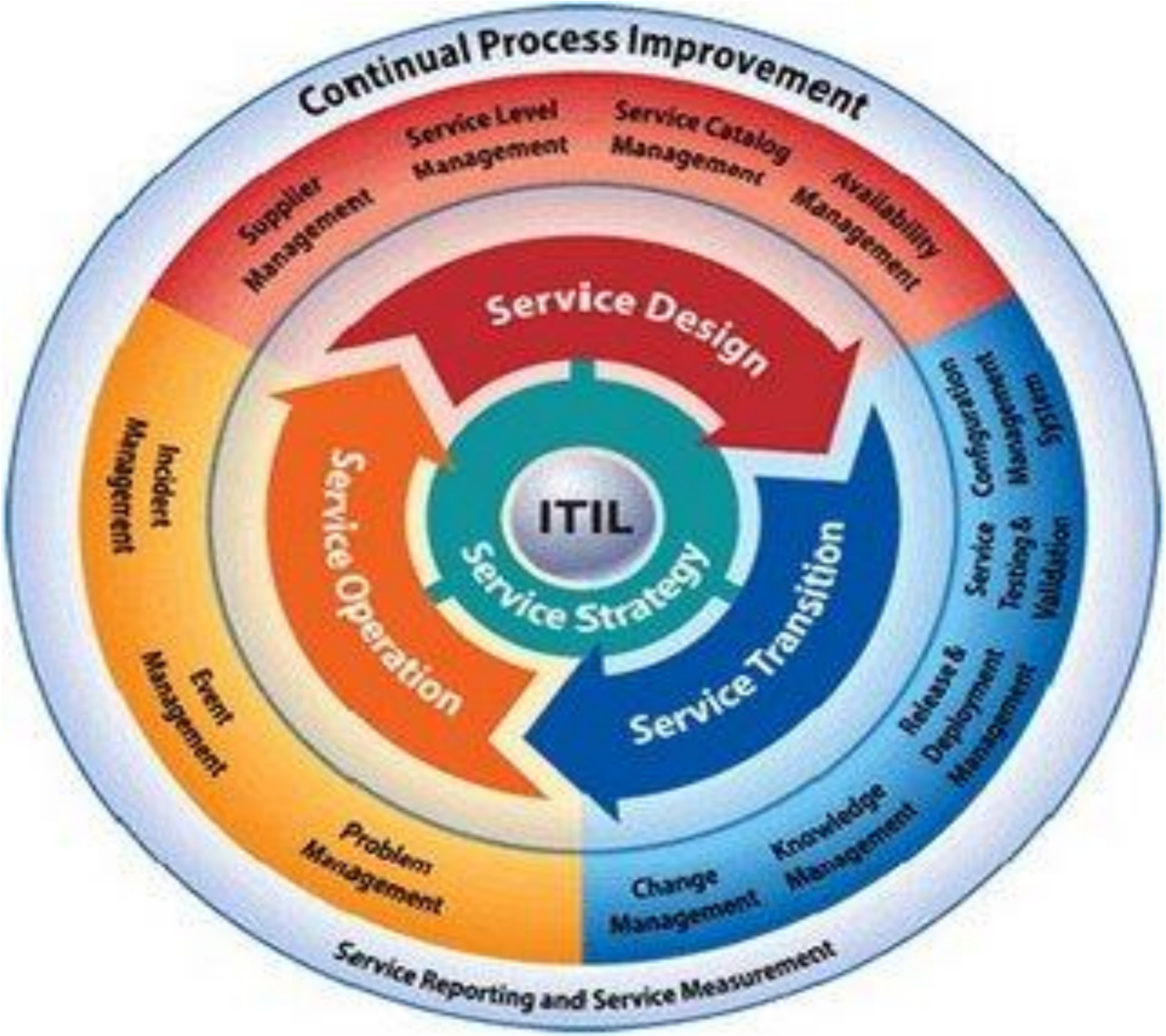
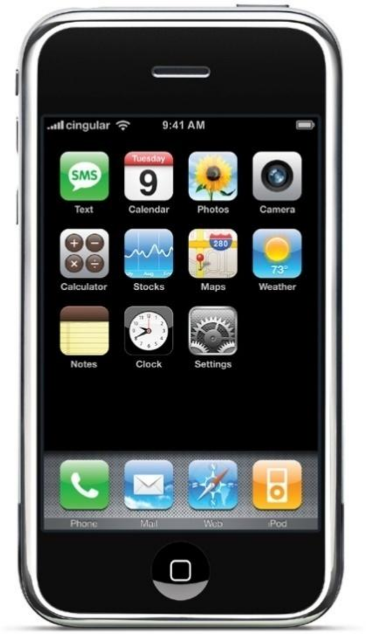
1. Service Design – Overview

Recap on ITIL Service Cycle



Product vs Service

- Good Product Design?



- Good Service Design?

What makes a good business process? (Davis R. 2009)

- Effective (can achieve its objective?)
- Relevant (crucial to the business)
- Efficient (no waste, unnecessary steps, multiple handover)
- Usable (Is it practical ?)
- Reused (modularity)
- Managed (process owner ?)
- Measured (process metrics)

2. Service Design Principles

-Goals

- Identifying service requirements

-Design activities

- Design aspects

- Service Design models

Goal of Service Design

- The design of appropriate and innovative IT services, including their architectures, processes, policies and documentation, to meet current and future agreed business requirements

Design process activities

- Requirements collection, analysis and engineering
- Design of appropriate services, technology, processes and process measurement
- Production and maintenance of IT policies and design documents
- Plan for the deployment and implementation of IT strategies using roadmaps, programmes
- Risk assessment

Service Design Aspects

- Design of the services (requirement, resources, capacities)
- Design of the management system and tools
- Technology architectures design
- Design of the processes and roles
- Design of the measurement methods and metrics of the services

Design Aspects for Incident Management

- Collect requirements to set up a service desk, examine required resources and capabilities
- Management system: e.g. need a knowledge database, escalation matrix, service level
- Technology architecture: solicit service management software, monitoring tools
- The process flow, e.g. process lifecycle from incident opening to closure
- Measurement: How to measure service level

Main Service Delivery Strategy

- In-sourcing
- Outsourcing
- Co-sourcing
- Partnership (multi-sourcing)
- Business Process Outsourcing (BPO), e.g. outsourcing whole data centre
- Application Service Provision (utility service model)
- Knowledge Process Outsourcing (e.g. outsource R&D)

4 Ps for good IT service design

- People (skills and competencies)
- Products (Technology and management systems employed in IT service delivery)
- Processes (Processes, Roles and Activities)
- Partners(Vendors, Manufacturers and Suppliers)

Service Design Processes

- Service Catalogue Management (SCM)
- Service Level Management (SLM)
- Availability Management
- Capacity Management
- IT Service Continuity Management
- Information Security Management
- Supplier Management

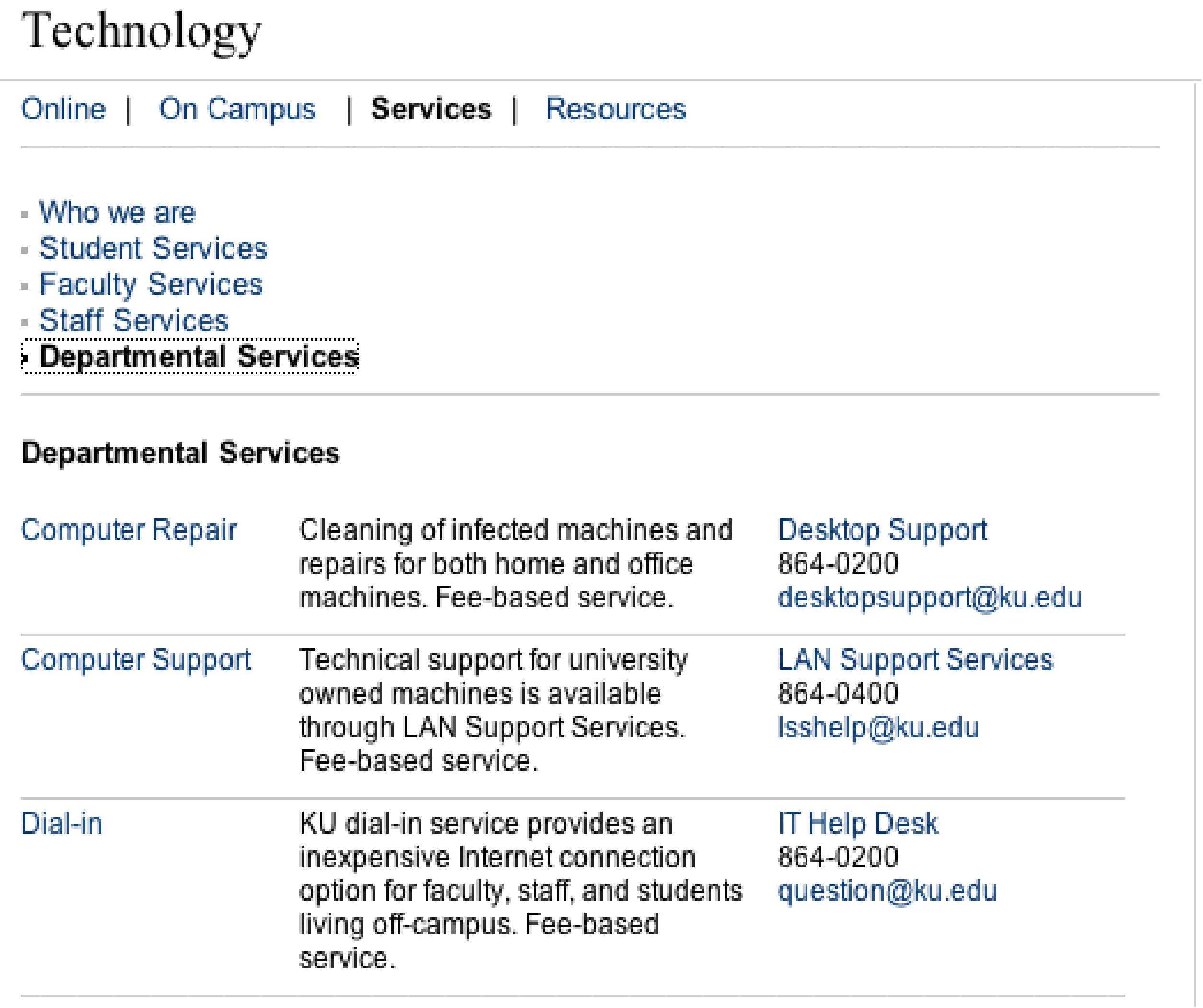
3. Service Design – Service Catalogue Management & Service Level Management

By Dr. Franklin Leung

What is a Service Catalog?

- A Service Catalog provides a clear description of all services offered and includes details about each service, e.g. who is able to use the service, who to contact for support, services hours and location, etc.

How about a better alternative?



Technology

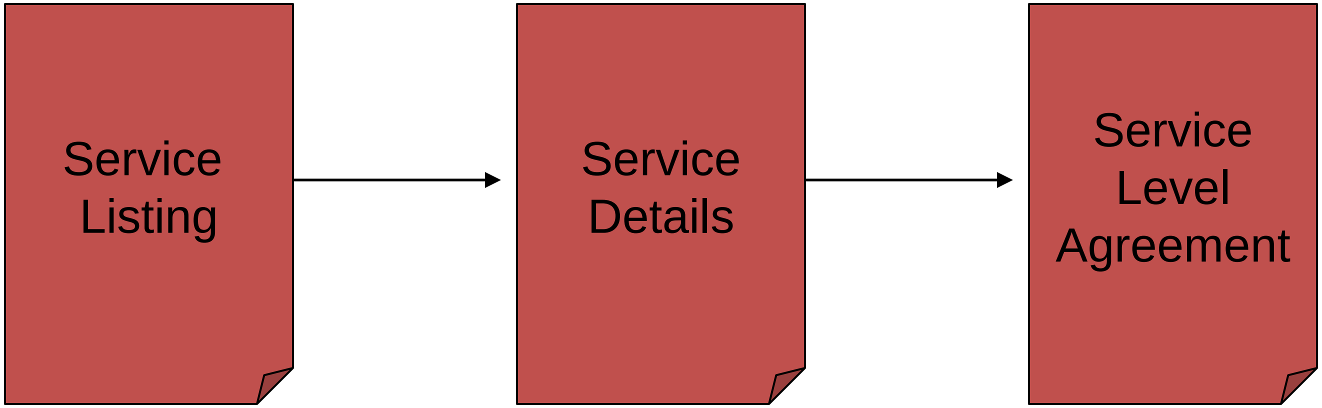
Online | On Campus | **Services** | Resources

- Who we are
- Student Services
- Faculty Services
- Staff Services
- **Departmental Services**

Departmental Services

Computer Repair	Cleaning of infected machines and repairs for both home and office machines. Fee-based service.	Desktop Support 864-0200 desktopsupport@ku.edu
Computer Support	Technical support for university owned machines is available through LAN Support Services. Fee-based service.	LAN Support Services 864-0400 lsshelp@ku.edu
Dial-in	KU dial-in service provides an inexpensive Internet connection option for faculty, staff, and students living off-campus. Fee-based service.	IT Help Desk 864-0200 question@ku.edu

Components of a Service Catalog



Service details

Service details contains items such as:

- Description (what service is this?)
- Main users (who use this service?)
- Functionality (what does this service provide?)
- Availability (when is this service available?)
- User support (when and from who is support available?)
- Any associated costs

Simple example of a Service Catalogue

Service	Customer	Accounts	Legal	Sales	HR	Retail
Payroll		X	X			
E-mail		X	X	X	X	X
Invoicing		X		X		X
Internet		X	X	X	X	
Intranet		X	X	X	X	X

Service Catalog Project Highlights

- Create service details for all currently displayed services.
- Create Service Level Agreements for all services
- Gather information on other services not currently listed.
- Promote, revise, and improve.

Service Level Management

- Service Level Management (SLM) is an IT “best practice” process that includes an ongoing review of services.
- SLM includes agreeing, monitoring , measuring, reporting, and reviewing services.
- Service Level is customer-centric (user-centric) and users should be able to understand the service level metrics (that is, the metrics should not be technical)

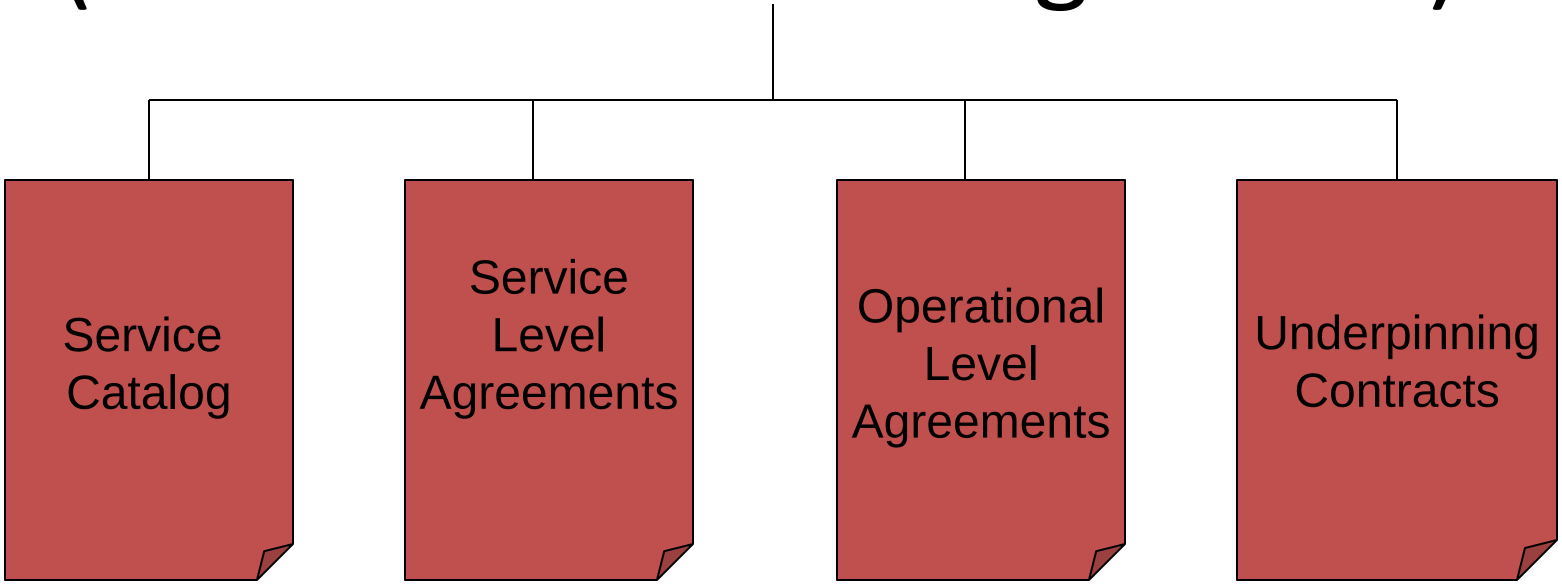
Example of Service Level

- Time periods when service is available
- Response time provided by a computer system
- Turnaround time from order to delivery after placing an order request for computer equipment
- Number of rings (or number of seconds) within which the call for service desk will be picked
- Recovery time for critical failure

Goals of Service Level Management

- Maintain & improve IT Service Quality, through a constant cycle of agreeing, monitoring & reporting upon IT Service achievement & instigation of actions to eradicate poor service

Components of SLM
(Service Level Management)

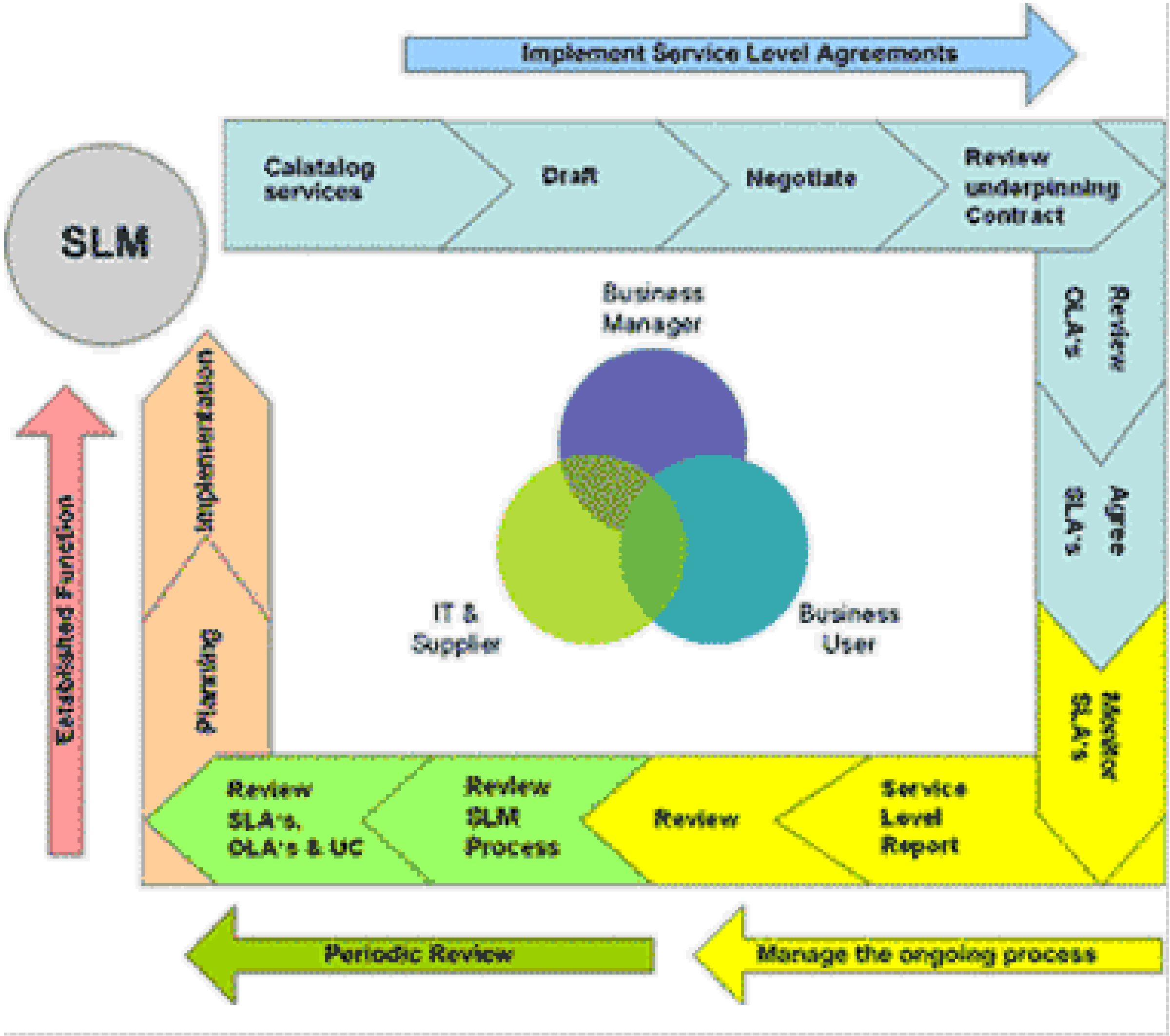


Service Catalog

Service Level Agreements

Operational Level Agreements

Underpinning Contracts


Implement Service Level Agreements
SLM
Calatalog services
Draft
Negotiate
Review underpinning Contract
Review OLA's
Agree SLA's
Monitor SLA's
Business Manager
IT & Supplier
Business User
Established Function
Implementation
Planning
Review SLA's, OLA's & UC
Review SLM Process
Review
Service Level Report
Periodic Review
Manage the ongoing process

Scope of Service Level Management

- SLAs should be established for all IT Services being provided. Underpinning Contracts & Operational Level Agreements (OLAs) should also be in place with all those suppliers upon who the delivery of service is dependent

What is a Service Level Agreement

- An SLA is a formal negotiated agreement that defines in quantitative terms the service being offered to a customer. Any metrics included should be capable of being measured on a regular basis. SLAs should be renegotiated whenever a business service is subject to change.

Components (Cont’d)

- Service Level Agreements (SLA) – established between users and IT service provider
- Operational Level Agreement (OLA) – established between the Service Manager and internal teams within IT (e.g. Network Team agrees to provide networks service at a defined service level, Database Team agrees to provide database service at a defined level
- Underpinning Contracts (OC) – established between Service Provider and External Vendors

Accenture Case

- A new initiative for IT Department to provide clear and verifiable service levels for each of the IT Products and Services offered. The optimal service levels would be determined by what users require.

SLM Key Concepts

- Planning the Process
- Implementing the Process
- Ongoing Process

SLM – Planning the process

- Initial planning activities
- Plan monitoring activites
- Establish perception of the services
- Underpinning contacts and OLAs

SLM – Implementing the Process

- Managing the expectation of the customer
- Seeking agreement on the service level
- SLA structures

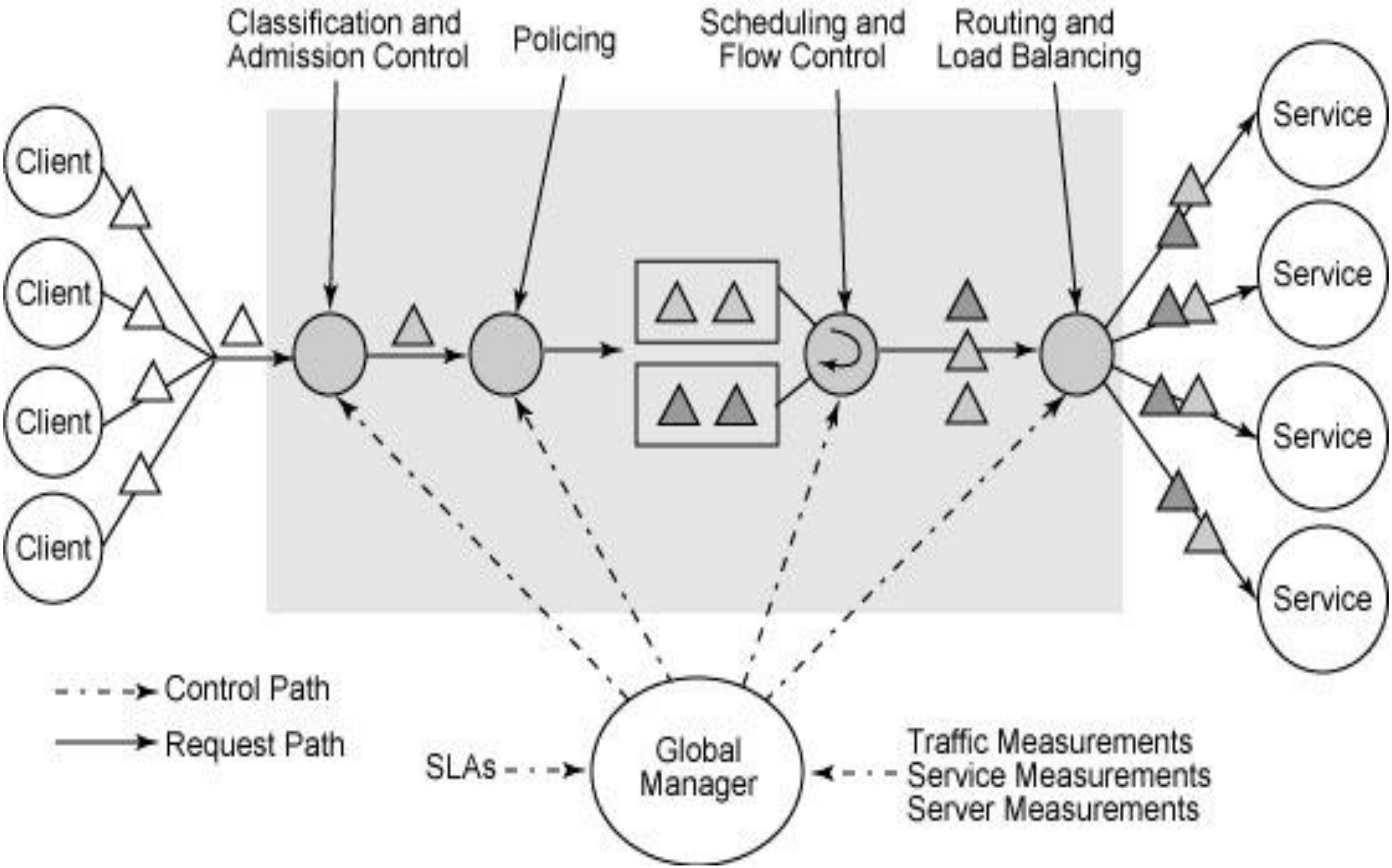
  - Service based

  - Customer based

  - Multi-level SLAs

Multi-Level SLAs Structure

- Corporate Level SLA (covering generic SLM issues appropriate to every customers)
  - Customer Level SLA (covering all SLM issues relevant to a particular customer)
    - Service Level SLA (Covering all SML issues relevant to the specific service)


Classification and Admission Control
Policing
Scheduling and Flow Control
Routing and Load Balancing
Client
Client
Client
Client
Service
Service
Service
Service
Control Path
Request Path
SLAs
Global Manager
Traffic Measurements
Service Measurements
Server Measurements

SLM – Implementing the Process

- SLRs – Service Level Requirements & draft SLA (interactive process)
- Wording of SLAs (terminology, style & culture)
- Seek agreement (with the owner & negotiate)
- Establish monitoring capabilities (if you can’t measure it, don’t put it in)

SLM – Implementing the Process (Cont’d)

- Review underpinning contracts (agreements with external suppliers)
- Review OLAs – Operational Level Agreements (agreements with internal support groups)
- Define reporting & review procedures (agreed with customer)

Examples of SLA Content

- Service Description
- Introduction & Administration
- Service Hours
- List of Services and expected performance levels (Throughput, availability & reliability)
- Transaction responses times & batch turnaround time
- Change Procedures
- Charging

Examples of SLA content (Cont’d)

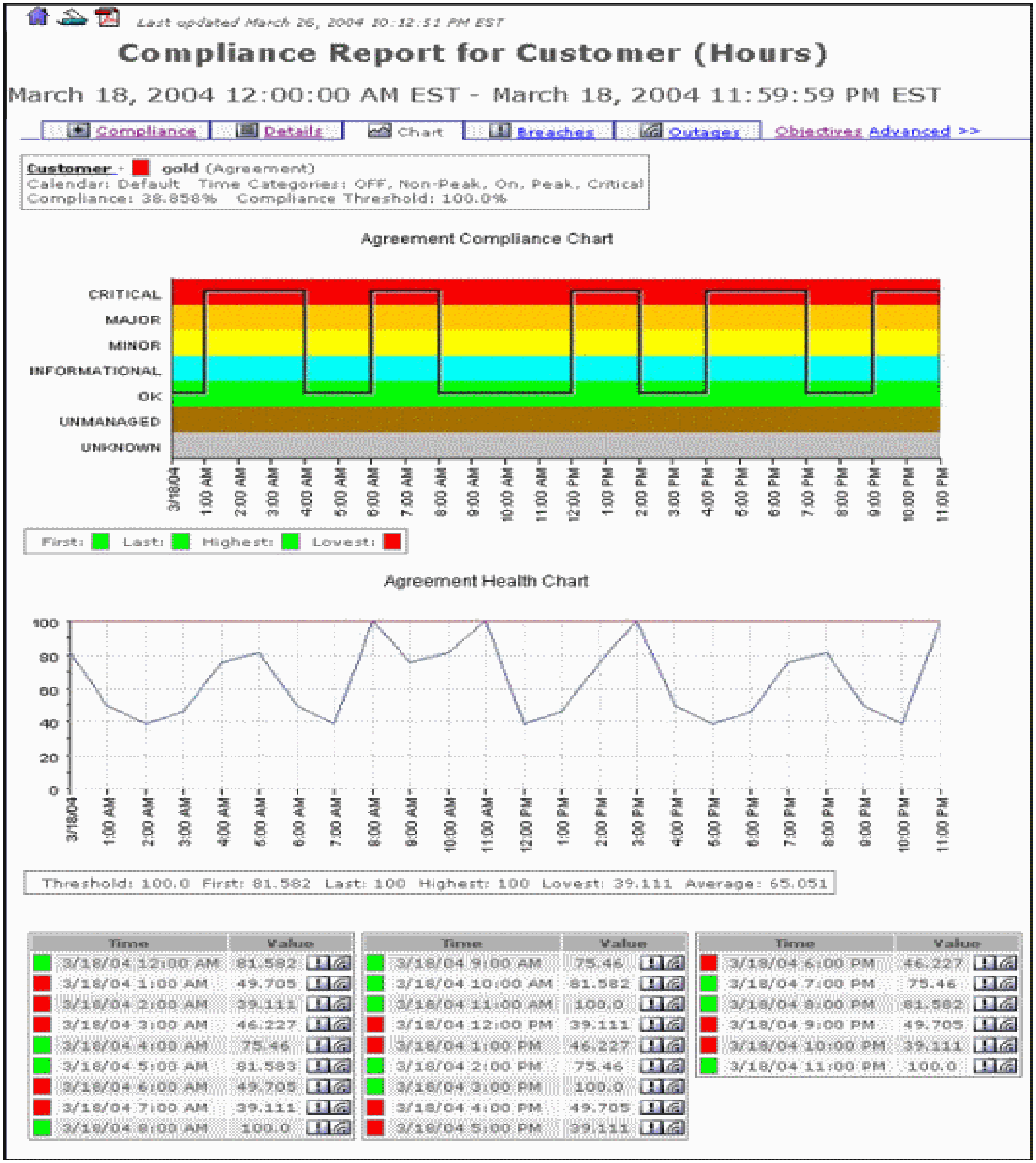
- Roles and responsibilities clearly defined (both client and service providers)
- Pricing and performance based incentives and penalties
- Security, Service Continuity (e.g. backup data centre) and Intellectual Property Ownership
- Frequency of performance reporting and review
- Clearly defined processes for changes to SLA
- Terms of voiding and termination of agreement

Awareness campaign

- Existence of new SLA must be advertised
- Service Desk & support groups must receive details of when new SLAs become operational
- Display tables of key targets for support areas
- Customers must also be included in the awareness campaign

SLM – Ongoing Process

- Monitoring SLAs (Automated tools can include thresholds to allow alerts & escalations)
- Reporting SLAs
- Service Review Meetings
- Maintenance of SLAs, contracts & OLAs



Last updated March 26, 2004 10:12:51 PM EST

Compliance Report for Customer (Hours)

March 18, 2004 12:00:00 AM EST - March 18, 2004 11:59:59 PM EST

Customer - 🟥 **gold** (Agreement)
Calendar: Default   Time Categories: OFF, Non-Peak, On, Peak, Critical
Compliance: 38.858%   Compliance Threshold: 100.0%

Agreement Compliance Chart

First: 🟩 Last: 🟩 Highest: 🟩 Lowest: 🟥

Agreement Health Chart

Threshold: 100.0 First: 81.582 Last: 100 Highest: 100 Lowest: 39.111 Average: 65.051

	Time	Value
🟩	3/18/04 12:00 AM	81.582
🟥	3/18/04 1:00 AM	49.705
🟥	3/18/04 2:00 AM	39.111
🟥	3/18/04 3:00 AM	46.227
🟩	3/18/04 4:00 AM	75.46
🟩	3/18/04 5:00 AM	81.583
🟥	3/18/04 6:00 AM	49.705
🟥	3/18/04 7:00 AM	39.111
🟩	3/18/04 8:00 AM	100.0
🟩	3/18/04 9:00 AM	75.46
🟩	3/18/04 10:00 AM	81.582
🟩	3/18/04 11:00 AM	100.0
🟥	3/18/04 12:00 PM	39.111
🟥	3/18/04 1:00 PM	46.227
🟩	3/18/04 2:00 PM	75.46
🟩	3/18/04 3:00 PM	100.0
🟥	3/18/04 4:00 PM	49.705
🟥	3/18/04 5:00 PM	39.111
🟥	3/18/04 6:00 PM	46.227
🟩	3/18/04 7:00 PM	75.46
🟩	3/18/04 8:00 PM	81.582
🟥	3/18/04 9:00 PM	49.705
🟥	3/18/04 10:00 PM	39.111
🟩	3/18/04 11:00 PM	100.0

Metrics
(Key Performance Indicator KPI)

- Percentage reduction in SLA targets missed
- Percentage increase in customer perception and satisfaction of SLA achievements, via service reviews and Customer Satisfaction Survey
- Percentage reduction in SLA breaches caused because of third-party support contracts (underpinning contracts)
- Total number and percentage increase in fully documented SLAs in place
- Percentage reduction in the costs associated with service provision

SLM Benefits

- Improvement in Service Quality
- Improved relationships including relationship building
- Clearer views of roles & responsibilities
- Clear & Consistent expectation of the level of service required
- Measurable targets

SLM – Possible Problems

- No monitoring of pre-SLA achievements
- Does not ensure (verify) that targets are achievable before committing to them
- Inadequate focus, resources, time & seniority/authorityToo technical or lengthy & not properly communicated

4. Service Design
- Availability Management

By Dr. Franklin Leung

4.1 Availability Management – General Concepts

Examples of reasons for system down

- Service Operational errors (e.g. type in wrong command, execute wrong program)
- Inadequate change management (e.g. migrate the wrong program to production environment, unauthorized change, unexpected impact due to lack of planning and communication)
- Inadequate capacity management (running of disk space, insufficient bandwidth)
- Hardware failures (network, server, storage)
- External factors (e.g. slow Internet due to damage of cross-continent cables)

Unavailability

- Availability does not come free
- Unavailability also has a cost…therefore, is not free either
- Consider the cost of failure of a highly critical business system
- Cost of failure can be measured against the VBF (Vital Business Function)

Unavailability

- Tangible costs include lost revenue, overtime payments, loss of IT staff productivity etc.
- Intangible costs include loss of customer goodwill due to dissatisfaction, loss of customer, reputation, staff morale

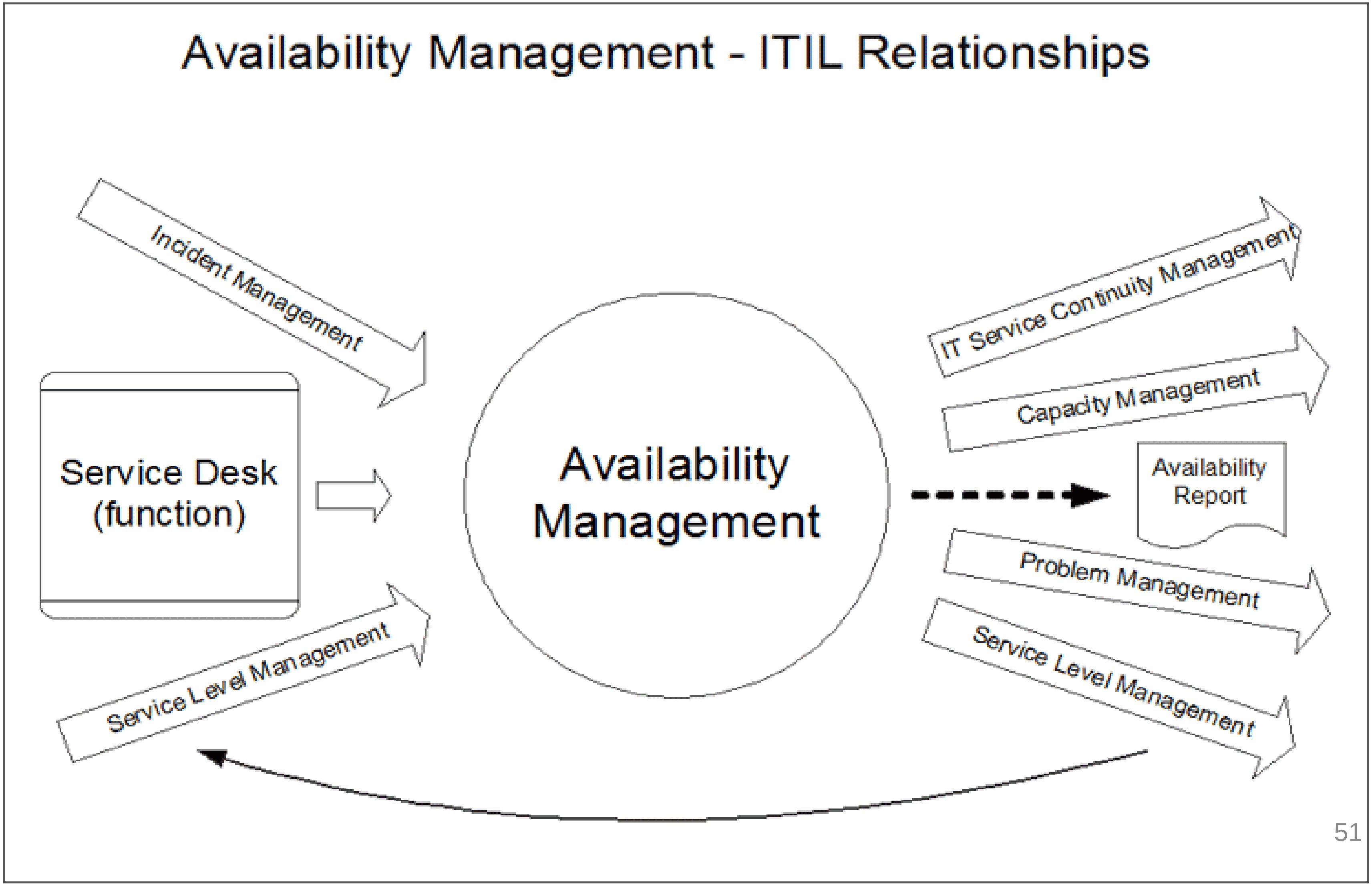
Objectives of Availability Management

- Provides a point of a focus and management for all availability-related issues, relating to services, components and resources
- Ensuring that availability targets in all areas are measured and monitored, and that they match or exceed the current and future agreed needs of the business in a cost-effective manner
- Continually optimize and proactively improve the availability of the IT infrastructure

Reactive vs Proactive

- Reactive activities of availability management
  - Monitoring, measuring, analyzing, reporting and reviewing all aspects of component and service availability
- Proactive activities of availability management
  - 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

AM’s relationship with other processes



5 key elements

- Availability
- Reliability
- Maintainability
- Serviceability
- Security
- Plus – Vital Business Functions (VBF)

Availability

- The ability of an IT service or component to perform its required function at a stated instant or over a stated period of time
- Availability is underpinned by the reliability and maintainability of the IT infrastructure & the effectiveness of IT support
- Availability is expressed as a percentage

Calculating Availability

- Availability = (AST – DT)x 100/AST

  = service or component availability (%)

  Where

  AST = Agreed Service Time

  DT = Actual downtime during agreed service time

Example

- Agreed Service Time Period:
  - Monday to Saturday , 24 hours
- Scheduled down time: Sunday (24 hours)
- Down time this week: 3 hours
- Availability this week = (6 x 24 - 3) / (6 x 24) = 97.9%

5 key elements - reliability

- The reliability of an IT service can be qualitatively stated as freedom from operational failure
- The reliability of each component & the level of resilience designed and built into the IT infrastructure will determine the reliability
- Reliability is measured as MTBSI (Mean Time between Service Incidents) or MTBF (Mean Time between Failures)

Availability versus Reliability

- Do “availability” and “reliability” mean the same?
- Any differences between “ a boss who is available” and “a boss who is reliable”?
- Think of your boyfriend/girlfriend:

  “Is he/she highly available to you” (can he/she spend a lot of time with you?)

  vs

  “Is he/she reliable”(Will he/she suddenly disappear for short periods frequently?)

5 key elements - Maintainability

- Maintainability refers to the ability of an IT Infrastructure component to be retained in, or restored to, an operational state
- Maintainability is measured as MTRS (Mean time to Restore Service) or MTTR (Mean Time to Repair)
- It is sometimes ambiguous as the time to repair a system may not be the same as the time to recover the service
- ITIL recommends the use of MTRS which covers the time to repair, the time to resolve, the time to recover.

Formula

Reliability (MTBSI in hours) =

Agreed Service Time / number of incidents

(breaks)

Reliability (MTFB in hours)

= (Agreed Service Time – Down Time)/

number of incidents

Maintainability (MTRS in hours)

= Total Down Time / number of incidents

Availability, Reliability, Maintainability

- Reliability reflects the frequency of failures (the more failures, the lower reliability)
- Maintainability reflects how long the service is down for each incident
- Availability reflects both reliability and maintainability as it measures the total uptime again the agreed service period.

5 key elements - Serviceability

- Serviceability
  - Serviceability describes the contractual arrangements made with Third Party IT Service Providers for Availability
  - There is no specific metric for Serviceability. The Availability, reliability & maintainability of IT service & components under their care is measured.

5 key elements - Security

- Security
  - Security is the Confidentiality, Integrity & Availability (CIA) of the data associated with a service
  - It is an aspect of overall availability

5 key elements (cont’d)

- Vital business functions (VBF) is the term used to reflect the business critical elements of the business process supported by an IT service
  - e.g. ATM dispenses cash & receipts, dispensing cash would be considered the Vital Business Function of the ATM service

Supplier & Maintainer relationship

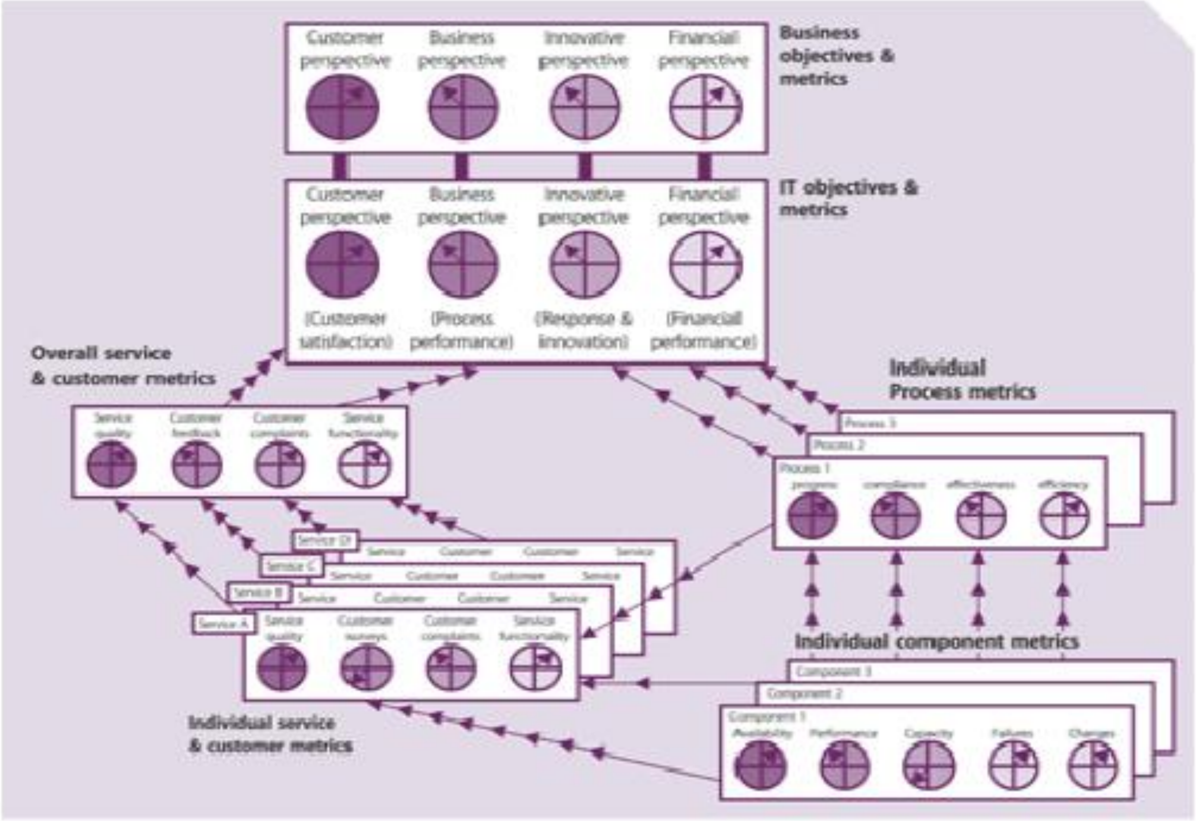
- The required levels of Availability for the IT services should be documented within a formal SLA (Service Level Agreement)
- The IT provider needs to formally agree with each Infrastructure supplier & maintainer the appropriate conditions & controls for the SLA to be met (Underpinning Contract)

IT availability Metrics Model

- Measurements need to be meaningful & add value if availability measurements & reporting are to ultimately deliver benefit to the IT & business organisation
- This is influenced strongly by the combination of ‘what you measure’ & ‘how you report it’

Metrics Tree

- Based on balanced scorecard
- Business Metrics -> IT metrics -> service & customer metrics -> individual process metrics -> individual component metrics


Customer perspective
Business perspective
Innovative perspective
Financial perspective
Business objectives & metrics
Customer perspective
Business perspective
Innovative perspective
Financial perspective
(Customer satisfaction)
(Process performance)
(Response & innovation)
(Financial performance)
IT objectives & metrics
Overall service & customer metrics
Service quality
Customer feedback
Customer complaints
Service functionality
Individual Process metrics
Process 3
Process 2
Process 1
progress
compliance
effectiveness
efficiency
Service D
Service C
Service B
Service A
Service
Customer
Customer
Service
Service quality
Customer surveys
Customer complaints
Service functionality
Individual component metrics
Component 3
Component 2
Component 1
Availability
Performance
Capacity
Failures
Changes
Individual service & customer metrics

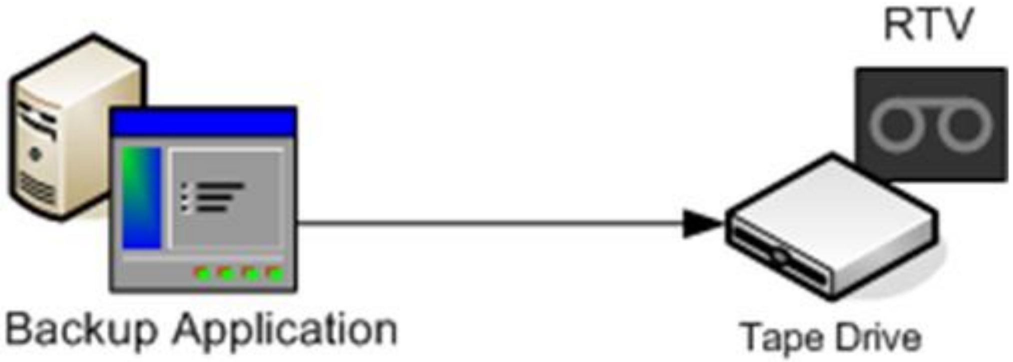
Availability Plan

- Availability plan should have aims, objectives & deliverables
- It considers the wider issues of people, process, tools & techniques
- Is “technology focused” (versus “user focused in service level management”)

4.2 Increase availability in the scenario of hardware failure

Background information on backup (making copies of data)

- Tape Drive is attached to a machine
- The simplest picture is that data in the hard disk of the computer is backed up to a tape drive



Tape Library

- Since we have to do daily/weekly backup and sometimes over one tape is needed for each backup, a tape library is needed. It also saves the routine task of changing tape daily as the machine will switch to use a new tape within the library automatically.

TLS-58132 / 264

TLS-5433 / 66

Recovery Options and High Availability Design

- Modes: Cold Standby (Recovery), Warm Standby, Hot Standby
- Challenges:
  - Time for Recovery (great impact to availability)
  - Manual versus automatic invocation
  - Data and Parameters Integrity (for database server, what happens if the database is being updated during the time of failure? Will the update become effective or not after recovery?)
  - Network routing (traffic may have to be re-routed to backup machines)

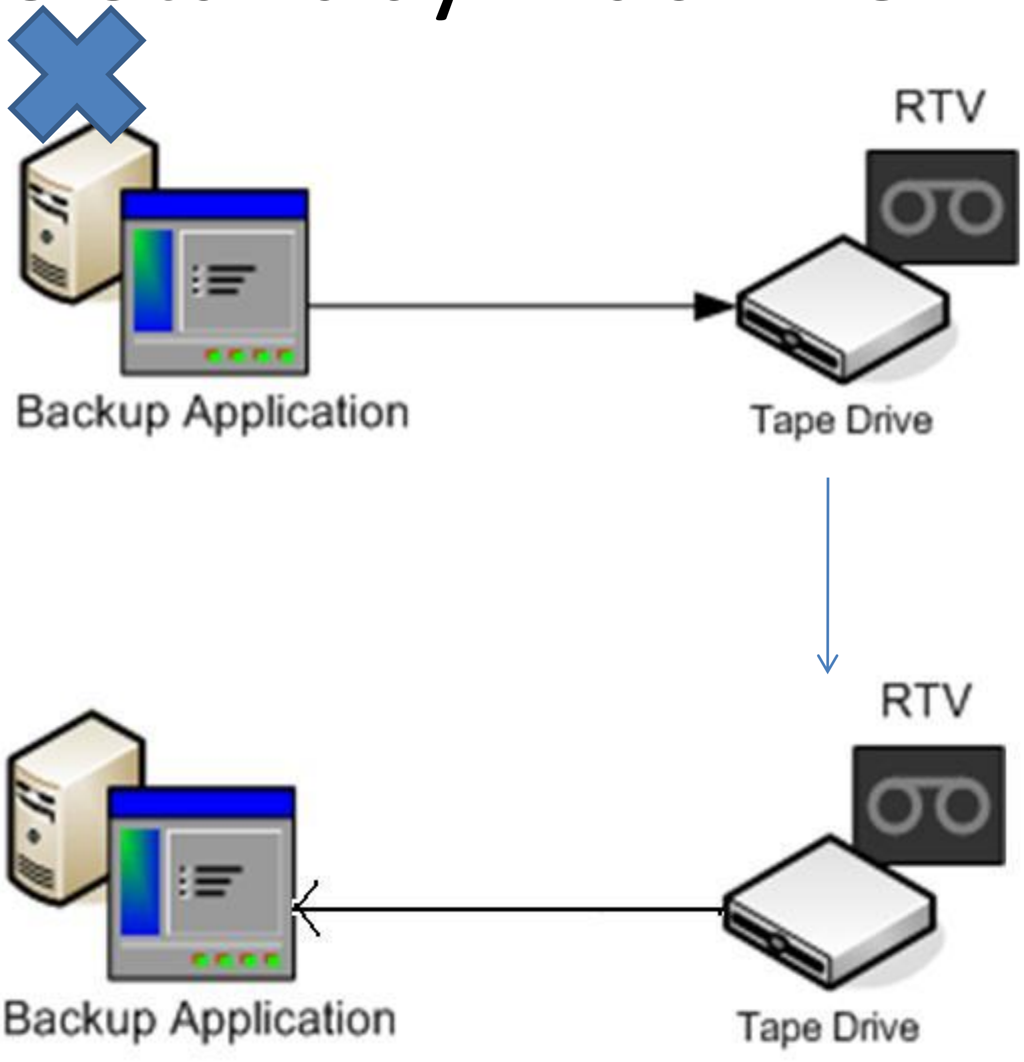
Cold Standby

- Database: A separate standby machine is idle. During failure of the primary machine, the content stored in backup media (e.g. tape) are restored from the primary machine to the standby machine.
- Others (e.g. web server, firewall, router): A separate standby machine with equivalent parameters is idle and ready. During failure of the primary machine, the standby machine will take over through manual invocation.

Primary machine fails, data restored to Standby machine via tape

Primary Server

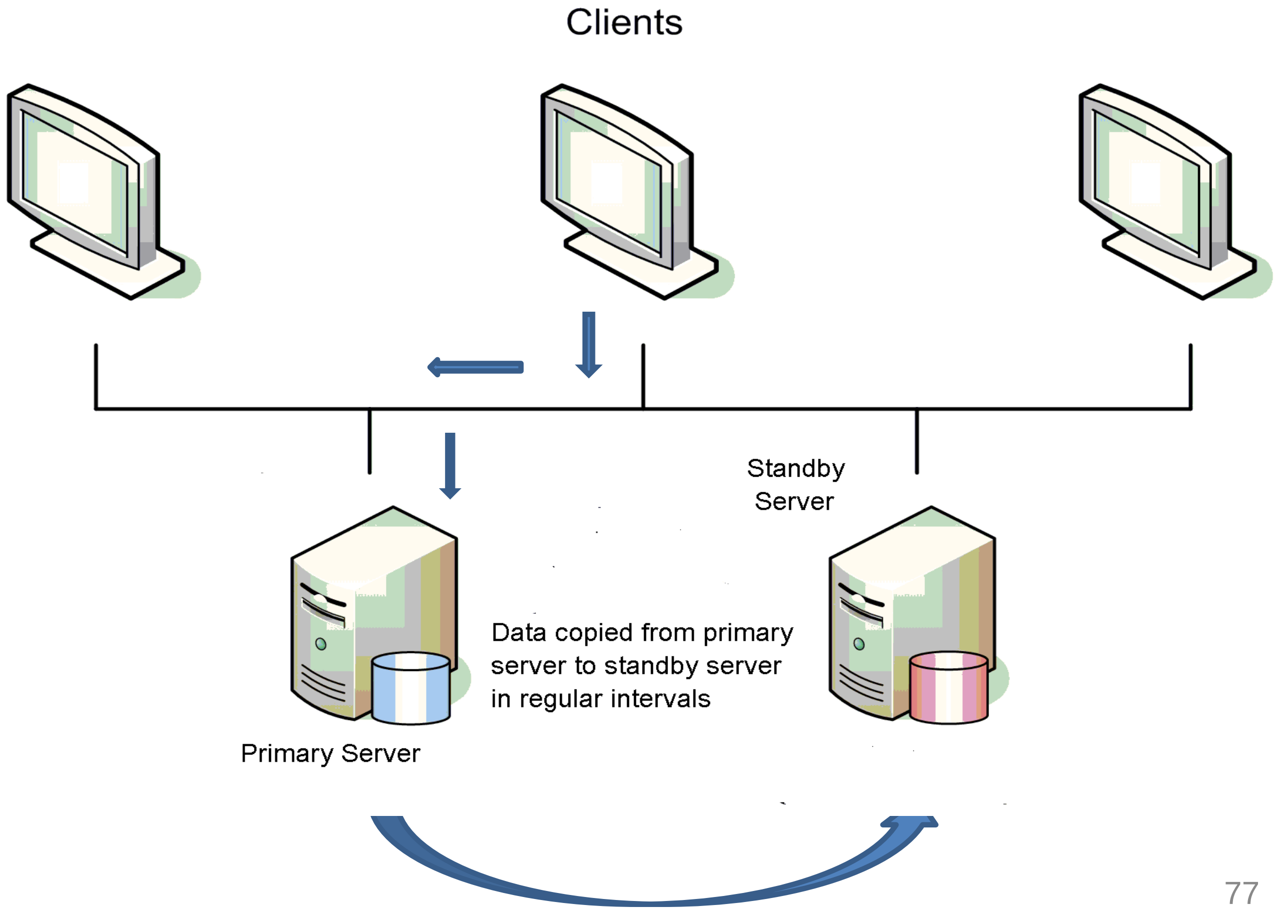
Standby Server



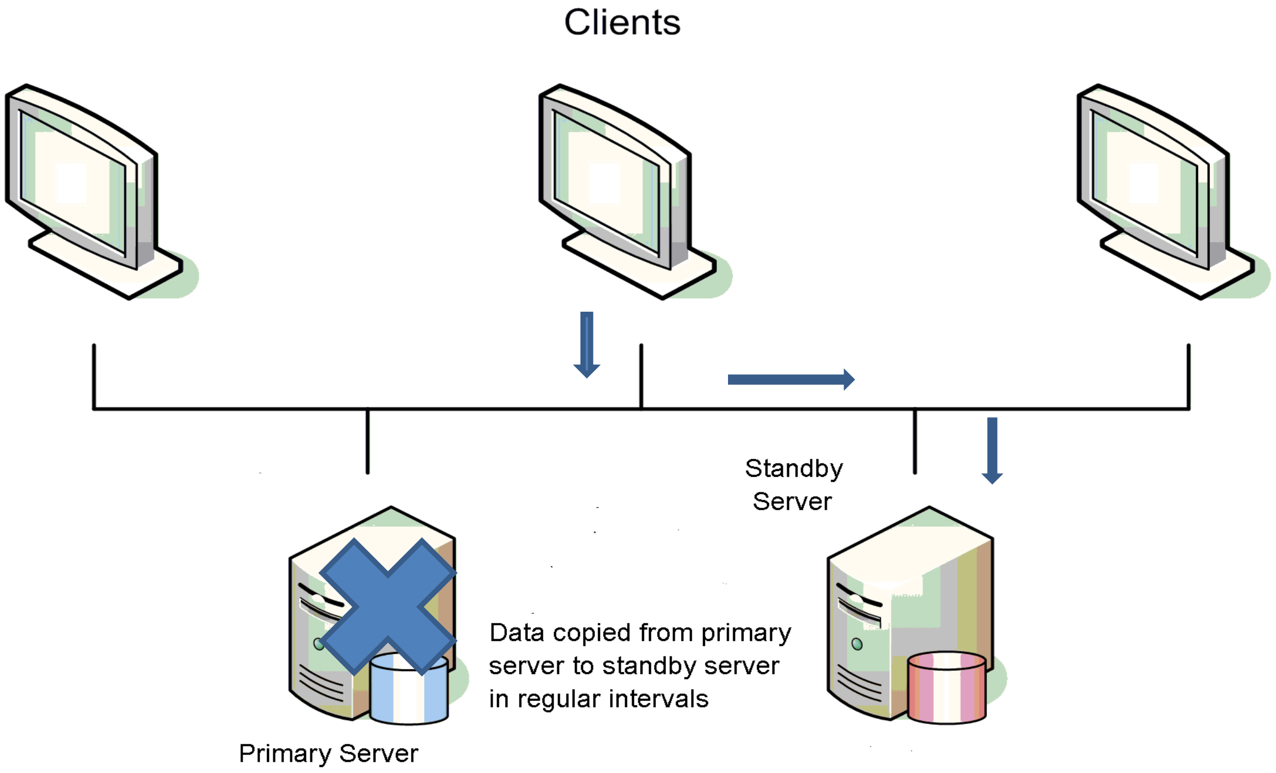
Warm Standby

- Database: The change log of the primary machine is copied regularly to the standby machine. The standby machine also have a static copy of database in the beginning. In case of machine failure, the database in the standby machine will be brought back to the image just to the point before failure by executing the changes in the change log.
- Others: Real-time monitoring of the primary machine allows automatic failover to the backup machine (i.e. backup machine can take up automatically)

Warm Standby – Data copied from primary to standby machine regularly



Warm Standby – If primary server is down, standby server will take over.


Clients
Standby Server
Data copied from primary server to standby server in regular intervals
Primary Server

Hot Standby (High Availability)

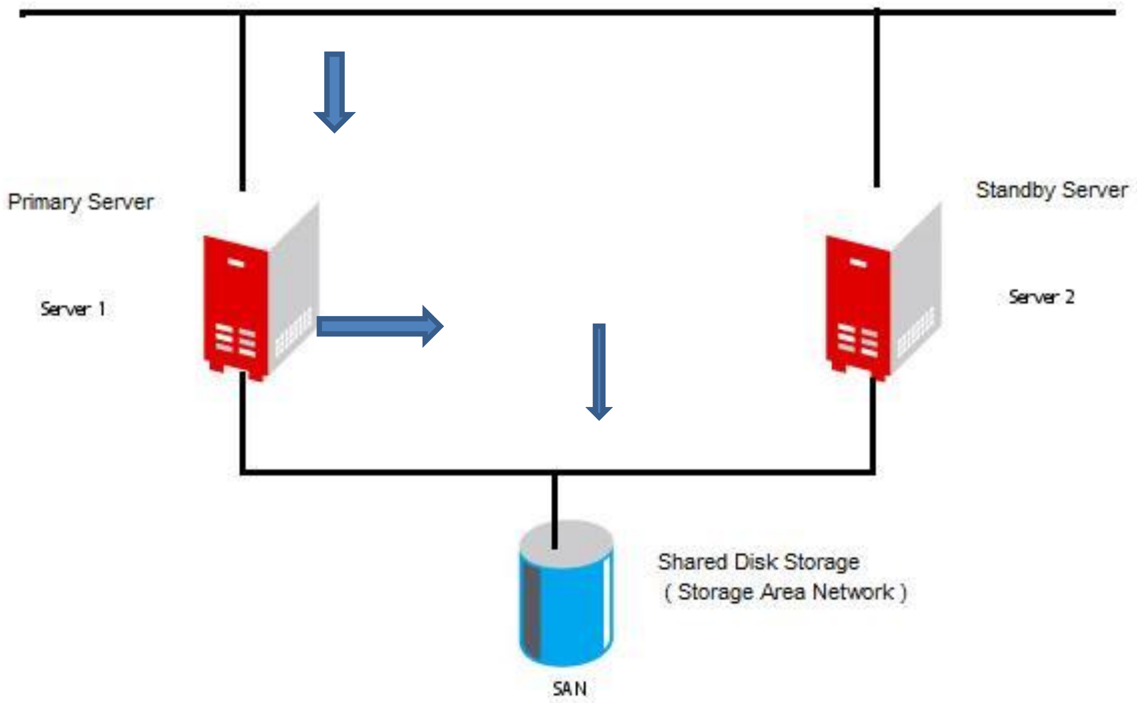
- high-availability configuration : two same-function components to form a cluster; in case one fails, another can continue to function as normal so that the system will not be down)
- Router, firewalls, file server, database servers can all have high availability configuration

High Availability as in airplane engine configuration

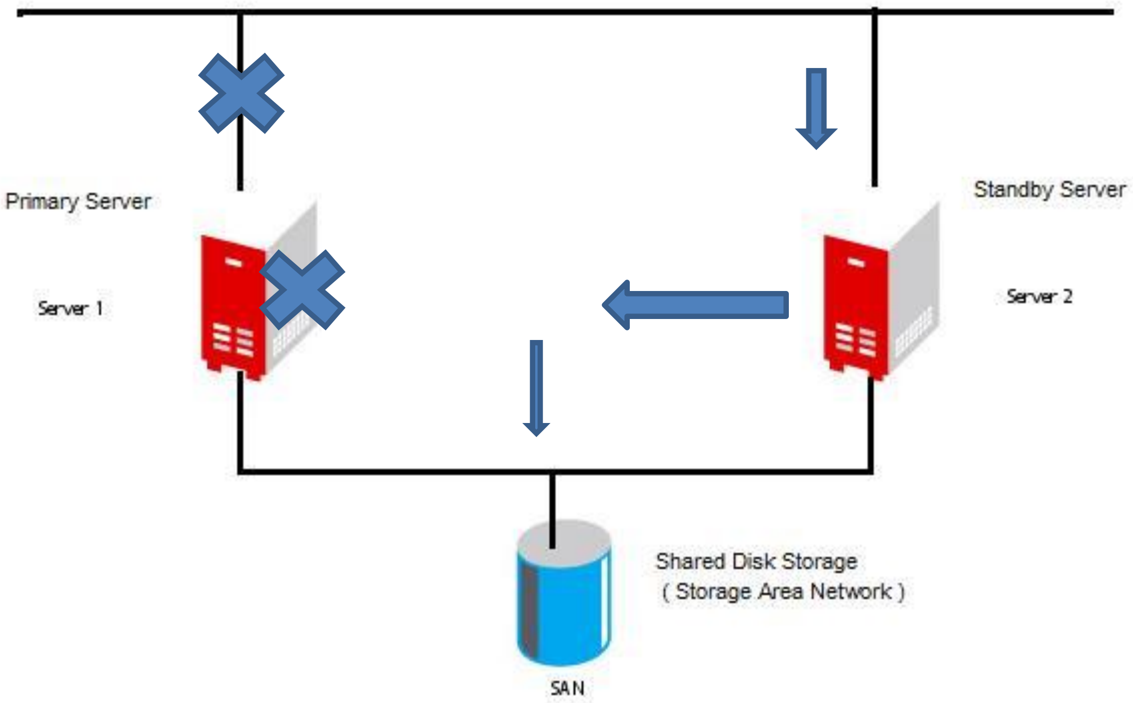
Hot Standby (Cont’d)

- Database: The database image is mirrored on real-time from primary to stand-by machine or both primary and stand-by machine share the same storage. In case of primary machine failure, the stand-by machine can become active and provide the services as if uninterrupted.
- Others: (Active-Active mode for router, firewall, web server) Both the primary and backup servers are actively running during normal operation. If the primary fails, the service is uninterrupted as the requests will be handled by backup machines as usual.

Normal operation – request is handled by primary server and data update is sent to the shared storage



When primary server (or network to primary server fails), standby server can take over and uses the same data storage



4.3 Component Failure Impact Analysis (CFIA)

Component Failure Impact Analysis

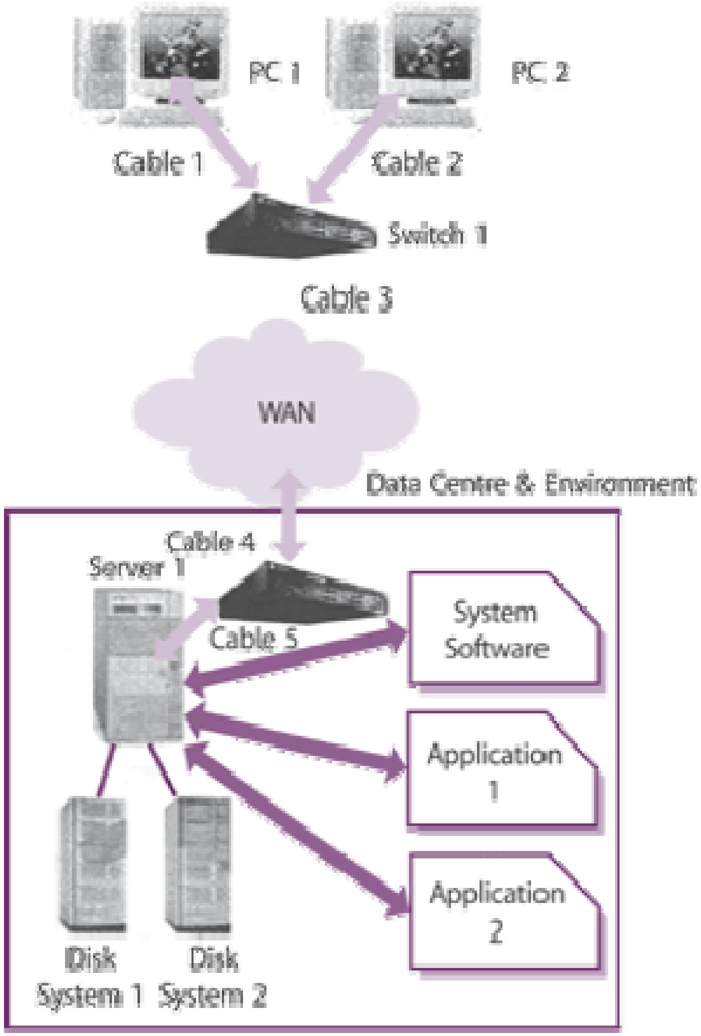
- Used to predict and evaluate the impact arising from component failures
- Devised by IBM in the early 1970s , originally on hardware design but can be used in wider context covering the full IT infrastructure

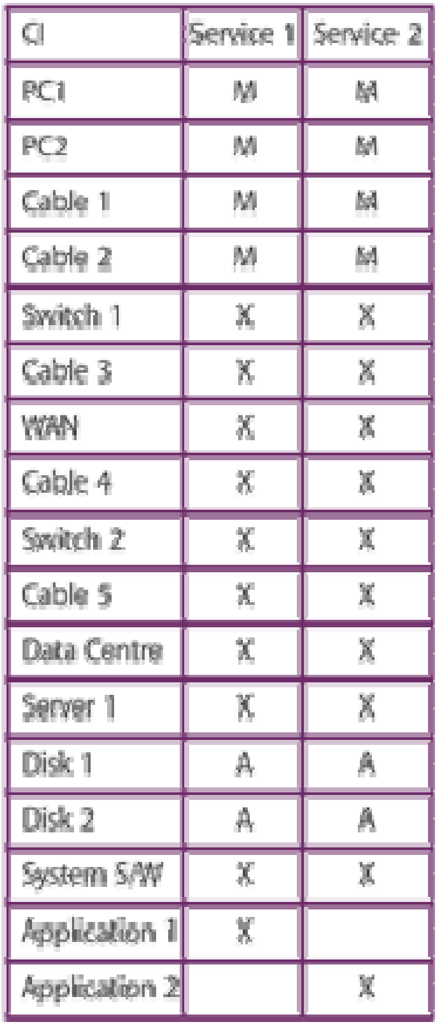
CFIA steps

- Create a grid with CI (Configuration Item) on one axis and the IT services on the other;
- Populate the grid as follows:
  - Leave a blank when the failure of the CI does not impact the service
  - 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 requires manual intervention to recover

CFIA (Cont’d)

- Having built the grid, CIs that have a large number of Xs are critical to many services and can result in high impact should the CI fail
- Equally, IT services have high count of Xs are complex and vulnerable to failure.





CI	Service 1	Service 2
PC1	M	M
PC2	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

Advanced CFIA

- The matrix can be expanded with additional fields such as
  - Component availability weighting
  - Probability of failure
  - Recovery time
  - Recovery procedures

Advanced CFIA – Calculating End-to-End Availability

Two levels of Availability Management

- Service Availability (including all aspects of service availability)
- Component Availability (individual components)

End-to-End Availability

- That is, from the user-end all the way to the end of the chain of the components providing the service.
- For example, an online securities trading system may involve a router, firewall, web server and database server. Either component fails will lead to unavailability of the service.
- End-to-end availability measures the availability of the whole service, comprising all the components

Availability for a network with multiple components

- Question:

  If a network has a router and a firewall, the router has availability 99% (i.e. 0.99) and the firewall has availability 98% (i.e. 0.98), what the overall availability for the whole network?

(Hints: if either component fails, the whole network fails.)

Availability for a network with multiple components

- Answer:

(a) (0.99 + 0.98) / 2

(b) 0.99 x 0.98

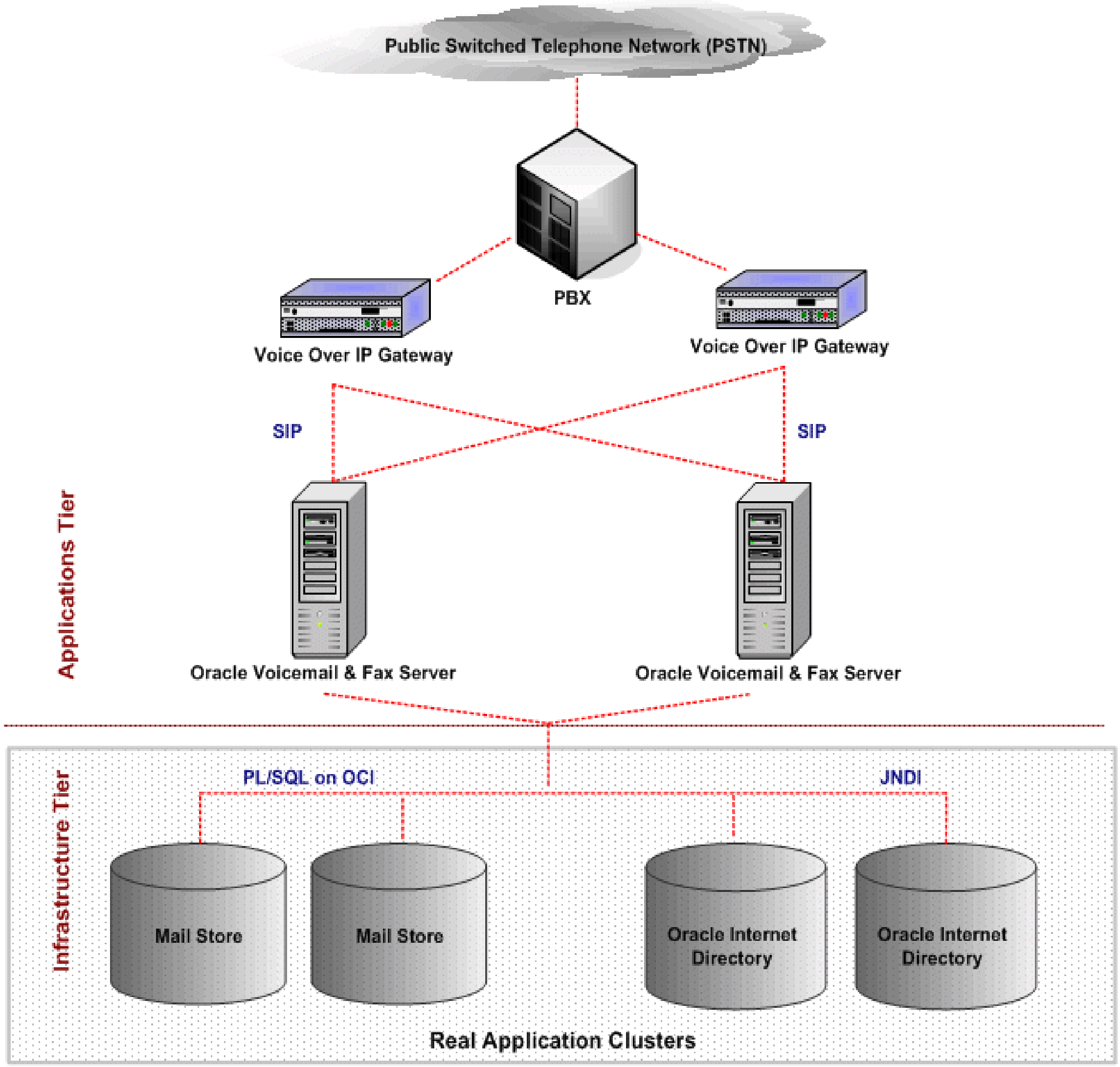
(c) 1 – (0.99 x 0.98)

(d) 1 – (1 – 0.99) x (1 – 0.98)

Question

- If you have a high availability (hot standby) configuration for a firewall and each firewall component has availability of 99% (0.99), what is the overall availability for this HA firewall configuration?
- (Hints : This is different from last question. The firewall will fail if both its firewall components fail - double failure. If just one of the firewall component fails, the firewall as a whole is still functioning)

High Availability Configuration Example


Public Switched Telephone Network (PSTN)
PBX
Voice Over IP Gateway
Voice Over IP Gateway
SIP
SIP
Applications Tier
Oracle Voicemail & Fax Server
Oracle Voicemail & Fax Server
Infrastructure Tier
PL/SQL on OCI
JNDI
Mail Store
Mail Store
Oracle Internet Directory
Oracle Internet Directory
Real Application Clusters

Answer

- (a) 0.99
- (b) 0.99 x 0.99
- (c) 1 – (0.99 x 0.99)
- (d) 1 – (1 – 0.99) x (1 – 0.99)

4.4 Availability Management – Benefits and Problems

Availability Manager

- Responsible for ensuring that the monitoring and reporting mechanism is in place
- Responsible for ensuring that escalation is made when services do not meet their agreed availability targets.
- If the targets are not met, will Availability Manager carry all the responsibilities? Not necessarily as the service down may have many reasons (e.g. inadequate capacity management, operational errors). However, availability management is primarily responsible to ensure that availability is measured and provide a point of focus for management.

Availability Management Benefits

- Single point of accountability via process owner – Someone is accountable for measuring and monitoring the system !!!

Benefits

The frequency & duration of IT service failures is reduced over time

Possible Problems

Lack of tools to underpin & support the process

Reference

- Davis, R. (2009), “What makes a good process”, BP Trend, November 2009
- itSMF (2007), An Introductory Overview of ITIL V.3, itSMF
- itSMF(2007), ITIL V3 – Service Design, itSMF