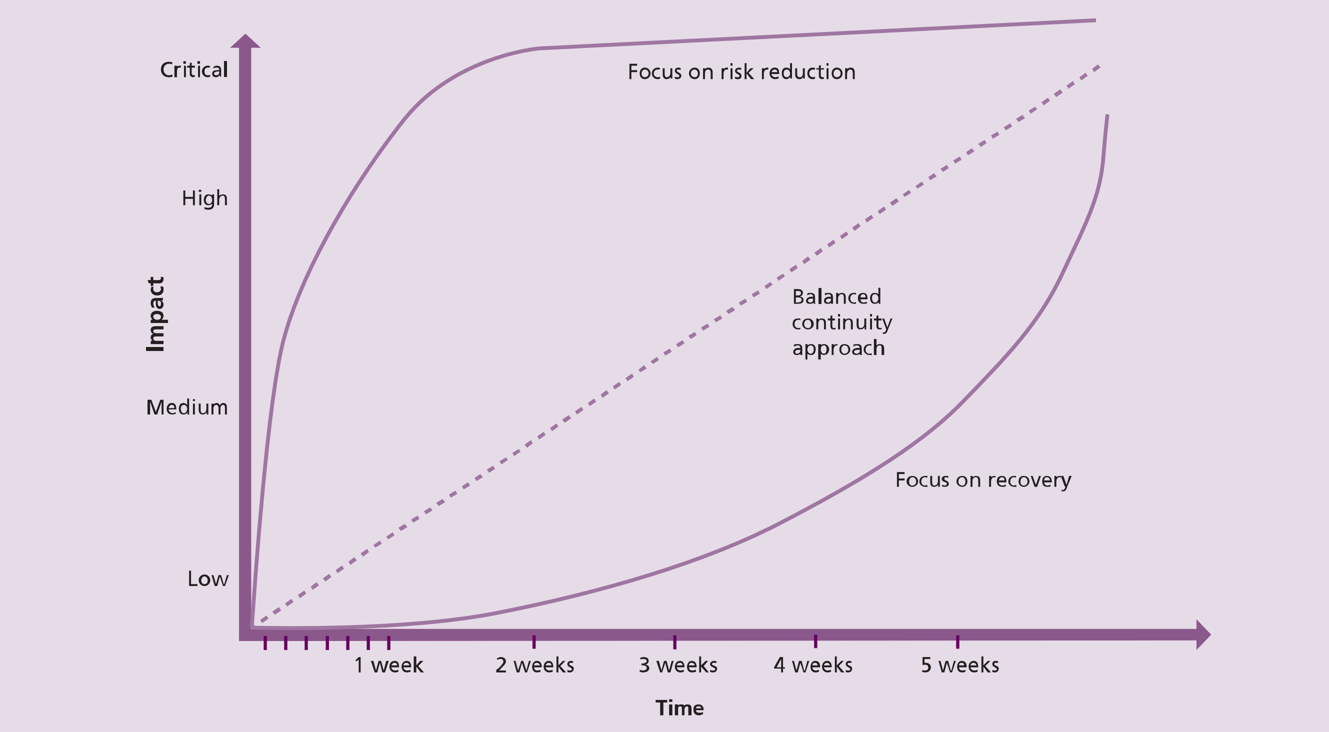


Figure 4.22 Graphical representation of business impacts

Impacts should be 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.

Example of business impact

An example is a money market dealing environment where loss of market data information could mean that the organization starts to lose money immediately as trading cannot continue. In addition, customers may go to another organization, which would mean potential loss of core business. Loss of the settlement system does not prevent trading from taking place, but if trades already conducted cannot be settled within a specified period of time, the organization may be in breach of regulatory rules or settlement periods and suffer fines and damaged reputation. 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. In a balanced scenario, impacts to the business will occur and become greater over time. However, not all organizations are affected in this way. In some organizations, impacts are not apparent immediately. At some point, however, for any organization, the impacts will accrue to such a level that the business can no longer operate. ITSCM ensures 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.

When conducting a BIA, it is important that senior business area representatives' views are sought on the impact following loss of service. It is also equally important that the views of supervisory staff and more junior staff are sought to ensure all aspects of the impact following loss of service are ascertained. Often different levels of staff will have different views on the impact, and all will have to be taken into account when producing the overall strategy.

In many organizations it will be impossible, or it will not be cost-justifiable, to recover the total service in a very short timescale. In many cases, business processes can be re-established without a full complement of staff, systems and other facilities, and 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
- 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.

Requirements – 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 organization is vulnerable to that threat. Risk assessment can also be used in assessing and reducing the chance of normal operational incidents and is a technique used by availability management to ensure the required availability and reliability levels can be maintained. Risk assessment is also a key aspect of information security management. A diagram on risk assessment and management (Figure 4.15) is contained within the availability management process in section 4.4.

A number of risk assessment and management methods are available for both the commercial and government sectors. Risk assessment is the assessment of the risks that may give rise to service disruption or security violation. Risk management is concerned with identifying appropriate risk responses or cost-justifiable countermeasures to combat those risks.

A standard methodology, such as the Management of Risk (M_o_R), should be used to assess and manage risks within an organization. The M_o_R framework is described in greater detail in Appendix M.

Conducting a formal risk assessment using M_o_R or another structured method will typically result in a risk profile, containing many risks that are outside the defined level of 'acceptable risk'. Following the risk assessment it is possible to determine appropriate risk responses or risk reduction measures (ITSCM mechanisms) to manage the risks, i.e. reduce the risk to an acceptable level or mitigate the risk. Wherever possible, appropriate risk responses should be implemented to reduce either the impact or the likelihood, or both, of these risks from manifesting themselves. In the context of ITSCM, there are a number of risks that need to be taken into consideration. Table 4.2 is not a comprehensive list but does give some examples of risks and threats that need to be addressed by the ITSCM process.

IT service continuity strategy

The results of the BIA and the risk assessment will enable appropriate business and IT service continuity strategies to be produced in line with the business needs. The strategy will be an optimum balance of risk reduction and recovery or continuity options. This includes 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. Those services that have been identified as high impacts in the short term within the BIA will want to concentrate efforts on preventive risk reduction methods – for example, through full resilience and fault tolerance – while an organization that has low short-term impacts would be better suited to comprehensive recovery options, as described in the following sections. Similar advice and guidance can be found in the Business Continuity Institute's *BCI Good Practice Guidelines*.

Risk response measures

Most organizations will 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 is usually 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 organization should assess all of the risks to reduce the potential

Table 4.2 Examples of risks and threats

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	All of the above Excessive demand for services Denial of service attack, e.g. against an internet firewall Technology failure, e.g. cryptographic system
Loss of data	Technology failure Human error Viruses, malicious software, e.g. attack applets
Loss of network services	Damage or denial of access to network service provider's premises Loss of service provider's IT systems/networks Loss of service provider's data Failure of the service provider
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

requirement to recover the business, which is likely to include the IT services.

The risk reduction measures need to be implemented and should be instigated in conjunction with availability management, as many of these reduce the probability of failure affecting the availability of service. Typical risk reduction measures include:

- Installation of uninterruptible power supply and backup power to the computer

- Fault-tolerant systems for critical applications where even minimal downtime is unacceptable – for example, a banking system
- RAID arrays and disk mirroring for LAN servers to prevent 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
- The elimination of SPOFs, such as single access network points or a single power supply into a building
- Resilient IT systems and networks
- Outsourcing services to more than one provider
- Greater physical and IT-based security controls
- Better controls to detect service disruptions, such as fire detection systems, coupled with suppression systems
- A comprehensive backup and recovery strategy, including off-site storage.

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. The detailed IT service continuity strategy will be developed to meet the agreed business needs, reflected in such specifications as recovery point objectives and recovery time objectives.

Off-site storage

One risk response method is to ensure all vital data is backed up and stored off-site. Once the recovery strategy has been defined, an appropriate backup strategy should be adopted and implemented to support it. The backup strategy must include regular (probably daily) removal of data (including the CMS to ease recovery) from the main data centres to a suitable off-site storage location. This will ensure retrieval of data following relatively minor operational failure as well as total and complete disasters. As well as the electronic data, all other important information and documents should be stored off-site, with the main example being the ITSCM plans.

ITSCM recovery options

An organization'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. The following is a list of the potential IT recovery options that need to be considered when developing the strategy.

Manual workarounds

For certain types of service, manual workarounds can be an effective interim measure for a limited timeframe until the IT service is resumed. For instance, a service desk call-logging service could survive for a limited time using paper forms linked to a laptop computer with a spreadsheet.

Reciprocal arrangements

In the past, reciprocal arrangements were typical contingency measures where agreements were put in place with another organization using similar technology. This is no longer effective or possible for most types of IT system, but can still be used in specific cases – for example, setting up an agreement to share high-speed printing facilities. Reciprocal arrangements can also be used for the off-site storage of backups and other critical information.

Gradual recovery

This option (sometimes referred to as 'cold standby') includes 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 organization to install its own computer equipment. It does not include the actual computing equipment, so is not applicable for services requiring speedy recovery, as setup time is required before recovery of services can begin. This recovery option is only recommended for services that can bear a delay of recovery time in days or weeks, not hours. Any non-critical service that can bear this type of delay should take into account the cost of this option versus the benefit to the business before determining if a gradual recovery option should be included in the ITSCM options for the organization.

The accommodation may be provided commercially by a third party (for a fee) or may be private (established by the organization itself) and provided as either a fixed or portable facility.

A portable facility is typically a prefabricated building provided by a third party and located, when needed, at a predetermined site agreed with the organization. This may be in another location some distance from the home site, perhaps another owned building. The replacement computer equipment will need to be planned, but suppliers of computing equipment do not always guarantee replacement equipment within a fixed deadline, though they would normally do so under their best efforts.

Intermediate recovery

This option (sometimes referred to as 'warm standby') is selected by organizations that need to recover IT facilities within a predetermined time or recovery time objective to prevent impacts to the business process. The predetermined time will have been agreed with the business during the BIA.

Most common is the use of commercial facilities, which are offered by third-party recovery organizations 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.

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 a disaster. It must be understood, however, that the restoration of services at the site may take some time, as delays may be encountered while the site is reconfigured for the organization that invokes the service, and the organization's applications and data will need to be restored from backups.

One potentially major disadvantage is the security implications of running IT services at a third party's data centre. This must be taken into account when planning to use this type of facility. For some organizations, the external intermediate recovery option may not be appropriate for this reason. 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.

Commercial recovery services can be provided in self-contained, portable or mobile form where an agreed system is delivered to a customer's site, within an agreed time.

Fast recovery

This option (sometimes referred to as 'hot standby') provides for fast recovery and restoration of services and is sometimes provided as an extension to the intermediate recovery provided by a third-party recovery provider. Some organizations will provide their own facilities within the organization, but not on an alternative site to the one used for the normal operations. Others implement their own internal second locations on an alternative site to provide more resilient recovery.

Where there is a need for a fast restoration of a service, it is possible to 'rent' floor space 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 customers can then recover and switch over to the backup facility with little loss of service. This typically involves the re-establishment of the critical systems and services within a 24-hour period.

Immediate recovery

This option (also often referred to as 'hot standby', 'mirroring', 'load balancing' or 'split site') provides for immediate restoration of services, with no significant loss of service to the customer. For business critical services, organizations requiring continuous operation will provide their own facilities within the organization, but not on the same site as the normal operations. Sufficient IT equipment will be 'dual located' in either an owned or hosted location to run the compete service from either location in the event of loss of one facility, with no loss of service to the customer. The second site can then be recovered while the service is provided from the single operable location. This is an expensive option, but may be justified for critical business processes or VBFs where non-availability for a short period could result in a significant impact, or where it would not be appropriate to be running IT services on a third party's premises for security or other reasons. The facility needs to be located separately and far enough away from the home site so that it will not

Table 4.3 Example set of recovery options

	Manual	Immediate	Fast	Intermediate	Gradual
Service desk	Yes		Yes	Yes	Yes
Mainframe payroll	Yes			Yes	Yes
Financial system			Yes		Yes
Dealer system		Yes		Yes	Yes

be affected by a disaster affecting that location. However, these mirrored servers and sites options should be implemented in close liaison with availability management as they support services with high levels of availability.

The strategy is likely to include a combination of risk response measures and a combination of the above recovery options, as illustrated in Table 4.3.

Table 4.3 shows that a number of options may be used to provide continuity of service and that, initially, continuity of the service desk is provided using manual processes such as a set of forms, and maybe a spreadsheet operating from a laptop computer, while recovery plans for the service are completed on an alternative 'fast recovery' site. Once the alternative site has become operational, the service desk can switch back to using the IT service. However, use of the external 'fast recovery' alternative site is probably limited in duration, so while running temporarily from this site, the 'intermediate site' can be made operational and long-term operations can be transferred there.

Different services within an organization require different in-built resilience and different recovery options. Whatever option is chosen, the solution will need to be cost-justified. As a general rule, the longer the business can survive without a service, the cheaper the solution will be. For example, a critical healthcare system that requires continuous operation will be very costly, as potential loss of service will need to be eliminated by the use of immediate recovery, whereas a service whose absence does not severely affect the business for a week or so could be supported by a much cheaper solution, such as intermediate recovery.

As well as the recovery of the computing equipment, planning needs to include the recovery of accommodation and infrastructure for both IT and user staff. Other areas to be taken into account include critical services such as power, telecommunications, water, couriers, post, paper records and reference material.

4.6.5.3 Stage 3 – Implementation

Once the strategy has been approved, the detailed IT service continuity plans need to be produced in line with the business continuity plans and the measures to implement the strategy need to be put in place. The measures to implement the strategy will include putting in place both the defined risk reduction and recovery option arrangements and performing initial testing to ensure that what was planned has been achieved.

Develop IT service continuity plans and procedures

ITSCM 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 reinstated within an acceptable period to the business. An example of an ITSCM recovery plan, a key part of the overall IT service continuity plan, is given in Appendix K. Generally the business continuity plans rely on the availability of IT services, facilities and resources. As a consequence of this, ITSCM plans need to address all activities to ensure that the required services, facilities and resources are delivered in an acceptable operational state and are 'fit for purpose' when accepted by the business. This entails not only the restoration of services 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.

It should be noted that the continuity plans are more than just recovery plans, and should include documentation of the resilience measures and the measures that have been put into place to enable recovery, together with explanations of why a particular approach has been taken (this facilitates decisions should invocation determine that the particular situation requires a modification to the plan). However, the format of the plan should enable rapid access to the recovery information itself, perhaps as an appendix that can be accessed directly. All key staff should have access to copies of all the necessary recovery documentation.

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 formalized documents maintained under the control of change management and service asset and configuration management) to ensure that only the latest versions are in circulation and each recipient should ensure that a personal copy is maintained off-site.

The plan should ensure that all details regarding recovery of the IT services following a disaster are fully documented. It should have sufficient details to enable a technical person unfamiliar with the systems to be able to follow the procedures. The recovery plans include key details 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 service and systems.

It is a good idea to include a checklist that covers specific actions required during all stages of recovery for the service and 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 service over to the business.

There are a number of technical plans that may already exist within an organization, documenting recovery procedures from a normal operational failure. The development and maintenance of these plans will be the responsibility of the specialist teams, but will be coordinated by the BCM team. These will be useful additions or appendices to the main plan. Additionally, plans that will need to be integrated with the main business continuity plan are:

- **Emergency response plan** To interface to all emergency services and activities
- **Damage assessment plan** Containing details of damage assessment contacts, processes and plans
- **Salvage plan** Containing information on salvage contacts, activities and processes
- **Vital records plan** Details of all vital records and information, together with their location, that are critical to the continued operation of the business
- **Crisis management and public relations plan** The plans on the command and control of different crisis situations and management of the media and public relations
- **Accommodation and services plan** Detailing the management of accommodation, facilities and the services necessary for their continued operation
- **Security plan** Showing how all aspects of security will be managed on all home sites and recovery sites
- **Personnel plan** Containing details of how all personnel issues will be managed during a major incident
- **Communication plan** Showing how all aspects of communication will be handled and managed with all relevant areas and parties involved during a major incident
- **Finance and administration plan** Containing details of alternative methods and processes for obtaining possible emergency authorization and access to essential funds during a major incident.

Finally, each critical business area is responsible for the development of a plan detailing the individuals who will be in the recovery teams and the tasks to be undertaken on invocation of recovery arrangements.

The ITSCM plan must contain all the information needed to recover the IT systems, networks 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. One of the most important inputs into the plan development is the results of the BIA. Additionally, other areas will need to be analysed, such as SLA, security requirements, operating instructions and procedures, and external contracts. It is likely that a separate SLA with alternative targets will have been agreed if running at a recovery site following a disaster.

Other areas that will need to be implemented following the approval of the strategy are as follows.

Organization planning

During the disaster recovery process, the organizational structure will inevitably be different from normal operation and will be based around:

- **Executive** This will include senior management/executive board, with overall authority and control within the organization and responsible for crisis management and liaison with other departments, divisions, organizations, the media, regulators, emergency services etc.
- **Coordination** Typically one level below the executive group, this is responsible for coordinating the overall recovery effort within the organization
- **Recovery** A series of business and service recovery teams should represent the vital business functions and the services that need to be established to support these functions. Each team is responsible for executing the plans within its 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. and 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 BIA should be documented within the recovery plans and applied during their execution.

Risk reduction/recovery arrangement implementation

The specific actions necessary to enable the strategy must be implemented. Risk reduction arrangements are usually undertaken in conjunction with availability management. Specific examples include those described in the 'risk response measures' section above.

It is important to remember that the recovery is also based around a series of standby arrangements including accommodation, procedures and people, as well as systems and telecommunications. Certain actions are necessary to implement the standby arrangements, as called for in the strategy. For example:

- Negotiating for third-party recovery facilities and entering into a contractual arrangement
- Preparing and equipping the standby accommodation
- Purchasing and installing standby computer systems.

Initial testing

Experience has shown that recovery plans that have not been fully tested do not work as intended, if at all. Testing is therefore a critical part of the overall ITSCM process and the only way of ensuring that the selected strategy, standby arrangements, logistics, business recovery plans and procedures will actually work in practice.

The IT service provider is responsible for ensuring that the IT services can be recovered in the required timescales with the required functionality and the required performance following a disaster.

Four basic types of tests can be undertaken:

- **Walk-through tests** These can be conducted when the plan has been produced simply by getting the relevant people together to see if the plan(s) at least work in a simulated way.
- **Full tests** These should be conducted as soon as possible after the plan production and at regular intervals of at least annually thereafter. They should involve the business units to assist in proving the capability to recover the services appropriately. They should, as far as possible, replicate an actual invocation of all standby arrangements and should involve external parties if they are planned to be involved in an actual invocation. The tests must not only prove recovery of the IT services but also the recovery of the business processes. It is recommended that an independent observer records all the activities of the tests and the timings of the service recovery. The observer's documentation of the tests will be vital input into the subsequent post-mortem review. The full tests may be announced or unannounced. The first test of the plan is likely to be announced and carefully planned, but subsequent tests may be 'sprung' on key players without warning. It is also essential that many different people get

involved, including those not very familiar with the IT service and systems, as the people with the most knowledge may not be available when a disaster actually occurs.

- **Partial tests** These can also be undertaken where recovery of certain elements of the overall plan is tested, such as single services or servers. These types of test should be in addition to the full test not instead of the full test. The full test is the best way of testing that all services can be recovered in required timescales and can run together on the recovery systems.
- **Scenario tests** These can be used to test reactions and plans to specific conditions, events and scenarios. They can include testing that business continuity plans and IT service continuity plans interface with each other, as well as interfacing with all other plans involved in the handling and management of a major incident.

All tests need to be undertaken against defined test scenarios, which are described as realistically as possible. It should be noted, however, that even the most comprehensive test does not cover everything. For example, in a service disruption where there has been injury to, or even death of, 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 CSFs, which will be used to determine the success or otherwise of the exercise.

4.6.5.4 Stage 4 – Ongoing operation

This stage consists of the activities necessary to firmly establish the ITSCM capabilities and maintain them in an accurate and reliable state as time goes on. It should be noted that maintaining the relevance of ITSCM will require ongoing participation in regular business impact analysis and risk assessment and management activities in cooperation with BCM, and action to implement any needed changes based on any resulting strategy revisions. The activities of ongoing operation are set out below.

Education, awareness and training

Education, awareness and training should cover the organization and, in particular, the IT organization, for service continuity-specific items. This ensures that all staff are aware of the implications of business continuity and of service continuity and consider these as part of their normal working, and that everyone involved in the plan has been trained in how to implement their actions.

Review and audit

Regular review of all of the deliverables from the ITSCM process needs to be undertaken to ensure that they remain current. With service providers struggling to do everything they must to serve their customers, it may be difficult to set aside the time needed for reviews and audits, but the work is necessary. The time when the plans need to be invoked in response to a real continuity event is not the time to discover it has become obsolete.

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, preferably at least annually, although testing of IT service continuity plans should be arranged in line with business needs and the needs of the business continuity plans.

All plans should also be tested after every major business change. It is important that any changes to the IT technology are also included in the strategy, implemented in an appropriate fashion and tested to ensure that they function correctly within the overall provision of IT following a disaster. The backup and recovery of IT service should also be monitored and tested to ensure that when they are needed during a major incident, they will operate as needed. This aspect is covered more fully in *ITIL Service Operation*.

Change management

The change management process should ensure that all changes are assessed for their potential impact on the ITSCM plans. If the planned change will invalidate the plans, then the plan must be updated before the change is implemented, and it should be tested as part of the change testing.

The plans themselves must be under very strict change management and service asset and configuration management control. Inaccurate plans and inadequate recovery capabilities may result in the failure of business continuity plans. Also, on an ongoing basis, whenever there are new services or where services have major changes, it is essential that a BIA and a risk assessment is conducted on the new or changed service and the strategy and plans updated accordingly.

4.6.5.5 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, but if the plans have not been tested, failures can be expected. It is important that due consideration is given to the design of all invocation processes, to ensure that they are fit for purpose and interface with all other relevant invocation processes.

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, so it is essential that guidance on the invocation process is readily available. Plans must be available to key staff in the office and away from the office.

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 serious building fire, the decision may be fairly easy to make. However, in the case of power failure or hardware fault, where a resolution is expected within a short period, a deadline should be set by which time if the incident has not been resolved, invocation will take place. If using external services providers, they should be warned immediately if there is a chance that invocation might take place.

The decision to invoke needs to take into account the:

- Extent of the damage and scope of the potential invocation
- Likely length of the disruption and unavailability of premises and/or services
- 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.

Therefore the design of the invocation process must provide guidance on how all of these areas and circumstances should be assessed to assist the person invoking the continuity plan.

The ITSCM plan should include details of activities that need to be undertaken, including:

- Retrieval of backup media or use of data vaulting to retrieve data
- Retrieval of essential documentation, procedures, workstation images etc. stored off-site
- Mobilization 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.

The invocation and initial recovery is likely to be a time of high activity, involving long hours for many individuals. This must be recognized 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. 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.

Once the recovery has been completed, the business should be able to operate from the recovery site at the level determined and agreed in the strategy and relevant SLA. The objective, however, will be to build up the business to normal levels, maintain operation from the recovery site in the short term and vacate the recovery site in the shortest possible time. Details of all these activities need to be contained within the plans. If using external services, there will be a finite contractual period for using the facility. 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.

4.6.6 Triggers, inputs, outputs and interfaces

4.6.6.1 Triggers

Many events may trigger ITSCM activity, including:

- New or changed business needs, or new or changed services
- New or changed targets within agreements, such as SLRs, SLAs, OLAs or contracts
- The occurrence of a major incident that requires assessment for potential invocation of either business or IT continuity plans
- Periodic activities such as the BIA or risk assessment activities, maintenance of continuity plans or other reviewing, revising or reporting activities
- Assessment of changes and attendance at change advisory board meetings
- Review and revision of business and IT plans and strategies
- Review and revision of designs and strategies
- Recognition or notification of a change of risk or impact of a business process or VBF, an IT service or component
- Initiation of tests of continuity and recovery plans
- Lessons learned from previous continuity events and associated recovery activities.

4.6.6.2 Inputs

There are many sources of input required by the ITSCM process:

- Business information: from the organization's business strategy, plans and financial plans, and information on their current and future requirements
- IT information: from the IT strategy and plans and current budgets
- A business continuity strategy and a set of business continuity plans: from all areas of the business
- Service information: from the SLM process, with details of the services from the service portfolio and the service catalogue and service level targets within SLAs and SLRs
- Financial information: from financial management for IT services, the cost of service provision, the cost of resources and components
- Change information: from the change management process, with a change schedule and a need to assess all changes for their impact on all ITSCM plans
- CMS: containing information on the relationships between the business, the services, the supporting services and the technology
- Business continuity management and availability management testing schedules
- Capacity management information identifying the resources required to run the critical services in the event of a continuity event
- IT service continuity plans and test reports from supplier and partners, where appropriate.

4.6.6.3 Outputs

The outputs from the ITSCM process include:

- A revised ITSCM policy and strategy
- A set of ITSCM plans, including all crisis management plans, emergency response plans and disaster recovery plans, together with a set of supporting plans and contracts with recovery service providers
- BIA exercises and reports, in conjunction with BCM and the business
- Risk assessment and management reviews and reports, in conjunction with the business, availability management and information security management
- An ITSCM testing schedule
- ITSCM test scenarios
- ITSCM test reports and reviews.

Forecasts and predictive reports are used by all areas to analyse, predict and forecast particular business and IT scenarios and their potential solutions.

4.6.6.4 Interfaces

Integration and interfaces exist from ITSCM to all other processes. Important examples are as follows:

- **Change management** All changes need to be considered for their impact on the continuity plans, and if amendments are required to the plan, updates to the plan need to be part of the change. The plan itself must be under change management control.
- **Incident and problem management** Incidents can easily evolve into major incidents or disasters. Clear criteria need to be agreed and documented for the invocation of the ITSCM plans.
- **Availability management** Undertaking risk assessment and implementing risk responses should be closely coordinated with the availability process to optimize risk mitigation.
- **Service level management** Recovery requirements will be agreed and documented in the SLAs. Different service levels could be agreed and documented that could be acceptable in a disaster situation.
- **Capacity management** Ensuring that there are sufficient resources to enable recovery onto replacement computers following a disaster.
- **Service asset and configuration management** The CMS documents the components that make up the infrastructure and the relationship between the components. This information is invaluable for all the stages of the ITSCM lifecycle, the maintenance of plans and recovery facilities.
- **Information security management** A very close relationship exists between ITSCM and information security management. A major security breach could be considered a disaster, so when conducting BIA and risk assessment, security will be a very important consideration.

4.6.7 Information management

ITSCM needs to record all of the information necessary to maintain a comprehensive set of ITSCM plans. This information base should include:

- Information from the latest version of the BIA
- Comprehensive information on risk within a risk register, including risk assessment and risk responses
- The latest version of the BCM strategy and business continuity plans
- Details relating to all completed tests and a schedule of all planned tests
- Details of all ITSCM plans and their contents
- Details of all other plans associated with ITSCM plans
- Details of all existing recovery facilities, recovery suppliers and partners, recovery agreements and contracts, spare and alternative equipment
- Details of all backup and recovery processes, schedules, systems and media and their respective locations.

All the above information needs to be integrated and aligned with all BCM information and all the other information required by ITSCM. Interfaces to many other processes are required to ensure that this alignment is maintained.

4.6.8 Critical success factors and key performance indicators

The following list includes some sample CSFs for ITSCM. Each organization should identify appropriate CSFs based on its objectives for the process. Each sample CSF is followed by a small number of typical KPIs that support the CSF. These KPIs should not be adopted without careful consideration. Each organization should develop KPIs that are appropriate for its level of maturity, its CSFs and its particular circumstances. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

- **CSF** IT services are delivered and can be recovered to meet business objectives
  - **KPI** Increase in success of regular audits of the ITSCM plans to ensure that, at all times, the agreed recovery requirements of the business can be achieved
  - **KPI** Regular successful validation that all service recovery targets are agreed and documented in SLAs and are achievable within the ITSCM plans
  - **KPI** Regular and comprehensive testing of ITSCM plans achieved consistently
  - **KPI** Regular reviews are undertaken, at least annually, of the business and IT continuity plans with the business areas
  - **KPI** Regular successful validation that IT negotiates and manages all necessary ITSCM contracts with third parties
  - **KPI** Overall reduction in the risk and impact of possible failure of IT services

- **CSF** Awareness throughout the organization of the business and IT service continuity plans
  - **KPI** Increase in validated awareness of business impact, needs and requirements throughout IT
  - **KPI** Increase in successful test results ensuring that all IT service areas and staff are prepared and able to respond to an invocation of the ITSCM plans
  - **KPI** Validated regular communication of the ITSCM objectives and responsibilities within the appropriate business and IT service areas.

4.6.9 Challenges and risks

4.6.9.1 Challenges

One of the major challenges facing ITSCM is to provide appropriate plans when there is no BCM process. If there is no BCM process, then IT is likely to make incorrect assumptions about business criticality of business processes and therefore adopt the wrong continuity strategies and options. Without BCM, expensive ITSCM solutions and plans will be rendered useless by the absence of corresponding plans and arrangements within the business. Also, if BCM is absent, then the business may fail to identify inexpensive non-IT solutions and waste money on ineffective, expensive IT solutions.

In some organizations, the business perception is that continuity is an IT responsibility, and therefore the business assumes that IT will be responsible for disaster recovery and that IT services will continue to run under any circumstances. This is especially true in some outsourced situations where the business may be reluctant to share its BCM information with an external service provider.

If there is a BCM process established, then the challenge becomes one of alignment and integration. ITSCM must ensure that accurate information is obtained from the BCM process on the needs, impact and priorities of the business, and that the ITSCM information and plans are aligned and integrated with those of the business. Having achieved that alignment, the challenge becomes one of keeping them aligned by management and control of business and IT change. It is essential, therefore, that all documents and plans are maintained under the strict control of change management and service asset and configuration management.

4.6.9.2 Risks

Some of the major risks associated with ITSCM include:

- Lack of a BCM process
- Lack of commitment from the business to the ITSCM processes and procedures
- Lack of appropriate information on future business plans and strategies
- Lack of senior management commitment or a lack of resources and/or budget for the ITSCM process
- The processes focus too much on the technology issues and not enough on the IT services and the needs and priorities of the business
- Risk assessment and management are conducted in isolation and not in conjunction with availability management and information security management
- ITSCM plans and information become out of date and lose alignment with the information and plans of the business and BCM.

4.7 INFORMATION SECURITY MANAGEMENT

Information security is a management process within the corporate governance framework, which provides the strategic direction for security activities and ensures objectives are achieved. It further ensures that the information security risks are appropriately managed and that enterprise information resources are used responsibly. Information security management provides a focus for all aspects of IT security and manages all IT security activities.

In this context, the term 'information' is used as a general term and includes data stores, databases and metadata.

Information security is a critical part of the warranty of a service. If the security of a service's information and information processing cannot be maintained at the levels required by the business, then the business will not experience the value that has been promised. Without information security the utility of the service cannot be accessed.

Information security management needs to be considered within the overall corporate governance framework. Corporate governance is the set of responsibilities and practices exercised by the board and executive management with the goal of providing strategic direction, ensuring the objectives are achieved, ascertaining the risks are being managed appropriately and verifying that the enterprise's resources are used effectively.

4.7.1 Purpose and objectives

The purpose of the information security management process is to align IT security with business security and ensure that the confidentiality, integrity and availability of the organization's assets, information, data and IT services always matches the agreed needs of the business.

The objective of information security management is to protect the interests of those relying on information, and the systems and communications that deliver the information, from harm resulting from failures of confidentiality, integrity and availability.

For most organizations, the security objective is met when:

- Information is observed by or disclosed to only those who have a right to know (confidentiality)
- Information is complete, accurate and protected against unauthorized modification (integrity)
- Information is available and usable when required, and the systems that provide it can appropriately resist attacks and recover from or prevent failures (availability)
- Business transactions, as well as information exchanges between enterprises, or with partners, can be trusted (authenticity and non-repudiation).

4.7.2 Scope

The information security management process should be the focal point for all IT security issues, and must ensure that an information security policy is produced, maintained and enforced that covers the use and misuse of all IT systems and services. Information security management needs to understand the total IT and business security environment, including the:

- Business security policy and plans
- Current business operation and its security requirements
- Future business plans and requirements
- Legislative and regulatory requirements
- Obligations and responsibilities with regard to security contained within SLAs
- The business and IT risks and their management.

Understanding all of this will enable information security management to ensure that all the current and future security aspects and risks of the business are cost-effectively managed.

Prioritization of confidentiality, integrity and availability must be considered in the context of business and business processes. The primary guide to defining what must be protected and the level of protection has to come from the business. To be effective, security must address entire business processes from end to end and cover the physical and technical aspects. Only within the context of business needs and risks can management define security.

The information security management process should include:

- The production, maintenance, distribution and enforcement of an information security policy and supporting security policies
- Understanding the agreed current and future security requirements of the business and the existing business security policy and plans
- Implementation of a set of security controls that support the information security policy and manage risks associated with access to services, information and systems
- Documentation of all security controls, together with the operation and maintenance of the controls and their associated risks
- Management of suppliers and contracts regarding access to systems and services, in conjunction with supplier management
- Management of all security breaches, incidents and problems associated with all systems and services
- The proactive improvement of security controls, and security risk management and the reduction of security risks
- Integration of security aspects within all other ITSM processes.

To achieve effective information security governance, management must establish and maintain an information security management system (ISMS) to guide the development and management of a comprehensive information security programme that supports the business objectives.

4.7.3 Value to the business

Information security management ensures that an information security policy is maintained and enforced that fulfils the needs of the business security policy and the requirements of corporate governance. It raises awareness of the need for security within all IT services and assets throughout the organization, ensuring that the policy is appropriate for the needs of the organization. It manages all aspects of IT and information security within all areas of IT and service management activity.

Information security management provides assurance of business processes by enforcing appropriate security controls in all areas of IT and by managing IT risk in line with business and corporate risk management processes and guidelines.

4.7.4 Policies, principles and basic concepts

Prudent business practices require that IT processes and initiatives align with business processes and objectives. This is critical when it comes to information security, which must be closely aligned with business security and business needs. Additionally, all processes within the IT organization must include security considerations.

Executive management is ultimately responsible for the organization's information and is tasked with responding to issues that affect its protection. In addition, boards of directors are expected to make information security an integral part of corporate governance. All IT service provider organizations must therefore ensure that they have a comprehensive information security management policy(s) and the necessary security controls in place to monitor and enforce the policies.

4.7.4.1 Policies

Information security management activities should be focused on and driven by an overall information security policy and a set of underpinning specific security policies. The information security policy should have the full support of top executive IT management and ideally the support and commitment of top executive business management. The policy should cover all areas of security, be appropriate, meet the needs of the business and should include:

- An overall information security policy
- Use and misuse of IT assets policy
- An access control policy
- A password control policy
- An email policy
- An internet policy
- An anti-virus policy
- An information classification policy
- A document classification policy
- A remote access policy
- A policy with regard to supplier access to IT service, information and components
- A copyright infringement policy for electronic material
- An asset disposal policy
- A records retention policy.

In most cases, these policies should be widely available to all customers and users, and their compliance should be referred to in all SLRs, SLAs, OLAs, underpinning contracts and agreements.

> **Exception**
>
> The only exception to this approach is in the case of Type III service providers where the information security policies related to one external customer should be confidential from other customers, and the provider's own detailed policies are likely to be confidential from the customers for intellectual property rights reasons. The only sharing of security policies in this case should be the aspects that relate directly to the provision of service to that specific customer.

The policies should be authorized by top executive management within the business and IT, and compliance with them should be endorsed on a regular basis. All security policies should be reviewed – and, where necessary, revised – on at least an annual basis.

4.7.4.2 Risk assessment and management in information security management

To achieve the objectives of information security management, formal risk assessment and management relating to security of information and information processing is fundamental. Indeed, it is difficult to identify any part of this process that does not relate to risk management in some way. The information security management process frequently collaborates not only with the business but also with the ITSCM and availability management processes to conduct risk assessments at various levels. See Appendix M for more detail on risk assessment and management methods. Performing accurate assessment of risk and active management of risk to acceptable levels is a core competency that every organization should develop and maintain.

4.7.4.3 Information security management system

The information security management process will have a formal system to establish policy and objectives and to achieve those objectives. This system will generally consist of:

- An information security policy and specific security policies that address each aspect of strategy, controls and regulation
- A security management information system (SMIS), containing the standards, management procedures and guidelines supporting the information security policies
- A comprehensive security strategy, closely linked to the business objectives, strategies and plans
- An effective security organizational structure
- A set of security controls to support the policy
- The management of security risks
- Monitoring processes to ensure compliance and provide feedback on effectiveness
- Communications strategy and plan for security
- Training and awareness strategy and plan.

Elements of the information security management system

The information security management system (ISMS) provides a basis for the development of a cost-effective information security programme that supports the business objectives. It will involve the four Ps of people, process, products (technology) and partners (suppliers) to ensure high levels of security are in place wherever it is appropriate.

ISO/IEC 27001 is the formal standard against which organizations may seek independent certification of their ISMS (meaning their frameworks to design, implement, manage, maintain and enforce information security processes and controls systematically and consistently throughout the organizations). The ISMS shown in Figure 4.23 shows an approach that is widely used and is based on the advice and guidance described in many sources, including ISO/IEC 27001.

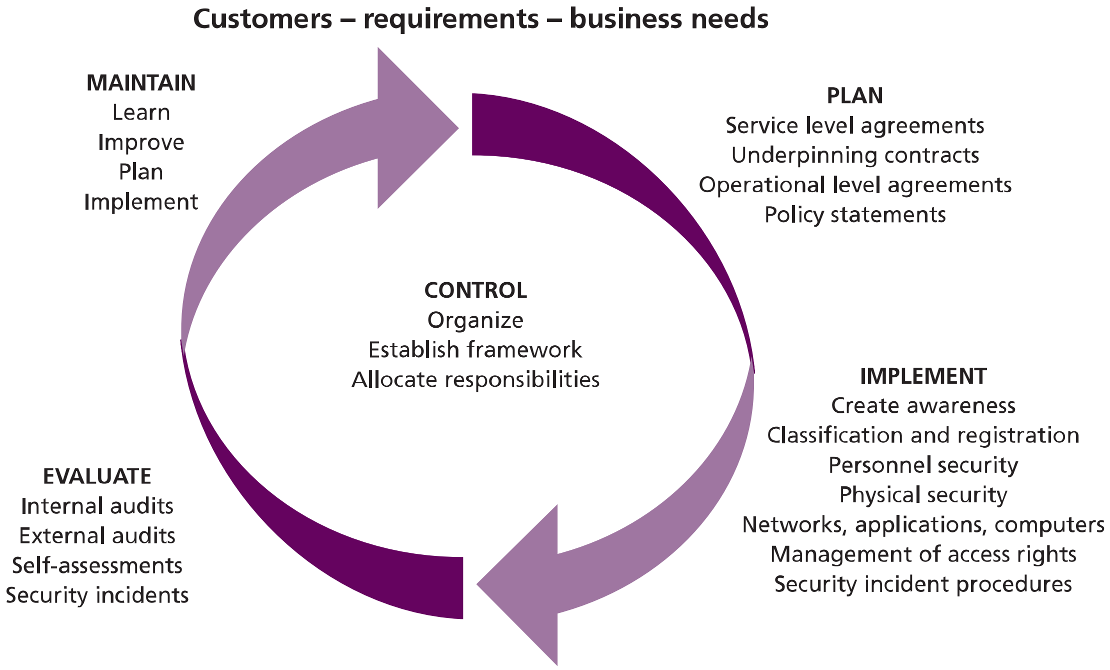


Figure 4.23 Elements of an ISMS for managing IT security

The five elements within this structure are as follows.

Control

The objectives of the control element of the ISMS are to:

- Establish a management framework to initiate and manage information security in the organization
- Establish an organizational structure to prepare, approve and implement the information security policy
- Allocate responsibilities
- Establish and control documentation.

Plan

The objective of the plan element of the ISMS is to devise and recommend the appropriate security measures, based on an understanding of the requirements of the organization.

The requirements will be gathered from such sources as business and service risk, plans and strategies, SLAs and OLAs and the legal, moral and ethical responsibilities for information security. Other factors such as the amount of funding available and the prevailing organization culture and attitudes to security must be considered.

The information security policy defines the organization's attitude and stance on security matters. This should be an organization-wide document, not just applicable to the IT service provider. Responsibility for the upkeep of the document rests with the information security manager.

Implement

The objective of the implementation element of the ISMS is to ensure that appropriate procedures, tools and controls are in place to underpin the information security policy. Measures include:

- Accountability for assets – service asset and configuration management and the CMS are invaluable here
- Information classification – information and repositories should be classified according to the sensitivity and the impact of disclosure.

The successful implementation of the security controls and measures is dependent on a number of factors:

- The determination of a clear and agreed policy, integrated with the needs of the business
- Security procedures that are justified, appropriate and supported by senior management
- Effective marketing and education in security requirements
- A mechanism for improvement.

Evaluate

The objectives of the evaluate element of the ISMS are to:

- Supervise and check compliance with the security policy and security requirements in SLAs and OLAs, and in underpinning contracts in conjunction with supplier management
- Carry out regular audits of the technical security of IT systems
- Provide information to external auditors and regulators, if required.

Maintain

The objectives of this maintain element of the ISMS are to:

- Improve security agreements as specified in, for example, SLAs and OLAs
- Improve the implementation of security measures and controls.

This should be achieved using a PDCA (Plan-Do-Check-Act) cycle, which is a formal approach suggested by ISO/IEC 27001 for the establishment of the ISMS. This cycle is described in more detail in *ITIL Continual Service Improvement*.

Security governance

Information security governance, when properly implemented, should provide six basic outcomes:

- Strategic alignment:
  - Security requirements should be driven by enterprise requirements
  - Security solutions need to fit enterprise processes
  - Investment in information security should be aligned with the enterprise strategy and agreed-on risk profile
- Value delivery:
  - A standard set of security practices, i.e. baseline security requirements following best practices

  - Properly prioritized and distributed effort to areas with greatest impact and business benefit
  - Institutionalized and commoditized solutions
  - Complete solutions, covering organization and process as well as technology
  - A culture of continual improvement
- Risk management:
  - Agreed-on risk profile
  - Understanding of risk exposure
  - Awareness of risk management priorities
  - Risk mitigation
  - Risk acceptance/deference
- Performance management:
  - Defined, agreed and meaningful set of metrics
  - Measurement process that will help identify shortcomings and provide feedback on progress made resolving issues
  - Independent assurance
- Resource management:
  - Knowledge is captured and available
  - Documented security processes and practices, including explicitly defined the interfaces between ISM and other processes
  - Developed security architecture(s) to efficiently utilize infrastructure resources
- Business process assurance.

4.7.5 Process activities, methods and techniques

The information security management process ensures that the security aspects with regard to services and all service management activities are appropriately managed and controlled in line with business needs and risks.

The key activities within the information security management process are:

- Production and maintenance of an overall information security policy and a set of supporting specific policies
- Communication, implementation and enforcement of the security policies, including:
  - Provision of advice and guidance to all other areas of the business and IT on all information security-related issues
- Assessment and classification of all information assets and documentation
- Implementation, review, revision and improvement of a set of security controls and risk assessment and responses, including:
  - Assessment of the impact of all changes on information security policies, controls and measures
  - Implementation of proactive measures to improve information security wherever it is in the business interest and cost-justifiable to do so
- Monitoring and management of all security breaches and major security incidents
- Analysis, reporting and reduction of the volumes and impact of security breaches and incidents
- Schedule and completion of security reviews, audits and penetration tests.

The interactions between these key activities are illustrated in Figure 4.24.

The developed information security management process, together with the procedures, methods, tools and techniques, constitute the security strategy. The security manager should ensure that technologies, products and services are in place and that the overall policy is developed and well published. The security manager is also responsible for security architecture, authentication, authorization, administration and recovery.

The security strategy also needs to consider how it will embed good security practices into every area of the business. Training and awareness are vital in the overall strategy, as security is often weakest at the end-user stage. It is here, as well, that there is a need to develop methods and processes that enable the policies and standards to be more easily followed and implemented.

Resources need to be assigned to track developments in these enabling technologies and the products they support. For example, privacy continues to be important and, increasingly, the focus of government regulation, making privacy compliance technologies an important enabling technology.

4.7.5.1 Security controls

All parties involved must understand that security is not a step in the lifecycle of services and systems and that security cannot be solved through technology. Rather, information security must be

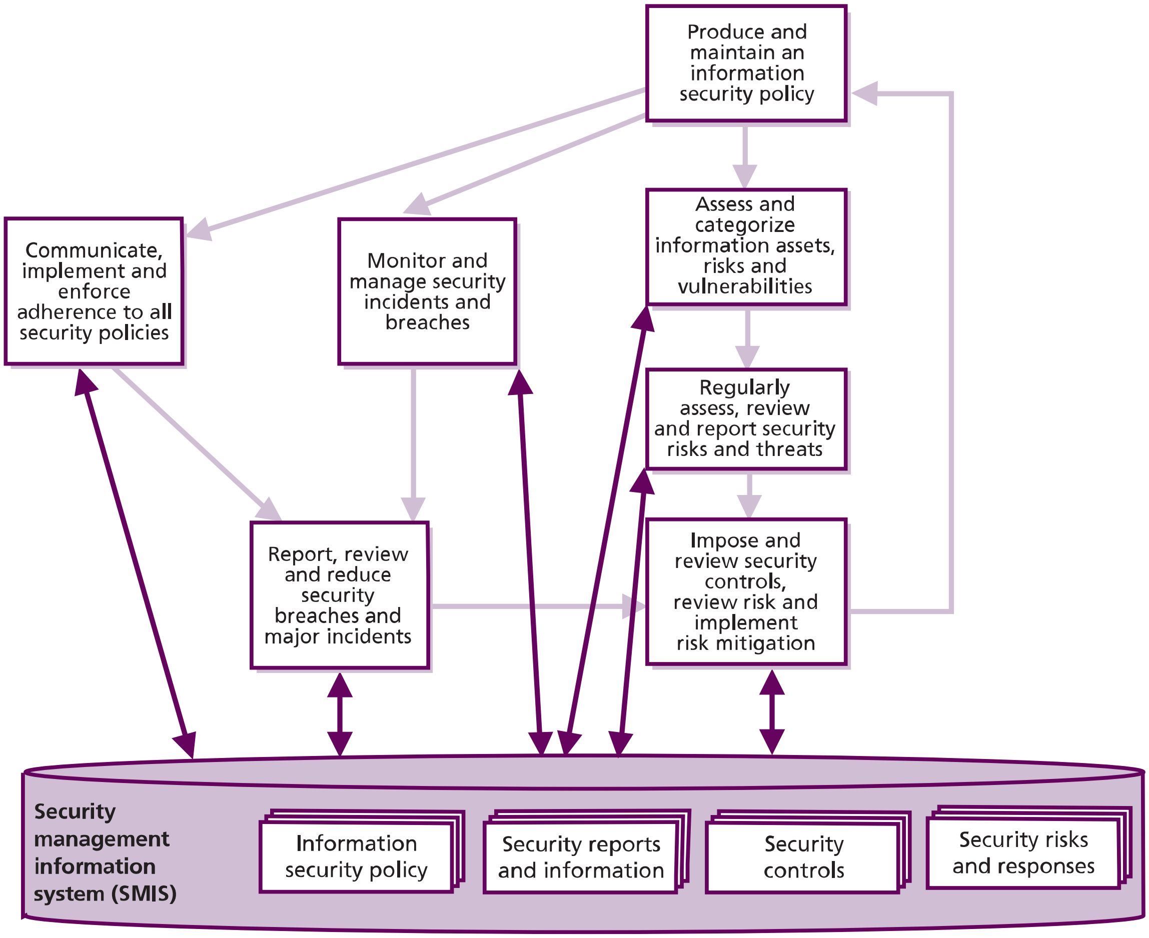

Produce and maintain an information security policy
Communicate, implement and enforce adherence to all security policies
Monitor and manage security incidents and breaches
Assess and categorize information assets, risks and vulnerabilities
Regularly assess, review and report security risks and threats
Report, review and reduce security breaches and major incidents
Impose and review security controls, review risk and implement risk mitigation
Security management information system (SMIS)
Information security policy
Security reports and information
Security controls
Security risks and responses

Figure 4.24 Information security management process

an integral part of all services and systems and is an ongoing process that needs to be continuously managed using a set of security controls.

The set of security controls should be designed to support and enforce the information security policy and to minimize all recognized and identified threats. The controls will be considerably more cost-effective if included within the design of all services. This will ensure the continued protection of all existing services and that new services and access to them are in line with the policy. The security controls and associated procedures for granting and preventing access to services by individuals will typically be executed on a day-to-day basis through the access management process.

Security measures can be used at a specific stage in the prevention and handling of security incidents, as illustrated in Figure 4.25. Security incidents are not solely caused by technical threats – statistics show that, for example, the large majority stem from human errors (intended or not) or procedural errors, and often have implications in other fields such as safety, legality or health.

The following stages can be identified. At the start there is a risk that a threat will materialize. A threat can be anything that disrupts the business process or has negative impact on the business. When a threat materializes, we speak of a security incident. This security incident may result in damage (to information or to assets) that has to be repaired or otherwise corrected. Suitable measures can be selected for each of these stages. The choice of measures will depend on the importance attached to the information.

- **Preventive** Security measures are used to prevent a security incident from occurring. The best-known example of preventive measures is the allocation of access rights to a limited group of authorized people. The further requirements associated with this measure include the control of access rights (granting, maintenance and withdrawal of rights), authorization (identifying

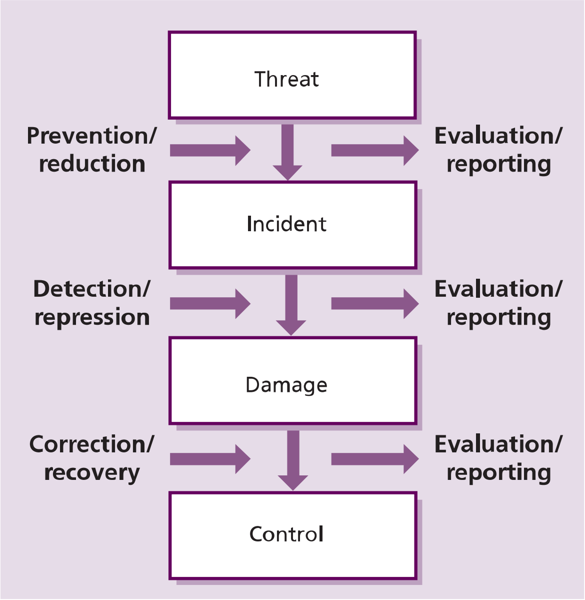


Figure 4.25 Security controls for threats and incidents

who is allowed access to which information and using which tools), identification and authentication (confirming who is seeking access) and access control (ensuring that only authorized personnel can gain access).

- **Reductive** Further measures can be taken in advance to minimize any possible damage that may occur. These are 'reductive' measures. Familiar examples of reductive measures are making regular backups and the development, testing and maintenance of contingency plans.
- **Detective** If a security incident occurs, it is important to discover it as soon as possible – detection. A familiar example of this is monitoring, linked to an alert procedure. Another example is virus-checking software.
- **Repressive** Measures are then used to counteract any continuation or repetition of the security incident. For example, an account or network address is temporarily blocked after numerous failed attempts to log on or the retention of a card when multiple attempts are made with a wrong PIN number.
- **Corrective** The damage is repaired as far as possible using corrective measures. For example, corrective measures include restoring the backup, or returning to a previous stable situation (roll-back, back-out). Fallback can also been seen as a corrective measure.

The documentation of all controls should be maintained to reflect accurately their operation, maintenance and method of operation.

4.7.5.2 Management of security breaches and incidents

In the case of serious security breaches or incidents, an evaluation is necessary in due course, to determine what went wrong, what caused it and how it can be prevented in the future. However, this process should not be limited to serious security incidents. All breaches of security and security incidents need to be studied in order to gain a full picture of the effectiveness of the security measures as a whole. A reporting procedure for security incidents is required to be able to evaluate the effectiveness and efficiency of the present security measures based on an insight into all security incidents. This is facilitated by the maintenance of log files and audit files and, of course, the incident records from the incident management process. The analysis of these statistics on security issues should lead to improvement actions focused on the reduction of the impact and volume of all security breaches and incidents, in conjunction with problem management.

4.7.6 Triggers, inputs, outputs and interfaces

4.7.6.1 Triggers

Information security management activity can be triggered by many events, including:

- New or changed corporate governance guidelines
- New or changed business security policy
- New or changed corporate risk management processes and guidelines
- New or changed business needs or new or changed services
- New or changed requirements within agreements, such as SLRs, SLAs, OLAs or contracts
- Review and revision of business and IT plans and strategies
- Review and revision of designs and strategies
- Service or component security breaches or warnings, events and alerts, including threshold events, exception reports

- Periodic activities, such as reviewing, revising or reporting, including review and revision of information security management policies, reports and plans
- Recognition or notification of a change of risk or impact of a business process or VBF, an IT service or component
- Requests from other areas, particularly SLM for assistance with security issues.

4.7.6.2 Inputs

Information security management will need to obtain input from many areas, including:

- **Business information** From the organization's business strategy, plans and financial plans, and information on its current and future requirements
- **Governance and security** From corporate governance and business security policies and guidelines, security plans, risk assessment and responses
- **IT information** From the IT strategy and plans and current budgets
- **Service information** From the SLM process with details of the services from the service portfolio and the service catalogue and service level targets within SLAs and SLRs, and possibly from the monitoring of SLAs, service reviews and breaches of the SLAs
- **Risk assessment processes and reports** From ISM, availability management and ITSCM
- **Details of all security events and breaches** From all areas of IT and ITSM, especially incident management and problem management
- **Change information** From the change management process with a change schedule and a need to assess all changes for their impact on all security policies, plans and controls
- **CMS** Containing information on the relationships between the business, the services, supporting services and the technology
- **Details of partner and supplier access** From supplier management and availability management on external access to services and systems.

4.7.6.3 Outputs

The outputs produced by the information security management process are used in all areas and should include:

- An overall information security management policy, together with a set of specific security policies
- A security management information system (SMIS), containing all the information relating to information security management
- Revised security risk assessment processes and reports
- A set of security controls, together with details of the operation and maintenance and their associated risks
- Security audits and audit reports
- Security test schedules and plans, including security penetration tests and other security tests and reports
- A set of security classifications and a set of classified information assets
- Reviews and reports of security breaches and major incidents
- Policies, processes and procedures for managing partners and suppliers and their access to services and information.

4.7.6.4 Interfaces

The effective and efficient implementation of an information security policy within an organization will, to a large extent, be dependent on good service management processes. Indeed, the effective implementation of some processes can be seen as a pre-requisite for effective security control. The key interfaces that information security management has with other processes are as follows:

- **Service level management** Information security management provides assistance with the determining of security requirements and responsibilities and their inclusion within SLRs and SLAs, together with the investigation and resolution of service and component security breaches.
- **Access management** This process performs the actions to grant and revoke access and applies the policies defined by information security management and included in the service design by availability management.
- **Change management** Information security management should assist with the assessment of every change for impact on security

and security controls. Also ISM can provide information on unauthorized changes that resulted from security breaches.

- **Incident and problem management** Information security management provides assistance with the resolution and subsequent justification and correction of security incidents and problems. The incident management process must include the ability to identify and deal with security incidents. Service desk and service operations staff must 'recognize' a security incident.
- **IT service continuity management** Information security management works collaboratively with ITSCM on the assessment of business impact and risk, and the provision of resilience, fail-over and recovery mechanisms. Security is a major issue when continuity plans are tested or invoked. A working ITSCM plan is a mandatory requirement for ISO/IEC 27001.
- **Service asset and configuration management** This will give the ability to provide accurate asset information to assist with security classifications. Having an accurate CMS is therefore an extremely useful information security management input.
- **Availability management** If data is unavailable or lacks integrity, then the ability of the service to perform its agreed function is compromised. This makes ISM a critical enabler of availability management. ISM is the process that is accountable for ensuring compliance with security policies in all services. Availability management is responsible for ensuring security requirements are defined and incorporated within the overall availability design. ISM operates collaboratively with both availability management and ITSCM to conduct integrated risk assessment and management exercises.
- **Capacity management** This must consider security implications when selecting and introducing new technology. Security is an important consideration when procuring any new technology or software.
- **Financial management for IT services** This should provide adequate funds to finance security requirements.
- **Supplier management** This should assist with the joint management of suppliers and their access to services and systems, and the terms and conditions to be included within contracts concerning supplier security responsibilities.
- **Legal and human resources issues** These must be considered when investigating security issues. Accordingly, ISM activity should be integrated with these corporate processes and functions.

4.7.7 Information management

All the information required by information security management should be contained within the SMIS. This should include all security controls, risks, breaches, processes and reports necessary to support and maintain the information security policy and the SMIS. This information should cover all IT services and components, and needs to be integrated and maintained in alignment with all other management information systems, particularly the service portfolio and the CMS. The SMIS will also provide the input to security audits and reviews and to the continual improvement activities so important to all SMISs. The SMIS will also provide invaluable input to the design of new systems and services.

4.7.8 Critical success factors and key performance indicators

The following list includes some sample CSFs for information security management. Each organization should identify appropriate CSFs based on its objectives for the process. Each sample CSF is followed by a small number of typical KPIs that support the CSF. These KPIs should not be adopted without careful consideration. Each organization should develop KPIs that are appropriate for its level of maturity, its CSFs and its particular circumstances. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

- **CSF** Business is protected against security violations
  - **KPI** Percentage decrease in security breaches reported to the service desk
  - **KPI** Percentage decrease in the impact of security breaches and incidents
  - **KPI** Percentage increase in SLA conformance to security clauses
- **CSF** The determination of a clear and agreed policy, integrated with the needs of the business

  - **KPI** Decrease in the number of non-conformances of the information security management process with the business security policy and process
- **CSF** Security procedures that are justified, appropriate and supported by senior management
  - **KPI** Increase in the acceptance and conformance of security procedures
  - **KPI** Increased support and commitment of senior management
- **CSF** Effective marketing and education in security requirements, and IT staff awareness of the technology supporting the services
  - **KPI** Increased awareness of the security policy and its contents, throughout the organization
  - **KPI** Percentage increase in completeness of supporting services against the IT components that make up those services
  - **KPI** Service desk supporting all services
- **CSF** A mechanism for improvement
  - **KPI** The number of suggested improvements to security procedures and controls
  - **KPI** Decrease in the number of security non-conformance detected during audits and security testing.
- **CSF** Information security is an integral part of all IT services and all ITSM processes
  - **KPI** Increase in the number of services and processes conformant with security procedures and controls
- **CSF** The availability of services is not compromised by security incidents
  - **KPI** Percentage decrease in the impact of security breaches and incidents
  - **KPI** Percentage reduction in the number of incidents of service unavailability linked to security breaches
- **CSF** Clear ownership and awareness of the security policies among the customer community
  - **KPI** Percentage increase in acceptable scores on security awareness questionnaires completed by customers and users.

4.7.9 Challenges and risks

4.7.9.1 Challenges

Information security management faces many challenges in establishing an appropriate information security policy with an effective supporting process and controls. One of the biggest challenges is to ensure that there is adequate support from the business, business security and senior management. If these are not available, it will be impossible to establish an effective information security management process. If there is senior IT management support, but there is no support from the business, IT security controls and risk assessment and management will be severely limited in what they can achieve. It is pointless implementing security policies, procedures and controls in IT if these cannot be enforced throughout the business. The major use of IT services and assets is outside of IT, and so are the majority of security threats and risks.

In some organizations the business perception is that security is an IT responsibility, and therefore the business assumes that IT will be responsible for all aspects of IT security and that IT services will be adequately protected. However, without the commitment and support of the business and business personnel, money invested in expensive security controls and procedures will be largely wasted and they will mostly be ineffective.

If there is a business security process established, then the challenge becomes one of alignment and integration. Information security management must ensure that accurate information is obtained from the business security process on the needs, risks, impact and priorities of the business and that the information security management policies, information and plans are aligned and integrated with those of the business. Having achieved that alignment, the challenge becomes one of keeping them aligned by management and control of business and IT change using strict change management and service asset and configuration management control. Again, this requires support and commitment from the business and senior management.

4.7.9.2 Risks

Information systems can generate many direct and indirect benefits, and as many direct and indirect risks. These risks have led to a gap between the need to protect systems and services and the degree of protection applied. The gap is caused by internal and external factors, including the widespread use of technology, increasing dependence of the business on IT, increasing complexity and interconnectivity of systems, disappearance of the traditional organizational boundaries, and increasingly onerous regulatory requirements.

This means that there are new risk areas that could have a significant impact on critical business operations, such as:

- Increasing requirements for availability and robustness
- Growing potential for misuse and abuse of information systems affecting privacy and ethical values
- External dangers from hackers, leading to denial-of-service and virus attacks, extortion, industrial espionage and leakage of organizational information or private data.

Because new technology provides the potential for dramatically enhanced business performance, improved and demonstrated information security can add real value to the organization by contributing to interaction with trading partners, closer customer relationships, improved competitive advantage and protected reputation. It can also enable new and easier ways to process electronic transactions and generate trust. In today's competitive global economy, if an organization wants to do business, it may well be asked to present details of its security posture and results of its past performance in terms of tests conducted to ensure security of its information resources.

Other areas of major risks associated with information security management include:

- A lack of commitment from the business to the information security management process and procedures
- Lack of commitment from the business and a lack of appropriate information on future plans and strategies
- A lack of senior management commitment or a lack of resources and/or budget for the information security management process
- The processes focusing too much on technology issues and not enough on the IT services and the needs and priorities of the business
- Risk assessment and management being conducted in isolation and not in conjunction with availability management and ITSCM
- Information security management policies, plans, risks and information becoming out of date and losing alignment with the corresponding relevant information and plans of the business and business security
- Security policies becoming bureaucratic and/ or excessively difficult to follow, discouraging compliance
- Security policies adding no value to business.

4.8 SUPPLIER MANAGEMENT

The supplier management process ensures that suppliers and the services they provide are managed to support IT service targets and business expectations. The aim of this section is to raise awareness of the business context of working with partners and suppliers, and how this work can best be directed toward realising business benefit for the organization.

It is essential that supplier management processes and planning are involved in all stages of the service lifecycle, from strategy and design, through transition and operation, to improvement. Complex business demands require the complete breadth of skills and capability to support provision of a comprehensive set of IT services to a business; therefore the use of value networks and the suppliers and the services they provide are an integral part of any end-to-end solution. Suppliers and the management of suppliers and partners are essential to the provision of quality IT services.

4.8.1 Purpose and objectives

The purpose of the supplier management process is to obtain value for money from suppliers and to provide seamless quality of IT service to the business by ensuring that all contracts and agreements with suppliers support the needs of the business and that all suppliers meet their contractual commitments.

The main objectives of the supplier management process are to:

- Obtain value for money from suppliers and contracts
- Ensure that contracts with suppliers are aligned to business needs, and support and align with agreed targets in SLRs and SLAs, in conjunction with SLM
- Manage relationships with suppliers

- Manage supplier performance
- Negotiate and agree contracts with suppliers and manage them through their lifecycle
- Maintain a supplier policy and a supporting supplier and contract management information system (SCMIS).

Note on terminology

The terms 'contract', 'underpinning contract' and 'agreement' can be confusing. For specific details, see section 4.3.4.2 in service level management which describes how these terms are used in this document. In this section specifically, because it focuses on the relationship with suppliers whose work naturally 'underpins' the delivery of IT services to the customer, the word 'contract', wherever it is used, should be understood to mean 'underpinning contracts and/or agreements'.

4.8.2 Scope

The supplier management process should include the management of all suppliers and contracts needed to support the provision of IT services to the business. Each service provider should have formal processes for the management of all suppliers and contracts. However, the processes should adapt to cater for the importance of the supplier and/or the contract and the potential business impact on the provision of services. Many suppliers provide support services and products that independently have a relatively minor, and fairly indirect, role in value generation, but collectively make a direct and important contribution to value generation and the implementation of the overall business strategy. The greater the contribution the supplier makes to business value, the more effort the service provider should put into the management of the supplier and the more that supplier should be involved in the development and realization of the business strategy. The smaller the supplier's value contribution, the more likely it is that the relationship will be managed mainly at an operational level, with limited interaction with the business. It may be appropriate in some organizations, particularly large ones, to manage internal teams and suppliers, where different business units may provide support of key elements.

The supplier management process should include:

- Implementation and enforcement of the supplier policy
- Maintenance of an SCMIS
- Supplier and contract categorization and risk assessment
- Supplier and contract evaluation and selection
- Development, negotiation and agreement of contracts
- Contract review, renewal and termination
- Management of suppliers and supplier performance
- Identification of improvement opportunities for inclusion in the CSI register, and the implementation of service and supplier improvement plans
- Maintenance of standard contracts, terms and conditions
- Management of contractual dispute resolution
- Management of sub-contracted suppliers.

IT supplier management often has to comply with organizational or corporate standards, guidelines and requirements, particularly those of corporate legal, finance and purchasing, as illustrated in Figure 4.26.

In order to ensure that suppliers provide value for money and meet their service targets, the relationship between each supplier should be owned by an individual within the service provider organization. However, a single individual may own the relationship for one or many suppliers, as illustrated in Figure 4.26. To ensure that relationships are developed in a consistent manner and that suppliers' performance is appropriately reviewed and managed, roles need to be established for a supplier management process owner and a contracts manager. In smaller organizations, these separate roles may be combined into a single responsibility.

4.8.3 Value to the business

The main objectives of the supplier management process are to provide value for money from suppliers and contracts and to ensure that all targets in underpinning supplier contracts and agreements are aligned to business needs and agreed targets within SLAs. This is to ensure the delivery to the business of end-to-end, seamless, quality IT services that are aligned to the business's

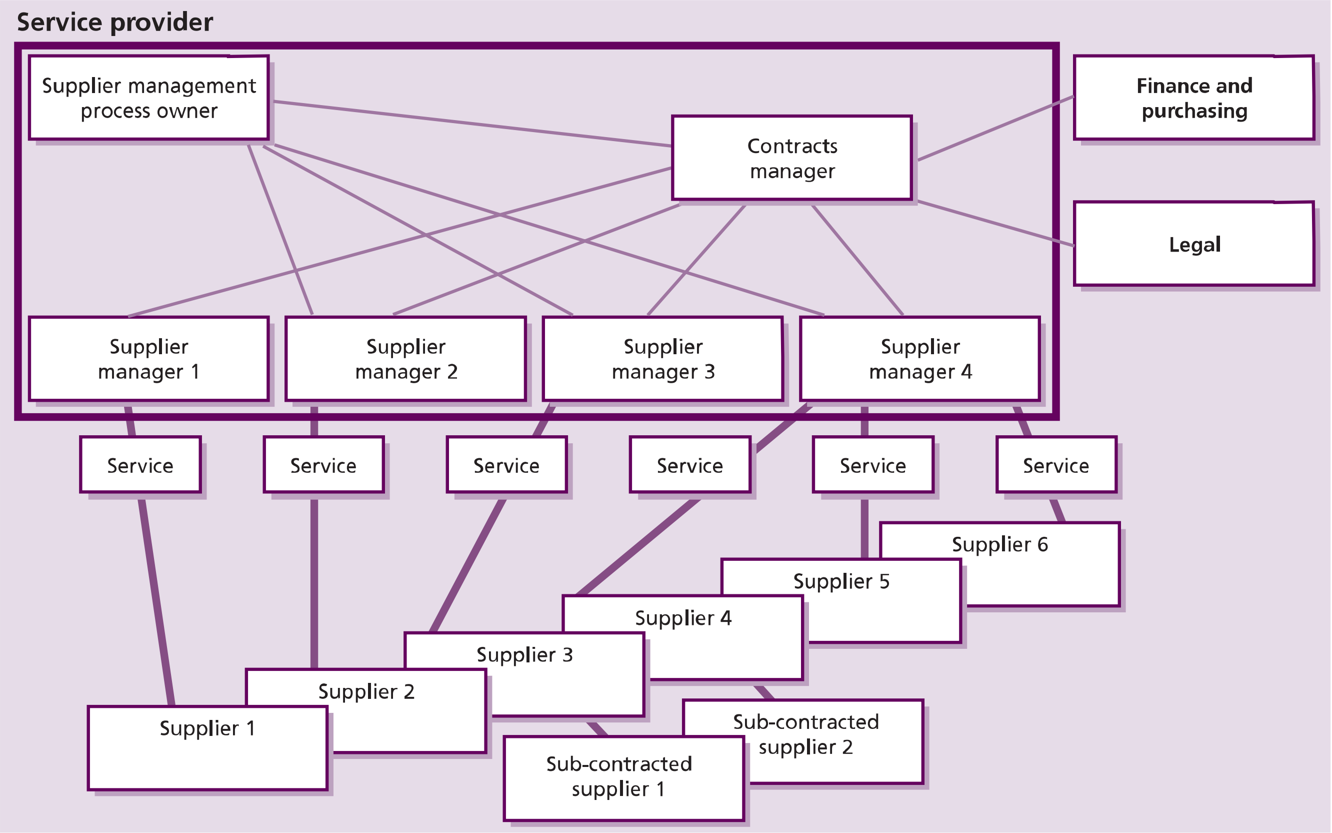


Figure 4.26 Supplier management – roles and interfaces

expectation. The supplier management process should align with all corporate requirements and the requirements of all other IT and service management processes, particularly ISM and ITSCM. This ensures that the business obtains value from supporting supplier services and that they are aligned with business needs.

4.8.4 Policies, principles and basic concepts

The supplier management process attempts to ensure that suppliers meet the terms, conditions and targets of their contracts, while trying to increase the value for money obtained from suppliers and the services they provide. All supplier management process activity should be driven by a supplier strategy and policy from service strategy. The supplier strategy, sometimes called the sourcing strategy, defines the service provider's plan for how it will leverage the contribution of suppliers in the achievement of the overall service strategy. Some organizations might adopt a strategy that dictates the use of suppliers only in very specific and limited circumstances, while another organization might choose to make extensive use of suppliers in IT service provision.

4.8.4.1 Policies

The supplier policies adopted by an organization are the documented management directions that will guide supplier-related decisions and ensure the correct execution of the defined strategy. Supplier policies may cover such areas as:

- The acceptable methods for communication with potential suppliers before and during the solicitation, bidding and procurement processes
- Allocation of roles and responsibilities – who is authorized to interact with suppliers and who is not
- Rules regarding accepting gifts or promotional items from suppliers
- Supplier standards – for example, all suppliers for a hospital in the US must be compliant with the Health Insurance Portability and Accountability Act
- Standards and guidelines for various supplier contract types and/or agreement types

- Data ownership and access policies when suppliers are involved. These policies are developed in collaboration with Information security management.

4.8.4.2 Underpinning contracts and agreements

The nature and extent of an agreement between a service provider and supplier depends on the relationship type and an assessment of the risks involved. A pre-agreement risk assessment is a vital stage in establishing any external supplier agreement. For each party, it exposes the risks that need to be addressed and must be comprehensive and practical, covering a wide variety of risks, including financial, business reputation, operational, regulatory and legal.

A comprehensive agreement minimizes the risk of disputes arising from a difference of expectations. A flexible agreement, which adequately caters for its adaptation across the term of the agreement, is maintainable and supports change with a minimum amount of renegotiation.

The contents of a basic underpinning contract or service agreement are:

- **Basic terms and conditions** The term (duration) of the contract, the parties, locations, scope, definitions and commercial basis.
- **Service description and scope** The functionality of the services being provided and its extent, along with constraints on the service delivery, such as performance, availability, capacity, technical interface and security. Service functionality may be explicitly defined, or in the case of well-established services, included by reference to other established documents, such as the service portfolio and the service catalogue.
- **Service standards** The service measures and the minimum levels that constitute acceptable performance and quality – for example, IT may have a performance requirement to respond to a request for a new desktop system in 24 hours, with acceptable service deemed to have occurred where this performance requirement is met in 95% of cases. Service levels must be realistic, measurable and aligned with the organization's business priorities and underpin the agreed targets within SLRs and SLAs.
- **Workload ranges** The volume ranges within which service standards apply, or for which particular pricing regimes apply.
- **Management information** The data that must be reported by the supplier on operational performance – take care to ensure that management information is focused on the most important or headline reporting measures on which the relationship will be assessed. KPIs related to supplier CSFs and balanced scorecards may form the core of reported performance data.
- **Responsibilities and dependencies** Description of the obligations of the organization (in supporting the supplier in the service delivery efforts) and of the supplier (in its provision of the service), including communication, contacts and escalation.

An extended service agreement may also contain:

- Service debit and credit regime (incentives and penalties)
- Additional performance criteria.

The following is a limited sample of the legal and commercial topics typically covered by a service or contractual agreement:

- Scope of services to be provided
- Service performance requirements
- Division and agreement of responsibilities
- Contact points, communication and reporting frequency and content
- Contract review and dispute resolution processes
- Price structure
- Payment terms
- Commitments to change and investment
- Agreement change process
- Confidentiality and announcements
- Intellectual property rights and copyright
- Liability limitations
- Termination rights of each party
- Obligations at termination and beyond.

The final form of an agreement, and some of the terminology, may be dictated by the views and preferences of the procurement and legal departments, or by specialist legal firms.

Hints and tips

Seek legal advice when formalizing external supplier agreements.

Formal contracts

Formal contracts are appropriate for external supply arrangements that make a significant contribution to the delivery and development of the business. Contracts provide for binding legal commitments between IT service provider and supplier, and cover the obligations each organization has to the other from the first day of the contract, often extending beyond its termination. A contract is used as the basis for external supplier agreements where an enforceable commitment is required. High-value and/or strategic relationships are underpinned by a formal contract. The formality and binding nature of a contract are not at odds with the culture of a partnering agreement, but rather form the basis on which trust in the relationship may be founded.

A contract is likely to be structured with a main body containing the commercial and legal clauses, and with the elements of a service agreement, as described earlier, attached as schedules. Contracts may also include a number of other related documents as schedules, for example:

- Security requirements
- Business continuity requirements
- Mandated technical standards
- Migration plans (agreed pre-scheduled change)
- Disclosure agreements.

Most large organizations have procurement and legal departments specializing in sourcing contracts. Specialist legal firms may be employed to support the internal procurement and legal function when establishing significant formal contracts.

Underpinning agreements

In ITIL an SLA is defined as an agreement between an IT service provider and a customer. A service level agreement describes the IT service, documents service level targets, and specifies the responsibilities of the IT service provider and the customer. Service providers should be aware that SLAs are widely used to formalize service-based relationships, both internally and externally, and that while conforming to the definition above, these agreements vary considerably in the detail covered.

Key messages

The views of some organizations, such as the Chartered Institute of Purchase and Supply and various specialist lawyers, are that SLAs ought not to be used to manage external relationships unless they form part of an underlying contract. *The Complete Guide to Preparing and Implementing Service Level Agreements*[5] emphasizes that a stand-alone SLA may not be legally enforceable but instead 'represents the goodwill and faith of the parties signing it'. This is important in the context of an SLA between a Type III IT service provider and its external customer.

Although in this publication the term SLA is not applied to the IT service provider–supplier relationship, the same principle applies to the service level targets agreed between these two parties. It is in service providers' and suppliers' interests to ensure that agreed service levels are incorporated into an appropriate contractual framework to ensure that these commitments are legally binding.

SLAs between the IT service provider and their customer(s) and the agreements that underpin them, including formal contracts, should be reviewed on a regular basis to ensure performance conforms to the service levels that have been agreed.

The organization is likely to be dependent on its own internal support groups to some extent. To be able to achieve SLA targets, it is advisable to have formal arrangements in place called operational level agreements (OLAs) with these groups. (For more on OLAs and their negotiation, see section 4.3 on the SLM process.) It is important in supplier management that the contributions of suppliers towards service provision are coordinated and aligned with the contributions of internal support groups to ensure that the combined efforts of these groups will ensure achievement of SLA targets and that there are no gaps or duplicated efforts. The activities of the SLM and supplier management processes need to be interfaced appropriately to achieve this.

[5] Pantry, S. and Griffiths, P. (2001). *The Complete Guide to Preparing and Implementing Service Level Agreements*. Library Association Publishing.

4.8.4.3 Supplier and contract management information system

In order to achieve consistency and effectiveness in the implementation of the supplier policy, an SCMIS should be established, together with clearly defined roles and responsibilities.

Ideally the SCMIS should form an integrated element of a comprehensive CMS or SKMS, recording all supplier and contract details, together with details of the type of service(s) or product(s) provided by each supplier, and all other information and relationships with other associated CIs. The services provided by suppliers will also form a key part of the service portfolio and the service catalogue. The relationship between the supporting services and the IT and business services they support are key to providing quality IT services.

This information within the SCMIS will provide a complete set of reference information for all supplier management procedures and activities:

- Definition of new supplier and contract requirements
- Evaluation and set up of new suppliers and contracts
- Supplier categorization and maintenance of the SCMIS
- Establishing new suppliers
- Management of suppliers and their performance and of the associated contracts
- Contract renewal or termination.

The first three elements within the above list are covered within the service design stage. The fourth element is part of service transition, and the last two are part of the service operation stage and are covered in more detail in *ITIL Service Transition* and *ITIL Service Operation*, respectively.

4.8.5 Process activities, methods and techniques

When dealing with external suppliers, it is strongly recommended that a formal contract with clearly defined, agreed and documented responsibilities and targets is established and managed through the stages of its lifecycle, from the identification of the business need to the operation and cessation of the contract. This activity is performed in line with the established supplier strategy and supplier policies. For a more detailed discussion of supplier strategy (otherwise called sourcing strategy), see *ITIL Service Strategy*, section 3.7.

The activities of supplier management can be summarized in this way:

- Definition of new supplier and contract requirements:
  - Identify business need and prepare of the business case, including options (internal and external), costs, timescales, targets, benefits, risk assessment
  - Produce a statement of requirement (SoR) and/or invitation to tender (ITT)
  - Ensure conformance to strategy/policy
- Evaluation of new suppliers and contracts:
  - Identify method of purchase or procurement
  - Establish evaluation criteria – for example, services, capability (both personnel and organization), quality and cost
  - Evaluate alternative options
  - Select
  - Negotiate contracts, targets and the terms and conditions, including responsibilities, closure, renewal, extension, dispute, transfer
  - Agree and award the contract
- Supplier and contract categorization and maintenance of the SCMIS:
  - Assess or reassess the supplier and contract
  - Ensure changes progressed through service transition
  - Categorize the supplier
  - Update SCMIS
  - Ongoing maintenance of the SCMIS
- Establishment of new suppliers and contracts:
  - Set up the supplier service and contract, within the SCMIS and any other associated corporate systems
  - Transition the service
  - Establish contacts and relationships
- Supplier, contract and performance management:
  - Manage and control the operation and delivery of service/products
  - Monitor and report (service, quality and costs)
  - Review and improve (service, quality and costs)

  - Manage the supplier and the relationship (communication, risks, changes, failures, improvements, contacts, interfaces)
  - Review, at least annually, service scope against business need, targets and agreements
  - Plan for possible closure/renewal/extension
- Contract renewal or termination:
  - Review (determine benefits delivered, ongoing requirement)
  - Renegotiate and renew or terminate and/or transfer
  - Transition to new supplier(s) or to internal resources.

The business, IT, finance, purchasing and procurement need to work together to ensure that all stages of the contract lifecycle are managed effectively. All areas need to be jointly involved in selecting the solution and managing the ongoing performance of the supplier, with each area taking responsibility for the interests of their own area, while being aware of the implications on the organization as a whole.

The activities involved in the stages of the contract lifecycle are explained in detail in the following sections and illustrated in Figure 4.27, along with a representation of the SCMIS.

4.8.5.1 Definition of new supplier and contract requirements

The activities associated with the identification of business needs and the subsequent evaluation of new suppliers and contracts are part of the service design stage. As part of the design of a new or changed service, the IT service provider will determine if and to what extent the contribution of suppliers will be required for a sound design and subsequent successful service provision. Once this decision has been made, the detailed requirements for new suppliers or new contracts with existing suppliers can be developed.

The outputs from service design provide the inputs to all other stages of the contract lifecycle. It is vital to the ongoing success of the contract and the relationship that the business is closely involved in all aspects of these activities. Every organization should have templates and a formal method for the production of business cases and their approval and sign-off. The detailing of the business's needs and the content of the business case should be agreed, approved and signed off by both the business and IT.

Both the original decision and the subsequent requirements should be developed to fulfil the defined supplier strategy and to ensure compliance with supplier policies.

4.8.5.2 Evaluation of new suppliers and contracts

When selecting a new supplier or contract, a number of factors need to be taken into consideration, including track record, capability, references, credit rating and size relative to the business being placed. In addition, depending on the type of supplier relationship, there may be personnel issues that need to be considered. Each organization should have processes and procedures for establishing new suppliers and contracts.

While it is recognized that factors may exist that influence the decision on type of relationship or choice of supplier (e.g. politics within the organization, existing relationships), it is essential that in such cases the reasoning is identified and the impact fully assessed to ensure costly mistakes are avoided.

Services may be sourced from a single supplier or multi-sourced. Services are most likely to be sourced from two or more competing suppliers where the requirement is for standard services or products that are readily available 'off-the-shelf'. Multi-sourcing is most likely to be used where cost is the prime determinant, and requirements for developing variants of the services are low, but may also be undertaken to spread risk. Suppliers on a multi-source list may be designated with 'preferred supplier' status within the organization, limiting or removing scope for use of other suppliers.

Partnering relationships are established at an executive level and are dependent on a willingness to exchange strategic information to align business strategies. Many strategically important supplier relationships are now positioned as partnering relationships. This reflects a move away from traditionally hierarchical relationships, where the supplier acts subordinately to the customer organization, to one characterized by:

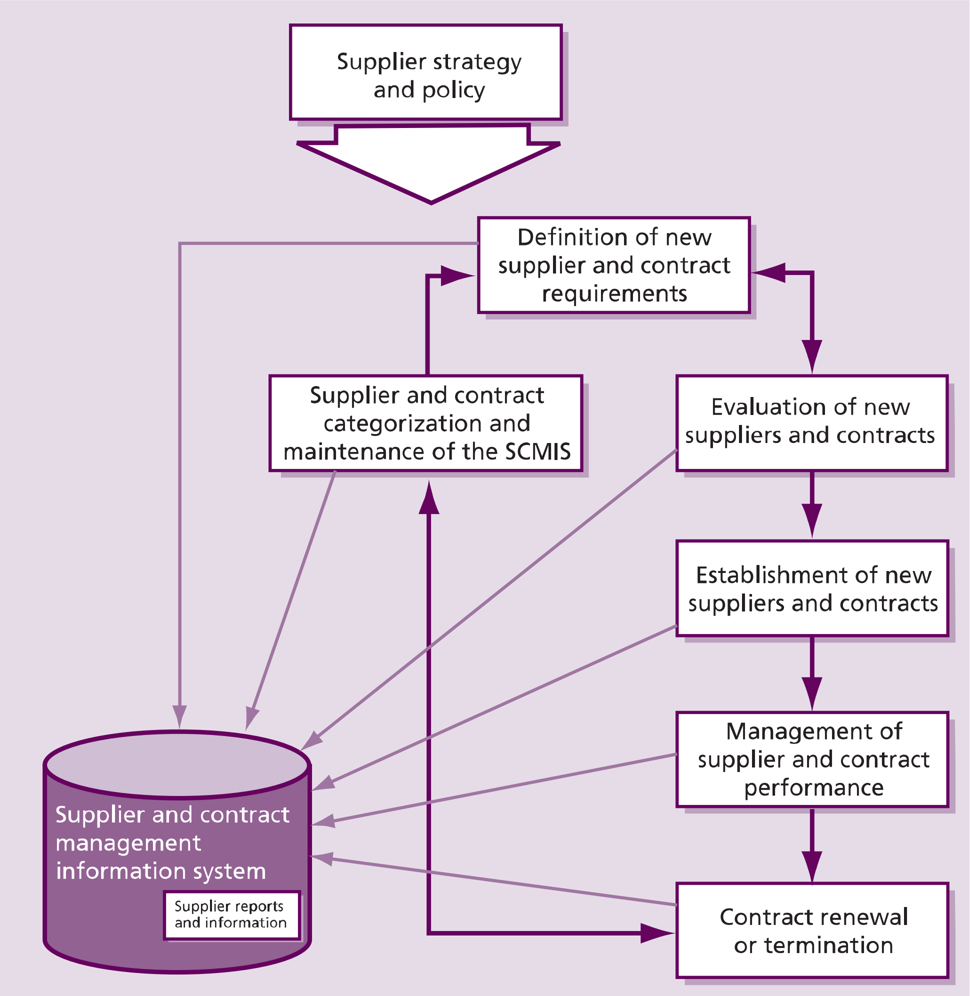

Supplier strategy and policy
Definition of new supplier and contract requirements
Supplier and contract categorization and maintenance of the SCMIS
Evaluation of new suppliers and contracts
Establishment of new suppliers and contracts
Management of supplier and contract performance
Contract renewal or termination
Supplier and contract management information system
Supplier reports and information

Figure 4.27 Supplier management process

- **Strategic alignment** Good alignment of culture, values and objectives, leading to an alignment of business strategies
- **Integration** A close integration of the processes of the two organizations
- **Information flow** Good communication and information exchange at all levels, especially at the strategic level, leading to close understanding
- **Mutual trust** A relationship built on mutual trust between the organizations and their staff
- **Openness** When reporting on service performance, costs and risk assessment
- **Collective responsibility** Joint partnership teams taking collective responsibility for current performance and future development of the relationship
- **Shared risk and reward** For example, agreeing how investment costs and resultant efficiency benefits are shared, or how risks and rewards from fluctuations in material costs are shared.

Both parties derive benefits from partnering. An organization derives progressively more value from a supplier relationship as the supplier's understanding of the organization as a whole increases, from its IT inventory architectures through to its corporate culture, values and business objectives. With time, the supplier is able to respond more quickly and more appropriately to the organization's needs. The supplier benefits from a longer-term commitment from the organization, providing it with greater financial stability and enabling it to finance longer-term investments, which benefit its customers.

A partnership makes it possible for the parties to align their IT infrastructures. Joint architecture and risk control agreements allow the partners to implement a range of compatible solutions from security, networking, data/information interchange, to workflow and application processing systems. This integration can provide service improvements and lowered costs. Such moves also reduce risks and costs associated with one-off tactical solutions, put in place to bridge a supplier's IT with that of the organization.

The key to a successful partnering relationship is being absolutely clear about the benefits and costs such a relationship will deliver before entering into it. Both parties then know what is expected of them at the outset. The success of the partnership may involve agreeing the transfer of staff to the partner or outsourcing organization as part of the agreement and relationship.

Service provider organizations should have documented and formal processes for evaluating and selecting suppliers based on:

- **Importance and impact** The importance of the service to the business, provided by the supplier
- **Risk** The risks associated with using the service
- **Costs** The cost of the service and its provision.

Often other areas of the service provider organization, such as legal, finance and purchasing, will get involved with this aspect of the process. Service provider organizations should have processes covering:

- Production of business case documents
- Production of SoR and ITT or proposal documents
- Formal evaluation and selection of suppliers and contracts
- The inclusion of standard clauses, terms and conditions within contracts, including early termination, benchmarking, exit or transfer of contracts, dispute resolution, management of sub-contracted suppliers and normal termination
- Transitioning of new contracts and suppliers.

These processes may, and should be, different, based on the type, size and category of the supplier and the contract.

4.8.5.3 Supplier categorization and maintenance of the supplier and contract management information system

The supplier management process should be adaptive and managers should spend more time and effort managing key suppliers than less important suppliers. This means that some form of categorization scheme should exist within the supplier management process to categorize the supplier and their importance to the service provider and the services provided to the business. Suppliers can be categorized in many ways, but one of the best methods for categorizing suppliers is based on assessing the risk and impact associated with using the supplier, and the value and importance of the supplier and its services to the business, as illustrated in Figure 4.28.

The amount of time and effort spent managing the supplier and the relationship can then be appropriate to its categorization:

- **Strategic** For significant 'partnering' relationships that involve senior managers sharing confidential strategic information to facilitate long-term plans. These relationships would normally be managed and owned at a senior management level within the service provider organization, and would involve regular and frequent contact and performance reviews. These relationships would probably require involvement of service strategy and service design resources, and would include ongoing specific improvement programmes (e.g. a network service provider supplying worldwide networks service and their support).
- **Tactical** For relationships involving significant commercial activity and business interaction. These relationships would normally be managed by middle management and would involve regular contact and performance reviews, often including ongoing improvement programmes (e.g. a hardware maintenance organization providing resolution of server hardware failures).
- **Operational** For suppliers of operational products or services. These relationships would normally be managed by junior operational management and would involve infrequent but regular contact and performance reviews (e.g.

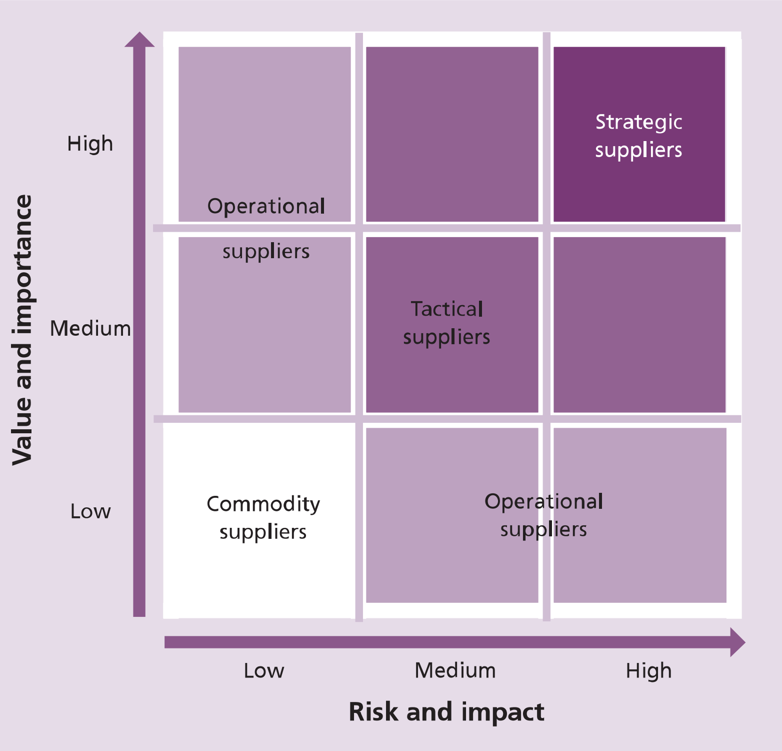


Figure 4.28 Supplier categorization

an internet hosting service provider, supplying hosting space for a low-usage, low-impact website or internally used IT service).

- **Commodity** For suppliers providing low-value and/or readily available products and services, which could be alternatively sourced relatively easily (e.g. paper or printer cartridge suppliers).

Strategically important supplier relationships are given the greatest focus. It is in these cases that supplier managers have to ensure that the culture of the service provider organization is extended into the supplier domain so that the relationship works beyond the initial contract. The rise in popularity of outsourcing, and the increase in the scope and complexity of some sourcing arrangements, has resulted in a diversification of types of supplier relationship. At a strategic level, it is important to understand the options that are available so that the most suitable type of supplier relationship can be established to gain maximum business benefit and evolves in line with business needs.

Hints and tips

To successfully select the most appropriate type of supplier relationship, there needs to be a clear understanding of the business objectives that are to be achieved.

A number of factors, from the nature of the service to the overall cost, determine the importance of a supplier from a business perspective. As shown later, the greater the business significance of a supplier relationship, the more the business needs to be involved in the management and development of a relationship. A formal categorization approach can help to establish this importance.

Standardized versus customized services

The business value, measured as the contribution made to the business value chain, provides a more business-aligned assessment than pure contract price. Also, the more standard the services being procured, the lower the dependence the organization has on the supplier, and the more readily the supplier could be replaced (if necessary). Standardized services support the business through minimal time to market when deploying new or changed business services, and in pursuing cost-reduction strategies. More information on this subject can be found in *ITIL Service Strategy*.

The more customized those services are, the greater the difficulty in moving to an alternative supplier. Customization may benefit the business, contributing to competitive advantage through differentiated service, or may be the result of operational evolution.

Tailored services increase the dependence on the supplier, increase risk and can result in increased cost. From a supplier perspective, tailored services may decrease their ability to achieve economies of scale through common operations, resulting in narrowed margins, and reduced capital available for future investment.

Standard products and services are the preferred approach unless a clear business advantage exists, in which case a strategic supplier delivers the tailored service.

Hints and tips

High-value or high-dependence relationships involve greater risks for the organization. These relationships need comprehensive contracts and active relationship management.

Having established the type of supplier, the relationship then needs to be formalized. In the discussion below, the term 'agreement' is used generically to refer to any formalization of a relationship between customer and supplier organizations, and may range from the informal to comprehensive legally binding contracts. Simple, low-value relationships may be covered by a supplier's standard terms and conditions, and be managed wholly by IT. A relationship of strategic importance to the business, on the other hand, requires a comprehensive contract that ensures that the supplier supports evolving business needs throughout the life of the contract. A contract needs to be managed and developed in conjunction with procurement and legal departments and business stakeholders.

Hints and tips

The agreement is the foundation for the relationship. The more suitable and complete the agreement, the more likely it is that the relationship will deliver business benefit to both parties.

The quality of the relationship between the service provider and its supplier(s) is often dependent on the individuals involved from both sides. It is therefore vital that individuals with the right attributes, skills, competencies and personalities are selected to be involved in these relationships.

Supplier relationships

A business service may depend on a number of internal and/or external suppliers for its delivery. These may include a mixture of strategic suppliers and commodity suppliers. Some suppliers supply directly to the organization; others are indirect or sub-contracted suppliers working via another supplier. Direct suppliers are directly managed by the service provider; indirect or sub-contracted suppliers are managed by the leading supplier. Any one supplier may provide products or services used to support a number of different business services.

Supply chain analysis shows the mapping between business services and supplier services. Analysis of business processes will reveal the suppliers involved in each process and the points of hand-off between them. Management of the supply chain ensures that functional boundaries and performance requirements are clearly established for each supplier to ensure that overall business service levels are achieved. Business services are most likely to meet their targets consistently where there are a small number of suppliers in the supply chain, and where the interfaces between the suppliers in the chain are limited, simple and well defined.

Reducing the number of direct suppliers reduces the number of relationships that need to be managed, the number of peer-to-peer supplier issues that need to be resolved, and the complexity of the supplier management activities. Some organizations may successfully reduce or collapse the whole supply chain around a single service provider, often referred to as a 'prime' supplier. Facilities management is often outsourced to a single specialist partner or supplier, who may in turn sub-contract restaurant services, vending machine maintenance and cleaning.

Outsourcing entire business services to a single 'prime supplier' may run additional risks. For these reasons, organizations need to consider carefully their supply chain strategies ahead of major outsourcing activity. The scope of outsourced services needs to be considered to reduce the number of suppliers, while ensuring that risk is managed and fits with typical competencies in the supply market.

The supplier and contract management information system

The SCMIS is a set of tools, data and information that is used to support supplier management. The SCMIS contains details of the organization's suppliers, together with details of the products and services that they provide to the business (e.g. email service, PC supply and installation, service desk), together with details of the contracts. The SCMIS contains supplier details, a summary of each product/service (including support arrangements), information on the ordering process and, where applicable, contract details. Ideally the SCMIS should be contained within the overall CMS. In most organizations, the SCMIS is owned by the supplier management process or the procurement or purchasing department.

An SCMIS is beneficial because it can be used to promote preferred suppliers and to prevent purchasing of unapproved or incompatible items. By coordinating and controlling the buying activity, the organization is more likely to be able to negotiate preferential rates.

4.8.5.4 Establishment of new suppliers and contracts

The SCMIS provides a single, central focal set of information for the management of all suppliers and contracts. When establishing new suppliers and contracts, adding them to the SCMIS needs to be handled via the change management process, to ensure that any impact is assessed and understood.

Risk management, in relation to working with suppliers, centres on assessing vulnerabilities in each supplier arrangement or contract that pose threats to any aspect of the business, including business impact, probability, customer satisfaction, brand image, market share, profitability, share price or regulatory impacts or penalties (in some industries).

The nature of the relationship affects the degree of risk to the business. Risks associated with an outsourced or strategic supplier are likely to be greater in number, and more complex to manage, than with internal supply. It is rarely possible to 'outsource' risk, although sometimes some of the risk may be transferred to the outsourcing organization. Blaming a supplier does not impress customers or internal users affected by a security incident or a lengthy system failure. New risks arising from the relationship need to be identified and managed, with communication and escalation as appropriate.

A substantial risk assessment should have been undertaken pre-contract, but this needs to be maintained in the light of changing business needs, changes to the contract scope, or changes in the operational environment.

The service provider organization and the supplier must consider the threats posed by the relationship to their own assets, and have their own risk profile. Each must identify their respective risk owners. In a well-functioning relationship, it is possible for much or all of the assessment to be openly shared with the other party. By involving supplier experts in risk assessments, especially in operational risk assessments, the organization may gain valuable insights into how best to mitigate risks, as well as improving the coverage of the assessment.

When evaluating risks of disruption to business services or functions, the business may have different priorities for service/function restoration. BIA is a method used to assess the impacts on different areas of the business, resulting from a loss of service. Risk assessment and BIA activities relating to suppliers and contracts should be performed in close conjunction with ITSCM, availability management and information security management, with a view to reducing the impact and probability of service failure as a result of supplier or supplier service failure.

Once these activities have been completed and the supplier and contract information has been input into the SCMIS, including the nominated individuals responsible for managing the new supplier and/or contracts, frequency of service/supplier review meetings and contractual review meetings needs to be established, with appropriate break points, automated thresholds and warnings in place. The introduction of new suppliers and contracts should be handled as major changes through transition and into operation. This will ensure that appropriate contacts and communication points are established.

4.8.5.5 Supplier, contract and performance management

At an operational level, integrated processes need to be in place between an organization and its suppliers to ensure efficient day-to-day working practices. For example:

- Is the supplier expected to conform to the organization's change management process or any other processes?
- How does the service desk notify the supplier of incidents?
- How is CMS information updated when CIs change as a result of supplier actions? Who is responsible?

There may be a conflict of interest between the service provider organization and its supplier, especially with regard to the change management, incident management, problem management, and service asset and configuration management processes. The supplier may want to use its processes and systems, whereas the service provider organization will want to use its own processes and systems. If this is the case, clear responsibilities and interfaces will need to be defined and agreed.

These and many other areas need to be addressed to ensure smooth and effective working at an operational level. To do so, all touch points and contacts need to be identified and procedures put in place so that everyone understands their roles and responsibilities. This should include identification of the single, nominated individual responsible for ownership of each supplier and contract. However, an organization should take care not to automatically impose its own processes, but to take the opportunity to learn from its suppliers.

In addition to process interfaces, it is essential to identify how issues are handled at an operational level. By having clearly defined and communicated escalation routes, issues are likely to be identified and resolved earlier, minimizing the impact. Both the organization and the supplier benefit from the early capture and resolution of issues.

Example of learning from a supplier

A contract had been awarded for a customized stores control system for which the IT organization had developed processes to support the live service once it was installed. This included procedures for recording and documenting work done on the service by field engineers (e.g. changes, repairs, enhancement and reconfigurations). At a project progress meeting, the supplier confirmed that they had looked at the procedures and could follow them if required. However, having been in this situation many times before, they had already developed a set of procedures to deal with such events. These procedures were considerably more elegant, effective and easier to follow than those developed and proposed by the IT organization.

Both sides should strive to establish good communication links. The supplier learns more about the organization's business, its requirements and its plans, helping the supplier to understand and meet the organization's needs. In turn, the organization benefits from a more responsive supplier who is aware of the business drivers and any issues, and is therefore more able to provide appropriate solutions. Close day-to-day links can help each party to be aware of the other's culture and ways of working, resulting in fewer misunderstandings and leading to a more successful and long-lasting relationship.

Formal reviews

Two levels of formal review need to take place throughout the contract lifecycle to minimize risk and ensure the business realizes maximum benefit from the contract:

- **Service/supplier performance reviews** Reports on performance should be produced on a regular basis, based on the category of supplier, and should form the basis of service review meetings. The more important the supplier, the more frequent and extensive the reports and reviews should be.
- **Service, service scope and contract reviews** These should also be conducted on a regular basis, at least annually for all major suppliers. The objective of these should be to review the service, overall performance, service scope and

targets and the contract, together with any associated agreements. This should be compared with the original business needs and the current business needs to ensure that supplier and contracts remain aligned to business needs and continue to deliver value for money.

Formal performance review meetings must be held on a regular basis to review the supplier's performance against service levels, at a detailed operational level. These meetings provide an opportunity to check that the ongoing service performance management remains focused on supporting business needs. When appropriate, the SLM process may also be represented in supplier performance review meetings. Typical topics include:

- Service performance against targets
- Incident and problem reviews, including any escalated issues
- Business and customer feedback
- Expected major changes that will (or may) affect service during the next service period, as well as failed changes and changes that caused incidents
- Key business events over the next service period that need particular attention from the supplier (e.g. quarter-end processing)
- Best practice
- Opportunities for improvement
- Progression of SIPs.

Major service improvement initiatives and actions are controlled through SIPs with each supplier, including any actions for dealing with any failures or weaknesses. Progress of existing SIPs, or the need for a new initiative, is reviewed at service review meetings. Proactive or forward-thinking organizations use SIPs not only to deal with failures but also to improve a consistently achieved service. It is important that a contract provides suitable incentives to both parties to invest in service improvement. These aspects are covered in more detail in *ITIL Continual Service Improvement*.

The governance mechanisms for suppliers and contracts are drawn from the needs of appropriate stakeholders at different levels within each organization, and are structured so that the organization's representatives face-off to their counterparts in the supplier's organization. Defining the responsibilities for each representative, meeting forums and processes ensure that each person is involved at the right time in influencing or directing the right activities.

The scale and importance of the service and/or supplier influence the governance arrangements needed. The more significant the dependency, the greater the commitment and effort involved in managing the relationship. The effort needed on the service provider side to govern an outsourcing contract should not be underestimated, especially in closely regulated industries, such as the finance and pharmaceutical sectors.

A key objective for supplier management is to ensure that the value of a supplier to the organization is fully realized. Value is realized through all aspects of the relationship, from operational performance assurance, responsiveness to change requests and demand fluctuations, through to contribution of knowledge and experience to the organization's capability. The service provider must also ensure that the supplier's priorities match the business's priorities. The supplier must understand which of its service levels are most significant to the business.

Example of supplier value

A large multinational company had software agreements in place with the same supplier in no fewer than 24 countries. By arranging a single global licensing deal with the supplier, the company made annual savings of £5 million.

Satisfaction surveys and benefits assessments

Satisfaction surveys also play an important role in revealing how well supplier service levels are aligned to business needs. If the customer does not have visibility into what is being delivered by the supplier versus what is being done by internal support groups, the service provider will need to structure satisfaction surveys carefully to be able to differentiate between the two contributions. A survey may reveal instances where there is dissatisfaction with the service, yet the supplier is apparently performing well against its targets (and vice versa). This may happen where service levels are inappropriately defined and should result in a review of the contracts, agreements and targets. Some service providers publish supplier league tables based on their survey results, stimulating competition between suppliers.

For those significant supplier relationships in which the business has a direct interest, both the business (in conjunction with the procurement department) and IT will have established their objectives for the relationship, and defined the benefits they expect to realize. This forms a major part of the business case for entering into the relationship.

These benefits must be linked and complementary, and must be measured and managed. Where the business is seeking improvements in customer service, IT supplier relationships contributing to those customer services must be able to demonstrate improved service in their own domain, and how much this has contributed to improved customer service.

For benefits assessments to remain valid during the life of the contract, changes in circumstances that have occurred since the original benefits case was prepared must be taken into account. A supplier may have been selected on its ability to deliver a 5% saving of annual operational cost compared with other options, but after two years has delivered no savings. However, where this is due to changes to contract, or general industry costs that would have also affected the other options, it is likely that a relative cost saving is still being realized. A maintained benefits case shows that saving.

Benefits assessments often receive lower priority than cost-saving initiatives, and are given less priority in performance reports than issues and problem summaries, but it is important to the long-term relationship that achievements are recognized. A benefits report must make objective assessments against the original objectives, but may also include morale-boosting anecdotal evidence of achievements and added value.

Hints and tips

It is important for both organizations, and for the longevity of the relationship, that the benefits being derived from the relationship are regularly reviewed and reported.

Balanced, managed relationships

To ensure that all activities and contacts for a supplier are consistent and coordinated, each supplier relationship should have a single nominated individual accountable for all aspects of the relationship.

Example of consistent supplier relationship management

A nationwide retail organization had an individual owning the overall management of its major network services supplier. However, services, contracts and billing were managed by several individuals spread throughout the organization. The individual owner put forward a business case for single ownership of the supplier and all the various contracts, together with consolidation of all the individual invoices into a single quarterly bill. The estimated cost savings to the organization were in excess of £600,000 per annum.

An assessment of the success of a supplier relationship, from a business perspective, is likely to be substantially based on financial performance. Even where a service is performing well, it may not be meeting one or both parties' financial targets. It is important that both parties continue to benefit financially from the relationship. A contract that squeezes the margins of a supplier too tightly may lead to under-investment by the supplier, resulting in a gradual degradation of service, or even threaten the viability of the supplier. In either case this may result in adverse business impacts to the organization.

The key to the successful long-term financial management of the contract is a joint effort directed towards maintaining the financial equilibrium, rather than a confrontational relationship delivering short-term benefits to only one party.

Building relationships takes time and effort. As a result, the organization may only be able to build long-term relationships with a few key suppliers. The experience, culture and commitment of those involved in running a supplier relationship are at least as important as having a good contract and governance regime. The right people with the right attitudes in the relationship team can make a poor contract work, but a good contract does not ensure that a team with poor relationships will deliver.

A considerable amount of time and money is normally invested in negotiating major supplier deals, with more again at risk for both parties if the relationship is not successful. Both organizations must ensure that they invest suitably in the human resources allocated to managing the relationship. The personality, behaviours and culture of the relationship representatives all influence the relationship. For a partnering relationship, all those involved need to be able to respect and work closely and productively with their opposite numbers.

4.8.5.6 Contract renewal or termination

Contract reviews must be undertaken on a regular basis to ensure each contract is continuing to meet business needs. Contract reviews assess the contract operation holistically and at a more senior level than the service reviews that are undertaken at an operational level. These reviews should consider:

- How well the contract is working and its relevance for the future
- Whether changes are needed: services, products, contracts, agreements, targets
- What is the future outlook for the relationship – growth, shrinkage, change, termination, transfer etc.?
- Commercial performance of the contract, reviews against benchmarks or market assessments, suitability of the pricing structure and charging arrangements
- Guidance on future contract direction and ensuring best-practice management processes are established
- Supplier and contract governance.

For high-value, lengthy or complex supply arrangements, the period of contract negotiation and agreement can be lengthy, costly and may involve a protracted period of negotiation. It can be a natural inclination to wish to avoid further changes to a contract for as long as possible. However, for the business to derive full value from the supplier relationship, the contract must be able to be regularly and quickly amended to allow the business to benefit from service developments.

Benchmarking provides an assessment against the marketplace. The supplier may be committed by the contract to maintaining charges against a market price. To maintain the same margin, the supplier is obliged to improve its operational efficiency in line with its competitors. Collectively, these methods help provide an assessment of an improving or deteriorating efficiency.

The point of responsibility within the organization for deciding to change a supplier relationship is likely to depend on the type of relationship. The service provider may have identified a need to change the supplier, based on the existing supplier's performance, but for a contractual relationship the decision needs to be taken in conjunction with the organization's procurement and legal departments.

The organization should take careful steps to:

- Perform a thorough impact and risk assessment of a change of supplier on the organization and its business, especially during a period of transition. This could be particularly significant in the case of a strategic relationship.
- Make a commercial assessment of the exit costs. This may include contractual termination costs if supplier liability is not clear, but the largest costs are likely to be associated with a transition project. For any significant-sized relationship, this typically includes a period of dual-supply as services are migrated. Any change associated with a change in supplier will increase costs, either immediately as fixed costs or over time where borne by the supplier and reflected back in service charges.
- Take legal advice on termination terms, applicable notice period and mechanisms, and any other consequences, particularly if the contract is to be terminated early.
- Reassess the market to identify potential benefits in changing supplier.

A prudent organization undertakes most of these steps at the time the original contract is established, to ensure the right provisions and clauses are included, but this review activity needs to be reassessed when a change of supplier is being considered.

4.8.6 Triggers, inputs, outputs and interfaces

4.8.6.1 Triggers

There are many events that could trigger supplier management activity. These include:

- New or changed corporate governance guidelines
- New or changed business and IT strategies, policies or plans
- New or changed business needs or new or changed services

- New or changed requirements within agreements, such as SLRs, SLAs, OLAs or contracts
- Review and revision of designs and strategies
- Periodic activities such as reviewing, revising or reporting, including review and revision of supplier management policies, reports and plans
- Requests from other areas, particularly SLM and information security management, for assistance with supplier issues
- Requirements for new contracts, contract renewal or contract termination
- Re-categorization of suppliers and/or contracts.

4.8.6.2 Inputs

Inputs comprise:

- **Business information** From the organization's business strategy, plans and financial plans, and information on its current and future requirements
- **Supplier and contracts strategy** This covers the sourcing policy of the service provider and the types of supplier and contract used. It is produced by the service strategy processes
- **Supplier plans and strategies** Details of the business plans and strategies of suppliers, together with details of their technology developments, plans and statements and information on their current financial status and projected business viability
- **Supplier contracts, agreements and targets** Of both existing and new contracts and agreements from suppliers
- **Supplier and contract performance information** Of both existing and new contracts and suppliers
- **IT information** From the IT strategy and plans and current budgets
- **Performance issues** The incident and problem management processes, with incidents and problems relating to poor contract or supplier performance
- **Financial information** From financial management for IT services, the cost of supplier service(s) and service provision, the cost of contracts and the resultant business benefit; and the financial plans and budgets, together with the costs associated with service and supplier failure
- **Service information** From the SLM process, with details of the services from the service portfolio and the service catalogue, service level targets within SLAs and SLRs, and possibly from the monitoring of SLAs, service reviews and breaches of the SLAs. Also customer satisfaction data on service quality
- **CMS** Containing information on the relationships between the business, the services, the supporting services and the technology.

4.8.6.3 Outputs

The outputs of supplier management are used within all other parts of the process, by many other processes and by other parts of the organization. Often this information is supplied as electronic reports or displays on shared areas or as pages on intranet servers to ensure the most up-to-date information is always used. The information provided is as follows:

- **SCMIS** This holds the information needed to execute the activities within supplier management – for example, the data monitored and collected as part of supplier management. This is then invariably used as an input to all other parts of the supplier management process.
- **Supplier and contract performance information and reports** These are used as input to supplier and contract review meetings to manage the quality of service provided by suppliers and partners. This should include information on shared risk where appropriate.
- **Supplier and contract review meeting minutes** These are produced to record the minutes and actions of all review meetings with suppliers.
- **Supplier SIPs** These are used to record all improvement actions and plans agreed between service providers and their suppliers, wherever they are needed, and should be used to manage the progress of agreed improvement actions, including risk reduction measures.
- **Supplier survey reports** Often many people within a service provider organization have dealings with suppliers. Feedback from these individuals should be collated to ensure consistency in the quality of service provided by suppliers in all areas. These can be published as league tables to encourage competition between suppliers.

4.8.6.4 Interfaces

The key interfaces that supplier management has with other processes are:

- **SLM** Supplier management provides assistance with the determining of targets, requirements and responsibilities for suppliers. It then sees to their inclusion within underpinning agreements and contracts to ensure that they support all SLR and SLA targets. SLM assists supplier management in the investigation of SLA and SLR breaches caused by poor supplier performance. SLM also provides invaluable input into the supplier management review process.
- **Change management** Supplier contracts and agreements are controlled documents and therefore subject to appropriate change management procedures. When changes are proposed, the involvement of suppliers should be assessed and reflected in planning.
- **ISM** ISM relies on supplier management for the management of suppliers and their access to services and systems, and their responsibilities with regard to conformance to the service provider's ISM policies and requirements.
- **Financial management for IT services** This process provides adequate funds to finance supplier management requirements and contracts and provides financial advice and guidance on purchase and procurement matters.
- **Service portfolio management** This process looks to supplier management input to ensure that all supporting services and their details and relationships are accurately reflected within the service portfolio.
- **ITSCM** This process works with supplier management with regard to the management of continuity service suppliers.

4.8.7 Information management

All the information required by supplier management should be contained within the SCMIS. This should include all information relating to suppliers and contracts, as well as all the information relating to the operation of the supporting services provided by suppliers. Information relating to these supporting services should also be contained within the service portfolio, together with the relationships to all other services and components. This information should be integrated and maintained in alignment with all other IT management information systems, particularly the service portfolio and the CMS.

4.8.8 Critical success factors and key performance indicators

The following list includes some sample CSFs for supplier management. Each organization should identify appropriate CSFs based on its objectives for the process. Each sample CSF is followed by a small number of typical KPIs that support the CSF. These KPIs should not be adopted without careful consideration. Each organization should develop KPIs that are appropriate for its level of maturity, its CSFs and its particular circumstances. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

- **CSF** Business protected from poor supplier performance or disruption
  - **KPI** Increase in the number of suppliers meeting the targets within the contract
  - **KPI** Reduction in the number of breaches of contractual targets
- **CSF** Supporting services and their targets align with business needs and targets
  - **KPI** Increase in the number of service and contractual reviews held with suppliers
  - **KPI** Increase in the number of supplier and contractual targets aligned with SLA and SLR targets
- **CSF** Availability of services is not compromised by supplier performance
  - **KPI** Reduction in the number of service breaches caused by suppliers
  - **KPI** Reduction in the number of threatened service breaches caused by suppliers
- **CSF** Clear ownership and awareness of supplier and contractual issues
  - **KPI** Increase in the number of suppliers with nominated supplier managers
  - **KPI** Increase in the number of contracts with nominated contract managers.

4.8.9 Challenges and risks

4.8.9.1 Challenges

Supplier management faces many challenges, which could include:

- Continually changing business and IT needs and managing significant change in parallel with delivering existing service
- Working with an imposed non-ideal contract, a contract that has poor targets or terms and conditions, or poor or non-existent definition of service or supplier performance targets
- Legacy issues, especially with services recently outsourced
- Insufficient expertise retained within the organization
- Being tied into long-term contracts, with no possibility of improvement, which have punitive penalty charges for early exit
- Situations where the supplier depends on the organization in fulfilling the service delivery (e.g. for a data feed) can lead to issues over accountability for poor service performance
- Disputes over charges
- Interference by either party in the running of the other's operation
- Being caught in a daily fire-fighting mode, losing the proactive approach
- Poor communication – not interacting often enough or quickly enough or not focusing on the right issues
- Personality conflicts and/or cultural conflicts
- One party using the contract to the detriment of the other party, resulting in win–lose changes rather than joint win–win changes
- Losing the strategic perspective, focusing on operational issues, causing a lack of focus on strategic relationship objectives and issues.

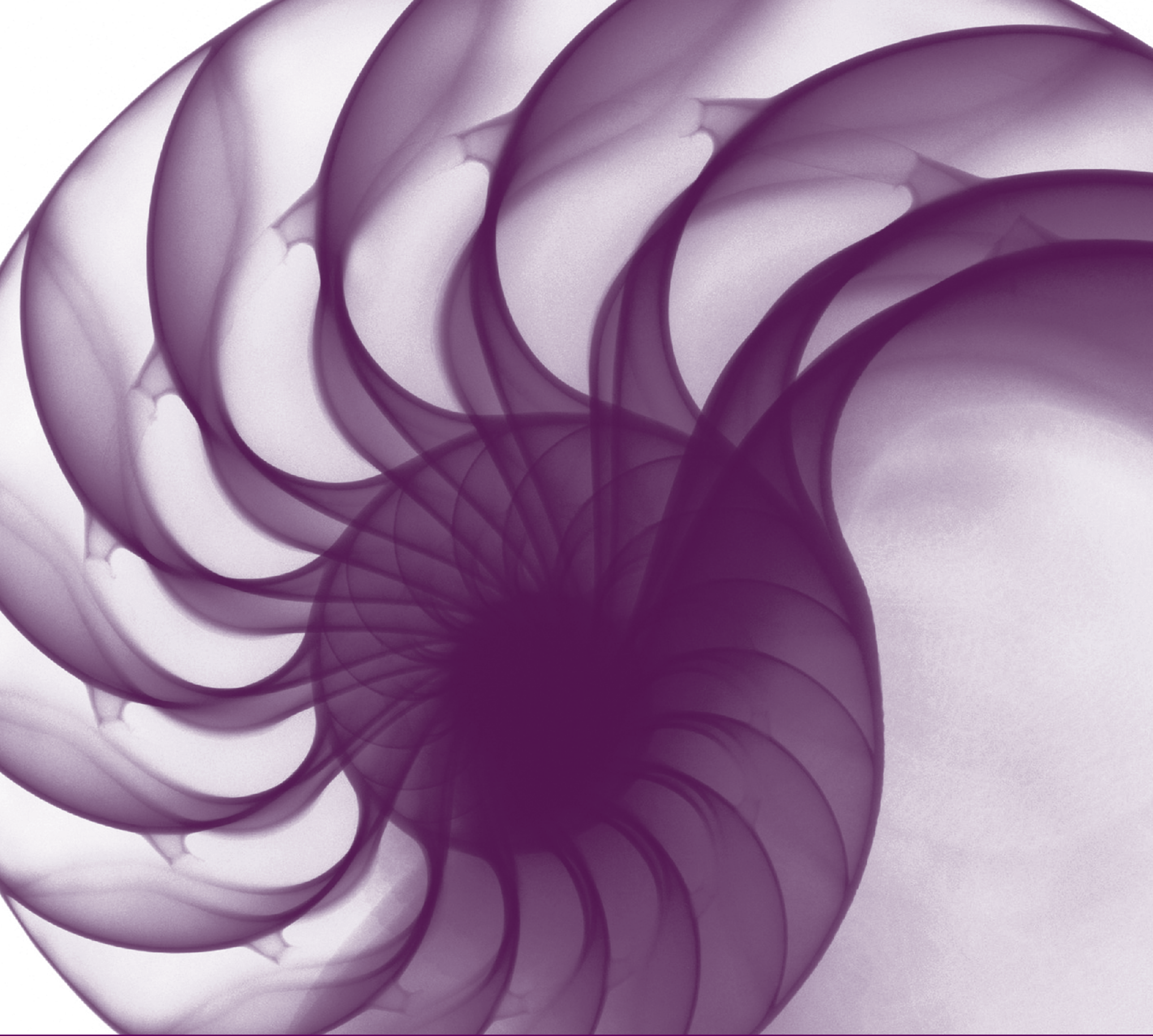
Key elements that can help to avoid the above issues are:

- A clearly written, well-defined and well-managed contract
- A mutually beneficial relationship
- Clearly defined (and communicated) roles and responsibilities on both sides
- Good interfaces and communications between the parties
- Well-defined service management processes on both sides
- Selecting suppliers who have achieved certification against internationally recognized certifications, such as ISO 9001 and ISO/IEC 20000.

4.8.9.2 Risks

The major areas of risk associated with supplier management include:

- Lack of commitment from the business and senior management to the supplier management process and procedures
- Lack of appropriate information on future business and IT policies, plans and strategies
- Lack of resources and/or budget for the supplier management process
- Legacy of badly written and agreed contracts that do not underpin or support business needs or SLA and SLR targets
- Suppliers agree to targets and service levels within contracts that are impossible to meet, or suppliers fail or are incapable of meeting the terms and conditions of the contract
- Supplier personnel or organizational culture are not aligned with that of the service provider or the business
- Lack of clarity and integration by supplier with service management processes, policies and procedures of the service provider
- Suppliers are not cooperative and are not willing to partake in and support the required supplier management process
- Suppliers are taken over and relationships, personnel and contracts are changed
- The demands of corporate supplier and contract procedures are excessive and bureaucratic
- Poor corporate financial processes, such as procurement and purchasing, do not support good supplier management.



5 Service design technology-related activities

5 Service design technology-related activities

This chapter considers the technology-related activities of requirement engineering and the development of technology architectures for management of data and information and for management of applications.

5.1 REQUIREMENTS ENGINEERING

Requirements engineering is the approach by which sufficient rigour is introduced into the process of understanding and documenting the requirements of the business, users and all other stakeholders, and ensuring traceability of changes to each requirement. This process comprises the stages of elicitation, analysis (which feeds back into the elicitation) and validation. All these contribute to the production of a rigorous, complete requirements document. The core of this document is a repository of individual requirements that is developed and managed. Often these requirements are instigated by IT but ultimately they need to be documented and agreed with the business.

There are many guidelines on requirements engineering, including the Recommended Practice for Software Requirements Specifications (IEEE 830), the Software Engineering Body of Knowledge (SWEBOK), capability maturity model integration (CMMI) and the V-Model, which is described in detail in *ITIL Service Transition*. Information about several international standards that may also be of use may be found in Appendix N.

It is important to remember that the guidance in this chapter focuses on the requirements for the technology related to a service. There many other areas around which an organization will need to define requirements to ensure successful design, transition and operation of a complete service, such as requirements for:

- User training
- Support staff training
- Marketing and communication related to the service and its deployment
- Service documentation
- Organizational and cultural readiness.

The standards and methods for defining requirements of this sort will be developed as part of the design coordination process, which will also ensure ongoing adherence during the service design stage.

5.1.1 Different requirement types

A fundamental assumption here is that the analysis of the current and required business processes results in functional requirements met through IT services (comprising applications, data, infrastructure, environment and support skills).

There are commonly said to be three major types of requirements for any system. These are:

- **Functional requirements** For a service these requirements are those necessary to support a particular business function, business process or to remove a customer or user constraint. These requirements describe the utility aspects of a service.
- **Management and operational requirements** (sometimes referred to as non-functional requirements) These define the requirements and constraints on the service and address the need for a responsive, available and secure service, and deal with such issues as ease of deployment, operability, management needs and security. These requirements describe the warranty aspects of a service.
- **Usability requirements** These requirements are those that relate to how easy it is for the user to access and use the service to achieve the desired outcomes, including addressing the 'look and feel' needs of the user. This requirement type is often seen as part of management and operational requirements, but for the purposes of this section it will be addressed separately. Depending on the context, these requirements enable utility, support warranty and influence user perceptions of a service.

5.1.1.1 Functional requirements

Functional requirements describe the things a service is intended to do – in other words, the utility it will provide – and can be expressed as tasks or functions that the component is required to perform. One approach for specifying functional requirements is through such methods as a system context diagram or a use case model. Other approaches show how the inputs are to be transformed into the outputs (data flow or object diagrams) and textual descriptions.

A system context diagram, for instance, captures all information exchanges between, on the one hand, the IT service and its environment and, on the other, sources or destinations of data used by the service. These information exchanges and data sources represent constraints on the service under development.

A use case model defines a goal-oriented set of interactions between external actors and the service under consideration. Actors are parties outside the service that interact with the service. An actor may represent a class of user, roles that users can play, or other services and their requirements. The main purpose of use case modelling is to establish the boundary of the proposed system and fully state the functional capabilities to be delivered to the users. Use cases are also helpful for establishing communication between business and application developers. They provide a basis for sizing and feed the definition of usability requirements. Use cases define all scenarios that an application has to support and can therefore easily be expanded into test cases. Since use cases describe a service's functionality on a level that is understandable for both business and IT, they can serve as a vehicle to specify the functional elements of a service level agreement (SLA), such as the actual business deliverables from the service.

One level 'below' the use case and the context diagram, many other modelling techniques can be applied. These models depict the static and dynamic characteristics of the services under development. A conceptual data model (whether called object or data) describes the different 'objects' in the service, their mutual relationships and their internal structure. Dynamics of the service can be described using state models (e.g. state transition diagrams) that show the various states of the entities or objects, together with events that may cause state changes. Interactions between the different application components can be described using interaction diagrams (e.g. object interaction diagrams).

> **Hints and tips**
>
> Alongside a mature requirements modelling process, computer-aided software engineering (CASE) tools can help in getting and keeping these models consistent, correct and complete.

5.1.1.2 Management and operational requirements

Management and operational requirements (or non-functional requirements) are used to define requirements and constraints on the IT service. The requirements serve as a basis for early systems and service sizing and estimates of cost, and can support the assessment of the viability of the proposed IT service. Management and operational requirements should also encourage developers to take a broader view of project goals.

Categories of management and operational requirements include:

- **Manageability** Does it run? Does it fail? How does it fail?
- **Efficiency** How many resources does it consume?
- **Availability and reliability** How reliable does it need to be?
- **Capacity and performance** What level of capacity do we need to support storage and throughput requirements?
- **Security** What classification of security is required?
- **Installation** How much effort does it take to install the application? Is it using automated installation procedures?
- **Continuity** What level of resilience and recovery is required?
- **Controllability** Can it be monitored, managed and adjusted?
- **Maintainability** How well can the application be adjusted, corrected, maintained and changed for future requirements?
- **Operability** Do the applications disturb other applications in their functionalities?
- **Measurability and reportability** Can we measure and report on all of the required aspects of the application?

The management and operational requirements can be used to prescribe the quality or warranty attributes of the application or service being built. These quality attributes can be used to design test plans for testing the applications on their compliance with management and operational requirements.

5.1.1.3 Usability requirements

The primary purpose of usability requirements is to ensure that the service meets the expectations of its users with regard to its ease of use. To achieve this:

- Establish performance standards for usability evaluations
- Define test scenarios for usability test plans and usability testing.

Like the management and operational requirements, usability requirements can also be used as the quality attributes of the application or service being built. These quality attributes can be used to design test plans for testing the applications on their compliance to usability requirements.

In order to establish usability requirements, care must be taken to establish the types of likely users and to understand their varied needs. For example, users who are colour-blind would not find a service that relied heavily on colour differentiation easy to use, or users working in a second language may have difficulty with screen terminology that does not translate well.

5.1.2 Requirements for support – the user view

Users have formally defined roles and activities as user representatives in requirements definition and acceptance testing. They should be actively involved in identifying all aspects of service requirements, including the three categories in section 5.1.1 above, and also in:

- User training procedures and facilities
- Support activities and service desk procedures.

5.1.3 Requirements investigation techniques

A range of techniques may be used to investigate business situations and elicit service requirements. Sometimes the customers and the business are not completely sure of what their requirements actually are and will need some assistance and prompting from the designer or requirements gatherer. This must be completed in a sensitive way to ensure that it is not seen as IT dictating business requirements. The two most commonly used techniques are interviewing and workshops, but these are usually supplemented by other techniques, such as observation and scenarios.

5.1.3.1 Interviews

The interview is a key tool and can be vital in achieving a number of objectives, such as:

- Making initial contact with key stakeholders and establishing a basis for progress
- Building and developing rapport with different users and managers
- Acquiring information about the business situation, including issues and problems.

There are three areas that are considered during interviews:

- Current business processes that need to be fulfilled in any new business systems and services
- Problems with the current operations that need to be addressed
- New features required from the new business system or service and any supporting IT service.

The interviewing process is improved when the interviewer has prepared thoroughly as this saves time by avoiding unnecessary explanations and demonstrates interest and professionalism. The classic questioning structure of 'Why, What, Who, When, Where, How' provides an excellent framework for preparing for interviews.

It is equally important to formally close the interview by:

- Summarizing the points covered and the actions agreed
- Explaining what happens next, both following the interview and beyond
- Asking the interviewee how any further contact should be made.

It is always a good idea to write up the notes of the interview as soon as possible – ideally straight away and usually by the next day. The advantages of interviewing are:

- Builds a relationship with the users

- Can yield important information
- Opportunity to understand different viewpoints and attitudes across the user group
- Opportunity to investigate new areas that arise
- Collection of examples of documents and reports
- Appreciation of political factors
- Study of the environment in which the new service will operate.

The disadvantages of interviewing are:

- Expensive in elapsed time
- No opportunity for conflict resolution
- No opportunity for consensus building.

5.1.3.2 Workshops

Workshops provide a forum in which issues can be discussed, conflicts resolved and requirements elicited. Workshops are especially valuable when time and budgets are tightly constrained, several viewpoints need to be canvassed and an iterative and incremental view of service development is being taken.

The advantages of the workshop are:

- Gain a broad view of the area under investigation – having a group of stakeholders in one room will allow a more complete understanding of the issues and problems
- Increase speed and productivity – it is much quicker to have one meeting with a group of people than interviewing them one by one
- Obtain buy-in and acceptance for the IT service
- Gain a consensus view – if all the stakeholders are involved, the chance of them taking ownership of the results is improved.

There are some disadvantages, including:

- It can be time-consuming to organize – for example, it is not always easy to get all the necessary people together at the same time
- It can be difficult to get all of the participants with the required level of authority
- It can be difficult to get a mix of business and operational people to understand the different requirements.

The success or failure of a workshop session depends, in large part, on the preparatory work by the facilitator and the business sponsor for the workshop. They should spend time before the event planning the following areas:

- The objective of the workshop – this has to be an objective that can be achieved within the time constraints of the workshop.
- Who will be invited to participate in the workshop – it is important that all stakeholders interested in the objective should be invited to attend or be represented.
- The structure of the workshop and the techniques to be used. These need to be geared towards achieving the defined objective (for example, requirements gathering or prioritization) and should take the needs of the participants into account.
- Arranging a suitable venue – this may be within the organization, but it is better to use a 'neutral' venue out of the office.

During the workshop, a facilitator needs to ensure that the issues are discussed, views are aired and progress is made towards achieving the stated objective. A record needs to be kept of the key points emerging from the discussion. At the end of

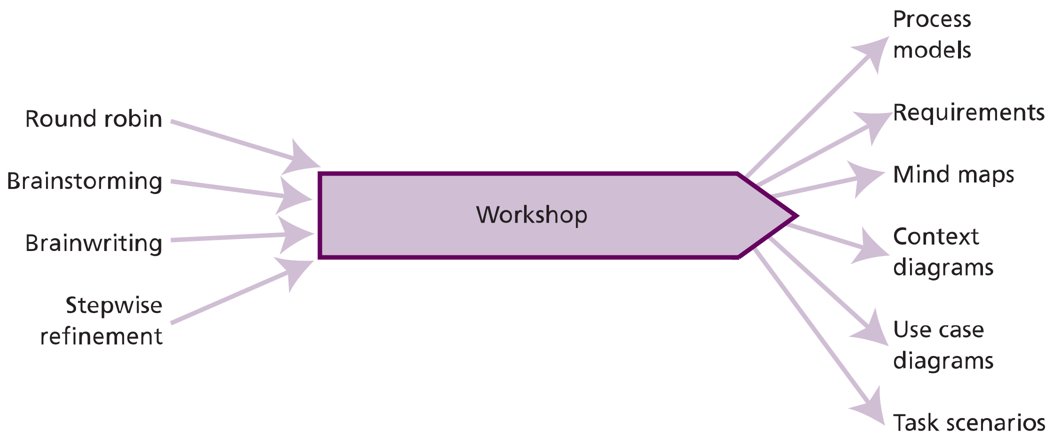


Figure 5.1 Requirements – workshop techniques

the workshop, the facilitator needs to summarize the key points and actions. Each action should be assigned to an owner.

There are two main categories of technique required for a requirements workshop – techniques for discovery and techniques for documentation, as shown in Figure 5.1.

5.1.3.3 Observation

Observing the workplace is very useful in obtaining information about the business environment and the work practices. This has two advantages:

- A much better understanding of the problems and difficulties faced by the business users
- It will help devise workable solutions that are more likely to be acceptable to the business.

Conversely, being observed can be rather unnerving, and the old saying 'you change when being observed' needs to be factored into your approach and findings.

Formal observation involves watching a specific task being performed. There is a danger of being shown just the 'front-story' without any of the everyday variances, but it is still a useful tool.

5.1.3.4 Protocol analysis

Protocol analysis is simply getting the users to perform a task, and for them to describe each step as they perform it.

5.1.3.5 Shadowing

Shadowing involves following a user for a period such as a day to find out about a particular job. It is a powerful way to understand a particular user role. Asking for explanations of how the work is done, or the workflow, clarifies some of the already assumed aspects.

5.1.3.6 Scenario analysis

Scenario analysis is essentially telling the story of a task or transaction. Its value is that it helps a user who is uncertain what is needed from a new service to realize it more clearly. Scenarios are also useful when analysing or redesigning business processes. A scenario will trace the course of a transaction from an initial business trigger through each of the steps needed to achieve a successful outcome.

Scenarios provide a framework for discovering alternative paths that may be followed to complete the transaction. This is extremely useful in requirements elicitation and analysis because real-life situations, including the exceptional circumstances, are debated.

Scenarios offer significant advantages:

- They force the user to include every step, so there are no taken-for-granted elements and the problem of tacit knowledge is addressed.
- By helping the user to visualize all contingencies, they help to cope with the uncertainty about future systems and services.
- A workshop group refining a scenario will identify those paths that do not suit the corporate culture.
- They provide a tool for preparing test scripts.

The disadvantages of scenarios are that they can be time-consuming to develop, and some scenarios can become very complex. Where this is the case, it is easier to analyse if each of the main alternative paths is considered as a separate scenario.

A popular approach to documenting scenario descriptions is to develop use case descriptions to support use case diagrams. However, there are also a number of graphical methods of documenting a scenario, such as storyboards, activity diagrams, task models and decision tree diagrams.

5.1.3.7 Prototyping

Prototyping is an important technique for eliciting, analysing, demonstrating and validating requirements. It is difficult for users to envisage the new service before it is actually built. Prototypes offer a way of showing the user how the new service might work and the ways in which it can be used. If a user is unclear what they need the service to do for them, utilizing a prototype often releases blocks to thinking and can produce a new wave of requirements. Incremental and iterative approaches to service development, such as the dynamic systems development method (DSDM), use evolutionary prototyping as an integral part of their development lifecycle.

There is a range of approaches to building prototypes. They may be built using an application development environment so that they mirror the service; images of the screens and navigations may be built using presentation software; or they may simply be 'mock-ups' on paper.

There are two basic methods of prototyping:

- The throw-away mock-up, which is only used to demonstrate the look and feel
- The incremental implementation, where the prototype is developed into the final system.

It is important to select consciously which is to be used, otherwise there is a danger that a poor-quality mock-up becomes the basis for the real system, causing problems later on.

There is a strong link between scenarios and prototyping because scenarios can be used as the basis for developing prototypes. In addition to confirming the users' requirements, prototyping can often help the users to identify new requirements. Prototypes are successfully used to:

- Clarify any uncertainty on the part of the service developers and confirm to the user that what they have asked for has been understood
- Open the user up to new requirements as they understand what the service will be able to do to support them
- Show users the 'look and feel' of the proposed service and elicit usability requirements
- Validate the requirements and identify any errors.

Potential problems include:

- Endless iteration
- A view that if the prototype works, the full service can be ready tomorrow.

5.1.3.8 Other techniques

Other techniques that could be used include:

- **Questionnaires** These can be useful to get a limited amount of information from a lot of people when interviewing them all would not be practical or cost-effective.
- **Special-purpose records** This technique involves the users in keeping a record about a specific issue or task. For example, they could keep a simple five-bar gate record about how often they need to transfer telephone calls – this could provide information about the problems with this business process.
- **Activity sampling** This is a rather more quantitative form of observation and can be used when it is necessary to know how people spend their time. For example: How much time is spent on invoicing? How much time is spent on reconciling payments? How much time is spent on sorting out queries?

5.1.4 Problems with requirements engineering

Requirements, seen by users as the uncomplicated bit of a new service development, are actually the most problematic aspect, and yet the time allocated is far less than for the other phases.

Tight timescales and tight budgets – both the result of constraints on the business – place pressures on the development team to deliver a service. The trouble is that without the due time to understand and define the requirements properly, the service that is delivered on time may not be the service that the business thought it was asking for.

Studies carried out into IT project failures tell a common story. Many of the projects and unsatisfactory IT services suggest the following conclusions:

- A large proportion of errors (over 80%) are introduced at the requirements phase.
- Very few faults (fewer than 10%) are introduced at design and development – developers are developing things right, but frequently not developing the right things.
- Most of the project time is allocated to the development and testing phases of the project.
- Less than 12% of the project time is allocated to requirements.

These findings are particularly significant because the cost of correcting errors in requirements increases dramatically the later into the development lifecycle they are found.

One of the main problems with requirements engineering is the lack of detailed skill and overall understanding of the area where people use it. If accurately performed, the work can integrate requirements from numerous areas in a few questions.

Other typical problems with requirements have been identified as:

- Lack of relevance to the objectives of the service
- Lack of clarity in the wording
- Ambiguity
- Duplication between requirements

- Conflicts between requirements
- Requirements expressed in such a way that it is difficult to assess whether or not they have been achieved
- Requirements that assume a solution rather than stating what is to be delivered by the service
- Uncertainty among users about what they need from the new service
- Users omitting to identify requirements
- Inconsistent levels of detail
- Failure to include requirements from service provider stakeholders such as service operations and support staff
- An assumption that user and IT staff have knowledge that they do not possess and therefore a failure to ensure that there is a common understanding
- Failure to include non-technical requirements critical to success of the service such as requirements for training, documentation, communications and marketing
- Requirements creep – the gradual addition of seemingly small requirements without taking the extra effort into account in the project plan.

Another problem is an apparent inability on the part of the users to articulate clearly what it is they wish the service to do for them. Very often they are deterred from doing so because the nature of the requirement under discussion is explained in a straightforward statement rather than in an open-ended manner that encourages a detailed exchange.

5.1.4.1 *Resolving requirements engineering problems*

Defining actors

There are some participants that must take part in the requirements engineering process. They represent three broad stakeholder groups:

- The business
- The user community
- The service development/management team.

The user community should be represented by the domain expert (or subject-matter expert) and end-users.

Dealing with tacit knowledge

When developing a new service, the users will pass on to IT their explicit knowledge, i.e. knowledge of procedures and data that is at the front of their minds and that they can easily articulate. A major problem when eliciting requirements is that of tacit knowledge, i.e. those other aspects of the work that a user is unable to articulate or explain.

Some common elements that cause problems and misunderstandings are:

- **Skills** Explaining how to carry out actions using words alone is extremely difficult.
- **Taken-for-granted information** Even experienced and expert business users may fail to mention information or clarify terminology, and the analyst may not realize that further questioning is required.
- **Front-story/back-story** This issue concerns a tendency to frame a description of current working practices, or a workplace, in order to give a more positive view than is actually the case.
- **Future systems knowledge** If the study is for a new service development, with no existing expertise or knowledge in the organization, how can the prospective users know what they want?
- **Common language** The difficulty of an outsider assuming a common language for discourse, and common norms of communication. (If they do not have this, then the potential for misrepresentation of the situation can grow considerably.)
- **Intuitive understanding** This is usually born of considerable experience. Decision makers are often thought to follow a logical, linear path of enquiry while making their decisions. In reality though, as improved decision-making skills and knowledge are acquired, the linear path is often abandoned in favour of intuitive pattern recognition.
- **Organizational culture** Without an understanding of the culture of an organization, the requirements exercise may be flawed.

Communities of practice are discrete groups of workers – maybe related by task, by department, by geographical location or some other factor – that have their own sets of norms and practices,

Table 5.1 Requirements engineering – tacit and explicit knowledge

	Tacit	Explicit
Individual	Skills, values, taken-for-granted, intuitiveness	Tasks, job descriptions, targets, volumes and frequencies
Corporate	Norms, back-story, culture, communities of practice	Procedures, style guides, processes, knowledge sharing

Table 5.2 Requirements engineering: examples of explicit and tacit knowledge[6]

Technique	Explicit knowledge	Tacit knowledge	Skills	Future requirements
Interviewing	✓✓	✓	X	✓
Shadowing	✓✓	✓✓	✓✓	X
Workshops	✓✓	✓✓	X	✓✓
Prototyping	✓✓	✓✓	✓✓	✓✓
Scenario analysis	✓✓	✓✓	X	✓✓
Protocol analysis	✓✓	✓✓	✓✓	X

Key: ✓✓ = suitable technique to discover this kind of knowledge; ✓ = less suitable technique to discover this kind of knowledge; X = unsuitable technique to discover this kind of knowledge.

[6] Rugg, G. and Maiden, N.A.M. (1995). Knowledge acquisition techniques for requirements engineering. Conference paper: 1994 workshop on requirements elicitation for software systems, Keele, Staffordshire.

distinct from other groups within the organization and the organization as a whole.

Table 5.1 provides some examples of tacit versus explicit knowledge that might be expected to come from both individuals and the corporation.

Table 5.2 shows the relationship between various techniques of requirements gathering and the kinds of knowledge they are most likely to develop. For example, the technique of shadowing is suitable for uncovering both explicit and tacit knowledge, as well as observing skills (how to carry something out); however, this technique is not likely to uncover future requirements.

5.1.5 Documenting requirements

The requirements document is at the heart of the process and can take a number of forms. Typically the document will include a catalogue of requirements, with each individual requirement documented using a standard template. One or more models showing specific aspects, such as the processing or data requirements, may supplement this catalogue.

Before they are formally entered into the catalogue, requirements are subject to careful scrutiny. This scrutiny may involve organizing the requirements into groupings and checking that each requirement is 'well-formed'.

Once the document is considered to be complete, it must be reviewed by business representatives and confirmed to be a true statement of the requirements at that point in time. During this stage the reviewers examine the requirements and question whether they are well-defined, clear and complete.

As we uncover the requirements from our various users, we need to document them. This is best done in two distinct phases – building the requirements list and, later, developing an organized requirements catalogue. The requirements list tends to be an informal document and can be presented in a table as four columns:

- Requirements
- Source
- Comment
- Detail level.

Each requirement in the list must be checked to see whether or not it is well formed and SMART (specific, measurable, achievable, relevant and time-bound).

When checking the individual and totality of requirements, the following checklist can be used:

- Are the requirements, as captured, unambiguous?
- Is the meaning clear?
- Is the requirement aligned to the service development and business objectives, or is it irrelevant?
- Is the requirement reasonable, or would it be expensive and time-consuming to satisfy?
- Do any requirements conflict with one another such that only one may be implemented?
- Do they imply a solution rather than state a requirement?
- Is each requirement entered separately, or are they really several requirements grouped into one entry?
- Do several requirements overlap or duplicate each other?

There are several potential outcomes from the exercise:

- Accept the requirement as it stands
- Re-word the requirement to remove jargon and ambiguity
- Merge duplicated/overlapping requirements
- Take unclear and ambiguous requirements back to the users for clarification.

5.1.5.1 The requirements catalogue

The requirements catalogue is the central repository of the users' requirements, and all the requirements should be documented here, following the analysis as described above. The requirements catalogue should form part of the service pipeline within the overall service portfolio. As the requirements take shape, requirements from all stakeholders can be documented here to provide a single source of complete information on the requirements for the new or changed service. Each requirement that has been analysed is documented using a standard template, such as that shown in Table 5.3.

Table 5.3 Requirements template

IT service	Author		Date
Requirement ID	Requirement name		
Source	Owner	Priority	Business process
Functional requirement description			
Management and operational and usability requirements	Description		
Justification			
Related documents			
Related requirements			
Comments			
Resolution			
Version no	Change history	Date	Change request

The key entries in the template are as follows:

- **Requirement ID** This is a unique ID that never changes and is used for traceability – for example, to reference the requirement in design documents, test specifications or implemented code. This ensures that all requirements have been met and that all implemented functions are based on requirements.
- **Source** The business area or users who requested the requirement or the document where the requirement was raised. Recording the source of a requirement helps ensure that questions can be answered or the need can be re-assessed in the future if necessary.
- **Owner** The user who accepts ownership of the individual requirement will agree that it is worded and documented correctly, and will sign it off at acceptance testing when satisfied.
- **Priority** The level of importance and need for a requirement. Usually approaches such as MoSCoW are used, where the following interpretation of the mnemonic applies:
  - **Must have** – a key requirement without which the service has no value.
  - **Should have** – an important requirement that must be delivered but, where time is short, could be delayed for a future delivery. This should be a short-term delay, but the service would still have value without it.
  - **Could have** – a requirement that would be beneficial to include if it does not cost too much or take too long to deliver, but it is not central to the service.
  - **Won't have** (but would like in the future) – a requirement that will be needed in the future but is not required for this delivery. In a future service release, this requirement may be upgraded to a 'must have'.
- **Requirement description** A succinct description of the requirement. A useful approach is to describe the requirement using the following structure:
  - Actor (or user role)
  - Verb phrase
  - Object (noun or noun phrase).
- **Related documents** Where the requirement incorporates complex business rules or data validation, a decision table or decision tree may be more useful to define complex business rules, while data validation rules may be defined in a repository. If a supplementary technique is used to specify or model the requirement, there should be a cross-reference to the related document.
- **Business process** A simple phrase to group together requirements that support a specific activity, such as sales, inventory, customer service, and so on.
- **Justification** Not all requirements that are requested will be met. This may be due to time and budget constraints, or may be because one requirement is dropped in favour of a conflicting requirement. Often a requirement is not met because it adds little value to the business. The justification sets out the reasons for requesting a requirement.
- **Related requirements** Requirements may be related to each other for several reasons. Sometimes there is a link between the functionality required by the requirements or a high-level requirement is clarified by a series of more detailed requirements.
- **Change history** The entries in this section provide a record of all the changes that have affected the requirement. This is required for configuration management and traceability purposes.

Guidance on assigning priority in the requirements template

The following should be clearly agreed:

- Requirement priorities can and do change over the life of a service development project.
- 'Should have' requirements need to be carefully considered because, if they are not delivered within the initial design stage, they may be impossible to implement later.
- Requirements are invariably more difficult and more expensive to meet later in the service lifecycle.
- It is not just the functional requirements that can be 'must haves' – some of the management and operational requirements should be 'must haves'.

5.1.5.2 Full requirements documentation

Effective requirements documentation should comprise the following elements:

- A glossary of terms, to define each organizational term used within the requirements document. This will help manage the problem of local jargon and will clarify synonyms and homonyms for anyone using the document
- A scoping model, such as a system context diagram
- The requirements catalogue, ideally maintained as part of an overall service portfolio
- Supporting models, such as business process models, data flow diagrams or interaction diagrams.

Managing changes to the documentation

Changes may come about because:

- The scope of the new service has altered through budget constraints
- The service must comply with new regulations or legislation
- Changes in business priorities
- Stakeholders have understood a requirement better after some detailed analysis (for example, using scenarios or prototyping) and amended the original requirement accordingly.

There are a number of specialist support tools on the market to support requirements processes. These are sometimes called CARE (computer-aided requirements engineering) or CASE. Features include:

- Maintaining cross-references between requirements
- Storing requirements documentation
- Managing changes to the requirements documentation
- Managing versions of the requirements documentation
- Producing formatted requirements specification documents from the database
- Ensuring documents delivered by any solution project are suitable to enable support.

5.1.6 Requirements and outsourcing

The aim is to select standard packaged solutions wherever possible to meet service requirements. However, whether solutions to IT requirements are to be purchased off the shelf, developed in-house or outsourced, all the activities up to the production of a specification of business requirements are done in-house. Many IT service development contracts assume it is possible to know what the requirements are at the start, and that it is possible to produce a specification that unambiguously expresses the requirements. For all but the simplest services this is almost never true. Requirements analysis is an iterative process – the requirements will change during the period the application and service are being developed. It will require user involvement throughout the development process, as in the DSDM and other 'agile' approaches.

5.1.6.1 Typical requirements outsourcing scenarios

Typical approaches to a contract for the development of IT systems to be delivered in support of an IT service are as follows:

- **Low-level requirements specification** The boundary between 'customer' and provider is drawn between the detailed requirements specification and any design activities. All the requirements that have an impact on the user have been specified in detail, giving the provider a very clear and precise implementation target. However, there is increased specification effort, and the added value of the provider is restricted to the less difficult aspects of development.
- **High-level requirements specification** The customer/provider boundary is between the high-level requirements and all other phases. The provider contract covers everything below the line. The customer is responsible for testing the delivered service against the business requirements. As it is easier to specify high-level requirements, there is reduced effort to develop contract inputs. However, there may be significant problems of increased cost and risk for both customer and provider, together with increased room for mistakes, instability of requirements and increased difficulty in knowing what information systems are wanted.

5.2 MANAGEMENT OF DATA AND INFORMATION

Data is one of the critical asset types that need to be managed in order to develop, deliver and support IT services effectively. Data/information management is how an organization plans, collects, creates, organizes, uses, controls, disseminates and disposes of its data/information, both structured records and unstructured data. It also ensures that the value of that data/information is identified and exploited, both in support of its internal operations and in adding value to its customer-facing business processes.

A number of terms are common in this area, including 'data management', 'information management' and 'information resource management'. For the purposes of this publication, the term 'data management' is used as shorthand for all of the three above.

Key message

The role of data management is not just about managing raw data: it is about managing all the contextual metadata – additional 'data about the data' – that goes with it, and when added to the raw data gives 'information' or 'data in context'.

Data, as the basis for the organization's information, has all the necessary attributes to be treated as an asset (or resource). For example, it is essential for 'the achievement of business objectives and the successful daily workings of an organization'. In addition, it can be 'obtained and preserved by an organization, but only at a financial cost'. Finally it can, along with other resources/assets, be used to 'further the achievement of the aims of an organization'.

Key factors for successful data management are as follows:

- All users have ready access through a variety of channels to the information they need to do their jobs.
- Data assets are fully exploited, through data sharing within the organization and with other bodies.
- Data assets are adequately protected and secured in accordance with corporate and IT security policies.
- The quality of the organization's data is maintained at an acceptable level, and the information used in the business is accurate, reliable and consistent.
- Legal requirements for maintaining the privacy, security, confidentiality and integrity of data are observed.
- The organization achieves a high level of efficiency and effectiveness in its data and information-handling activities.
- An enterprise data model is used to define the most important entities and their relationships – this helps to avoid redundancies and to avoid the deterioration of the architecture as it is changed over the years.

5.2.1 Managing data assets

If data is not managed effectively:

- People maintain and collect data that is not needed.
- The organization may have historic information that is no longer used.
- The organization may hold a lot of data that is inaccessible to potential users.
- Information may be disseminated to more people than it should be, or not to those people to whom it should.
- The organization may use inefficient and out-of-date methods to collect, analyse, store and retrieve the data.
- The organization may fail to adhere to regulatory requirements such as data retention or security.
- The organization may fail to collect the data that it needs, reducing data quality, and data integrity is lost, for example, between related data sources.

In addition, whether or not information is derived from good-quality data is a difficult question to answer, because there are no measurements in place against which to compare it. For example, poor data quality often arises because of poor checks on input and/or updating procedures. Once inaccurate or incomplete data has been stored in the IT system, any reports produced using this data will reflect these inaccuracies or gaps. There may also be a lack of consistency between internally generated management information from the operational systems, and from other internal, locally used systems, created because the central data is not trusted.

One way of improving the quality of data is to use a data management process that establishes policies and standards, provides expertise and makes it easier to handle the data aspects of new services. This should then allow full data/information asset management to:

- Add value to the services delivered to customers
- Reduce risks in the business
- Reduce the costs of business processes
- Stimulate innovation in internal business processes.

5.2.2 Scope of data management

There are four areas of management included within the scope of data/information management:

- **Management of data resources** The governance of information in the organization must ensure that all these resources are known and that responsibilities have been assigned for their management, including ownership of data and metadata. This process is normally referred to as data administration and includes responsibility for:
  - Defining information needs
  - Constructing a data inventory and an enterprise data model
  - Identifying data duplication and deficiencies
  - Maintaining a catalogue/index of data/information content
  - Measuring the cost and value of the organization's data.
- **Management of data/information technology** The management of the IT that underpins the organization's information systems; this includes processes such as database design and database administration. This aspect is normally handled by specialists within one of the IT functions – see *ITIL Service Operation* for more details.
- **Management of information processes** Business processes will lead to IT services involving one or other of the data resources of the organization. The activities of creating, collecting, accessing, modifying, storing, deleting and archiving data – i.e. the data lifecycle – must be properly controlled, often jointly with the application management activities.
- **Management of data standards and policies** The organization will need to define standards and policies for its data management as an element of an IT strategy. Policies will govern the procedures and responsibilities for data management in the organization, as well as technical policies, architectures and standards that will apply to the IT infrastructure that supports the organization's information systems.

The best-practices scope of data management activities includes managing non-structured data that is not held in conventional database systems – for example, using formats such as text, image and audio. It is also responsible for ensuring process quality at all stages of the data lifecycle, from requirements to retirement. The main focus in this publication will be on its role in the requirements, design and development phases of the asset and service lifecycle.

The team supporting data management activities may also provide a business information support service. In this case the team is able to answer questions about the meaning, format and availability of data internal to the organization because it manages the metadata. It is also able to understand and explain what external data might be needed in order to carry out necessary business processes and will take the necessary action to source this.

Critically, when creating or redesigning processes and supporting IT services, it is best practice to consider reusing data and metadata across different areas of the organization. The ability to do this may be supported by a corporate data model – sometimes known as a common information model – to help support reuse, often a major objective for data management.

5.2.3 Data management and the service lifecycle

It is recommended that a lifecycle approach be adopted in understanding the use of data in business processes. General issues include:

- What data is currently held and how can it be classified?
- What data needs to be collected or created by the business processes?
- How will the data be stored and maintained?
- How will the data be accessed, by whom and in what ways?
- How will the data be disposed of, and under whose authority?
- How will the quality of the data be maintained (accuracy, consistency, currency etc.)?

- How can the data be made more accessible/available?

5.2.4 Supporting the service lifecycle

During requirements and initial design, data management can assist design and development teams with service-specific data modelling and give advice on the use of various techniques to model data.

During detailed ('physical') design and development, the data management team (usually part of the technical management function) can provide technical expertise on database management systems and on how to convert initial 'logical' models of data into physical, product specific, implementations.

Many new services have failed because poor data quality has not been addressed during the development of the service, or because a particular development created its own data and metadata, without consultation with other service owners, or with data management.

5.2.5 Valuing data

Data is an asset and has value. Clearly in some organizations this is more obvious than in others. Organizations that are providers of data to others – for example, Yell, Dun & Bradstreet, and Reuters – can value data as an 'output' in terms of the price that they are charging external organizations to receive it. It is also possible to think of value in terms of what the internal data would be worth to another organization.

It is more common to value data in terms of what it is worth to the owner organization. A number of ways of doing this have been suggested:

- **Valuing data by availability** One approach often used is to consider which business processes would not be possible if a particular piece of data was unavailable, and how much that non-availability of data would cost the business.
- **Valuing lost data** Another approach is to think about the costs of obtaining some data if it were to be destroyed.
- **Valuing data by considering the data lifecycle** This involves thinking about how data is created or obtained in the first place, how it is made available to people to use, and how data is retired, either through archiving or physical destruction. It may be that some data is provided from an external source and then held internally, or it may be that data has to be created by the organization's internal systems. In these two cases, the lifecycle is different and the processes that take place for data capture will be entirely separate. In both cases the costs of redoing these stages can be evaluated.

The more highly valued the data, the more the effort that needs to be expended on ensuring its confidentiality, integrity and availability.

5.2.6 Classifying data

Data can be initially classified as operational, tactical or strategic:

- **Operational data** This data is necessary for the ongoing functioning of an organization and can be regarded as the lowest, most specific, level.
- **Tactical data** This data is usually needed by second-line management – or higher – and is typically concerned with summarized data and historical data, typically year-to-year data or quarterly data. Often the data that is used here appears in management information systems that require summary data from a number of operational systems in order to deal with an accounting requirement, for example.
- **Strategic data** This data is often concerned with longer-term trends and comparison with the outside world. Therefore providing the necessary data for a strategic support system involves bringing together the operational and tactical data from many different areas with relevant external data. Much more data is required from external sources.

An alternative method is to use a security classification of data and documents. This is normally adopted as a corporate policy within an organization. An orthogonal classification distinguishes between organization-wide data, functional-area data and service-specific data:

- Organization-wide data needs to be centrally managed.
- The next level of data is functional-area data, which should be shared across a complete business function. This involves sharing data 'instances' (for example, individual customer records) and also ensuring that consistent metadata across that functional area, such as standard address formats, are being used.

- The final level is IT service-specific, where the data and metadata are valid for one IT service and do not need to be shared with other services.

5.2.7 Setting data standards

One of the critical aspects of data administration is to ensure that standards for metadata are in place – for example, what metadata is to be kept for different underlying 'data types'. Different details are kept about structured tabular data than for other areas. 'Ownership' is a critical item of this metadata, some sort of unique identifier is another, a description in business meaningful terms another, and a format might be another. The custodian or steward, someone in the IT organization who takes responsibility for the day-to-day management of the data, is also recorded.

Another benefit of a data management process would be in the field of reference data. Certain types of data, such as postcodes or names of countries, may be needed across a variety of systems and need to be consistent. It is part of the responsibility of data administration to manage reference data on behalf of the whole business, and to make sure that the same reference data is used by all systems in the organization.

Hints and tips

Standards for naming must be in place, so, for example, if a new type of data is requested in a new service, then there is a need to use names that meet these standards. An example standard might be 'all capitals, no underlining and no abbreviations'.

5.2.8 Data ownership

Data administration can assist the service developer by making sure responsibilities for data ownership are taken seriously by the business and by the IT department. One of the most successful ways of doing this is to get the business and the IT organization to sign up to a data charter – a set of procedural standards and guidance for the careful management of data in the organization, by adherence to corporately defined standards. Responsibilities of a data owner are often defined here and may include:

- Agreeing a business description and a purpose for the data
- Defining who can create, amend, read and delete occurrences of the data
- Authorizing changes in the way data is captured or derived
- Approving any format, domain and value ranges
- Approving the relevant level of security, including making sure that legal requirements and internal policies about data security are adhered to.

5.2.9 Data migration

Data migration is an issue where a new service is replacing one or more existing services, and it is necessary to carry across, into the new service, good-quality data from the existing systems and services. There are two types of data migration of interest to projects here: one is the data migration into data warehouses etc., for business intelligence/analytics purposes; the other is data migration to a new transactional, operational service. In both cases it will be beneficial if data migration standards, procedures and processes are laid down by data management. Data migration tools may have already been purchased on behalf of the organization by the data management team. Without this support, it is very easy to underestimate the amount of effort that is required, particularly if data consolidation and cleaning has to take place between multiple source systems, and the quality of the existing services' data is known to be questionable.

5.2.10 Data capture

It is also very important to work with data management on effective measures for data capture. The aim here is to capture data as quickly and accurately as possible. There is a need to ensure that the data capture processes require the minimum amount of keying, and exploit the advantages that graphical user interfaces provide in terms of minimizing the number of keystrokes needed, also decreasing the opportunity for errors during data capture. It is reasonable to expect that the data management process has standards for, and can provide expertise on, effective methods of data capture in various environments, including 'non-structured' data capture using mechanisms such as scanning.

5.2.11 Data storage

One area where technology has moved on very rapidly is in the area of storage of data. There is a need to consider different storage media (for example, optical storage) and to be aware of the size and cost implications associated with this. The main reason for understanding the developments in this area is that they make possible many types of data management that were considered too expensive before. For example, to store real-time video, which uses an enormous bandwidth, has, until the last two to three years, been regarded as too expensive. The same is true of the scanning of large numbers of paper documents, particularly where those documents are not text-based but contain detailed diagrams or pictures. Understanding technology developments with regard to electronic storage of data is critical to understanding the opportunities for the business to exploit the information resource effectively by making the best use of new technology.

5.2.12 Data retrieval and usage

Once the data has been captured and stored, the next aspect to consider is the retrieval of information from the data. Services to allow easy access to structured data via query tools of various levels of sophistication are needed by all organizations, and generate their own specific architectural demands.

The whole area of searching within scanned text and other non-structured data such as video, still images or sound is a major area of expansion. Techniques such as automatic indexing, and the use of search engines to give efficient access via keywords to relevant parts of a document, are essential technologies that have been widely implemented, particularly on the internet. Expertise in the use of data or content within websites should exist within the data management as well as content management – standards and procedures that are vital for websites.

5.2.13 Data integrity and related issues

When defining requirements for IT services, it is vital that management and operational requirements related to data are considered. In particular, the following areas must be addressed:

- Recovery of lost or corrupted data
- Controlled access to data
- Implementation of policies on archiving of data, including compliance with regulatory retention periods
- Periodic data integrity checks.

Data integrity is concerned with ensuring that the data is of high quality and uncorrupted. It is also about preventing uncontrolled data duplication, and hence avoiding any confusion about what is the valid version of the data. There are several approaches that may assist with this. Various technology devices such as 'database locking' are used to prevent multiple, inconsistent, updating of data. In addition, prevention of illegal updating may be achieved through access control mechanisms.

5.3 MANAGEMENT OF APPLICATIONS

Definition: application

An application is defined as software that provides functions which are required by an IT service. Each application may be part of more than one IT service. An application runs on one or more servers or clients.

Applications, along with data and infrastructure components such as hardware, the operating system and middleware, make up the technology components that are part of a service. The application itself is only one component, albeit an important one of the service. Therefore it is important that the application delivered matches the agreed requirements of the business. However, too many organizations spend too much time focusing on the functional requirements of the new service and application, and insufficient time is spent designing the management and operational requirements (non-functional requirements) of the service. This means that when the service becomes operational, it meets all of the functionality required, but totally fails to meet the expectation of the business and the customers in terms of its quality and performance; it therefore becomes unusable.

Two alternative approaches are necessary to fully implement management of applications. One approach employs an extended service development lifecycle (SDLC) to support the development of a service. SDLC is a systematic approach to problem solving and is composed of the following steps:

- Feasibility study
- Analysis
- Design
- Testing
- Implementation
- Evaluation
- Maintenance.

The other approach takes a global view of all services to ensure the ongoing maintainability and manageability of the applications:

- All applications are described in a consistent manner, via an application portfolio that is managed and maintained to enable alignment with dynamic business needs.
- Consistency of approach to development is enforced through a limited number of application frameworks and design patterns and through a 'reuse first' philosophy.
- Common software components, usually to meet management and operational requirements, are created or acquired at an 'organizational' level and used by individual systems as they are designed and built.

5.3.1 The application portfolio

This is simply a full record of all applications within the organization and is dynamic in its content. Table 5.4 presents examples of the attributes an organization may wish to capture in the application portfolio for each application listed there.

5.3.2 Linking application and service portfolios

Some organizations maintain a separate application portfolio with separate attributes, while in other organizations the application portfolio is stored within the configuration management system (CMS), together with the appropriate relationships. Other organizations combine the application portfolio together with the service portfolio. It is for each organization to decide the most appropriate strategy for its own needs. What is clear is that there should be very close relationships and links between the applications and the services they support and the infrastructure components used.

5.3.3 Application frameworks

The concept of an application framework is a very powerful one. The application framework covers all management and operational aspects and actually provides solutions for all the management and operational requirements that surround an application.

Implied in the use of application frameworks is the concept of standardization. If an organization uses and has to maintain an application framework for every single application, there will not be many benefits of the use of an application framework.

Table 5.4 Examples of application portfolio attributes

Application name	IT operations owner	New development cost
Application identifier	IT development owner	Annual operational costs
Application description	Support contacts	Annual support cost
Business process supported	Database technologies	Annual maintenance costs
IT services supported	Dependent applications	Outsourced components
Executive sponsor	IT systems supported	Outsource partners
Geographies supported	User interfaces	Production metrics
Business criticality	IT architecture, including Network topology	OLA link
SLA link	Application technologies used	Support metrics
Business owner	Number of users	

An organization that wants to develop and maintain application frameworks, and to ensure the application frameworks comply with the needs of the application developers, must invest in doing so. It is essential that applications framework architectures are not developed in isolation, but are closely related and integrated with all other framework and architectural activities. The service, infrastructure, environment and data architectures must all be closely integrated with the application architecture and framework.

5.3.3.1 Architecture, application frameworks and standards

Architecture-related activities have to be planned and managed separately from individual system-based software projects. It is also important that architecture-related activities be performed for the benefit of more than just one application. Application developers should focus on a single application, while application framework developers should focus on more than one application, and on the common features of those applications in particular.

A common practice is to distinguish between various types of application. For instance, not every application can be built on top of a Microsoft® Windows operating system platform, connected to a UNIX server, using HTML, Java applets, JavaBeans and a relational database. The various types of application can be regarded as application families. All applications in the same family are based on the same application framework.

Utilizing the concept of an application framework, the first step of the application design phase is to identify the appropriate application framework. If the application framework is mature, a large number of the design decisions are given. If it is not mature, and all management and operational requirements cannot be met on top of an existing application framework, the preferred strategy is to collect and analyse the requirements that cannot be dealt with in the current version of the application framework. Based on the application requirements, new requirements can be defined for the application framework. Next, the application framework can be modified so that it can cope with the application requirements. In fact, the whole family of applications that corresponds to the application framework can then use the newly added or changed framework features.

Hints and tips

Developing and maintaining an application framework is a demanding task and, like all other design activities, should be performed by competent and experienced people. Alternatively, application frameworks can be acquired from third parties.

5.3.4 The need for CASE tools and repositories

One important aspect of that overall service design is the need to align applications with their underlying support structures. Application development environments traditionally have their own CASE tools that offer the means to specify requirements, draw design diagrams (according to particular modelling standards), or even generate complete applications, or nearly complete application skeletons, almost ready to be deployed. These environments also provide a central location for storing and managing all the elements that are created during application development, generally called a repository. Repository functionality includes version control and consistency checking across various models. The current approach is to use metaCASE tools to model domain-specific languages and use these to make the CASE-work more aligned to the needs of the business.

5.3.5 Design of specific applications

The requirements phase was addressed earlier in the requirements engineering section of this chapter. The design phase is one of the most important phases within the application lifecycle. It ensures that an application is conceived with operability and management of the application in mind. This phase takes the outputs from the requirements phase and turns them into the specification that will be used to develop the application.

The goal for designs should be satisfying the organization's requirements. Design includes the design of the application itself, and the design of the infrastructure and environment within which the application operates. Architectural considerations are the most important aspect of this phase, since they can impact on the structure and content of both application and operational model. Architectural considerations for the application (design of the application architecture) and architectural considerations for the environment (design of the IT architecture) are strongly related and need to be aligned. Application architecture and design should not be considered in isolation but should form an overall integrated component of service architecture and design. Ensuring that this overall integrated approach is used falls within the design coordination process.

Generally, in the design phase, the same models will be produced as have been delivered in the requirements phase, but during design many more details are added. New models include the architecture models, where the way in which the different functional components are mapped to the physical components (e.g. desktops, servers, databases and network) needs to be defined. The mapping, together with the estimated load of the system, should allow for the sizing of the infrastructure required.

Another important aspect of the architecture model is the embedding of the application in the existing environment. Which pieces of the existing infrastructure will be used to support the required new functions? Can existing servers or networks be used? With what impact? Are required functions available in existing applications that can be utilized? Do packages exist that offer the functionality needed or should the functions be built from scratch?

The design phase takes all requirements into consideration and starts assembling them into an initial design for the solution. Doing this not only gives developers a basis to begin working; it is also likely to bring up questions that need to be asked of the customers/users. If possible, application frameworks should be applied as a starting point.

It is not always possible to foresee every aspect of a solution's design ahead of time. As a solution is developed, new things will be learned about how to do things and also how not to.

The key is to create a flexible design, so that making a change does not send developers all the way back to the beginning of the design phase. There are a number of approaches that can minimize the chance of this happening, including:

- Designing for management and operational requirements
- Managing trade-offs
- Using application-independent design guidelines; using application frameworks
- Employing a structured design process/ manageability checklist.

Design for management and operational requirements means giving management and operational requirements a level of importance similar to that for the functional requirements, and including them as a mandatory part of the design phase. This includes a number of management and operational requirements such as availability, capacity, maintainability, reliability, continuity and security. It is now inconceivable in modern application development projects that user interface design (usability requirements) would be omitted as a key design activity. However, many organizations ignore or forget manageability. Details of the necessary management and operational requirements are contained within the service design package (SDP) and service acceptance criteria (SAC) in Appendices A and B, respectively.

5.3.6 Managing trade-offs

Managing trade-off decisions focuses on balancing the relationship among resources, the project schedule, and those features that need to be included in the application for the sake of quality.

When development teams try to complete this balancing, it is often at the expense of the management and operational requirements. One way to avoid that is to include management and operational requirements in the application-independent design guidelines – for example, in the form of an application framework. Operability and manageability effectively become standard components of all design processes (for example, in the form of an application framework) and get embedded into the working practices and culture of the development organization.

5.3.7 Typical design outputs

The following are examples of the outputs from an applications design forming part of the overall service design:

- Input and output design, including forms and reports
- A usable user interface (human/computer interaction) design

- A suitable data/object model
- A process flow or workflow model
- Detailed specifications for update and read-only processes
- Mechanisms for achieving audit controls, security, confidentiality and privacy
- A technology specific 'physical' design
- Scripts for testing the systems design
- Interfaces and dependencies on other applications.

There are guidelines and frameworks that can be adopted to determine and define design outputs within application management, such as CMMI.

5.3.8 Design patterns

A design pattern is a general, repeatable solution to a commonly occurring problem in software design. Object-oriented design patterns typically show relationships and interactions between classes or objects, without specifying the final application classes or objects that are involved. Design patterns describe both a problem and a solution for common issues encountered during application development.

An important design principle used as the basis for a large number of the design patterns found in recent literature is that of separation of concerns (SoC). Separation of concerns will lead to applications divided into components, with a strong cohesion and minimal coupling between components. The advantage of such an application is that modification can be made to individual components with little or no impact on other components.

In typical application development projects, more than 70% of the effort is spent on designing and developing generic functions and on satisfying the management and operational requirements. That is because each individual application needs to provide a solution for such generic features as printing, error handling and security.

Among others, the Object Management Group (OMG, www.omg.com) defined a large number of services that are needed in every application. OMG's object management architecture clearly distinguishes between functional and management and operational aspects of an application. It builds on the concept of providing a run-time environment that offers all sorts of facilities to an application.

In this concept, the application covers the functional aspects, and the environment covers all management and operational aspects. Application developers should, by definition, focus on the functional aspects of an application, while others can focus on the creation of the environment that provides the necessary management and operational services. This means that the application developers focus on the requirements of the business, while the architecture developers or application framework developers focus on the requirements of the application developers.

5.3.9 Developing individual applications

Once the design phase is completed, the application development team will take the designs that have been produced and move on to developing the application. Both the application and the related environment are made ready for deployment. Application components are coded or acquired, integrated and tested.

To ensure that the application is developed with management at the core, the development team needs to focus on ensuring that the developing phase continues to correctly address the management and operational aspects of the design (e.g. responsiveness, availability, security). Application development must also be done with clear understanding of how the application fits into the overall service solution. All service requirements should be found in the SDP.

The development phase guidance covers the following topics:

- Consistent coding conventions
- Application-independent building guidelines
- Operability testing
- Management checklist for the building phase
- Organization of the build team roles.

5.3.9.1 Consistent coding conventions

The main reason for using a consistent set of design and coding conventions is to standardize the structure and coding style of an application so that everyone can easily read, understand and manage the application development process. Good design and coding conventions result in precise, readable and unambiguous source code that is consistent with the organizational coding and management standards and is as intuitive to follow as possible. Adding application operability into this convention ensures that all applications are built in a way that ensures that they can be fully managed all the way through their lifecycles.

A coding convention itself can be a significant aid to managing the application, as consistency allows the management tools to interact with the application in a known way. It is better to introduce a minimum set of conventions that everyone will follow rather than to create an overly complex set that encompasses every facet but is not followed or used consistently across the organization.

5.3.10 Templates and code generation

A number of development tools provide a variety of templates for creating common application components. Rather than creating all the pieces of an application from scratch, developers can customize an existing template. They can also reuse custom components in multiple applications by creating their own templates. Other development tools will generate large pieces of code (skeletons) based on the design models and coding conventions. The code could include hooks at the code pieces that need to be added.

In this respect, templates and application frameworks should be considered IT assets. These assets not only guide the developing of applications, but also incorporate the lessons learned or intellectual capital from previous application development efforts. The more that standard components are designed into the solution, the faster applications can be developed, against lower costs in the long term (not ignoring the fact that development of templates, code generators and application frameworks requires significant investment).

5.3.11 Embedded application instrumentation

The development phase deals with incorporating instrumentation into the fabric of the application. Developers need a consistent way to provide instrumentation for application drivers/middleware components (e.g. database drivers) and applications that is efficient and easy to implement. To keep application developers from trying to start from the beginning for every new application they develop, the computer industry provides methods and technologies to simplify and facilitate the instrumentation process.

These include:

- Application Response Measurement (ARMS)
- IBM Application Management Specification (AMS)
- Common Information Model (CIM) and Web-Based Enterprise Management (WBEM) from the Distributed Management Task Force (DMTF)
- Desktop Management Interface (DMI)
- Microsoft Windows© Management Instrumentation (WMI)
- Java Management Extension (JMX).

Each of these technologies provides a consistent and richly descriptive model of the configuration, status and operational aspects of applications and services. These are provided through programming application program interfaces (APIs) that the developer incorporates into an application, normally through the use of standard programming templates.

It is important to ensure that all applications are built to conform to some level of compliance for the application instrumentation. Ways to do this could include:

- Provide access to management data through the instrumentation API
- Publish management data to other management systems, again through the instrumentation API
- Provide applications event handling
- Provide a diagnostic hook.

5.3.11.1 Diagnostic hooks

Diagnostic hooks are of greatest value during testing and when an error has been discovered in the production service. They mainly provide the information necessary to solve problems and application errors rapidly and restore service. They can also be used to provide measurement and management information of applications.

The four main categories are:

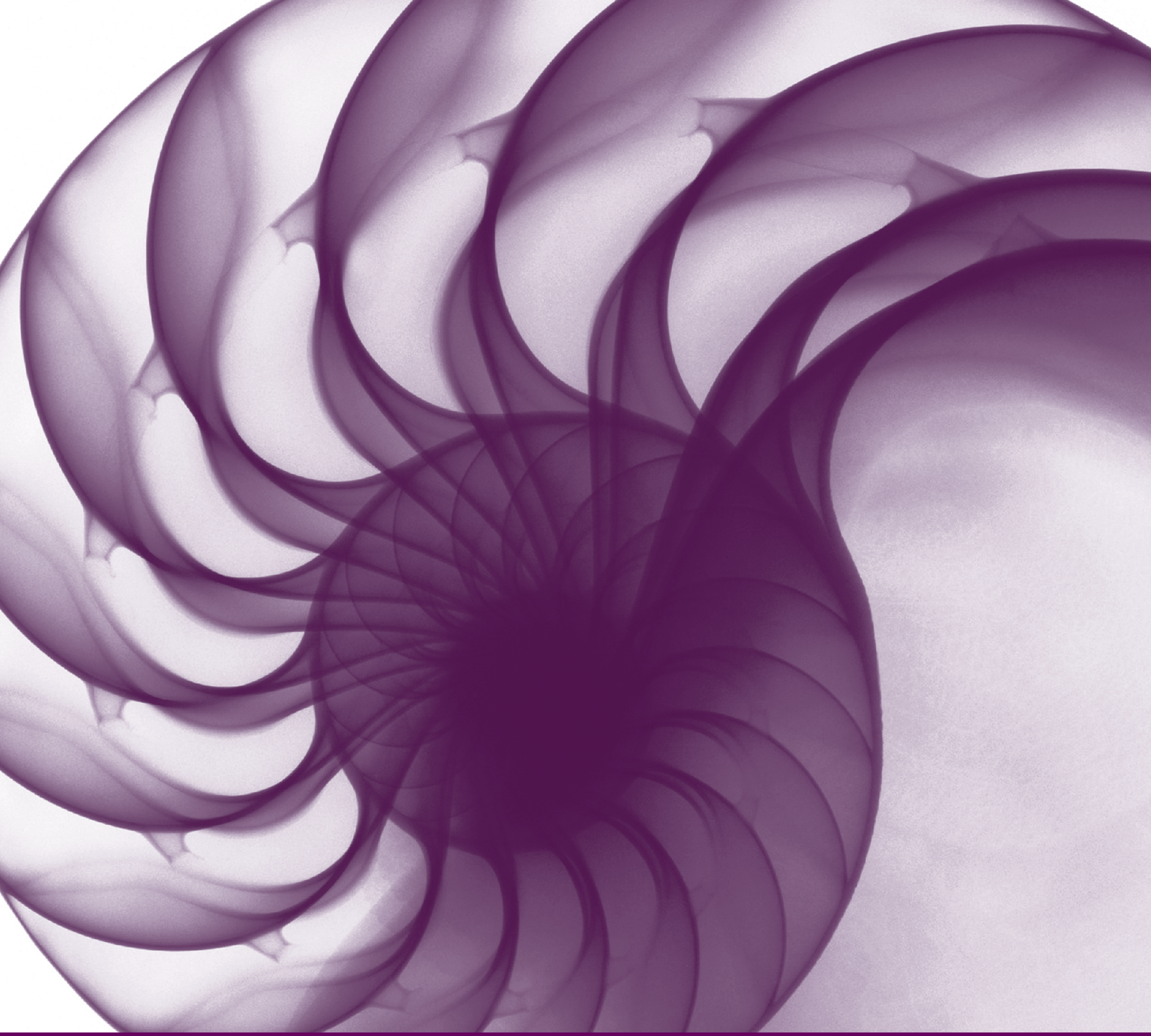
- System-level information provided by the operating systems and hardware
- Software-level information provided by the application infrastructure components such as database, web server or messaging systems

- Custom information provided by the applications
- Information on component and service performance.

5.3.12 Major outputs from development

The major outputs from the development phase are:

- Scripts to be run before or after deployment
- Scripts to start or stop the application
- Scripts to check hardware and software configurations of target environments before deployment or installation
- Specification of metrics and events that can be retrieved from the application and that indicate the performance status of the application
- Customized scripts initiated by service operation staff to manage the application (including the handling of application upgrades)
- Specification of access control information for the system resources used by an application
- Specification of the details required to track an application's major transactions
- SLA targets and requirements
- Operational requirements and documentation
- Support requirements
- Application recovery and backups
- Other IT service management requirements and targets.



6 Organizing for service design

6 Organizing for service design

This chapter describes the general concepts of organizing for service management in relation to service design and the related practices. It includes generic roles, responsibilities and competencies that apply across the service lifecycle and specific aspects for the processes described in this publication.

Section 2.2.3 describes the basic concepts of organization, function, group, team, department, division and role that are used in this chapter.

6.1 ORGANIZATIONAL DEVELOPMENT

There is no single best way to organize, and the best practices described in ITIL need to be tailored to suit individual organizations and situations. Any changes made will need to take into account resource constraints and the size, nature and needs of the business and customers. The starting point for organizational design is strategy. Organization development for service management is described in more detail in *ITIL Service Strategy*, Chapter 6.

6.2 FUNCTIONS

A function is a team or group of people and the tools or other resources they use to carry out one or more processes or activities. In larger organizations, a function may be broken out and performed by several departments, teams and groups, or it may be embodied within a single organizational unit (e.g. the service desk). In smaller organizations, one person or group can perform multiple functions – e.g. a technical management department could also incorporate the service desk function.

For service design to be successful, an organization will need to clearly define the roles and responsibilities required to undertake the processes and activities identified in Chapters 4 and 5. These roles will need to be assigned to individuals, and an appropriate organizational structure of teams, groups or functions established and managed.

ITIL Service Design does not define any specific functions of its own, but it does rely on the technical and application management functions described in *ITIL Service Operation*. Technical and application management provide the technical resources and expertise to manage the whole service lifecycle, and practitioner roles within service design may be performed by members of these functions.

An organization may already have one or both of the functions described in sections 6.2.1 and 6.2.2. If either or both of these functions exist, they will play a prominent role in service design activities.

6.2.1 Alignment with application development

While it is possible for an IT service provider to design, deploy, deliver and improve IT services without developing any applications in house, many if not most organizations perform some of their own software development. When this is the case, the organization will typically assign the work to a functional unit specializing in application development.

If an application development function exists, this team will focus on building functionality – that is, the utility – required by the business. Historically, what the application does is more important to this team than how the application is operated. This is why the input of application management, technical management, IT operations management and even the service desk should be sought during service design to ensure that the overall output of service design will meet all customer needs, not just those related to functionality.

Most application development work is performed as part of a project where the focus is on delivering specific units of work to specification, on time and within budget. Application development may, therefore, be overly focused on the narrow parameters of the project, particularly if they have little responsibility to support the application once the team has moved on to the next project. This problem can be compounded if the person(s) leading the overall service requirements definition are part of the application development function. In this situation they may be too focused on functionality of the most prominent application in the service and neglect the detailed requirements and design of manageability, training, documentation, marketing and other important elements needed for success in transition and ongoing service operation.

Hints and tips

While costs associated with application development may seem relatively easy to quantify, since they are frequently linked to carefully budgeted projects, time needed to diagnose and recover from application errors discovered in service operation should also be accounted for and budget reserved for these activities.

The application development function utilizes software development lifecycles to guide and provide formal structure to their work. This work must then, in turn, be integrated into the overall service lifecycle as the applications to be developed form a central part of the IT services provisioned by the IT service provider organization as a whole. As the various possible software development lifecycles that could be employed are well documented in other sources, they are not discussed here.

6.2.2 Alignment with project management

Another functional unit that may exist within the IT service provider organization is project management, sometimes called the project management office (PMO). The purpose of this team is to define and maintain the service provider's project management standards and to provide overall resources and management of IT projects. The project management function usually leverages the principles of project management as described in one or more of the recognized project management methodologies such as PRojects IN Controlled Environments (PRINCE2) from the Office of Government Commerce (OGC) or the Project Management Body of Knowledge (PMBOK) from the Project Management Institute (PMI).

If a project management function exists, they will be actively involved in the work of the service design as well as the service transition stages of the service lifecycle, as well as during any other temporary endeavour that would benefit from application of formal project management.

The project management function can provide value not only through the management of individual projects and through the propagation of consistent and repeatable project management methods, but also through providing project portfolio management. Project portfolio management interfaces with overall service portfolio management and ensures that resources are appropriately allocated across the complete set of projects being managed and maximizes project success.

For more information on how to establish, develop and maintain appropriate support structures for portfolios, programmes and projects, see *Portfolio, Programme and Project Offices* (OGC, 2008).

6.2.3 Example service design organization structures

The following example organization structures show how the various service design roles might be combined and structured. Each organization should consider all of the roles that they require and how these can be combined within their organizational constraints to create a structure that meets their needs.

6.2.3.1 Small organization

In the small organization illustrated in Figure 6.1, there is a service design manager who is the process owner, process manager and process practitioner for overall design coordination (see section 6.3.5), as well as serving as the process owner for the service level management process. This role may be fulfilled by a manager of the functional unit in which most of the responsibility for the design of new or changed services resides. In some organizations that may be the application development function or the department responsible for the infrastructure.

For each specific service design project, the activities that relate to service design under processes such as availability management, capacity management and IT service continuity management may be led by the project manager with regular involvement from the service owner, but the design of reliable and repeatable processes is still under the authority of the individual process owners and process managers. Practitioners are likely to be drawn as needed for each service design effort from the technical management, application management and application development functions.

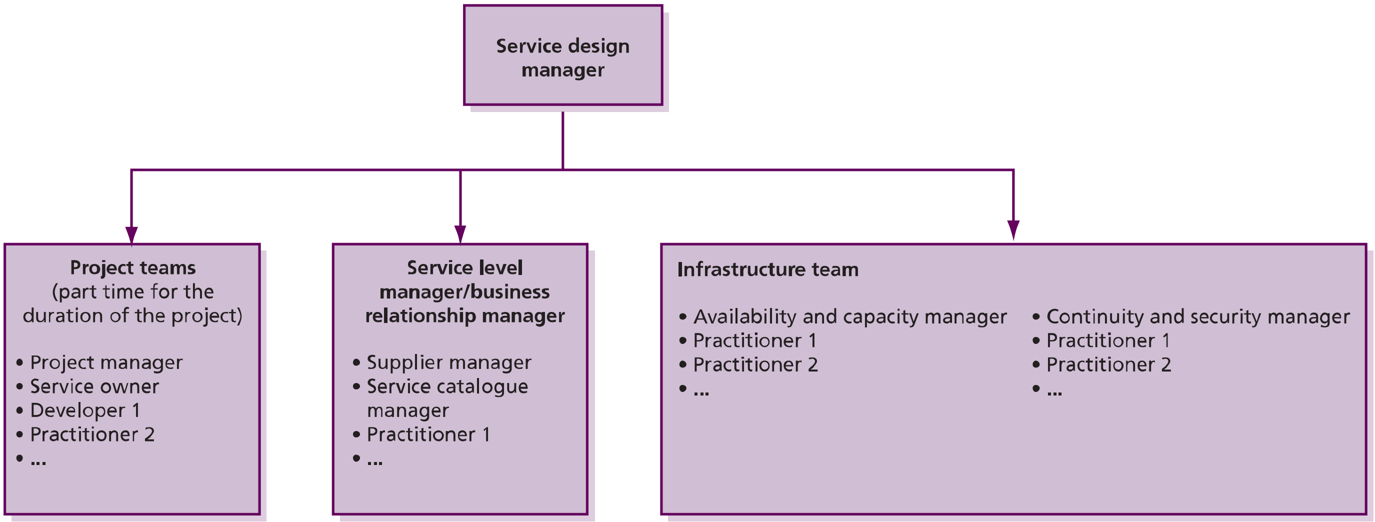


Figure 6.1 Example of a service design organization structure for a small organization

There may be a full-time service level manager who is the process manager for service level management and who also fills the role of process manager for the business relationship management process. In the small organization illustrated in Figure 6.1 most of the supplier management process is performed by the corporate procurement function, but the service level manager is the process owner from IT's perspective and a supplier manager ensures coordination and alignment with IT and business needs. A service catalogue manager reports to the service level manager who owns this process.

The roles of process owner and process manager in many instances are likely to be combined in a small organization, as well has having the same person fulfil several roles. For example, leadership of availability management and capacity management may be assigned to the same person, or leadership of IT service continuity management and information security management may be combined. Another possibility is the assignment of leadership of availability, capacity and IT service continuity management to the same person, although this is most likely in a very small organization. It is important, however, not to combine roles when there is a requirement for governance or compliance reasons to retain a separation. This may be to ensure checks and balances within a critical activity or process.

6.2.3.2 Larger organization

In the sample larger organization illustrated in Figure 6.2 there is a central headquarters (HQ) organization which includes all process owners, as well as a service design team to plan, coordinate and manage all service designs under the leadership of one or more service design managers or design coordination process managers. The HQ has a service management office (SMO) which oversees the adoption and deployment of service management methods, including guiding all process design and improvement. Obviously, this office will be active in all stages of the lifecycle, but is particularly critical to service design. There is also a project management office (PMO) which performs project/programme portfolio management and provides project management resources and capabilities. A global programmes group also resides in the HQ to lead programmes and projects of a global nature.

Each geographical region has its own process managers and practitioners for key processes such as service level management, change management, availability management and supplier management. (Note: Although the detailed discussion of change management resides in *ITIL Service Transition*, this process is very active in the service design stage, so it is mentioned in this example.)

Although clearly defined roles and responsibilities, as well as good communication, are critical to all organizations, these are particularly important in a large organization. Failure to clearly define the boundaries between what is done in the

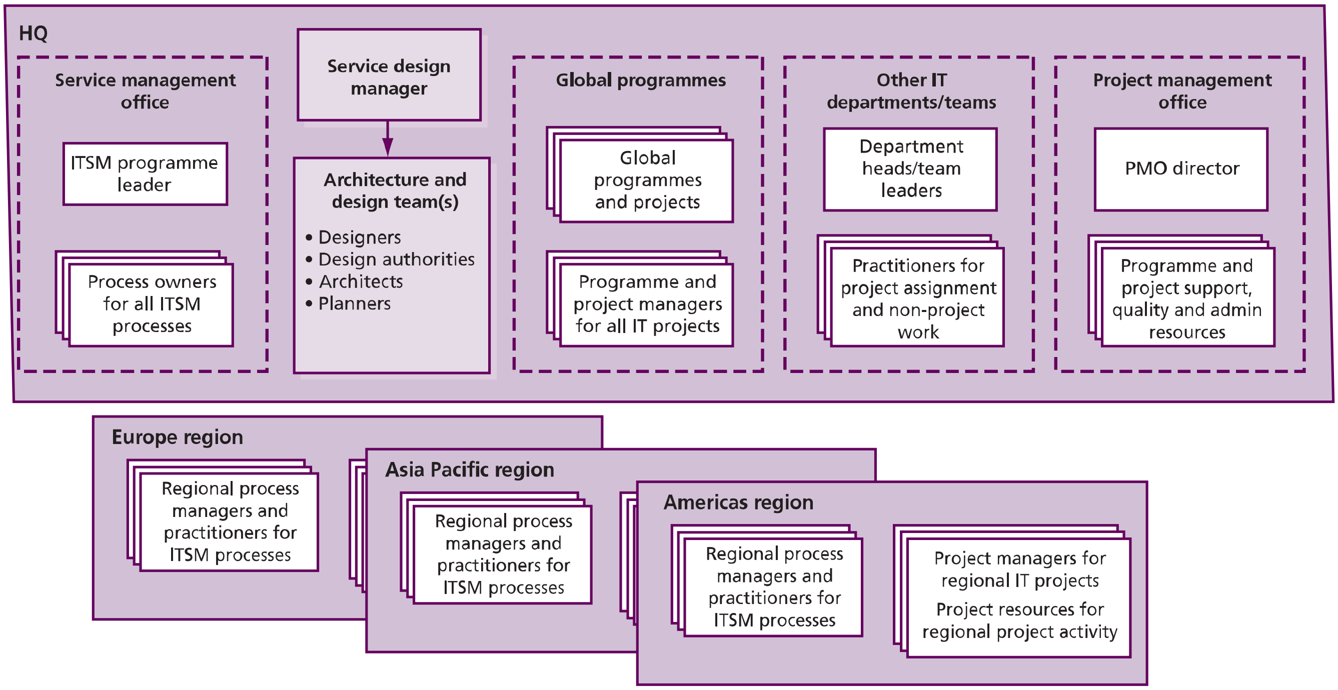

HQ
Service management office
ITSM programme leader
Process owners for all ITSM processes
Service design manager
Architecture and design team(s)
• Designers
• Design authorities
• Architects
• Planners
Global programmes
Global programmes and projects
Programme and project managers for all IT projects
Other IT departments/teams
Department heads/team leaders
Practitioners for project assignment and non-project work
Project management office
PMO director
Programme and project support, quality and admin resources
Europe region
Regional process managers and practitioners for ITSM processes
Asia Pacific region
Regional process managers and practitioners for ITSM processes
Americas region
Regional process managers and practitioners for ITSM processes
Project managers for regional IT projects
Project resources for regional project activity

Figure 6.2 Example of a service design organization structure for a large organization

various individual locations versus what is done at the corporate level can lead to unfilled gaps, duplicated efforts, delays, rework and unsatisfactory results. Authority must be clearly delineated and regular lines of communication established and maintained.

6.3 ROLES

A number of roles need to be performed in support of service design. Please note that this section provides guidelines and examples of role descriptions. These are not exhaustive or prescriptive, and in many cases roles will need to be combined or separated. Organizations should take care to apply this guidance in a way that suits their own structures and objectives.

A role is a set of responsibilities, activities and authorities granted to a person or team. A role is defined in a process or function. One person or team may have multiple roles – for example, the roles of configuration manager and change manager may be carried out by a single person.

Roles are often confused with job titles, but it is important to realize that they are not the same. Each organization will define appropriate job titles and job descriptions that suit its needs, and individuals holding these job titles can perform one or more of the required roles.

It should also be recognized that a person may, as part of their job assignment, perform a single task that represents participation in more than one process. For example, a technical analyst who submits a request for change (RFC) to add memory to a server to resolve a performance problem is participating in activities of the change management process at the same time as taking part in activities of the capacity management and problem management processes.

Roles fall into two main categories – generic roles such as process manager and process owner, and specific roles that are involved within a particular lifecycle stage or process such as a service design manager or IT designer/architect. Roles can be combined or divided in a number of different ways, depending on the organizational context. For example, in many organizations there will be one person who fulfils both the service catalogue process owner and service catalogue process manager roles. In a small organization the availability manager role may be combined with process manager roles from capacity management or IT service continuity management. In larger organizations there may be many different people carrying out each of these roles, split by geography, technology or other criteria. The exceptions to this are that there must be only one process owner for each process and one service owner for each service.

Roles are accountable or responsible for an activity. They may also be consulted or informed about something: for example a service owner may be consulted about a change during an impact assessment activity. The RACI model, described in section 6.4, provides a useful way of defining and communicating roles and responsibilities.

ITIL does not describe all the roles that could possibly exist in an organization, but provides representative examples to aid in an organization's definition of their own roles.

What is a service manager?

Service manager is a generic term for any manager within the service provider. The term is commonly used to refer to a business relationship manager, a process manager or a senior manager with responsibility for IT services overall. A service manager is often assigned several roles such as business relationship management, service level management and continual service improvement.

6.3.1 Generic service owner role

To ensure that a service is managed with a business focus, the definition of a single point of accountability is absolutely essential to provide the level of attention and focus required for its delivery.

The service owner is accountable for the delivery of a specific IT service. The service owner is responsible to the customer for the initiation, transition and ongoing maintenance and support of a particular service and accountable to the IT director or service management director for the delivery of the service. The service owner's accountability for a specific service within an organization is independent of where the underpinning technology components, processes or professional capabilities reside.

Service ownership is as critical to service management as establishing ownership for processes which cross multiple vertical silos or departments. It is possible that a single person may fulfil the service owner role for more than one service.

The service owner has the following responsibilities:

- Ensuring that the ongoing service delivery and support meet agreed customer requirements
- Working with business relationship management to understand and translate customer requirements into activities, measures or service components that will ensure that the service provider can meet those requirements
- Ensuring consistent and appropriate communication with customer(s) for service-related enquiries and issues
- Assisting in defining service models and in assessing the impact of new services or changes to existing services through the service portfolio management process
- Identifying opportunities for service improvements, discussing these with the customer and raising RFCs as appropriate
- Liaising with the appropriate process owners throughout the service lifecycle
- Soliciting required data, statistics and reports for analysis and to facilitate effective service monitoring and performance
- Providing input in service attributes such as performance, availability etc.
- Representing the service across the organization
- Understanding the service (components etc.)
- Serving as the point of escalation (notification) for major incidents relating to the service
- Representing the service in change advisory board (CAB) meetings
- Participating in internal service review meetings (within IT)
- Participating in external service review meetings (with the business)
- Ensuring that the service entry in the service catalogue is accurate and is maintained
- Participating in negotiating service level agreements (SLAs) and operational level agreements (OLAs) relating to the service
- Identifying improvement opportunities for inclusion in the continual service improvement (CSI) register
- Working with the CSI manager to review and prioritize improvements in the CSI register
- Making improvements to the service.

The service owner is responsible for continual improvement and the management of change affecting the service under their care. The service owner is a primary stakeholder in all of the underlying IT processes which enable or support the service they own. For example:

- **Incident management** Is involved in (or perhaps chairs) the crisis management team for high-priority incidents impacting the service owned
- **Problem management** Plays a major role in establishing the root cause and proposed permanent fix for the service being evaluated
- **Release and deployment management** Is a key stakeholder in determining whether a new release affecting a service in production is ready for promotion
- **Change management** Participates in CAB decisions, authorizing changes to the services they own
- **Service asset and configuration management** Ensures that all groups which maintain the data and relationships for the service architecture they are responsible for have done so with the level of integrity required
- **Service level management** Acts as the single point of contact for a specific service and ensures that the service portfolio and service catalogue are accurate in relation to their service
- **Availability management and capacity management** Reviews technical monitoring data from a domain perspective to ensure that the needs of the overall service are being met
- **IT service continuity management** Understands and is responsible for ensuring that all elements required to restore their service are known and in place in the event of a crisis
- **Information security management** Ensures that the service conforms to information security management policies
- **Financial management for IT services** Assists in defining and tracking the cost models in relation to how their service is costed and recovered.

6.3.2 Generic process owner role

The process owner role is accountable for ensuring that a process is fit for purpose. This role is often assigned to the same person who carries out the process manager role, but the two roles may be separate in larger organizations. The process owner role is accountable for ensuring that their process is performed according to the agreed and documented standard and meets the aims of the process definition.

The process owner's accountabilities include:

- Sponsoring, designing and change managing the process and its metrics
- Defining the process strategy
- Assisting with process design
- Ensuring that appropriate process documentation is available and current
- Defining appropriate policies and standards to be employed throughout the process
- Periodically auditing the process to ensure compliance to policy and standards
- Periodically reviewing the process strategy to ensure that it is still appropriate and change as required
- Communicating process information or changes as appropriate to ensure awareness
- Providing process resources to support activities required throughout the service lifecycle
- Ensuring that process technicians have the required knowledge and the required technical and business understanding to deliver the process, and understand their role in the process
- Reviewing opportunities for process enhancements and for improving the efficiency and effectiveness of the process
- Addressing issues with the running of the process
- Identifying improvement opportunities for inclusion in the CSI register
- Working with the CSI manager and process manager to review and prioritize improvements in the CSI register
- Making improvements to the process.

6.3.3 Generic process manager role

The process manager role is accountable for operational management of a process. There may be several process managers for one process, for example regional change managers or IT service continuity managers for each data centre. The process manager role is often assigned to the person who carries out the process owner role, but the two roles may be separate in larger organizations.

The process manager's accountabilities include:

- Working with the process owner to plan and coordinate all process activities
- Ensuring that all activities are carried out as required throughout the service lifecycle
- Appointing people to the required roles
- Managing resources assigned to the process
- Working with service owners and other process managers to ensure the smooth running of services
- Monitoring and reporting on process performance
- Identifying improvement opportunities for inclusion in the CSI register
- Working with the CSI manager and process owner to review and prioritize improvements in the CSI register
- Making improvements to the process implementation.

6.3.4 Generic process practitioner role

A process practitioner is responsible for carrying out one or more process activities.

In some organizations, and for some processes, the process practitioner role may be combined with the process manager role; in others there may be large numbers of practitioners carrying out different parts of the process.

The process practitioner's responsibilities typically include:

- Carrying out one or more activities of a process
- Understanding how their role contributes to the overall delivery of service and creation of value for the business
- Working with other stakeholders, such as their manager, co-workers, users and customers, to ensure that their contributions are effective
- Ensuring that inputs, outputs and interfaces for their activities are correct
- Creating or updating records to show that activities have been carried out correctly.

6.3.5 Design coordination roles

This section describes roles that need to be performed in support of the design coordination process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.5.1 Design coordination process owner

The design coordination process owner's responsibilities typically include:

- Carrying out the generic process owner role for the design coordination process (see section 6.3.2 for more detail)
- Setting the scope and policies for service design
- Overseeing the overall design of all service design processes to ensure that they will work together to meet the needs of the business.

6.3.5.2 Design coordination process manager

The design coordination process manager's responsibilities typically include:

- Carrying out the generic process manager role for the design coordination process (see section 6.3.3 for more detail)
- Coordinating interfaces between design coordination and other processes
- Ensuring that overall service strategies are reflected in the service design practice
- Ensuring the consistent design of appropriate services, service management information systems, architectures, technology, processes, information and metrics to meet current and evolving business outcomes and requirements
- Coordinating all design activities across projects, changes, suppliers and support teams, and managing schedules, resources and conflicts where required
- Planning and coordinating the resources and capabilities required to design new or changed services
- Producing service design packages (SDPs) based on service charters and change requests
- Ensuring that appropriate service designs and/ or SDPs are produced and that they are handed over to service transition as agreed
- Managing the quality criteria, requirements and handover points between the service design stage and service strategy and service transition
- Ensuring that all service models and service solution designs conform to strategic, architectural, governance and other corporate requirements
- Improving the effectiveness and efficiency of service design activities and processes

- Ensuring that all parties adopt a common framework of standard, reusable design practices in the form of activities, processes and supporting systems, whenever appropriate.

The service design manager

Many organizations will have a person with the job title 'service design manager'. This job typically combines the roles of design coordination process owner and design coordination process manager. It may also include some degree of line management of the people involved in service design.

6.3.6 Service catalogue management roles

This section describes a number of roles that need to be performed in support of the service catalogue management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.6.1 Service catalogue management process owner

The service catalogue management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the service catalogue management process (see section 6.3.2 for more detail)
- Working with other process owners to ensure there is an integrated approach to the design and implementation of service catalogue management, service portfolio management, service level management and business relationship management.

6.3.6.2 Service catalogue management process manager

The service catalogue management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the service catalogue management process (see section 6.3.3 for more detail)
- Coordinating interfaces between service catalogue management and other processes, especially service asset and configuration management, and release and deployment management
- Ensuring that all operational services and all services being prepared for operational running are recorded within the service catalogue
- Ensuring that all the information within the service catalogue is accurate and up to date
- Ensuring that appropriate views of the service catalogue are maintained and made available to those for whom they are targeted
- Ensuring that all the information within the service catalogue is consistent with the information within the service portfolio
- Ensuring that the information within the service catalogue is adequately protected and backed up.

6.3.7 Service level management roles

This section describes a number of roles that need to be performed in support of the service level management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.7.1 Service level management process owner

The service level management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the service level management process (see section 6.3.2 for more detail)
- Liaising with the business relationship management process owner to ensure proper coordination and communication between the two processes
- Working with other process owners to ensure there is an integrated approach to the design and implementation of service catalogue management, service portfolio management, service level management and business relationship management.

6.3.7.2 Service level management process manager

The service level management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the service level management process (see section 6.3.3 for more detail)

- Coordinating interfaces between service level management and other processes, especially service catalogue management, service portfolio management, business relationship management and supplier management
- Keeping aware of changing business needs
- Ensuring that the current and future service level requirements (service warranty) of customers are identified, understood and documented in SLA and service level requirements (SLR) documents
- Negotiating and agreeing levels of service to be delivered with the customer (either internal or external); formally documenting these levels of service in SLAs
- Negotiating and agreeing OLAs and, in some cases, other SLAs and agreements that underpin the SLAs with the customers of the service
- Assisting with the production and maintenance of an accurate service portfolio, service catalogue, application portfolio and the corresponding maintenance procedures
- Ensuring that targets agreed within underpinning contracts are aligned with SLA and SLR targets
- Ensuring that service reports are produced for each customer service and that breaches of SLA targets are highlighted, investigated and actions taken to prevent their recurrence
- Ensuring that service performance reviews are scheduled, carried out with customers regularly and documented, with agreed actions progressed
- Ensuring that improvement initiatives identified in service reviews are acted on and progress reports are provided to customers
- Reviewing service scope, SLAs, OLAs and other agreements on a regular basis, ideally at least annually
- Ensuring that all changes are assessed for their impact on service levels, including SLAs, OLAs and underpinning contracts, including attendance at change advisory board (CAB) meetings if appropriate
- Identifying all customers and other key stakeholders to involve in SLR, SLA and OLA negotiations
- Developing relationships and communication with customers, key users and other stakeholders
- Defining and agreeing complaints and their recording, management, escalation (where necessary) and resolution
- Definition recording and communication of all complaints
- Measuring, recording, analysing and improving customer satisfaction.

These next two roles, while not strictly speaking service level management roles, typically play a large part in the successful execution of the process.

6.3.7.3 Service owner role in service level management

Persons assigned to the role of service owner participate in the service level management process by:

- Ensuring that the ongoing service delivery and support meet agreed customer requirements
- Ensuring consistent and appropriate communication with customer(s) for service-related enquiries and issues
- Providing input in service attributes such as performance and availability
- Participating in external service review meetings (with the business)
- Soliciting required data, statistics and reports for analysis and to facilitate effective service monitoring and performance
- Participating in negotiating SLAs and OLAs relating to the service.

6.3.7.4 Business relationship manager role in service level management

Persons assigned to the role of business relationship manager participate in the service level management process by:

- Ensuring high levels of customer satisfaction
- Establishing and maintaining a constructive relationship between the service provider and the customer at a strategic level
- Confirming customer high-level requirements
- Facilitating service level agreement negotiations by ensuring that the correct customer representatives participate
- Identifying opportunities for improvement.

6.3.8 Availability management roles

This section describes a number of roles that need to be performed in support of the availability management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.8.1 Availability management process owner

The availability management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the availability management process (see section 6.3.2 for more detail)
- Working with managers of all functions to ensure acceptance of the availability management process as the single point of coordination for all availability-related issues, regardless of the specific technology involved
- Working with other process owners to ensure there is an integrated approach to the design and implementation of availability management, service level management, capacity management, IT service continuity management and information security management.

6.3.8.2 Availability management process manager

The availability management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the availability management process (see section 6.3.3 for more detail)
- Coordinating interfaces between availability management and other processes, especially service level management, capacity management, IT service continuity management and information security management
- Ensuring that all existing services deliver the levels of availability agreed with the business in SLAs
- Ensuring that all new services are designed to deliver the levels of availability required by the business, and validation of the final design to meet the minimum levels of availability as agreed by the business for IT services
- Assisting with the investigation and diagnosis of all incidents and problems that cause availability issues or unavailability of services or components
- Participating in the IT infrastructure design, including specifying the availability requirements for hardware and software
- Specifying the requirements for new or enhanced event management systems for automatic monitoring of availability of IT components
- Specifying the reliability, maintainability and serviceability requirements for components supplied by internal and external suppliers
- Being responsible for monitoring actual IT availability achieved against SLA targets, and providing a range of IT availability reporting to ensure that agreed levels of availability, reliability and maintainability are measured and monitored on an ongoing basis
- Proactively improving service availability wherever possible, and optimizing the availability of the IT infrastructure to deliver cost-effective improvements that deliver tangible benefits to the business
- Creating, maintaining and regularly reviewing an availability management information system and 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
- Ensuring that the availability management process, its associated techniques and methods are regularly reviewed and audited, and that all of these are subject to continual improvement and remain fit for purpose
- Creating availability and recovery design criteria to be applied to new or enhanced infrastructure design
- Working with financial management for IT services, ensuring the levels of IT availability required are cost-justified
- Maintaining and completing an availability testing schedule for all availability mechanisms
- Ensuring that all availability tests and plans are tested after every major business change
- Assisting security and IT service continuity management with the assessment and management of risk

- Assessing changes for their impact on all aspects of availability, including overall service availability and the availability plan
- Attending CAB meetings when appropriate.

6.3.9 Capacity management roles

This section describes a number of roles that need to be performed in support of the capacity management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.9.1 Capacity management process owner

The capacity management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the capacity management process (see section 6.3.2 for more detail)
- Working with managers of all functions to ensure acceptance of the capacity management process as the single point of coordination for all capacity and performance-related issues, regardless of the specific technology involved
- Working with other process owners to ensure there is an integrated approach to the design and implementation of capacity management, availability management, IT service continuity management and information security management.

6.3.9.2 Capacity management process manager

The capacity management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the capacity management process (see section 6.3.3 for more detail)
- Coordinating interfaces between capacity management and other processes, especially service level management, availability management, IT service continuity management and information security management
- Ensuring that there is adequate IT capacity to meet required levels of service, and that senior IT management is correctly advised on how to match capacity and demand and to ensure that use of existing capacity is optimized
- Identifying, with the service level manager, capacity requirements through discussions with the business users
- Understanding the current usage of the infrastructure and IT services, and the maximum capacity of each component
- Performing sizing on all proposed new services and systems, possibly using modelling techniques, to ascertain capacity requirements
- Forecasting future capacity requirements based on business plans, usage trends, sizing of new services etc.
- Production, regular review and revision of the capacity plan, in line with the organization's business planning cycle, identifying current usage and forecast requirements during the period covered by the plan
- Ensuring that appropriate levels of monitoring of resources and system performance are set
- Analysis of usage and performance data, and reporting on performance against targets contained in SLAs
- Raising incidents and problems when breaches of capacity or performance thresholds are detected, and assisting with the investigation and diagnosis of capacity-related incidents and problems
- Identifying and initiating any tuning to be carried out to optimize and improve capacity or performance
- Identifying and implementing initiatives to improve resource usage – for example, demand management techniques
- Assessing new technology and its relevance to the organization in terms of performance and cost
- Being familiar with potential future demand for IT services and assessing this on performance service levels
- Ensuring that all changes are assessed for their impact on capacity and performance and attending CAB meetings when appropriate
- Producing regular management reports that include current usage of resources, trends and forecasts
- Sizing all proposed new services and systems to determine the computer and network resources required, to determine hardware utilization, performance service levels and cost implications

- Assessing new techniques and hardware and software products for use by capacity management that might improve the efficiency and effectiveness of the process
- Testing performance of new services and systems
- Preparing reports on service and component performance against targets contained in SLAs
- Maintaining a knowledge of future demand for IT services and predicting the effects of demand on performance service levels
- Determining performance service levels that are maintainable and cost-justified
- Recommending tuning of services and systems, and making recommendations to IT management on the design and use of systems to help ensure optimum use of all hardware and operating system software resources
- Acting as a focal point for all capacity and performance issues.

6.3.10 IT service continuity management roles

This section describes a number of roles that need to be performed in support of the IT service continuity management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on their needs.

6.3.10.1 IT service continuity management process owner

The IT service continuity management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the IT service continuity management process (see section 6.3.2 for more detail)
- Working with the business to ensure proper coordination and communication between business continuity management and IT service continuity management
- Working with managers of all functions to ensure acceptance of the IT service continuity management process as the single point of coordination for all IT service continuity-related issues, regardless of the specific technology involved
- Working with other process owners to ensure there is an integrated approach to the design and implementation of IT service continuity management, information security management, availability management and business continuity management.

6.3.10.2 IT service continuity management process manager

The IT service continuity management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the IT service continuity management process (see section 6.3.3 for more detail)
- Coordinating interfaces between IT service continuity management and other processes, especially service level management, information security management, availability management, capacity management and business continuity management
- Performing business impact analyses for all existing and new services
- Implementing and maintaining the IT service continuity management process, in accordance with the overall requirements of the organization's business continuity management process, and representing the IT services function within the business continuity management process
- Ensuring that all IT service continuity management plans, risks and activities underpin and align with all business continuity management plans, risks and activities, and are capable of meeting the agreed and documented targets under any circumstances
- Performing risk assessment and risk management to prevent disasters where cost-justifiable and where practical
- Developing and maintaining the organization's continuity strategy
- Assessing potential service continuity issues and invoking the service continuity plan if necessary
- Managing the service continuity plan while it is in operation, including fail-over to a secondary location and restoration to the primary location
- Performing post-mortem reviews of service continuity tests and invocations, and instigating corrective actions where required
- Developing and managing the IT service continuity management plans to ensure that, at all times, the recovery objectives of the business can be achieved

- Ensuring that all IT service areas are prepared and able to respond to an invocation of the continuity plans
- Maintaining a comprehensive IT testing schedule, including testing all continuity plans in line with business requirements and after every major business change
- Undertaking quality reviews of all procedures and ensuring that these are incorporated into the testing schedule
- Communicating and maintaining awareness of IT service continuity management objectives within the business areas supported and IT service areas
- Undertaking regular reviews, at least annually, of the continuity plans with the business areas to ensure that they accurately reflect the business needs
- Negotiating and managing contracts with providers of third-party recovery services
- Assessing changes for their impact on service continuity and continuity plans
- Attending CAB meetings when appropriate.

6.3.11 Information security management roles

This section describes a number of roles that need to be performed in support of the information security management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.11.1 Information security management process owner

The information security management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the information security management process (see section 6.3.2 for more detail)
- Working with the business to ensure proper coordination and communication between organizational (business) security management and information security management
- Working with managers of all functions to ensure acceptance of the information security management process as the single point of coordination for all information security-related issues, regardless of the specific technology involved
- Working with other process owners to ensure there is an integrated approach to the design and implementation of information security management, availability management, IT service continuity management and organizational security management.

6.3.11.2 Information security management process manager

The information security management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the information security management process (see section 6.3.3 for more detail)
- Coordinating interfaces between information security management and other processes, especially service level management, availability management, IT service continuity management and organizational security management
- Developing and maintaining the information security policy and a supporting set of specific policies, ensuring appropriate authorization, commitment and endorsement from senior IT and business management
- Communicating and publicizing the information security policy to all appropriate parties
- Ensuring that the information security policy is enforced and adhered to
- Identifying and classifying IT and information assets (configuration items) and the level of control and protection required
- Assisting with business impact analyses
- Performing security risk assessment and risk management in conjunction with availability and IT service continuity management
- Designing security controls and developing security plans
- Developing and documenting procedures for operating and maintaining security controls
- Monitoring and managing all security breaches and handling security incidents, taking remedial action to prevent recurrence wherever possible
- Reporting, analysing and reducing the impact and volumes of all security incidents in conjunction with problem management
- Promoting education and awareness of security
- Maintaining a set of security controls and documentation, and regularly reviewing and auditing all security controls and procedures

- Ensuring all changes are assessed for impact on all security aspects, including the information security policy and security controls, and attending CAB meetings when appropriate
- Ensuring security tests are performed as required
- Participating in any security reviews arising from security breaches and instigating remedial actions
- Ensuring that the confidentiality, integrity and availability of the services are maintained at the levels agreed in the SLAs and that they conform to all relevant statutory requirements
- Ensuring that all access to services by external partners and suppliers is subject to contractual agreements and responsibilities
- Acting as a focal point for all security issues.

6.3.12 Supplier management roles

This section describes a number of roles that need to be performed in support of the supplier management process. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs.

6.3.12.1 Supplier management process owner

The supplier management process owner's responsibilities typically include:

- Carrying out the generic process owner role for the supplier management process (see section 6.3.2 for more detail)
- Working with the business to ensure proper coordination and communication between corporate vendor management and/or procurement and supplier management
- Working with other process owners to ensure there is an integrated approach to the design and implementation of supplier management, service level management and corporate vendor management and/or procurement processes.

6.3.12.2 Supplier management process manager

The supplier management process manager's responsibilities typically include:

- Carrying out the generic process manager role for the supplier management process (see section 6.3.3 for more detail)
- Coordinating interfaces between supplier management and other processes, especially service level management and corporate vendor management and/or procurement processes
- Providing assistance in the development and review of SLAs, contracts, agreements or any other documents for third-party suppliers
- Ensuring that value for money is obtained from all IT suppliers and contracts
- Ensuring that all IT supplier processes are consistent and interface with all corporate supplier strategies, processes and standard terms and conditions
- Maintaining and reviewing a supplier and contract management information system
- Reviewing and making risk assessments of all suppliers and contracts on a regular basis
- Ensuring that any underpinning contracts, agreements or SLAs developed are aligned with those of the business
- Ensuring that all supporting services are scoped and documented and that interfaces and dependencies between suppliers, supporting services and supplier processes are agreed and documented
- Ensuring that all roles and relationships between lead and any sub-contracted suppliers are documented, maintained and subject to contractual agreement
- Reviewing lead suppliers' processes to ensure that any sub-contracted suppliers are meeting their contractual obligations
- Performing contract or SLA reviews at least annually, and ensuring that all contracts are consistent with organizational requirements and standard terms and conditions wherever possible
- Updating contracts or SLAs when required, ensuring that the change management process is followed
- Maintaining a process for dealing with contractual disputes, and ensuring that any disputes are dealt with in an efficient and effective manner
- Maintaining a process for dealing with the expected end, early end or transfer of a service
- Monitoring, reporting and regularly reviewing supplier performance against targets, identifying improvement actions as appropriate and ensuring these actions are implemented

- Ensuring changes are assessed for their impact on suppliers, supporting services and contracts and attending CAB meetings when appropriate
- Coordinating and supporting all individual IT supplier and contract managers, ensuring that each supplier/contract has a nominated owner within the service provider organization.

6.3.13 Other service design roles

This section describes a number of roles that may exist in an organization to support the service design stage of the service lifecycle. Some of these roles may also include responsibilities that associate with other service lifecycle stages. These roles are not job titles, and each organization will have to define appropriate job titles and job descriptions depending on its needs. Responsibilities described may also be reorganized into other roles based on the organization's needs and objectives.

6.3.13.1 IT planner

An IT planner is responsible for the production and coordination of IT plans. The main objectives of the role are as follows:

- Developing IT plans that meet and continue to meet the IT requirements of the business
- Coordinating, measuring and reviewing the implementation progress of all IT strategies and plans
- Producing and maintaining the overall set of IT standards, policies, plans and strategies, encompassing all aspects of IT required to support an organization's business strategy. IT planning includes participation in the creation of SLAs and the planning of all aspects of infrastructure – internal and external, public or private, internet and intranet – necessary to ensure that the provision of IT services satisfies business
- Assuming responsibility for all aspects of IT standards, policy and strategy implementation for IT as a whole and for significant projects or major new strategic applications
- Recommending policy for the effective use of IT throughout the organization and working with IT designers to ensure that overall plans and strategies are developed in conjunction with IT design for all areas of IT
- Reviewing IT costs against budgets and new developments, initiating proposals to change IT plans and strategies where appropriate, in conjunction with financial management for IT services
- Assuming full responsibility for the management, planning and coordination of IT systems and services, including investigation, analysis, specification, design, development, testing, maintenance, upgrade, transition and operation. It is essential that while performing these activities, the business, IT management and all the service management processes are kept up to date with the progress of projects
- Obtaining and evaluating proposals from suppliers of equipment, software, transmission services and other services, ensuring that all business and IT requirements are satisfied
- Identifying internal and external influencing factors, forecasting future needs and setting plans for the effective use of IT within the organization
- Sponsoring and monitoring research, development and long-term planning for the provision and use of IT architectures, products and services
- Reviewing IT performance with all other areas and initiating any improvements in organization to ensure that service levels and targets continue to be met in all areas
- Taking ultimate responsibility for prioritizing and scheduling the implementation of new or changed services within IT
- Working with senior management and other senior specialists and planners in formulating plans and making procurement decisions applicable to all areas of IT
- Recognizing the key business drivers and those areas of business need that are not adequately supported by current and planned IT services, developing the plans and IT response to the business requirements
- Identifying suitable applications, services and products, together with their environments, to meet business needs within the required planning timeframe
- Developing the initial plans for the implementation of authorized new IT services, applications and infrastructure support,

identifying budgetary, technical and staffing constraints, and clearly listing costs and expected benefits

- Monitoring the existing IT plans in relation to business needs and IT strategy to determine opportunities for improving business processes through the use of new technology, and to identify unforeseen risks to the achievement of forecast business benefits
- Investigating major options for providing IT services effectively and efficiently and recommending new innovative solutions, based on new approaches to processes, provision, recruitment and retention, and global supply contracts
- Producing feasibility studies, business models, IT models, business cases, statements of requirements (SoRs) and invitations to tender (ITTs) for recommended new IT systems, identifying the business impact, the probability of satisfying business needs, the anticipated business benefits and the risks and consequences of failure
- Overseeing and coordinating the programme of planned IT project implementations and changes, taking appropriate action to identify and overcome problems and resolve conflict
- Conducting post-implementation reviews in conjunction with change management of those information systems introduced in pursuit of the plans, to assess the extent to which expected business benefits were realized
- Liaising with strategy, transition and operations teams and processes to plan for their immediate and future needs
- Providing authoritative advice and guidance on relevant national and international standards, regulations, protocols and tariffs
- Documenting all work using required standards, methods and tools
- Ensuring that all IT planning processes, roles, responsibilities and documentation are regularly reviewed and audited for efficiency, effectiveness and compliance
- Maintaining a good overall knowledge of all IT product capabilities and the technical frameworks in which they operate
- Where required, assessing changes for their conformance to the design strategies, including attendance at CAB meetings if appropriate.

6.3.13.2 IT designer/architect

An IT designer/architect is responsible for the overall coordination and design of the required technology. Often designers and architects within large organizations specialize in one of the five aspects of design (see Chapter 3). However, an integrated approach to design should always be adopted; therefore designers and architects need to work together within a formal method and framework to ensure consistent and compatible designs are produced. In smaller organizations, some or all of the roles are usually combined, and this is less of an issue, although a formal approach should still be used. Whenever designs are produced, they should always adopt an integrated approach, covering all areas, and should be accepted and signed off by all areas. All designers need to understand how architectures, strategies, designs and plans fit together and understand all the main aspects of design.

The designer/architect should produce a detailed process map that documents all the processes and their high-level interfaces. This ensures that the overall structure is not unnecessarily complex, that the process's central interfaces are part of the design, and provides an overview to everyone on how the customer and all other stakeholders interact with the processes.

To perform the role of designer or architect, it is necessary for staff to have good knowledge and practical experience of design philosophies and planning, including programme, project and service management, methods and principles. The main objectives of the IT designer/architect are as follows:

- Producing and reviewing the designs of all new or changed services, SLAs, OLAs and contracts
- Producing a process map of all of the processes and their high-level interfaces, to ensure integration, consistency and continuity across all processes
- Designing secure and resilient technology architectures that meet all the current and anticipated future IT requirements of the organization
- Ensuring that the design of all processes, roles, responsibilities and documentation is regularly reviewed and audited for efficiency, effectiveness and compliance

- Designing an appropriate and suitable service portfolio, supporting all activities within the complete service lifecycle
- Designing measurement methods and metrics to support the continual improvement of service provision and all supporting processes
- Producing and keep up to date all IT design, architectural, policy and specification documentation
- Producing and maintaining all aspects of IT specification, including the overall designs, architectures, topologies and configurations of the infrastructure, environment, applications and data, and the design documentation of all IT systems. This should include not just the technology, but also the management systems, processes, information flows and external services
- Recommending proactive, innovative IT solutions for the improvement of IT design and operation whenever and wherever possible
- Translating logical designs into physical designs, taking account of business requirements, target environments, processes, performance requirements, existing systems and services, and any potential safety-related aspects
- Creating and maintaining IT design policies, philosophies and criteria, covering all areas including connectivity, capacity, interfaces, security, resilience, recovery, access and remote access, and ensuring that all new services meet their service levels and targets
- Working with capacity management and reviewing IT traffic volumes and requirements, identifying trends in traffic flows and levels of service
- Proposing design enhancements to IT infrastructure, capacity changes, continuity, backup and recovery arrangements, as required, and being aware of operational requirements, especially in terms of service levels, availability, response times, security and repair times. All these activities are performed in liaison with all of the service management processes
- Reviewing IT costs against external service providers, new developments and new services, initiating proposals to change IT design where appropriate cost reductions and benefits can be achieved, in consultation with financial management for IT services
- Providing advice and guidance to management on the design and planning phases of IT systems, to ensure that requirements (particularly capacity, recovery, performance and security needs) are reflected in the overall specifications
- Providing advice and guidance to all areas of IT and business management, analysts, planners, designers and developers on all aspects of IT design and technology
- Interfacing with designers and planners from external suppliers and service providers, ensuring all external IT services are designed to meet their agreed service levels and targets
- Playing a major role in the selection of any new IT infrastructure or technology solutions
- Assuming technical responsibility for IT standards, policy and design for all significant projects or major application areas, assisting with the impact assessment and evaluation of major new IT design options
- Providing technical advice and guidance on relevant national and international standards, regulations, protocols and tariffs
- Taking full responsibility for the design aspects of all stages of the lifecycle of IT systems, including investigation, analysis, specification, design, development, construction, testing, maintenance, upgrade, transition, operation and improvement
- Working with IT colleagues where appropriate, producing or updating IT and corporate design documentation and models
- Updating or providing input to cost benefit analyses, risk assessments, business cases, SoRs and ITTs and development plans, to take account of design decisions
- Obtaining and assisting with the evaluation and selection of proposals and solutions from suppliers of equipment, software and other IT service and product providers
- Constructing, interpreting and monitoring test plans to verify correct operation of completed systems against their design objectives
- Documenting all work using required standards, methods and tools
- Maintaining a good technical knowledge of all IT product capabilities and the technical frameworks in which they operate
- Where required, assessing changes for their conformance to the design principles, including attendance at CAB meetings if appropriate.

6.4 RESPONSIBILITY MODEL – RACI

Clear definitions of accountability and responsibility are essential for effective service management. To help with this task the RACI model or 'authority matrix' is often used within organizations to define the roles and responsibilities in relation to processes and activities. The RACI matrix provides a compact, concise, easy method of tracking who does what in each process and it enables decisions to be made with pace and confidence.

The RACI model is described in more detail in section 3.7.4.

6.5 COMPETENCE AND TRAINING

6.5.1 Competence and skills for service management

Delivering service successfully depends on personnel involved in service management having the appropriate education, training, skills and experience. People need to understand their role and how they contribute to the overall organization, services and processes to be effective and motivated. As changes are made, job requirements, roles, responsibilities and competencies should be updated if necessary.

Each service lifecycle stage depends on appropriate skills and experience of people and their knowledge to make key decisions. In many organizations, personnel will deliver tasks appropriate to more than one lifecycle stage. They may well find themselves allocated (fully or partially) from operational tasks to support a design exercise and then follow that service through service transition. They may then, via early life support activities, move into support of the new or changed services that they have been involved in designing and implementing into the live environment.

The specific roles within ITIL service management all require specific skills, attributes and competences from the people involved to enable them to work effectively and efficiently. However, whatever the role, it is imperative that the person carrying out that role has the following attributes:

- Awareness of the business priorities, objectives and business drivers
- Awareness of the role IT plays in enabling the business objectives to be met
- Customer service skills
- Awareness of what IT can deliver to the business, including latest capabilities
- The competence, knowledge and information necessary to complete their role
- The ability to use, understand and interpret the best practice, policies and procedures to ensure adherence.

The following are examples of attributes required in many of the roles, dependent on the organization and the specific roles assigned:

- Management skills – both from a person management perspective and from the overall control of process
- Ability to handle meetings – organizing, chairing, and documenting meetings and ensuring that actions are followed up
- Communication skills – an important element of all roles is raising awareness of the processes in place to ensure buy-in and conformance. An ability to communicate at all levels within the organization will be imperative
- Articulateness – both written (e.g. for reports) and verbal
- Negotiation skills are required for several aspects, such as procurement and contracts
- An analytical mind – to analyse metrics produced from the activity.

Many people working in service management are involved with continual service improvement. *ITIL Continual Service Improvement* provides specific guidance on the skill levels needed for CSI activities.

6.5.2 Competence and skills framework

Standardizing job titles, functions, roles and responsibilities can simplify service management and human resource management. Many service providers use a common framework of reference for competence and skills to support activities such as skill audits, planning future skill requirements, organizational development programmes and resource allocation. For example, resource and cost models are simpler and easier to use if jobs and roles are standard.

The Skills Framework for the Information Age (SFIA) is an example of a common reference model for the identification of the skills needed to develop effective IT services, information systems and technology. SFIA defines seven generic levels at which tasks can be performed, with the associated professional skills required for each level. A second dimension defines core competencies that can be combined with the professional skills. SFIA is used by many IT service providers to identify career development opportunities.

More information on SFIA can be found at www.sfia.org.uk

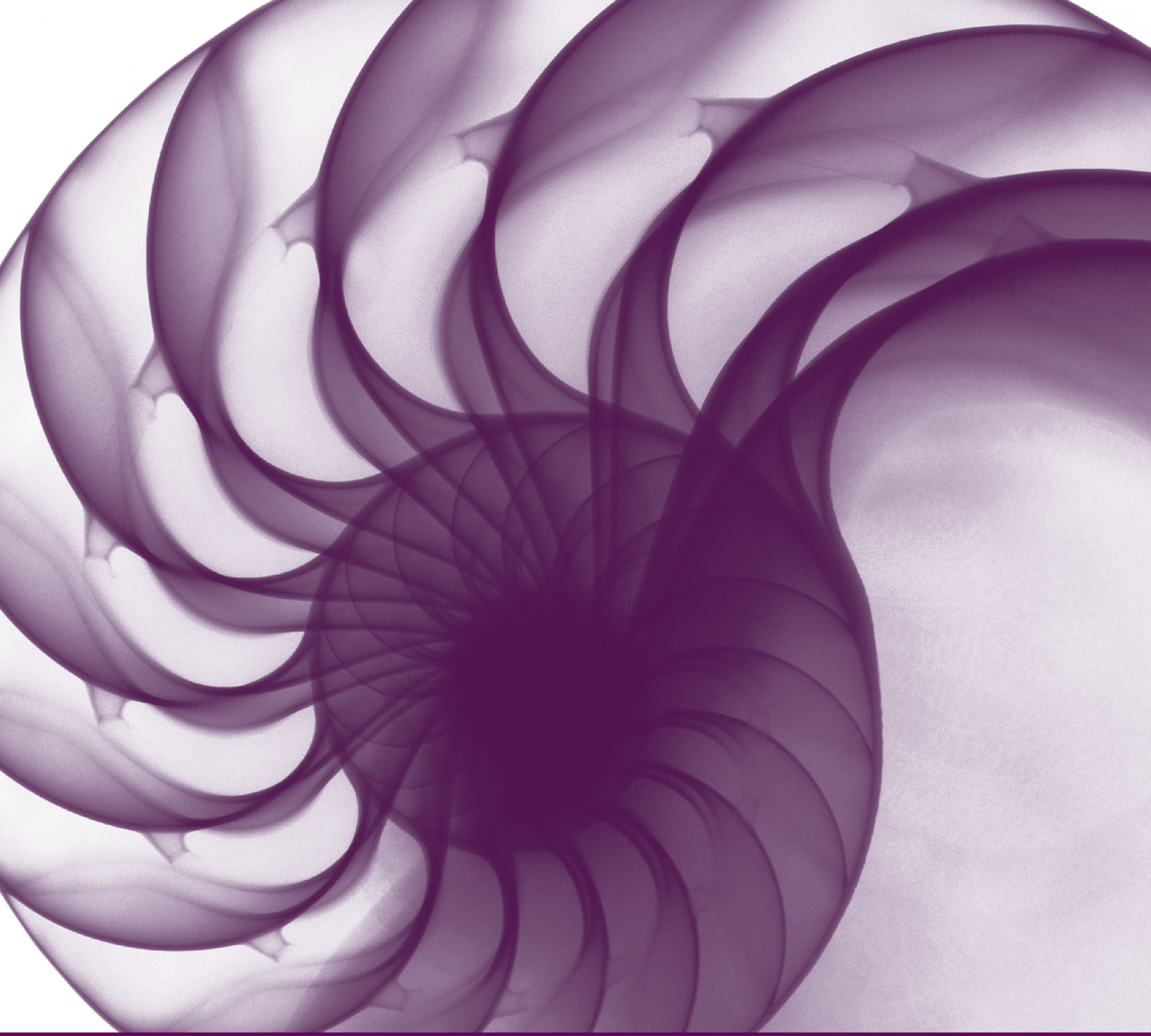
6.5.3 Training

Training in service management helps service providers to build and maintain their service management capability. Training needs must be matched to the requirements for competence and professional development.

The official ITIL qualification scheme enables organizations to develop the competence of their personnel through approved training courses. The courses help students to gain knowledge of ITIL best practices, develop their competencies and gain a recognized qualification. The scheme has four levels:

- Foundation level
- Intermediate level
- ITIL Expert
- ITIL Master.

More information on ITIL qualifications can be found at www.itil-officialsite.com



7 Technology considerations

7 Technology considerations

It is generally recognized that the use of service management tools is essential for the success of all but the very smallest process implementations. However, it is important that the tool being used supports the processes – not the other way around. As a general rule, do not modify processes to fit the tool. However, while striving to adhere to this principle, organizations need to be pragmatic and recognize that there may not be a tool that supports the designed process exactly – some degree of process re-design may be necessary.

Organizations should also not limit their tool requirements to functionality: consider the product's ability to perform, enlarge the size of the databases, recover from failure and maintain data integrity. Does the product conform to international standards? Is it efficient enough to enable you to meet your service management requirements?

Often organizations believe that by purchasing or developing a tool all their problems will be solved, and it is easy to forget that we are still dependent on the process, the function and, most importantly, the people. Remember:

'A fool with a tool is still a fool.'

7.1 SERVICE DESIGN TOOLS

There are many tools and techniques that can be used to assist with the design of services and their associated components. These tools and techniques enable:

- Hardware design
- Software design
- Environmental design
- Process design
- Data design.

The tools and techniques are many and varied, including both proprietary and non-proprietary, and are useful in:

- Speeding up the design process
- Ensuring that standards and conventions are followed
- Offering prototyping, modelling and simulation facilities
- Enabling 'What if?' scenarios to be examined
- Enabling interfaces and dependencies to be checked and correlated
- Validating designs before they are developed and implemented to ensure that they satisfy and fulfil their intended requirements.

Developing service designs can be simplified by the use of tools that provide graphical views of the service and its constituent components, from the business processes, through the service and service level agreement (SLA) to the infrastructure, environment, data and applications, processes, operational level agreements (OLAs), teams, contracts and suppliers. Some service asset and configuration management tools provide such facilities, and are sometimes part of an integrated ITSM tool. They can contain or be linked to 'auto-discovery' tools and mechanisms and allow the relationships between all of these elements to be graphically represented, providing the ability to drill down within each component and obtain detailed information if needed.

If these types of tool also contain financial information, and are then linked to a 'metrics tree' providing key performance indicators (KPIs) and metrics of the various aspects of the service, then the service can be monitored and managed through all stages of its lifecycle. Sharing this single, centralized set of service information allows everyone in the service provider organization and the business to access a single, consistent, 'real-world' view of the service and its performance, and provides a solid base for the development of good relationships and partnerships between the service provider and its customers.

These types of tool not only facilitate the design processes, but also greatly support and assist all ITSM methods and lifecycle stages, including:

- Management of all stages of the service lifecycle
- All aspects of the service and its performance
- Service achievement, SLA, OLA, contractual and supplier measurement, reporting and management

- Consolidated metrics and metrics trees, with views from management dashboards down to detailed component information, performance and fault analysis and identification
- Consistent and consolidated views across all processes, systems, technologies and groups
- Relationships and integration of the business and its processes with IT services, systems and processes
- A comprehensive set of search and reporting facilities, enabling accurate information and analysis for informed decision-making
- Management of service costs
- Management of relationships, interfaces and inter-dependencies
- Management of the service portfolio and service catalogue
- A configuration management system (CMS)
- A service knowledge management system (SKMS).

The following generic activities will be needed to implement such an approach:

- Establish the generic lifecycle for IT assets (requirements, design and develop, build, test, deploy, operate and optimize, dispose) and define the principal processes, policies, activities and technologies within each stage of the lifecycle for each type of asset
- Formalize the relationships between different types of IT asset, and the relationship between IT asset acquisition and management and other IT disciplines
- Define all roles and responsibilities involved in IT asset activities
- Establish measures for understanding the (total) cost of ownership of an IT service
- Establish policies for the reuse of IT assets across services – for example, at the corporate level
- Define a strategy for the acquisition and management of IT assets, including how it should be aligned with other IT and business strategies.

For the applications asset type, additionally:

- Document the role played by applications in the delivery of IT services to the business
- Ensure the generic IT asset lifecycle model is adapted to an applications lifecycle, tailored to different application types
- Set standards for the use of different approaches to developing applications, and recognize the role of development methodologies, including those based on 'reuse' (see section 3.11.3 for further discussion)
- Ensure that procedures are in place to consider all requirement types (such as operability, service performance, maintainability, security) in the early stages of application development
- Set standards for deciding on the optimal delivery of applications to the organization, such as the use of application service providers, customized developments, commercial off-the-shelf and package customization.

For the data/information asset type, additionally:

- Establish how the general principles of IT asset acquisition and management can help to manage the data/information resources of an organization.

Ensure that data designs are undertaken in the light of:

- The importance of standardized and reusable metadata
- The need for data quality
- The value of data to an organization
- The importance of legacy data and the need to carry data forward into new systems
- The need for data administration and database administration skills
- Understanding the 'corporate' (or common/cooperative) subject area and individual service ('system') views of data
- The need to manage data of non-traditional types such as text, scanned images, video and audio
- Awareness of the major storage, security and legal issues for data
- Specifying how the generic IT assets lifecycle model can be adapted to the data asset type.

For the IT infrastructure and environmental asset type, additionally:

- Establish standards for acquisition and management of the IT infrastructure and environmental equipment (including hardware, power, operating system software, database management system software, middleware and networks) and ensure they provide a stable yet adaptable foundation that underpins the provision of IT services to the business

- Establish how the generic IT assets lifecycle model should be adapted to a specific IT infrastructure lifecycle
- Establish activities to optimize the usage of IT infrastructure assets through their reuse
- Specify the need for tools and describe how their overall use and integration assists in the management of an effective IT infrastructure and related services
- Specify green IT/sustainability requirements in areas such as power consumption and recyclability of assets at the end of the asset lifecycle review.

For the skills (people, competencies), additionally:

- Formalize how the competencies of individuals responsible for the IT assets and related services can be regarded as an asset within the organization and are managed as such
- Specify how the IT asset lifecycle applies to people assets, particularly in terms of measurable competencies, such as skill, knowledge, understanding, qualifications, experience, attitude and behaviour
- Ensure the documentation of the competencies currently in place and specify how these can be reused or enhanced
- Ensure organization standards are compatible with existing standard competency frameworks for the IT sector, such as SFIA+ (Skills For The Information Age) skills, and competencies are incorporated into roles and responsibilities.

In addition, to establish effective interfaces and dependencies:

- Define the interfaces that IT asset acquisition and management has with IT-enabled business change, IT project management and IT security
- Formalize the interfaces that IT asset acquisition and management have with functions and processes outside IT
- Formalize measurement and reporting in this area by:
  - Identifying suitable metrics and the reports on IT assets for distribution throughout the organization as appropriate
  - Formalizing quality control and measurement in the acquisition and management of IT assets.

7.2 SERVICE MANAGEMENT TOOLS

Tools will enable the service design processes to work more effectively. Tools will increase efficiency and effectiveness, and provide a wealth of management information, leading to the identification of weaknesses and opportunities for improvement. The longer-term benefits to be gained from the use of tools are cost savings and increased productivity, which in turn can lead to an increase in the quality of the IT service provision.

The use of tools will enable the centralization of key processes and the automation and integration of core service management processes. The raw data collected by the tools can be analysed, resulting in the identification of 'trends'. Preventive measures can then be implemented, again improving the quality of the IT service provision.

7.2.1 Defining tool requirements

Some points that organizations should consider when evaluating service management tools include:

- Data structure, data handling and integration
- 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 ease of use permitted by the user interface
- Support for monitoring service levels
- Distributed clients with a centralized shared database (e.g. client server)
- Conversion requirements for previously tracked data
- Data backup, control and security
- Support options provided by the tool vendor
- Scalability at increasing of capacity (the number of users, volume of data and so on).

Consideration must be given to the exact requirements for the tool. What are the mandatory requirements and what are the desired requirements? Generally the tool should support the processes, not the other way round, so minimize modification of the processes to fit the tool. Where possible, it is better to purchase a fully integrated tool (although not at the expense of efficiency and effectiveness) to underpin many (if not all) service management processes. If this is not possible, consideration must be given to the interfaces between the various tools.

It is essential to have a statement of requirements (SoR) for use during the selection process – this statement can be used as a checklist. The tool requirements should be categorized using the MoSCoW analysis:

- M – **Must** have this
- S – **Should** have this if at all possible
- C – **Could** have this if it does not affect anything else
- W – **Won't** have this time but **would** like in the future.

For more information on the documentation of requirements, see section 5.1.5.

The tool must be adequately flexible to support your required access rights. You must be able to determine who is permitted to access what data and for what purpose – for example, read access to customers.

7.2.2 Tool selection

In the early stages, consideration must also be given to the platform on which the tool will be expected to operate – this may be on existing hardware and software or a new purchase. There may be restrictions laid down by IT strategy – for example, all new products may have to reside on specific servers. This might restrict which products could be included in the evaluation process. Make sure that the procurement fits within existing approved budgets.

Hints and tips

There are many service management tools available. Details can be found on the internet, in service management manuals, from asking other organizations, from asking consultants or by attending seminars and conferences to see what products are available.

During the early stages of the selection process, think about vendor and tool credibility. Are they still going to be supporting the purchase in a few months' or a year's time? Consider the past record of the supplier as well as that of the tool. Telephone the supplier's service desk to see how easy it is to get through, and ask some test questions to assess their technical competence. Ask the vendor to arrange a visit to a reference site to see what the experience is with the tool in practice – if possible without the vendor or supplier present. Make sure that the organization has similar requirements of the tool. See the tool in operation and speak to the users about their experiences, both initially and ongoing.

Assess the training needs of the organization and evaluate the capability of the supplier to provide the appropriate training. Also the ongoing training and tool update (upgrades and changes in user requirements) will need to be assessed to ascertain the support and training costs. In particular, consider training costs, training location, time required, and how soon after training the tool will be in use; and during the implementation project ensure that sufficient training is provided – think about how the new tool will impact both IT and customer. Also ensure that interfaces with other tools and telephony are functioning correctly. It is wise to identify whether the planned combination has been used (or tried) elsewhere and with what results. Consider periods of parallel running alongside existing solutions before finally going live.

When evaluating tools, a 100% fit to requirements should not be expected and will almost certainly not be found. The '80/20 rule' should be brought into effect instead. A tool is deemed to be fit for its purpose if it meets 80% or more of the business's operational requirements. Those operational requirements should be categorized as discussed earlier.

Any product should be rejected as unsuitable, however, if not all of the mandatory requirements ('must haves') are met. In some circumstances, it will be impossible to find an existing software product that will either meet all of the mandatory requirements or provide an 80% match. In this situation, the product offering the best functional design should be selected and the unsuitable elements re-written. This enhancement process should be done by the vendor if at all possible. In some cases, part of the enhancement costs may be met by the purchaser. Some products have been designed to include user hooks – this provides accessibility to site-written code at key procedural points, without the need for the package to be modified.

7.2.3 Implementation considerations

The work does not end when the product has been selected. In many ways this could be considered as only the beginning. The tool now has to be implemented. Once the hardware platform has been prepared and the software loaded, data population needs to be considered. What, where from, how and when? Timing is important to the testing, implementation and the go-live processes. Resources must be available to ensure success. In other words, do not schedule implementation during a known busy period, such as year-end processing.

Hints and tips

Today 'Software as a Service' (SaaS) products are available where hardware and software are not required (see *ITIL Service Strategy*, section C.2). These products give network-based access to and management of commercially available software. These types of product will still require planning and implementation, but this should simplify the process as no dedicated hardware is required.

Consideration should also be given to managed service providers and application service providers that may be able to provide the same functionality.

7.2.4 Evaluation process and criteria

Whatever tool or type of tool is chosen, the fulfilment of the requirements can be differentiated between:

- **Out of the box** The requirement is fulfilled.
- **Configuration** The tool can be configured with *x* days of effort to fulfil the requirement, and this will be preserved over product upgrades.
- **Customization** The tool must be reprogrammed with x days of effort to fulfil the requirement, and this may have to be repeated on every product upgrade.

Extensive customization of any product is always best avoided because of the high costs incurred at product upgrade. Vendors may be unwilling to support old releases, and purchasers may be unable to resource the necessary re-application of any bespoke customization. Customization may also release the vendor from much of its support obligations – this would be disastrous if, as a result, your service management system is unavailable for any length of time. Further costs would be incurred in providing the bespoke training that would be required. It would be impossible to take advantage of any cheap scheduled training courses being run by the software supplier.

The process of tool evaluation is shown in Figure 7.1.

Figure 7.1 shows the standard approach of identifying requirements before identifying products, but pragmatically there may be some element of overlap, where exploration of tools on the market opens one's eyes to new options that change the requirements. These stages are

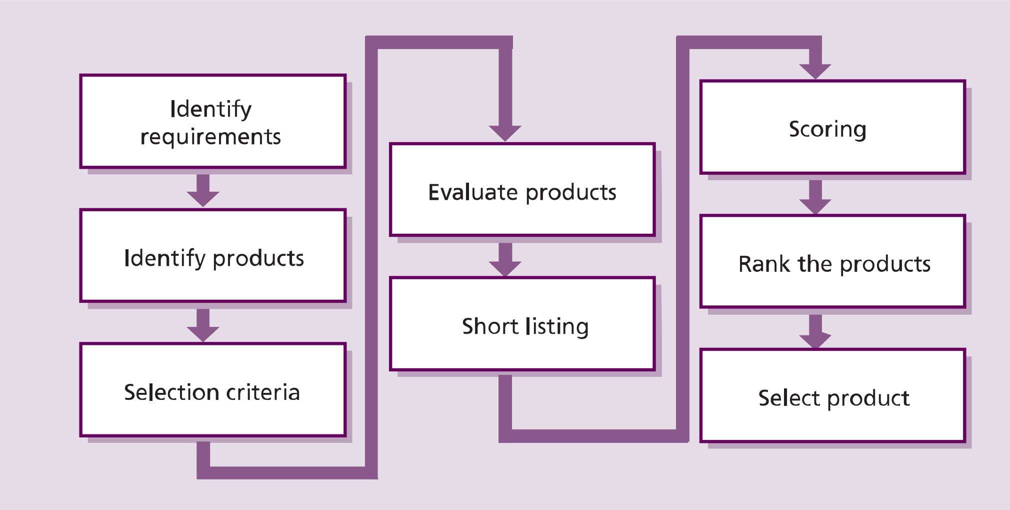


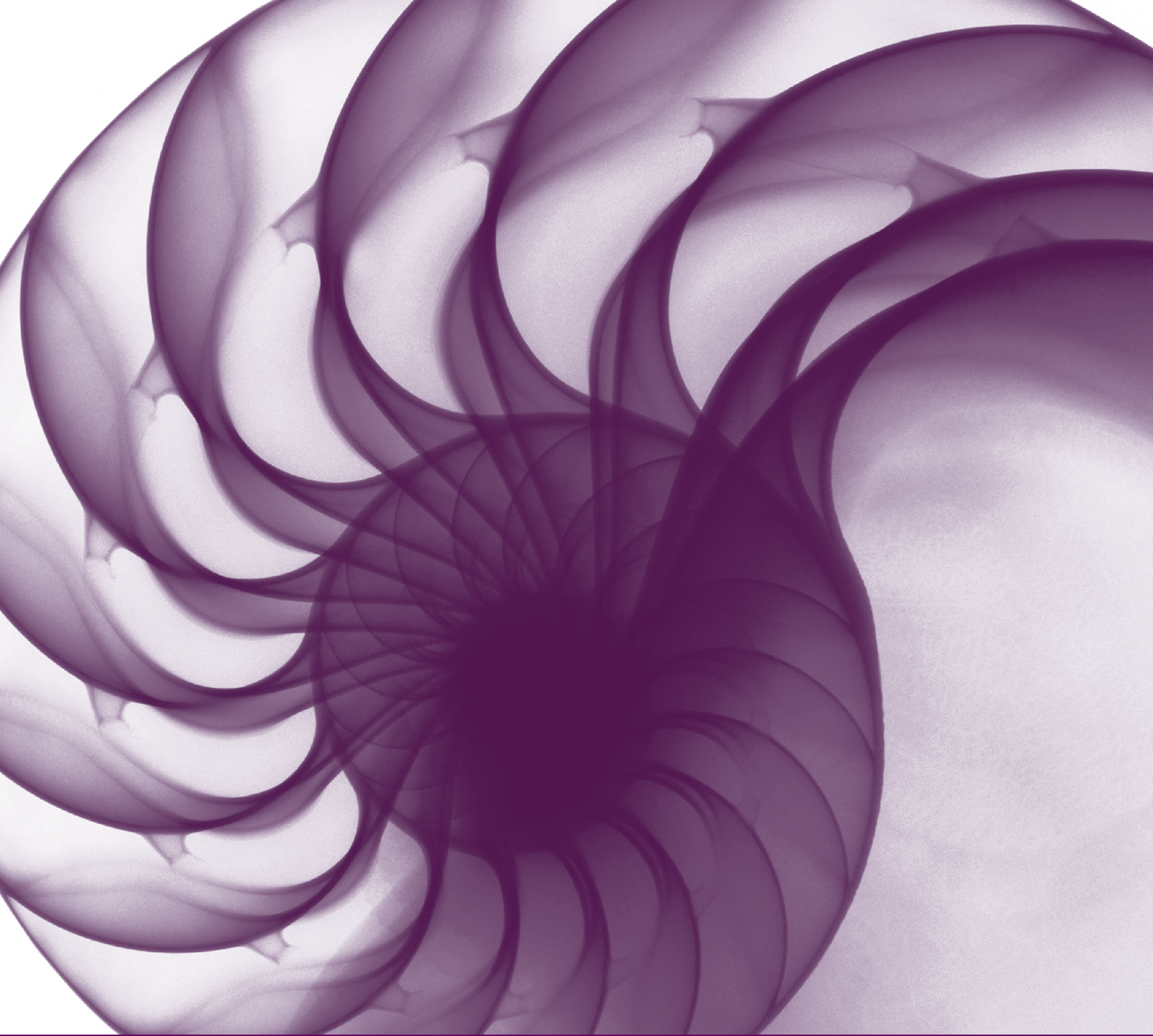
Figure 7.1 Service management tool evaluation process

targeted primarily at the evaluation of packaged software products, but a similar approach could also be used when evaluating custom-built software. Produce a clear SoR that identifies the business requirements together with the mandatory facilities and those features that it would be 'nice to have'. Also identify the site policies and standards to which the product must conform. Such standards may include it running under particular system software, or on specific hardware.

Remember the considerations about the supplier's suitability, and carry out a formal evaluation of the products under consideration.

If well-developed and appropriate tools are used to support the processes, the results achieved will be far greater and often the overall costs of service provision will be less. Selecting the right tool means paying attention to a number of issues:

- An 80% fit to all functional and technical requirements
- A meeting of all mandatory requirements
- Little (if any) product customization required
- Adherence of tool and supplier to service management best practice
- A sound data structure and handling
- Integration with other service management and operational management tools
- Support of open standards and interfaces
- Being business-driven not technology-driven
- Administration and maintenance costs within budget
- Acceptable levels of maintenance and release policies
- Security and integrity
- Availability of training and consultancy services
- Good report generation
- Scalability and growth.



8 Implementing service design

8 Implementing service design

This chapter considers the task of implementing the service design processes and tackles issues such as:

- Where do we start?
- How do we improve?
- How do we know we are making progress?

The activities of implementing and improving service design need to be focused on the needs and desires of the customer and the business. Therefore these activities should be driven and prioritized by:

- Business needs and business impacts
- Risks to the services and processes.

The activities will be influenced significantly by the requirements outlined in the SLRs and by the agreements made in the SLAs.

8.1 BUSINESS IMPACT ANALYSIS

A valuable source of input when trying to ascertain the business needs, impacts and risks is the business impact analysis (BIA). The BIA is an essential element of the overall business continuity process (see section 4.6) and will dictate the strategy for risk reduction and disaster recovery. Its normal purpose is to identify the effect a disaster would have on the business. It will show which parts of the organization will be most affected by a major incident and what effect it will have on the company as a whole. It therefore enables the recognition of the most critical business functions to the company's survival and where this criticality differs depending on the time of the day, week, month or year. Additionally, experience has shown that the results from the BIA can be an extremely useful input for a number of other areas as well, and will give a far greater understanding of the service than would otherwise be the case.

The BIA could be divided into two areas:

- One by business management, which has to investigate the impact of the loss (or partial loss) of a business process or a business function. This includes the knowledge of manual workarounds and their costs.
- A second role located in service management is essential to break down the effects of service loss to the business. This element of the BIA shows the impact of service disruption to the business. The services can be managed and influenced by service management. Other aspects also covered in 'business BIA' cannot be influenced by service management.

As part of the design phase of a new or changed service, a BIA should be conducted to help define the business continuity strategy and to enable a greater understanding about the function and importance of the service. This will enable the organization to define:

- Which are the critical services, what constitutes a major incident on these services, and the subsequent impact and disruption caused to the business – important in deciding when and how to implement changes
- Acceptable levels and times of service outage levels – again important in the consideration of change and implementation schedules
- Critical business and service periods – important periods to avoid
- The cost of loss of service – important for financial management for IT services
- The potential security implications of a loss of service – important considerations in the management of risk.

8.2 SERVICE LEVEL REQUIREMENTS

As part of the service level management process (see Chapter 4), the service level requirements for all services will be ascertained and the ability to deliver against these requirements will be assessed and finally agreed in a formal service level agreement (SLA). For new services, the requirements must be ascertained at the start of the development process, not after completion. Building the service with service level requirements uppermost in mind is essential from a service design perspective.

8.3 RISKS TO THE SERVICES AND PROCESSES

When implementing the service design and IT service management processes, business-as-usual practices must not be adversely affected. This aspect must be considered during the production and selection of the preferred solution to ensure that disruption to operational services is minimized. This assessment of risk should then be considered in detail in the service transition activities as part of the implementation process.

8.4 IMPLEMENTING SERVICE DESIGN

The process, policy and architecture for the design of IT services outlined in this publication will need to be documented and utilized to ensure the appropriate innovative IT services can be designed and implemented to meet current and future agreed business requirements.

The IT service management processes outlined in Chapter 4 and in the other ITIL publications in this series will also need to be implemented to ensure service delivery that matches the requirements of the business.

8.4.1 Where do we start?

The question often asked is 'Which process shall I implement first?' The real answer is all of them, as the true value of implementing all of the service management processes is far greater than the sum of the individual processes. All the processes are interrelated, and in some cases are totally dependent on others. What is ultimately required is a single, integrated set of processes, providing management and control of a set of IT services throughout their entire lifecycle.

While recognizing that, to get the complete benefit of implementing IT service management, all of the processes need to be addressed, it is also recognized that organizations cannot do everything at once. It is therefore recommended that the areas of greatest need be addressed first. A detailed assessment needs to be undertaken to ascertain the strengths and weaknesses of IT service provision. This should be undertaken by performing customer satisfaction surveys, talking to customers, talking to IT staff and analysing the processes in action. If desired, process and organizational maturity can also be assessed using established maturity scales. See section 8.4.2 and Appendix H for more information on maturity assessment. From the detailed assessment, short-, medium- and long-term strategies can be developed.

It may be that 'quick wins' need to be implemented in the short term to improve the current situation, but these improved processes may have to be discarded or amended as part of the medium- or long-term strategies. If 'quick wins' are implemented, it is important that they are not done at the expense of the long-term objectives, so these must be considered at all times. Every organization will have to start somewhere, and the starting point will be wherever the organization is now in terms of IT service management maturity. If the right 'quick wins' are selected, their achievement will not only improve the immediate situation, but will also build commitment to the adoption of a service management approach through demonstrated value of the principles in action.

Implementation priorities should be set against the goals of a service improvement plan (SIP). For example, if availability of IT services is a critical issue, focus on those processes aimed at maximizing availability (e.g. incident management, problem management, change management and availability management). Throughout the implementation process, key players must be involved in the decision-making process. These will include receivers as well as providers of the service. There can be a tendency, when analysing the areas of greatest need, to go straight for tools to improve the situation. Workshops or focus groups will be beneficial in understanding the requirements and the most suitable process for implementation that will include people, processes, products and partners.

8.4.2 How do we improve?

The first thing to do is to establish a formal process and method of implementation and improvement of service design, with the appropriate governance in place. This formal process should be based around the six-stage approach illustrated in Figure 8.1. More information can also be found on this approach in section 4.1.4.2 as well as in *ITIL Continual Service Improvement*.

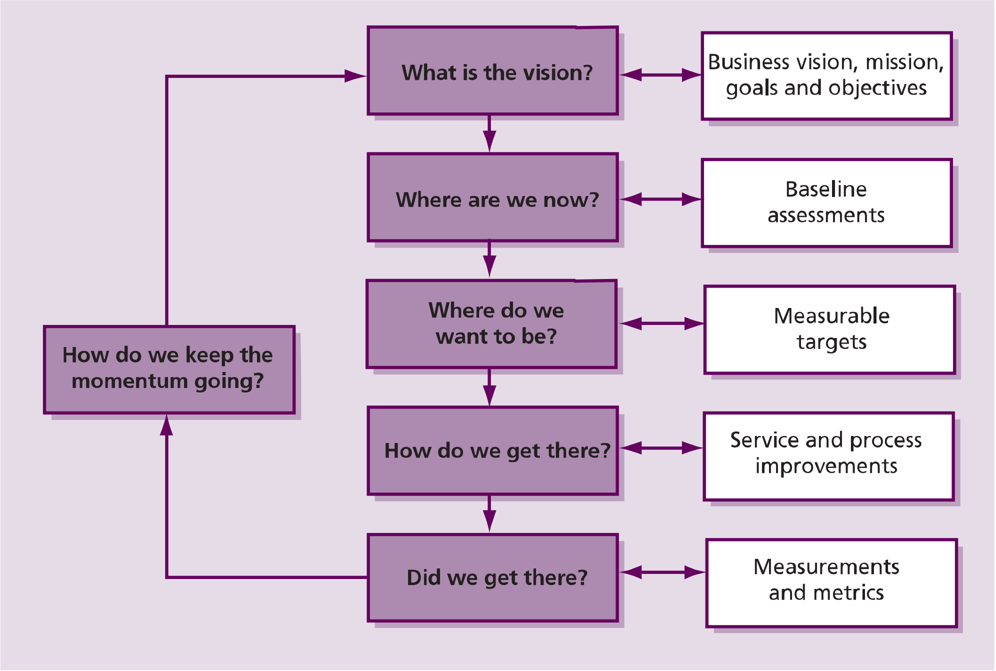


Figure 8.1 Implementation/continual service improvement approach

It is important that, when implementing or improving processes, a structured project management method is used. The improvement process can be summarized as:

- First, understanding the vision by ascertaining the high-level business objectives. The 'vision-setting' should set and align business and IT strategies.
- Second, assessing the current situation to identify strengths that can be built on and weaknesses that need to be addressed. So 'Where are we now?' is an analysis of the current position in terms of the business, organization, people and process.
- Third, 'Where do we want to be?' is a development of the principles defined in the vision-setting, agreeing the priorities for improvement.
- Fourth, detailing the SIP to achieve higher-quality service provision.
- Next, measurements and metrics need to be put in place to show that the milestones have been achieved and that the business objectives and business priorities have been met.
- Finally, the process should ensure that the momentum for quality improvement is maintained.

The implementation/continual service improvement approach is useful in checking the alignment between the business and IT, as shown in Figure 8.1.

The following are key elements for successful alignment of IT with business objectives:

- Vision and leadership in setting and maintaining strategic direction, clear goals, and measurement of goal realization in terms of strategic direction
- Acceptance of innovation and new ways of working
- Thorough understanding of the business, its stakeholders and its environment
- IT staff understanding the needs of the business
- The business understanding the potential of IT
- Information and communication available and accessible to everyone who needs it
- Separately allocated/dedicated time to familiarize with the material about the business
- Continuous tracking of technologies to identify opportunities for the business.

The implementation/continual service improvement approach may be used at any level – strategic, tactical or operational – depending on the focus of the implementation or improvement being addressed. In the following sections each step in the approach will be discussed in greater detail.

Further information on the continual service improvement approach can be found in Chapter 3 of *ITIL Continual Service Improvement*.

8.4.2.1 What is the vision?

The starting point for all of these activities is the culture and environment of the service provider organization. The people and the culture have to be appropriate and acceptable to improvement and change. Therefore, before attempting anything else, the culture within the service provider needs to be reviewed to ensure that it will accept and facilitate the implementation of the required changes and improvements. The following key steps need to be completed to achieve this stage of the cycle:

- Establish a vision, aligned with the business vision and objectives
- Establish the scope of the project/programme
- Establish a set of high-level objectives
- Establish governance, sponsorship and budget
- Obtain senior management commitment
- Establish a culture focused on:
  - Quality
  - Customer and business focus
  - A learning environment
  - Continual improvement
  - Commitment to the 'improvement cycle'
  - Ownership and accountability.

8.4.2.2 Where are we now?

Once the vision and high-level objectives have been defined, the service provider then needs to review the current situation, in terms of what processes are in place and the maturity of the organization. The activities that need to be completed here are a review, assessment or a more formal audit of the current situation, using a preferred technique such as:

- Internal review or audit
- Maturity assessment
- External assessment or benchmark
- ISO/IEC 20000 assessment or audit
- Audit against COBIT
- Strengths, weaknesses, opportunities and threats (SWOT) analysis
- Risk assessment and management methodology.

The review should include people, processes, products and partners, as well as cultural and other factors:

- The culture and maturity of the service provider organization
- The processes in place and their capability, maturity and adoption
- The skills and competence of the people
- The services and technology
- The suppliers, contracts and their capability
- The quality of service and the current measurements, metrics and key performance indicators (KPIs)

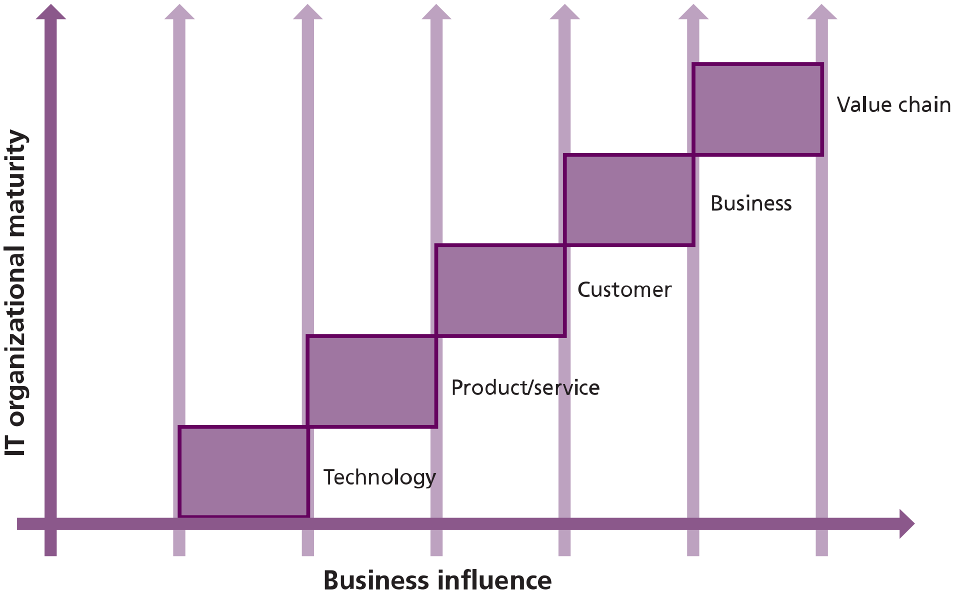


Figure 8.2 Cultural maturity assessment

- The alignment with business goals, objectives and business strategy
- A report summarizing the findings and recommendations.

Special attention should be paid to obtaining baseline measurements and metrics of the current state. These baselines will provide a valuable objectivity in assessing the best opportunities for improvement, inform the development of measurable targets for improvement, and provide a basis for later comparison after improvement efforts have been undertaken. It may be that the quality of the metrics available prior to improvement are of poor quality, with issues about accuracy and completeness, but any available metrics are better than none. Also, creating a baseline early will uncover measurement and reporting weaknesses for future improvement.

The review of the culture should include assessing it in terms of its capability and maturity within the IT service provider organization, as shown in Figure 8.2. This assessment should be based on the fact that each growth stage represents a transformation of the IT organization and as such will examine:

- Changes in people (skills and competencies)
- Processes and ways of working
- Technology and tools (to support and enable the people and processes)
- Steering (the visions, goals and results)
- Attitude (the values and beliefs)
- The appropriate level and degree of interaction with the business, customers, users and other stakeholders.

The assessment should also include a review of the capability and maturity of the service design processes, as shown in Figure 8.3. All aspects of the processes and their use should be examined, including:

- Vision: steering, objectives and plans
- Process maturity, functionality, usage, application, effectiveness and efficiency together with ownership, management and documentation
- People: the roles, responsibilities, skills and knowledge of the people
- Products, including the tools and technology used to automate the processes
- Culture: the focus, attitudes and beliefs.

The above framework can be used to provide consistency in process assessment. Assessing these two aspects will determine the current state of the organization and its service management capability and maturity. When starting out on the implementation or improvement of service design, or any set of processes, it is important to build on the strengths of the existing cultures and processes and rapidly identify and improve the weaknesses. A more detailed explanation of this framework is contained in Appendix H.

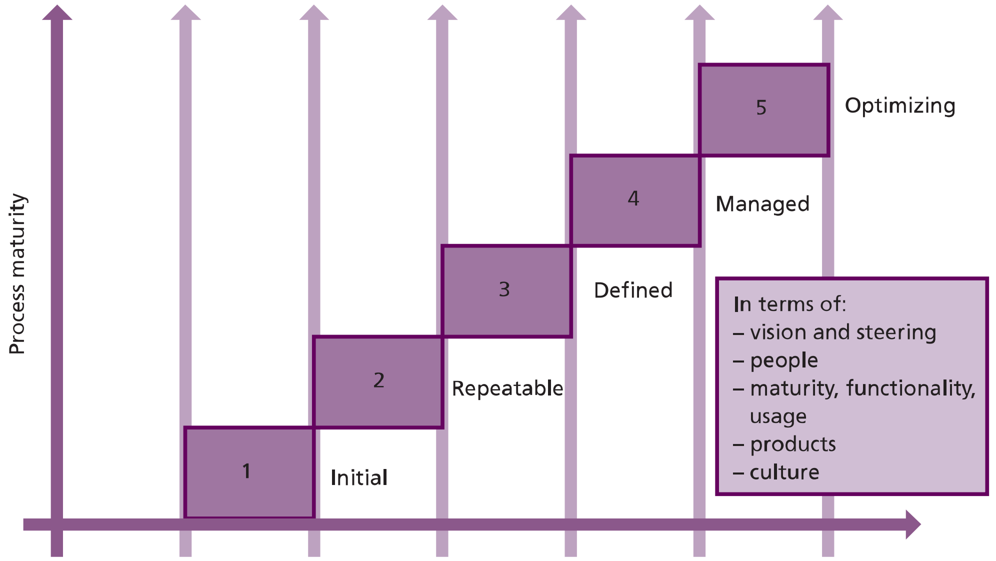


Figure 8.3 Process maturity framework

8.4.2.3 Where do we want to be?

Based on the current state assessment, and the vision and high-level objectives, a future desired state can be defined. This should be expressed in terms of planned outcomes, including some or all of:

- Improved IT service provision alignment with total business requirements
- Improved quality of service design
- Improvements in service levels and quality
- Increases in customer satisfaction
- Improvements in process performance.

The future desired state should be defined as specifically as possible to ensure success. The use of SMART objectives (specific, measurable, achievable, relevant and time-bound) is valuable in building clear and unambiguous expectations for the improvement.

8.4.2.4 How do we get there?

A set of improvements should then be identified to move forward from the current state to the agreed future state. A plan to implement these improvements should then be developed, incorporating service transition and service operation, and should include:

- The improvement actions
- The approach to be taken and the methods to be used
- Activities and timescales
- Risk assessment and management
- Resources and budgets
- Roles and responsibilities
- Monitoring, measurement and review.

Improvement plans should also take into consideration challenges, critical success factors and risks. See Chapter 9 for a discussion of these elements.

8.4.2.5 Did we get there?

Often organizations instigate improvement initiatives without considering or designing the measurement system from the outset. The success of the initiative cannot, therefore, be ascertained because we have no benchmark or baseline before, during or after the implementation. It is imperative that the measurements are designed before the implementation. A defined set of metrics needs to be utilized in order to ensure that the desired future state is achieved. This desired future state needs to be expressed in measurable terms (a central aspect of SMART objectives) such as:

- X% reduction in service design non-conformances
- X% increase in customer satisfaction
- X% increase in the service availability of critical services.

Thus once the improvement actions and plans have been completed, checks and reviews should be completed in order to determine:

- Did we achieve our desired new state and objectives?
- Are there any lessons learnt and could we do it better next time?
- Did we identify any other improvement actions?

For examples of specific metrics that might be used for process improvement, see the 'Critical success factors and key performance indicators' sections for each process in Chapter 4.

8.4.2.6 How do we keep the momentum going?

Having improved, the need is to consolidate and move on. The organization and the culture must recognize that they can always get better, and therefore must establish an environment of continual improvement. So, once they have achieved the new desired state, they must review the vision and objectives, identify more improvement actions, log them in the continual service improvement (CSI) register, and repeat the six-stage approach again. So this stage is all about:

- Developing a learning environment
- Establishing a desire to improve throughout the organization
- Recognizing and reinforcing the message that quality and improvement are everybody's job
- Maintaining the momentum on improvement and quality.

8.5 MEASUREMENT OF SERVICE DESIGN

The success of the service design and the success of the improvement to the processes around the service design must be measured, and the data must be analysed and reported on. Where the design or process does not meet the requirements of the business as a whole, changes to the process may be required and the results of those changes must also be measured. Continuous measurement, analysis and reporting are mandatory requirements for both the service design stage and the IT service management processes.

There are measurement methods available that enable the analysis of service improvement. The balanced scorecard is a method developed by Robert Kaplan and David Norton as a concept for measuring a company's activities in terms of its vision and strategies. It gives a comprehensive view of the performance of a business. The system forces managers to focus on the important performance metrics that drive success. It balances a financial perspective with customer, internal process, and learning and growth perspectives. More information can be found on the balanced scorecard method in *ITIL Continual Service Improvement*.

Six Sigma is a methodology developed by Bill Smith at Motorola Inc. in 1986, and was originally designed to manage process variations that cause defects, defined as unacceptable deviation from the mean or target, and to systematically work towards managing variation to eliminate those defects. Six Sigma has now grown beyond defect control and is often used to measure improvement in IT process execution. (Six Sigma is a registered service mark and trademark of Motorola Inc.)

Six Sigma (DMADV) is an improvement system used to develop new processes at Six Sigma quality levels and is defined as:

- **Define** Formally define the goals of the design activity that are consistent with customer demands and organization strategy
- **Measure** Identify critical success factors, capabilities, process capability and risk assessment
- **Analyse** Develop and design alternatives, create high-level design and evaluate design capability to select the best design
- **Design** Develop detailed design, optimize design and plan for design verification
- **Verify** Set up pilot runs, implement production process and hand over to process owners.

The Six Sigma (DMAIC) process (define, measure, analyse, improve, control) is an improvement system for existing processes falling below specification and looking for incremental improvement.

8.5.1 Prerequisites for success

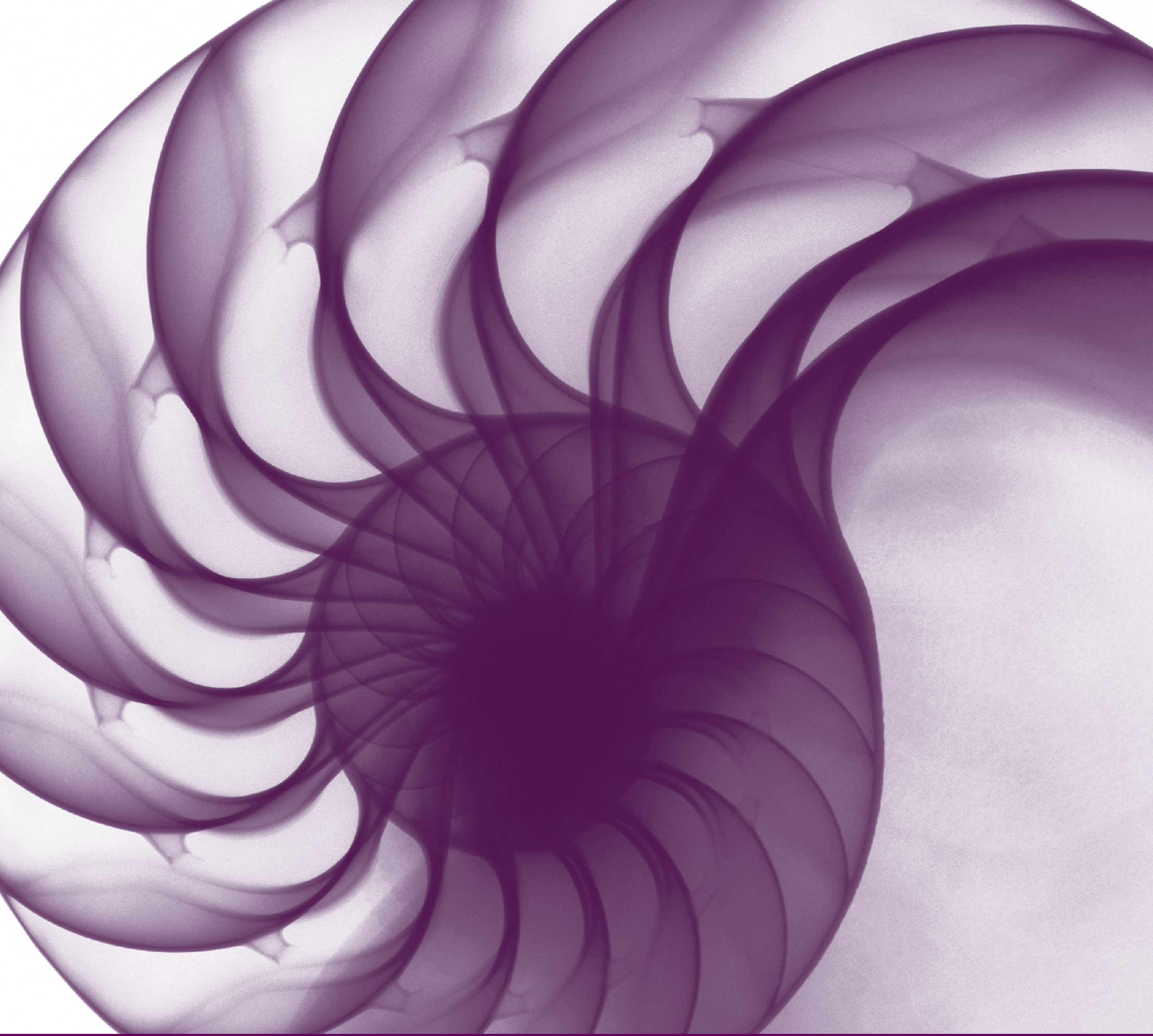
Introducing and managing an effective and efficient measurement programme for service design is dependent upon:

- Clearly defined goals and objectives for the service design stage
- A strong understanding of the processes, procedures, functions, roles and responsibilities associated with successful service design
- A strong understanding of the interfaces and dependencies between service design elements and the rest of the service lifecycle
- A strong understanding of and alignment with the needs of the business
- Development of appropriate measurement and analysis technology, methods and techniques to enable plans to be realized
- Alignment of measurements with the required metrics to accurately evaluate the health of service design, identify and implement opportunities for improvement and validate improvement accomplishments
- Regular review of the measurement programme to ensure ongoing alignment with overall service and service management requirements.

It is important to long-term success that those activities necessary for successful measurement of service design be automated wherever possible to free up human resources for the critical tasks of analysing the metrics and determining the true meaning of the information uncovered.

Measurement of service design must lead to well-prioritized and efficient improvement of service design results without unnecessary expenditure of resources to obtain the metrics.

For further information on service improvement practices, see *ITIL Continual Service Improvement*.



9

Challenges, risks and critical success factors

9 Challenges, risks and critical success factors

9.1 CHALLENGES

With every undertaking there will be challenges or difficulties to face and to overcome. This will be especially true when attempting to design new services and processes that meet the requirements of all stakeholders within the business. Experience has shown that the following will help to overcome the challenges:

- Understanding the business requirements and the business priorities and ensuring that these are uppermost in mind when designing the processes and the services
- Understanding the people and the organizational culture
- Effective communication both for explaining what is happening and how individuals will be affected and for listening to the requirements and needs of the individuals. It is vitally important to communicate with people about concerns that relate to their daily job
- Involving as many people as possible in the design. Setting up focus groups or steering groups can be very effective in getting the right solution as well as gaining wider support
- Gaining commitment from senior management as well as from all levels of staff.

Some examples of challenges that may be faced are:

- Organizational resistance to change
- Difficulty with documentation and adherence to agreed practices and processes
- Unclear or changing requirements from the business. This may be unavoidable in some cases because business needs are likely to change. The important thing is to ensure that there is a very close relationship between the IT service provider organization and the business customer of the service, so that any changing requirements can be identified as quickly as possible
- A lack of awareness and knowledge of service and business targets and requirements
- Linked to the above point, it may be that certain facilities are not built into the design. Again, it is imperative that representatives of every user of the designed service or process are involved throughout the process to reduce the chance of this happening. Details of service testing (an important element here) are contained within *ITIL Service Transition*
- A resistance to planning, or a lack of planning leading to unplanned initiatives and unplanned purchases
- Inefficient use of resources causing wasted time and money
- Lack of good knowledge and appreciation of the business impacts and priorities, as mentioned previously
- Poor relationships, communication or lack of cooperation between the IT service provider and the business may result in the design not achieving the business requirements
- Resistance to work within the agreed strategy
- Use of, and therefore the constraints of, old technology and legacy systems
- Required tools are too costly or too complex to implement or maintain with the current staff skills
- Lack of information, monitoring and measurements
- Unreasonable targets and timescales previously agreed in the SLAs and OLAs
- Over-commitment of available resources with an associated inability to deliver (e.g. projects always late or over budget)
- Poor supplier management and/or poor supplier performance
- Lack of focus on service availability
- The need to ensure alignment with current architectural directions, strategy and policies. An example of this may be that the procured infrastructure may have poor monitoring and control features
- The use of diverse and disparate technologies and applications
- Lack of awareness and adherence to the operational aspects of security policies and procedures
- Ensuring normal daily operation or business as usual is considered as part of the design
- Cost and budgetary constraints

- Difficulty ascertaining the return on investments and the realization of business benefit.

9.2 RISKS

There are a number of risks directly associated with the service design stage of the service lifecycle. These risks need to be identified to ensure that they are appropriately addressed. The risks include:

- If any of the CSFs for service design are not met, then the service design or service management process will not be successful.
- If maturity levels of one process are low, it will be impossible to achieve full maturity in other processes.
- Business requirements are not clear to IT staff.
- Business timescales are such that insufficient time is given for proper service design.
- Insufficient testing, resulting in poor design and therefore poor implementation.
- An incorrect balance is struck between innovation, risk and cost while seeking a competitive edge, where desired by the business.
- The fit between infrastructures, customers and partners is not sufficient to meet the overall business requirements.
- A coordinated interface is not provided between IT planners and business planners.
- The policies and strategies, especially the service management strategy, are not available from service strategy, or its content is not clearly understood.
- Over- or under-engineered processes. Processes with too little definition and control may not consistently meet the stated objectives. Processes with too much definition and control can become an impediment to efficiency and may actually produce a negative impact on business outcomes.
- There are insufficient resources and budget available for service design activities.
- Services being developed in isolation using their 'own' assets and infrastructure. This can appear to be cheaper in isolation, but can be much more costly in the long term because of the financial savings of corporate buying and the extra cost of supporting different architectures.
- Insufficient time given to the design stage, or insufficient training given to the staff tasked with the design.
- Insufficient engagement or commitment with the application's functional development, leading to insufficient attention to service design requirements.

9.3 CRITICAL SUCCESS FACTORS AND KEY PERFORMANCE INDICATORS

Critical success factor (CSF) is a term for an element that is necessary for an organization or project to achieve its mission. CSFs can be used as a means for identifying the important elements of success.

Key performance indicators (KPIs) are measures that quantify objectives and enable the measurement of performance. KPIs should be set and measured against the design and for each of the processes to ensure that the CSFs are met. Together, CSFs and KPIs establish the baseline and mechanisms for tracking performance. Achievement against KPIs should be monitored and used to identify opportunities for improvement, which should be logged in the CSI register for evaluation and possible implementation.

Hints and tips

It is recommended that each IT organization focuses on a small sub-set of CSFs and KPIs at any one time. The required CSFs and KPIs should be set at the beginning of any implementation or improvement activities.

It is important that CSFs are agreed during the design stage of a service and of the processes, and that KPIs are set, measured and reported on to indicate the quality of the service design and the service design processes. There is a requirement to be able to analyse how well the service infrastructure was designed. It is possible to arrive at a good design in a very resource-inefficient manner, and vice versa, so it is important to look at the quality as well as resources needed to achieve the required quality. KPIs around the success of delivery of the service indicate the effectiveness of the service design – for example, does the service meet the (defined) business requirements for availability, reliability, throughput, security, maintainability, serviceability, functionality etc.?

KPIs around the resource estimates, however, will show us how efficient the design is.

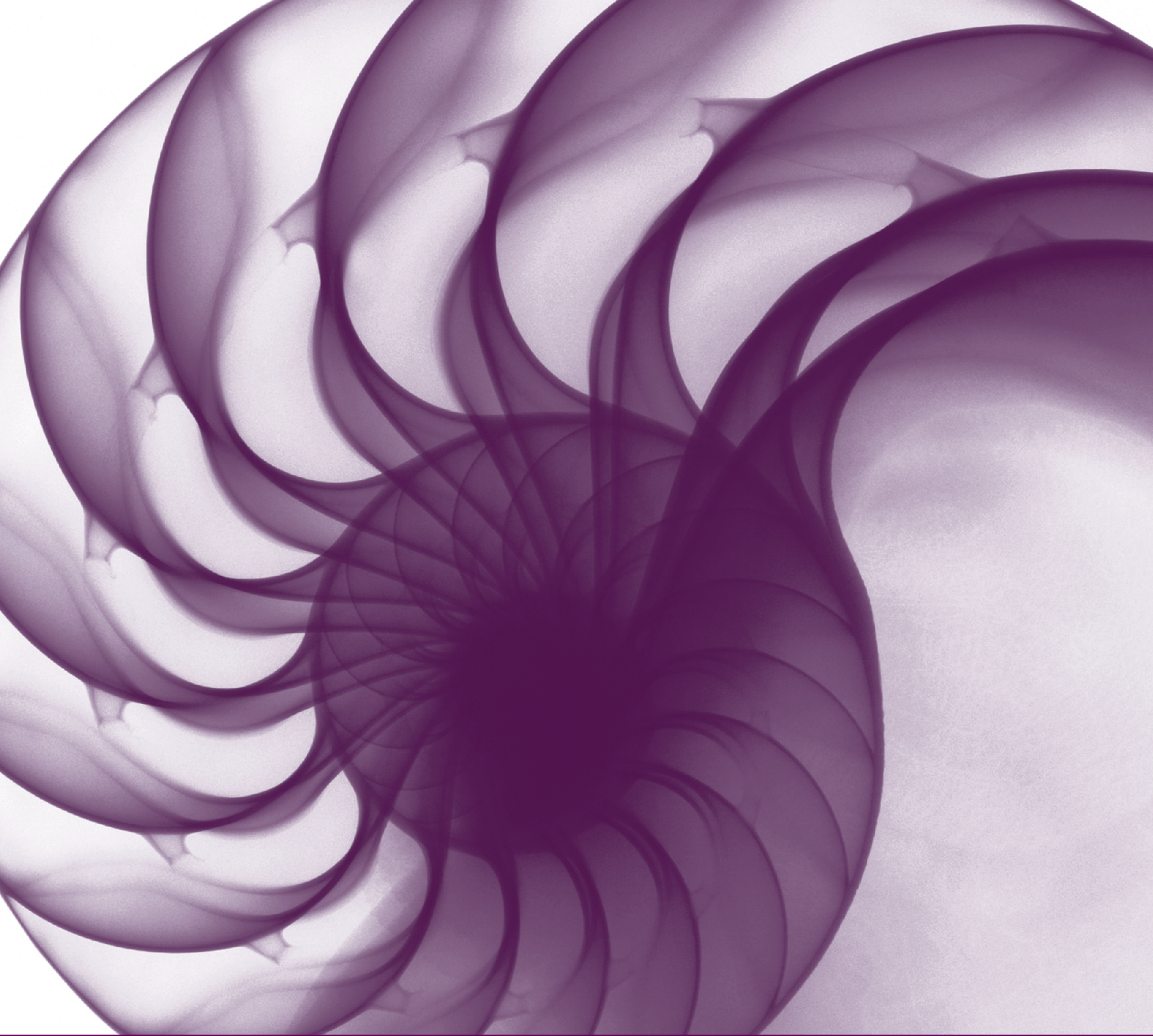
These should be defined as part of quality assurance (QA) planning and release acceptance. These KPIs could be supported by similar component metrics.

KPIs for the service design stage may include:

- Percentage of service design requirement specifications produced on time (and to budget)
- Percentage of service design plans produced on time
- Percentage of service design packs completed on time
- Percentage of QA and acceptance criteria plans produced on time
- Accuracy of service design – for example, was the correct infrastructure built to support the service?
- Percentage accuracy of the cost estimate of the whole service design stage
- Accuracy of service level agreement(s), operational level agreement(s) and contract(s) – do they really support the required level of service?

To judge service provision and ITSM process performance, clearly defined objectives with measurable targets should be set. Confirmation needs to be sought that these objectives and the milestones set in the continual service improvement (CSI) stage of the lifecycle have been reached and that the desired service quality or desired improvement in quality has been achieved. It is vital when designing services or processes that KPIs are designed from the outset and collected regularly and at important milestones. For example, at the completion of each significant stage of the programme, a post-implementation review (PIR) should be conducted to ensure the objectives have been met. The PIR will include a review of supporting documentation and the general awareness among staff of the refined processes.

A comparison is required of what has been achieved against the original goals set in the project. Once this has been confirmed, new improvement targets should be defined. To confirm that the milestones have been reached, KPIs need to be constantly monitored. These KPIs include customer satisfaction targets, so there will be a need to survey customers planned at various stages to confirm that changes made are improving the customer perception of the service quality. It is possible that the services have higher availability, that there are fewer incidents and that response times have improved, but at the same time the customer's perception of service quality has not improved. Clearly this is as important, and will need to be addressed by talking to customers to ascertain their concerns. Confirmation will need to be sought that improvements put in place are addressing the customer's primary needs.



Afterword

Afterword

Service design can be described as the design of appropriate and innovative IT services, including their architectures, processes, policies and documentation, to meet current and future agreed business requirements. This publication has explained that the better and more careful the design, the better the solution taken into live operation. It is also highly likely that the better the design, the less re-work time that will need to be undertaken during the transition and live stages of the service lifecycle.

Excellence in service design requires that the service provider moves beyond a focus on the purely technical aspects of a service and considers the non-technical aspects that can be just as critical to maximizing the value ultimately received by the customer. Proper service design does not merely allow for the activation of a new or changed service in the live environment, but also provides the basis for establishing effective use of the service by business users and customers and for effective and efficient service management, maintenance, support and ongoing improvement. As organizations move to adopt the principles of service management, the guidance in this publication will aid the evolution of their service design practices towards the holistic approach advocated in these pages.