

Textbook for
Fundamental Information Technology Engineers

NO. 5

CURRENT IT TOPICS

9	2	0	0	1	
U	I	O	P		
H	J	K	L		
V	B	N	M		

Contents

1. Computer Systems

1.1 Multimedia systems	2
1.1.1 Multimedia contents development	2

2. Information Processing And Security

2.1 Application fields of computer systems	14
2.1.1 e-business overview	14
2.1.2 e-learning overview	17
2.1.3 e-government overview	21
2.2 Security	28
2.2.1 Internet security	28

3. System Development And Operations

3.1 Program design	38
3.1.1 Object oriented program design	38

4. Network Technology

4.1 Technology advances	48
4.1.1 Evolution in telecommunications	48
4.1.2 Wireless LAN	52
4.2 XML (Extensible Markup Language)	61
4.2.1 XML (Extensible Markup Language)	61
4.2.2 Unicode and Non roman text	66
4.2.3 Namespace	69
4.2.4 DTD (Document Type Definition)	70

4.2.5	Schema definition	73
4.2.6	XSL (Extensible Stylesheet Language)	77
4.2.7	XML Document Object Model	80
4.2.8	Using XML in the HTML document	83
4.2.9	SQL Server OLE-DB XML interface	86
4.2.10	Multimedia front end (Flash)	90

5. Database Technology

5.1	Management of large databases	93
------------	--------------------------------------	-----------

Index	101
--------------	------------

- Microsoft®, MS-DOS®, Microsoft® Windows® and Microsoft® Windows NT® are registered trademarks of Microsoft Corporation of the United States in the United States and other countries.
- The product names appearing in this textbook are trademarks or registered trademarks of the respective manufacturers.

Textbook for Fundamental Information Technology Engineers

No. 5 CURRENT IT TOPICS

First edition first printed August 1, 2002
Second edition first printed August 1, 2003

Japan Information-Technology Engineers Examination Center
INFORMATION-TECHNOLOGY PROMOTION AGENCY, JAPAN
Center Office 15F, Bunkyo Green Court 2-28-8, Hon-Komagome, Bunkyo-Ku,
Tokyo, 113-8663, JAPAN

©Ministry of Economy, Trade and Industry 2001 – 2006
Authorized translation of the Japanese edition ©2001 Computer Age Co., Ltd. / 2006 Infotech Serve Inc.
This translation is published under license of Infotech Serve Inc., Tokyo, Japan

1

Computer Systems

Chapter Objectives

Understanding the principles of multimedia design. This shows the steps required to develop a multimedia system.

- ① Understanding what comprises a multimedia system
- ② Understanding the type of project roles found in a multimedia development team
- ③ Understanding steps in multimedia system development

Introduction

The increased computing power of the personal computer coupled with the availability of wide bandwidth means multimedia contents can be created easily in today's machines. End users expect the ability to access the combination of audio and video contents. In addition, the changes in infrastructure like e-learning means that the multimedia is used to deliver such contents.

1.1 Multimedia systems

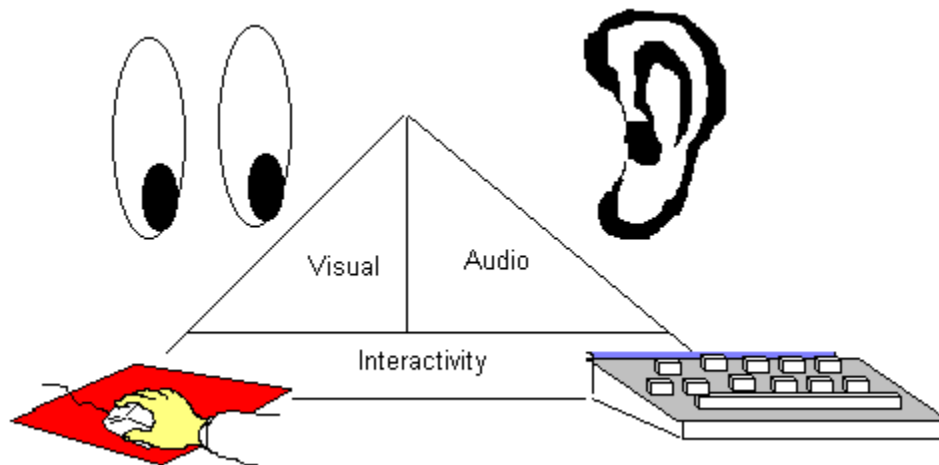
1.1.1 Multimedia contents development

(1) Components of multimedia

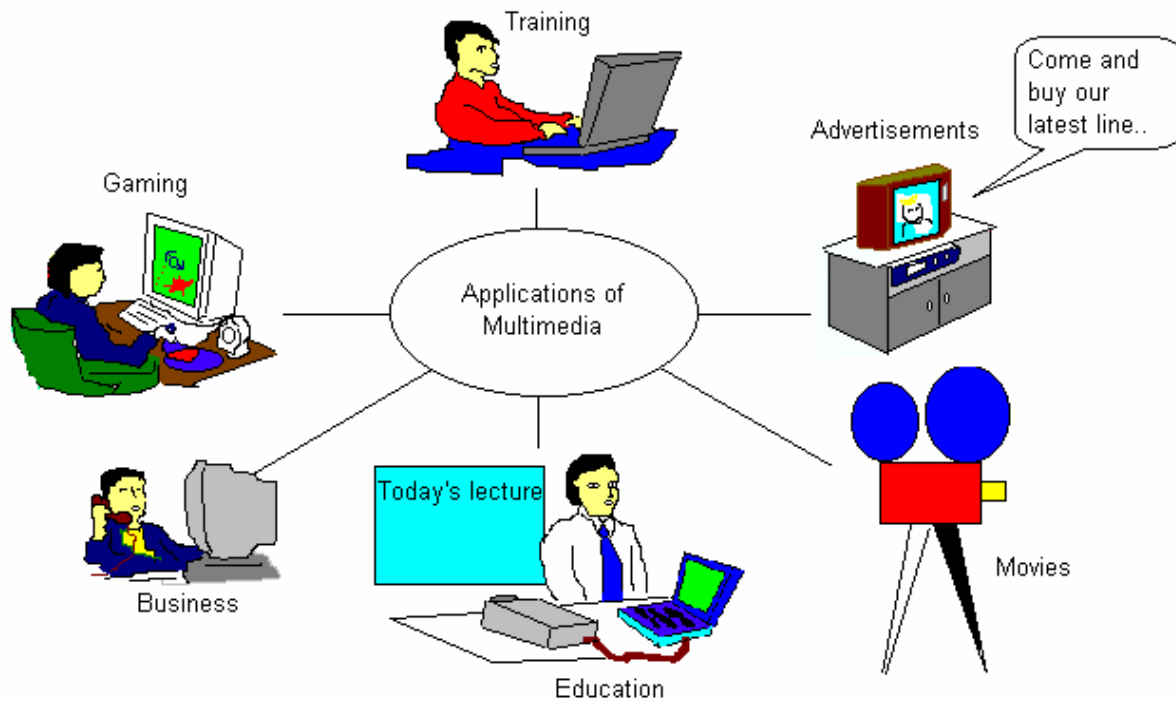
Multimedia comprises a combination of

- ① Audio
- ② Animation
- ③ Video
- ④ Static pictures
- ⑤ Text

Multimedia uses a combination of the visual, audio and tactile senses to deliver the message across.



(2) Types of multimedia applications



① Education

The benefits of using multimedia as a training tool are

- a. "Just in time" training can be implemented in multimedia.
- b. Courses can be taken any time and anywhere. Moreover, it allows the user to learn at their own pace.
- c. Interactive multimedia makes a good simulation tool.
- d. User interactions can be recorded, analyzed and reported.

② Training

The use of multimedia has expanded the dimension of training. The contents can be delivered over the web. The role of the trainer has been transformed to that of a content provider. Unlike a conventional classroom, the contents can be delivered anywhere and anytime. The combination of sight, sound and interactivity allows the learner to enhance their learning experience.

③ Entertainment

Movies, Net gaming and concerts can be streamed as multimedia contents. They are many examples of the use of multimedia effects in the movie industry. Character Graphics is also becoming more prominent with whole clips comprising of computer generated characters.

4 Chapter 1 Computer Systems

④ Advertisements

There is an increase in the use of multimedia in advertising. Computer characters have sometimes replaced the real world character in the advertisement.

⑤ Business

Multimedia allows the users to manipulate the products in 3 dimensional space. This is unlike a physical shop where it is possible to feel and touch the product. The use of multimedia allows the display of the products as a 3 dimensional figure that can be manipulated by the user.

In addition, the properties of the figure like color and pattern can be modified interactively.

Multimedia also helps in the visualization of the information.

(3) Project team roles

The following roles can be found in a multimedia project.

① Project manager

He or she is responsible for overseeing the project's timeline and resources.

② Graphic designer

He or she functions as a art director and production artist. This role is responsible for the look and feel of a project.

③ Programmer

This role is responsible for the mechanics of the project. Many multimedia packages support some form of scripting language. The use of the language allows the control of the flow of the package.

④ Script writer

This role includes the creative and technical writers. They are responsible for the development of characters and stories up to the technical text.

⑤ Sound designer

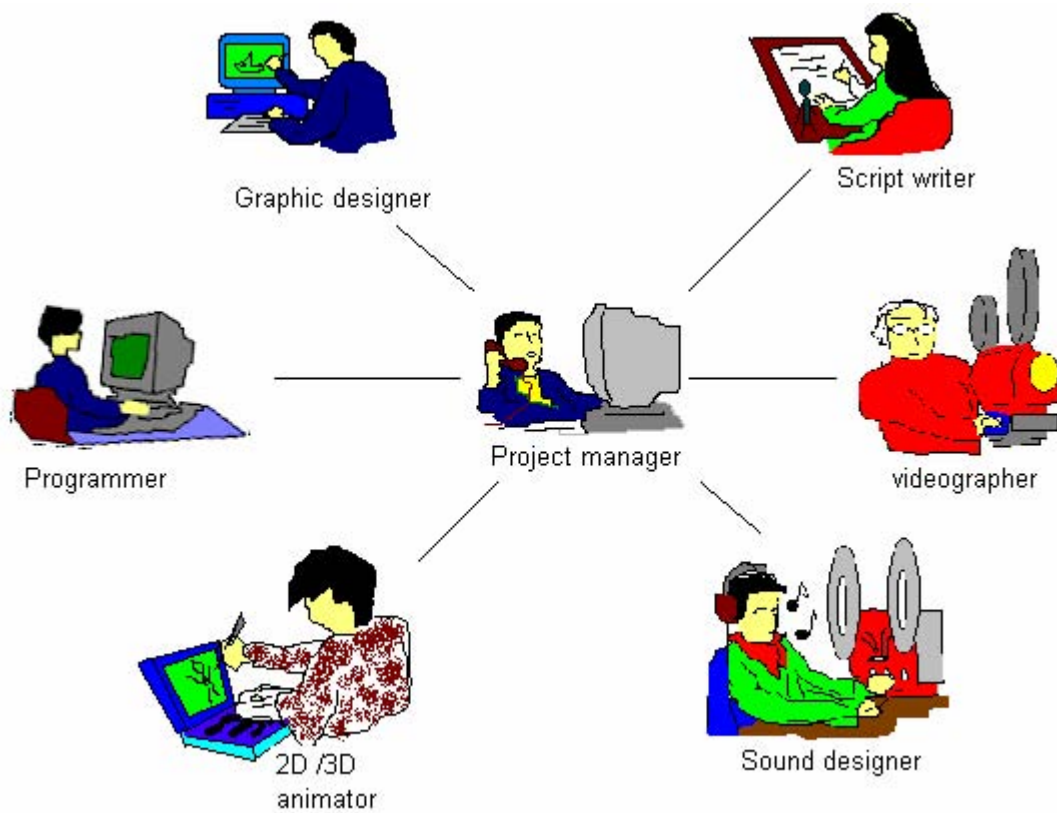
The use of well-executed voice-overs, original compositions, and vibrant sound effects designed expressly for a multimedia environment do wonders for a title. Royalty free packages can be purchased and used in the project.

⑥ 3D / 2D Animator

There is usually animation involved in multimedia. It may be as simple as the creation of rollover graphics or as complex as developing 3D visualisation animations

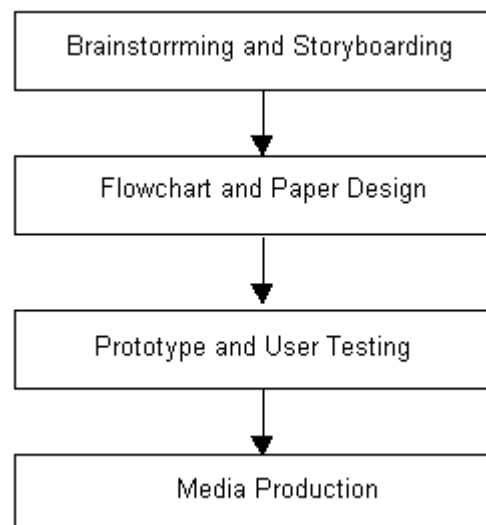
⑦ Videographer

This role is in charge of editing and creating the video output.



(4) Summary of the design steps

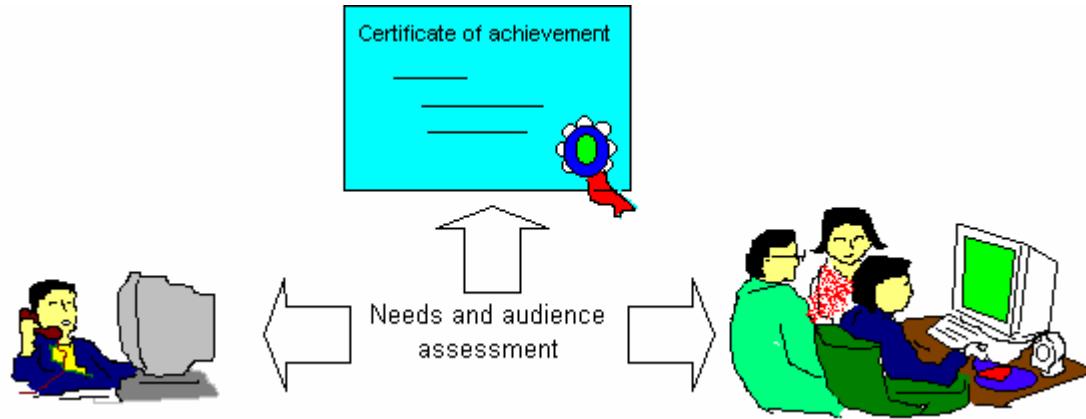
The steps for the design of a multimedia system is shown below



6 Chapter 1 Computer Systems

① Brainstorming and storyboarding

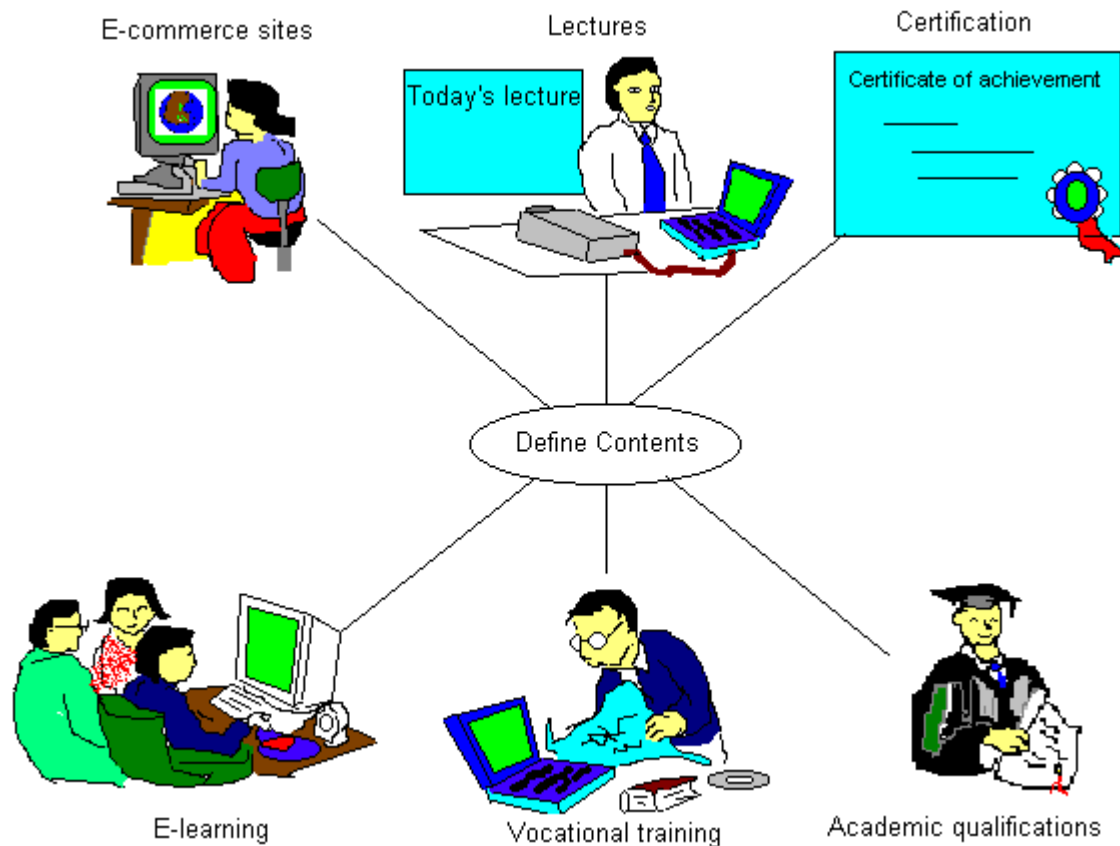
There is a need to determine Who the Target Audience is. Factors like age, sex, and nationality have a big impact on all aspects of the design from user interface to graphic design and structure.



Assess the purpose of the project. Interview the client and inquire what is the objective of the project. The points to focus on are the measures that the client sees as an indication of the attainment of the objectives.

The needs of the audience vary from one group to the next. Content for school going children need to be interesting enough to capture their attention and convert the learning experience into a fun filled experience.

Professionals will look for knowledge based content that allow skills to be learnt in as a short a time as possible.



Undergraduates focus on content that allows them to do their assignments easily and to practice their skills through mock examinations and exercises provided by the contents. Organizations will focus on improving their learning curve. The knowledge requirement to perform the various processes in the work flow is collected and analyzed. These knowledge requirements become the objectives.

a. Define contents

The nature of the contents is very important to the site. Clarification with the user to the objectives and the knowledge requirement to achieve the objectives are essential to allow the creation of the right content.

Operation based content can be recorded and played back.

The following are defined when creating contents

- Objectives
- Use the medium that best conveys the information
- Use the multimedia as a supportive and not just for decoration
- Synchronize the multimedia with the text or picture
- Make the interface interactive

Match the type of testing to the information covered.

8 Chapter 1 Computer Systems

b. Storyboard

Storyboards simply serve the purpose of roughly illustrating a concept along with a short write up detailing what each "screen" might look like. You do not have to be a brilliant artist to create a storyboard. Matchstick figures and crude drawings are fine. Initial storyboards for a title do not have to be anything fancy, or even accurate for that matter and in some cases you may choose to use PHOTO CD references or even scanned photographs / pictures to suggest the scene. Brainstorming and storyboarding are critical phases of a title's development



Narration: Welcome to this presentation

The relationship between individual screens and actions represented as a sequence of images is known as a Storyboard. It comprises a set of images showing the menus, dialogue boxes and windows. The navigation structure and functionality available in the system is shown by the story board. It is shown to the potential users, to allow the visualization of the composition. The scope of the interface can be seen clearly in the storyboard. The use of storyboards allows the design to be understood easily.

Each panel in the storyboard also shows what resources are in the frame. e.g. the narration, background music or movie to be shown. A preview is possible using a storyboard. It can be treated as a visual script for camera angles, lighting, sound effects, prop arrangement, and the narrative development and continuity.

The procedure for creating a storyboard is

- Give consideration to the scenarios of use which the storyboard will reflect. A storyboard may represent several activities such as entering, saving or printing information. Alternatively a separate storyboard may be created to represent each distinct theme.
- Construct the storyboard as a sequence of screen representations. using separate images to reflect changes in system appearance. Thus the storyboard indicates the availability and purpose of dialogue windows, menu items, toolbars and icons.
- The elements of a storyboard can be annotated with explanatory captions to aid audience understanding and evaluation.
- The completed storyboard can be shown to design teams as well as intended users to solicit evaluative feedback.

Several storyboards can be created and shown to an audience in order to explore different design options.

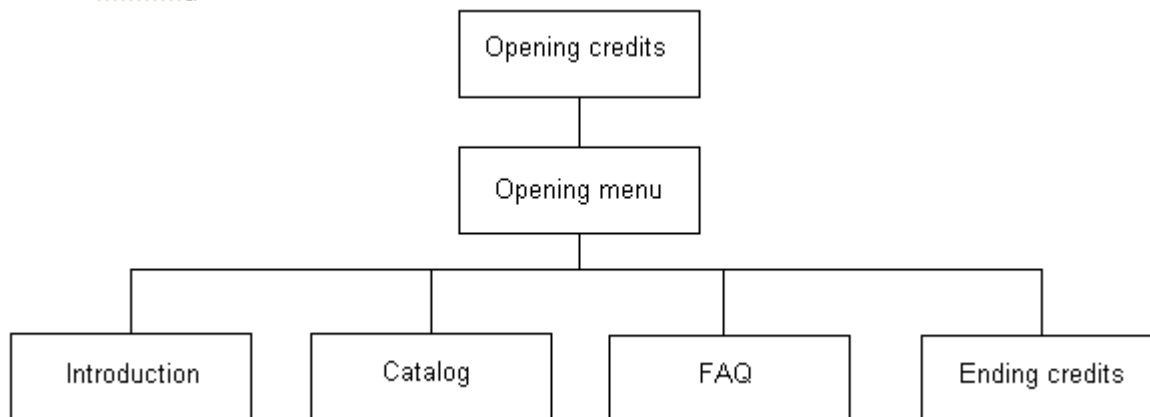
- It may be useful to video or audio record the feedback sessions for later review or to show to other colleagues.

② Flowchart and paper design

The concept of a series of "scenes" as in a movie is a good way to think about the organization of a multimedia title as you develop your storyboard. For instance, imagine a house as a multimedia title. There are many places within a house - bedrooms, bathrooms, a kitchen, etc. As the storyboard develops we must begin to consider how the places link to each other. To solve this part of the project we must start to consider how our information is to be organized and how it will be accessed from different parts of the project.

a. Creating the flowchart

- List all possible content categories.
- Start to group things by topic. Identify the obvious metacategories, which will become the top-level subjects that users see first. Then move topics from the master list to sublists.
- Refine the topic groups.
Move topics around on the sublists until they work.
- Arrange the groups into a linked, hierarchical structure. We are now starting to develop a "hierarchy" of information.



③ Prototype and User testing

The type of tools used for multimedia development can be classified as

a. Video authoring

These are used to edit video. There include tools like Premier or AfterEffects that allow

b. Multimedia authoring

These include tools like Macromedia Director or Flash

c. Landscape generators

These include tools like Bryce, Vue

d. Character modeling

These cover tools like Poser.

10 Chapter 1 Computer Systems

e. Audio editing

There include tools like SoundForge.

The various interactions are defined based on the storyboard development.

④ Media Production

a. Building blocks

The model upon which the building blocks of a multimedia application are based is referred to as an *activities model*. An activity can be defined as a series of related actions, happening one after another. Activities invariably commence at a single starting point, but they can have various outcomes.

The model for an application is designed by clustering groups of activities together using associative connections called links.

Broadly this model can consist of two types of activities: *presenting* activities and *controlling* activities. Multimedia applications of enormous detail and complexity can be designed from webs of these activities.

b. Presenting activity

The most basic activity is to display a single, non-interactive block of text, image or video sequence. A more complex activity could involve the display of a collection of individual images and video segments. A further refinement might involve the use of audio to cue the start of a video and then present a series of images each cued to certain sections of the audio track.

c. Controlling activity

The most basic variation of a one-after-the-other approach is the use of a selection structure based on menus. From an opening menu screen, the user chooses between a number of selections. On selecting the item a single activity is usually performed, after which the user is returned to the main menu. The menu activity in this respect is an example of a controlling activity.

Each item has a single starting point but has multiple end points. The nature of the input made by the user determines which of the multiple end points will be chosen.

Using this simple activity model, it is possible to identify various types of multimedia application structures: linear presentation, data driven engine, hierarchical menu, information retrieval and *hypermedia*.

d. Linear presentations

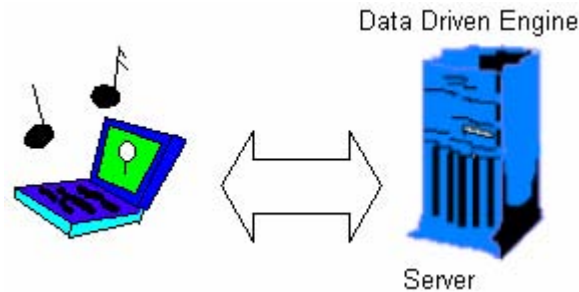
A *linear presentation* has one event after another but, for such a presentation to be of benefit to the user, there needs to be some element of control, with the presentation alternating between presenting and controlling activities. The controlling activities fix the duration of the presentation.

For example, in the event of a live paper presentation at a conference, the controlling activities would be programmed to accept cues from the speaker during their presentation. The cues could either be keystrokes (any key) or the use of a mouse device to move between the different screens. This allows the presentation to be synchronized with the messages being expressed by the speaker.

A more sophisticated speaker-support presentation might allow some other forms of information manipulation at each control point, such as the ability to go backwards or forwards through a presentation, or the option to jump out of the presentation into another application.

e. Data-driven engine (DDE)

One strategy to make linear presentations easier to develop, and simpler to control, is to use a data-driven engine approach. This involves using an engine, which interprets instructions from a file that defines what is to occur at each stage of the presentation. The file (which is usually a plain text — ASCII — file) has the items to be shown arranged in the most appropriate order; thereafter each presenting activity is included to show whatever is described. This is followed by a single controlling activity that awaits user commands about what should happen next. An example of this could involve the presentation of a series of slides in which the sequence list would merely be a list of the titles of the slides to be shown.



The composition of a data-driven engine (DDE) remains the same irrespective of how many items are shown: the controlling file just gets extended. Should the designer wish to rearrange the sequence of items all that needs to be done is to edit the items in the file. By programming using an authoring or coding system it is possible for the presentation activity to read and load different script files and, depending on the level of sophistication of the coding, it may also be able to interpret more advanced instructions from the file. Moreover, different presentations can be created simply by making new links in the file.

f. Hierarchical menus

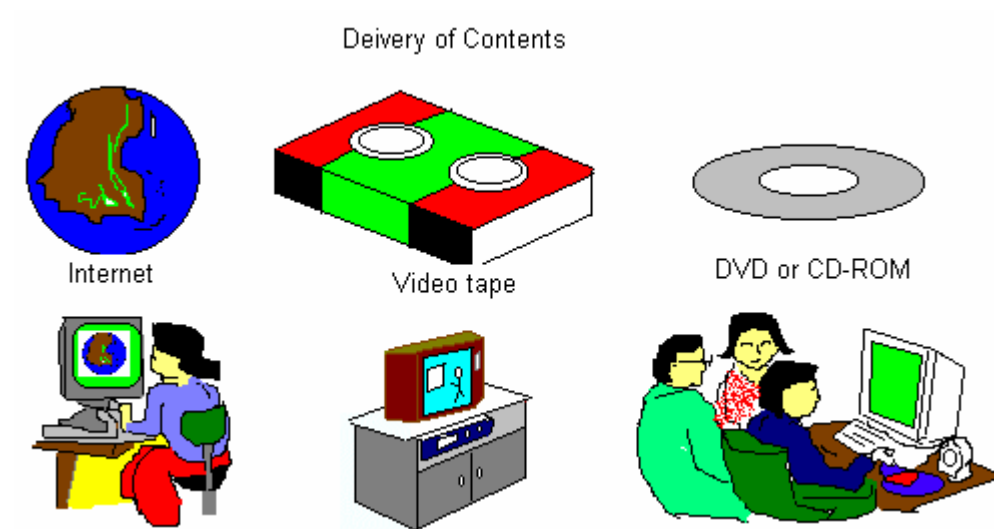
A menu is executed by using a presenting activity to reveal the menu, and then a controlling activity so that the user's selection is activated. This might take the user to another presenting activity or it might branch off to another level of the menu. When a menu selection is completed, the user is usually returned to the menu. Likewise, when the user exits a sub-menu, they return to the main menu. The designer needs to ensure that the user goes back to the main menu before closing the application.

Such a menu structure can quickly become difficult to use, especially if some of the activities (and menu items) appear in several menus. Using a data-engine approach may help to simplify the whole process by associating with each menu a script file that outlines the menu items and the titles of the presentations that will be executed when chosen. Code for a menu engine can be generated to display the menu from this file, which then sends the name of the selected item to a single-presentation engine program. This simplifies the menu structure down to two engine programs presentation and menu — and the detailed application structure is actually stored in a group of text files, which will be easy to modify and manage without conventional programming skills.

12 Chapter 1 Computer Systems

g. Delivery of contents

Finally, the actual production and the delivery of the media can be done.



2 Information Processing and Security

Chapter Objectives

Understanding the concepts of e-business, e-learning and e-government as well the security needed to protect the system:

- ① Understanding what comprises e-business, the standards involved and the type of e-business models
- ② Understanding the e-learning infra-structure , the standards used and the interfaces provided
- ③ Understanding what comprises e-government and the steps involved in achieving it
- ④ Understand the type of threats to security and the counter-measures needed to overcome them

Introduction

The ever increasing computing power and lower costs has led to an introduction of new areas of application. The expanding use of the internet has caused serious attention to be paid to the element of security. There is a need to examine the type of threats and the counter-measures to such threats.

In this chapter, we will cover the areas of e-business, e-learning and e-government. The type of threats and counter-measures used to combat them will also be discussed.

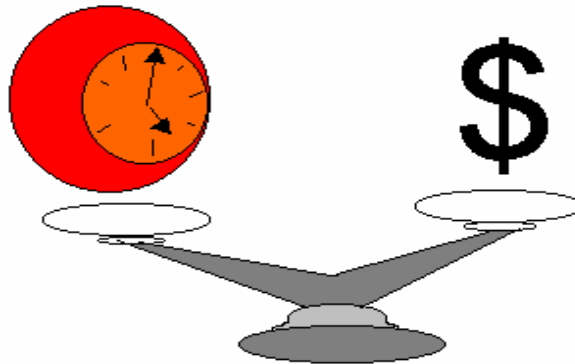
2.1 Application fields of computer systems

2.1.1 e-business overview

(1) Types of e-business

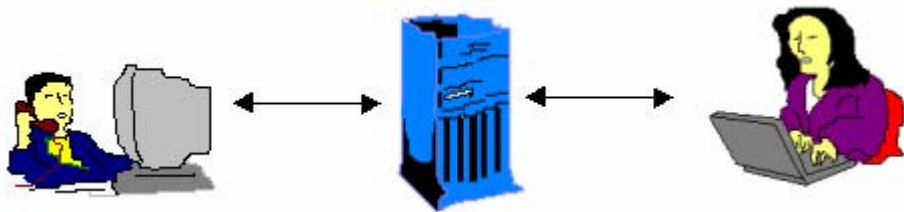
The use of e-business allows the saving of time and money.

The types of e-business can be classified as follows.



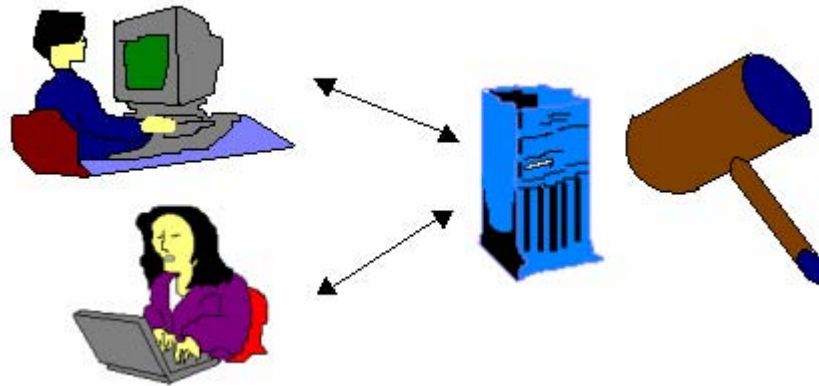
① B to B

These are targeted mostly towards companies. Some models like auction sites can also be used by the public. The types of B2B models are explained.



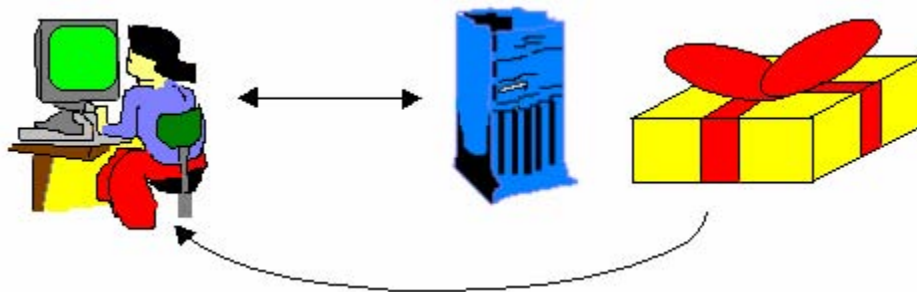
② Auctions

These sites allow users to bid for the product or service that they need.



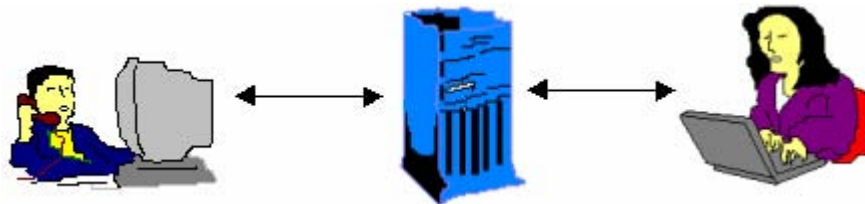
③ B to C(Business to Consumer)

The customers in a B2B model are corporations making volume purchases. In a B2C model, the consumer is targeted towards the public. It attempts to take the retail environment from a "bricks and mortar" environment and create an online presence. A combination of both is used is known as a "click and mortar" model where both a physical and an online store exists.



④ Business to government

These represent transactions like open bids or submission of relevant documents.



(2) Business processes

The following steps represent how companies conduct business

- ① Discover what products and services are being offered.
- ② Determine which shared process and information exchanges to use for obtaining these products or services.
- ③ Determine the contact points and establish a form of communication for the exchange of business documents.
- ④ Agree on the contractual terms, contact points and form of communication.

(3) Standards in e-business

ebXML is built on the following principles

- ① A robust, reliable, and secure communications infrastructure.
- ② A mechanism for modeling business information and processes.
- ③ A marketplace that all allows enterprises to find each other, agree to become trading partners and to conduct business with each other.

The ebXML is becoming a standard used in e-business transaction. This standard is sponsored by UN/CEFACT (United Nations Center For Trade Facilitation And Electronic Business) and OASIS (Organization for the Advancement of Structural Information Standards).

It allows a framework to be defined for global electronic business. The main features in ebXML are

- a. It allows businesses to find each other
- b. It allows the conduct of business
- c. It is based on XML

It was designed for a business to business interaction. It uses the UML (Unified Modeling language) notation.

(4) Communications Infrastructure

The communications infrastructure is provided through:

- ① A message transport mechanism with a standard, well defined interface, standard packaging rules, and predictable delivery and security.
- ② A "business service interface", i.e., a piece of software that sits at each business to handle Incoming and outgoing messages.
- ③ Integration to existing back-office applications. (This is currently out of the scope of ebXML, but it probably represents both the biggest challenge and the biggest opportunity for software providers.)

(5) Business Modeling

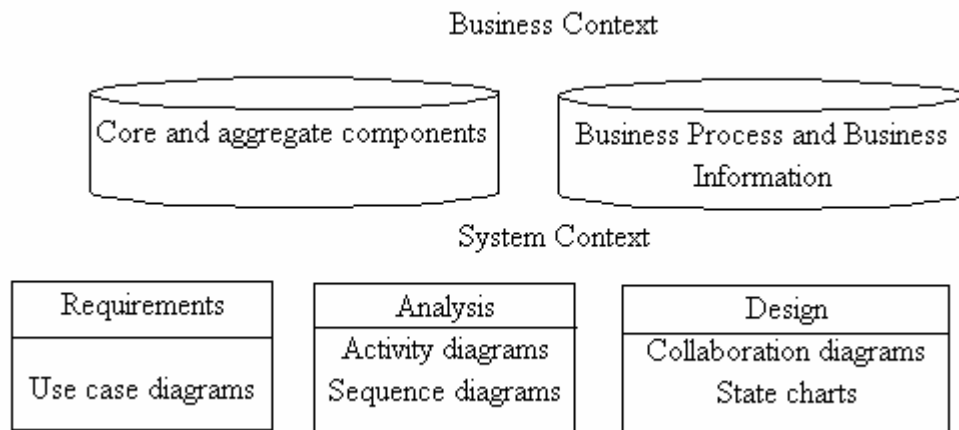
Business modeling is provided via the following elements of the ebXML standard:

- ① A specification for defining business processes and information models.
- ② A set of reusable business objects based on core components.
- ③ A process for defining message structures.

(6) Marketplace

The marketplace to allow enterprises to find each other and agree to become trading partners is provided by the following elements of the ebXML standard:

- ① A shared repository (registry) where enterprises can register and discover each other.
- ② A process for defining and agreeing to a formal Trading Partner Agreement (TPA).



2.1.2 e-learning overview

(1) Rationale for e-learning

① Personalized

The areas of study can be customized for a company, a department, or even an individual learner.

② Interactive

Simulations are possible and the ability to utilize support features like forums and online chat with the instructor and fellow students gives a broader and more interactive exposure to the student.

③ Just-in-time

Learning subjects of relevance when the students actually need it compare with traditional learning where they may get to apply the knowledge until much later.

④ Current

Content update is easy to do and therefore new materials and programs can be easily provided

18 Chapter 2 Information Processing And Security

⑤ User-centric

It focuses on the needs of the student, instead of on the abilities of the instructor.

(2) Basic definitions in e-learning

① Course

This is a unit of formal learning. It comprises the content represented by learning objects, a delivery mechanism and an assessment.

② Learning object

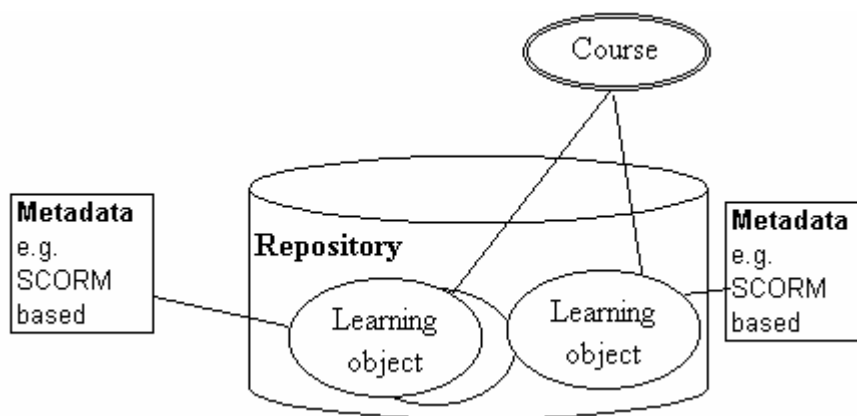
A standalone piece of learning. There may be mixed with other learning objects to form courses.

③ Repository

A database of learning objects

④ Metadata

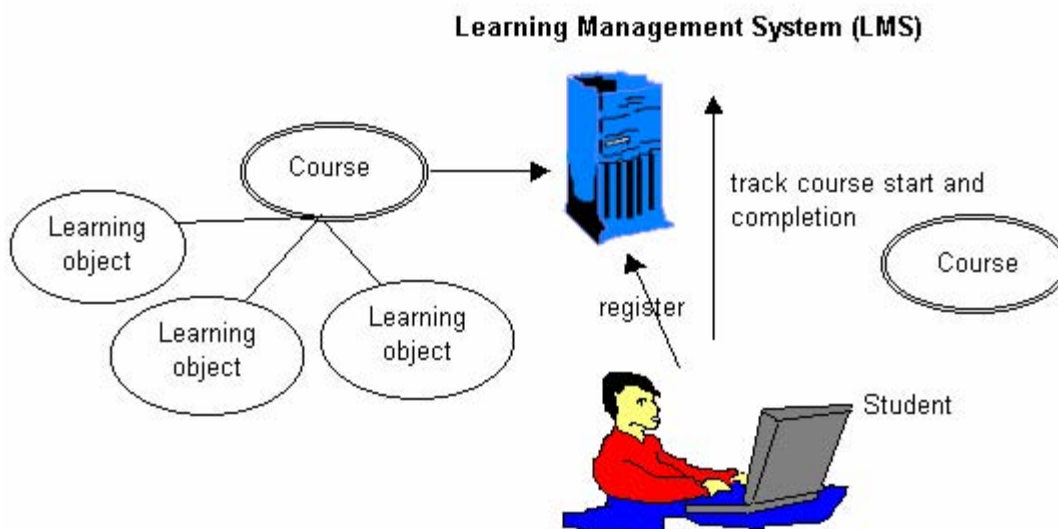
This contains data describing the structure of the learning objects.



(3) Types of management systems in e-learning

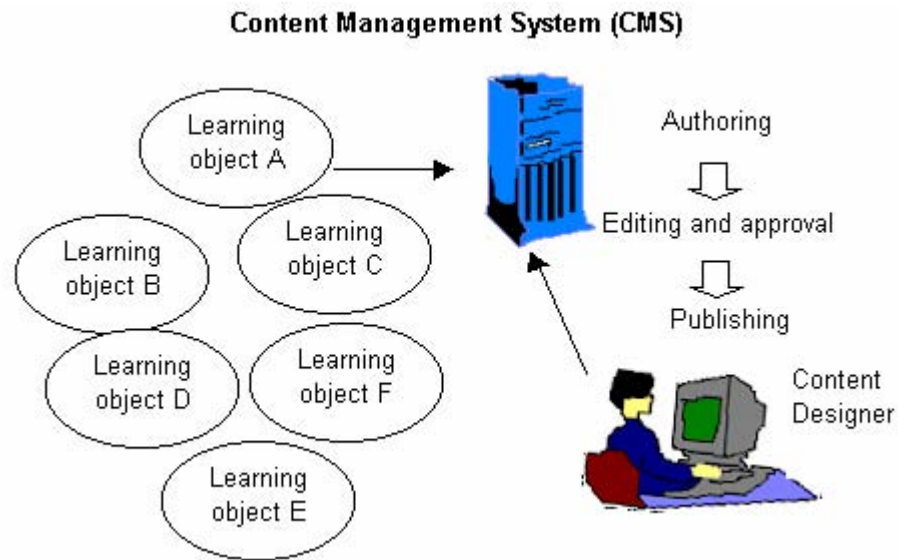
① Learning Management System (LMS)

The courses comprised pre-defined learning objects. The student is limited to what is offered. There is no customization possible.



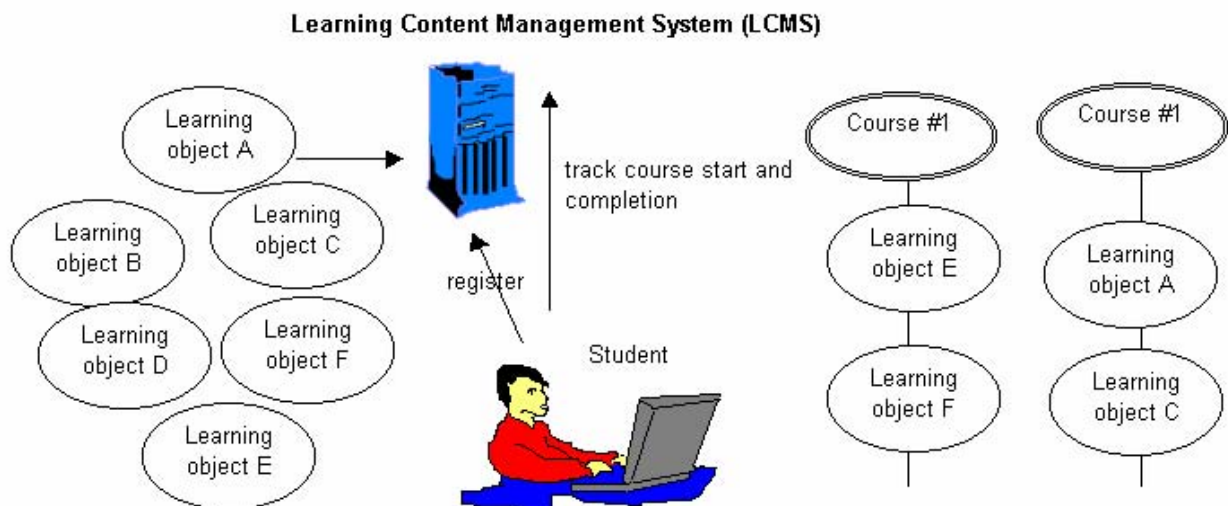
② Content Management System (CMS)

This allows the creation of the learning objects. A learning object may be as simple as a graphic or a video. The graphic designer and content developers will author and create the learning objects and store it in the CMS.



③ Learning Content Management System (LCMS)

The LCMS allows the assembly of learning objects into personalized courses for the student



(4) Standards in e-learning

The following represent organizations involved in standardizing the definition of learning standards.

- ① Advanced Distributed Learning (ADL) Initiative
 - a. Shareable Courseware Object Reference Model (SCORM):
This is a US government organization releasing the Shareable Courseware Object Reference Model (SCORM) standard. This is the most recently adopted set of standards.
 - b. IMS (Instructional Management System) Global Learning Consortium
The IMS Global Learning Consortium, has its headquarters in Burlington Massachusetts. It is developing and promoting open specifications for facilitating online distributed learning activities such as locating and using educational content, tracking learner progress, reporting learner performance, and exchanging student records between administrative systems. there are 2 major objectives:
 - Defining the technical standards for interoperability of applications and services in distributed learning.
 - Supporting the incorporation of IMS specifications into products and services worldwide. IMS promotes widespread adoption of specifications that will allow distributed learning environments and content from multiple authors to work together. IMS is a global consortium with members from educational, commercial, and government organizations. The funding comes from membership fees.

② AICC: The Aviation Industry CBT (Computer-Based Training) Committee

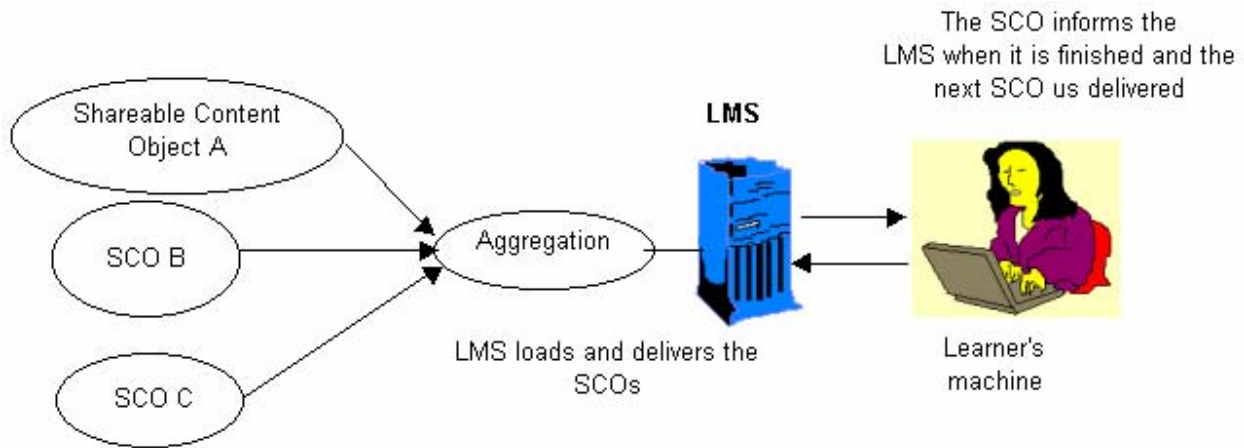
The Aviation Industry CBT Committee (AICC) is an international association of technology-based training professionals. Guidelines for the aviation industry in the development, delivery, and evaluation of CBT and related training technologies was developed by AICC. The main goals of AICC are:

- a. Development of guidelines that promote the economic and effective implementation of computer-based training (CBT) for airplane operators in.
- b. Develop guidelines to enable interoperability.
- c. Provide an open forum for the discussion of CBT and other training technologies.

Although AICC primarily attends to the aviation industry, over 12 years focus on the specifications required to meet this industry's needs has led to a very well Although it was originally targeted for the aviation industry, the specifications developed could also be used for computer managed instruction. As a result, a wide range of learning consortiums and accredited standards groups are in the process of adopting and adapting the AICC guidelines to their own industries. It was one of the earliest standards developed.

③ SCORM standard

a. Reusable Learning Objects



The content is broken down into chunks. . The requirements for each chunk are

1. Each chunk must be able to communicate with learning systems using a standardized method that does not depend on the system
2. What happens within a chunk is the chunk's business.
3. How a learner moves *between* chunks is controlled by the learning system.
4. Each chunk must have a description that enables designers to search for and find the right chunk for the right job. Such chunks are called **learning objects**. There is no standard for the size (or *granularity*) of a learning object. Larger learning objects are typically harder to reuse, and smaller learner objects save less work for those who are reuse them.

b. Content Aggregation

SCOs are self-contained units of learning. They can be used as building blocks (or legos) to create packages of SCOs, but they cannot be broken down into smaller units.. Three things must be done to create a larger unit of learning from SCOs.

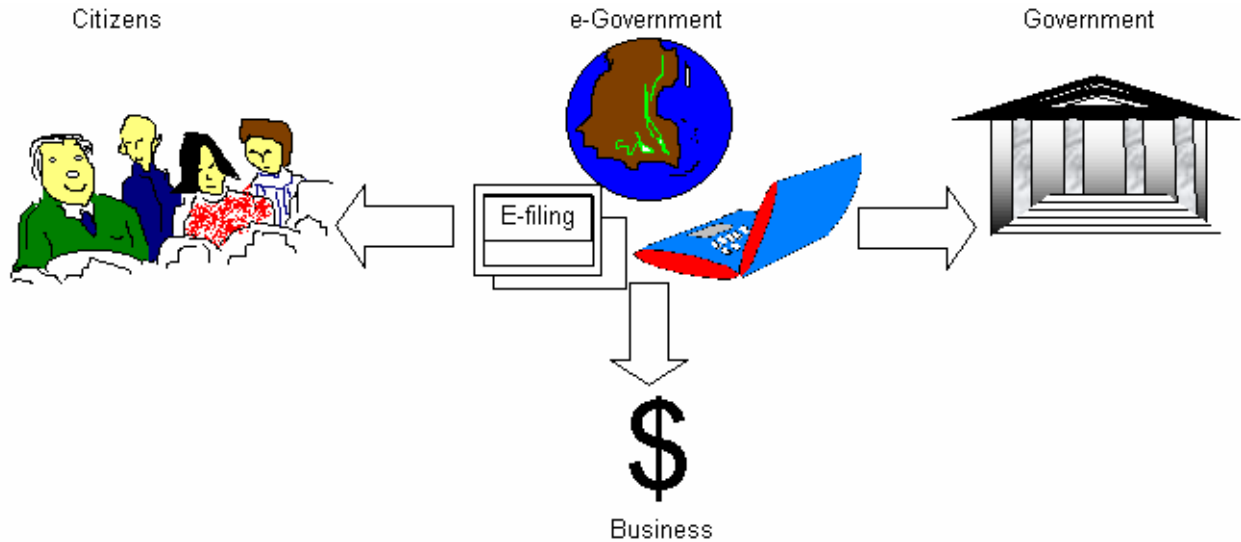
1. The SCOs must be found and organized into a structure.
2. Instructions must be written that tell an LMS which SCO comes after which.
3. The SCOs and instructions must be bundled into a portable package.

This process is called content aggregation. Note that content aggregation includes instructions for moving between SCOs but not for movement within individual SCOs. SCORM has adopted a content packaging format from the IMS Global Learning Consortium. A SCORM package contains a manifest file that declares the contents of the package and is set up to describe the order in which the SCOs are to be delivered. It also tells the LMS where to find the SCOs themselves. The physical resources represented by the SCO can be physically included in the package, or they can be referenced externally by the package.

2.1.3 e-government overview

E-government is the application of information technology to the processes of government. This results in greater transparency to citizens and businesses by allowing access to the information.

(1) Categories of e-government



① G to C (Government to Citizen)

This allows citizens to interact from their home. Besides acting as a source of information, transactions can also be conducted. Inquiries and searches can be done. Training can be provided online. e.g. training for driver's license. A diversity of services is offered online. One area where much interest is shown is e-voting.

② G to B (Government to Business)

This automates the interaction between government and businesses. 2 areas where services are provided are e-procurement and selling of government surplus.

③ e-procurement

In effect a reverse auction concept can be used. The government can put up their needs and bids can be done from the vendors. This allows for a more transparent view of the bidding process.

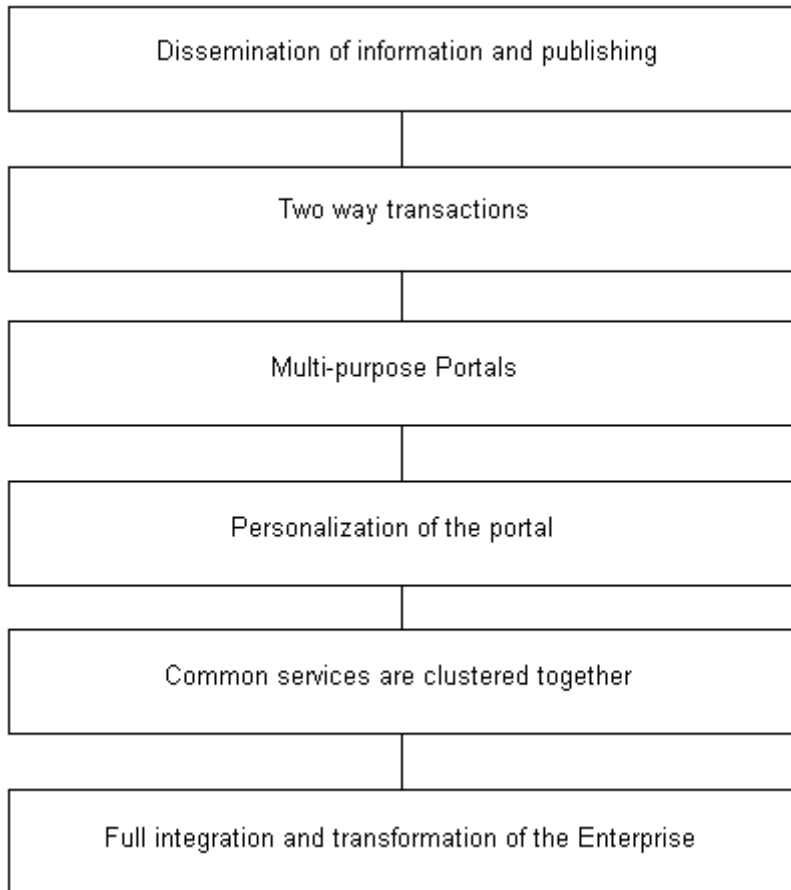
④ e-auction

This allows governments to utilize 3rd party auction sites or set up their own auction sites to get rid of surplus or seized materials or goods.

⑤ G to G (Government to Government)

Intra-government activities between agencies involving interactions between agencies.

(2) Stages in implementing e-government



① Dissemination of information and publishing

The individual agencies set up their own web sites providing information and a range of services and contacts. An electronic glossary can be constructed reducing the number of calls that the public has to make to interact with the relevant authorities.

② Two way transactions

The public can conduct monetary and secure transactions using a combination of digital signatures and secure sites. The public must be confident of the privacy of the exchange

③ Multi-purpose portals

A portal becomes a single point of entry that has links to the relevant departments and agencies.

④ Personalization of the portal

This allows the user to personalize their view of the site. They can choose the type of agencies and services that would appear when you connect to the portal. This becomes the equivalent of CRM (Customer Relationship Management) from government.

⑤ Common services are clustered

The public will be presented with a view of services without the need to draw the distinction between departments. Agency processes are streamlined with each other to allow for a smooth flow of processes

⑥ Full integration and transformation of the enterprise

This becomes a full service center with the removal of redundancy and merging of departments. New departments may rise.

(3) Benefits of e-government

It reduces the need for paper forms and manual data entry. It reduces the number of phone calls and face-to-face inquiries. It eliminates time consuming manual processes. Latest technology used to streamline operations. It allows agencies to give higher level of service in these times of reduced budgets.

(4) Privacy Issues

The communication revolution has also brought about limitless opportunity to share information. This has also increased concerns about privacy. Privacy issue sensitive when dealing with records compiled and maintained by government. While striving to better serve their citizens, agencies must be conscious of their responsibility as custodians of sensitive information.

(5) What Role does Government Play?

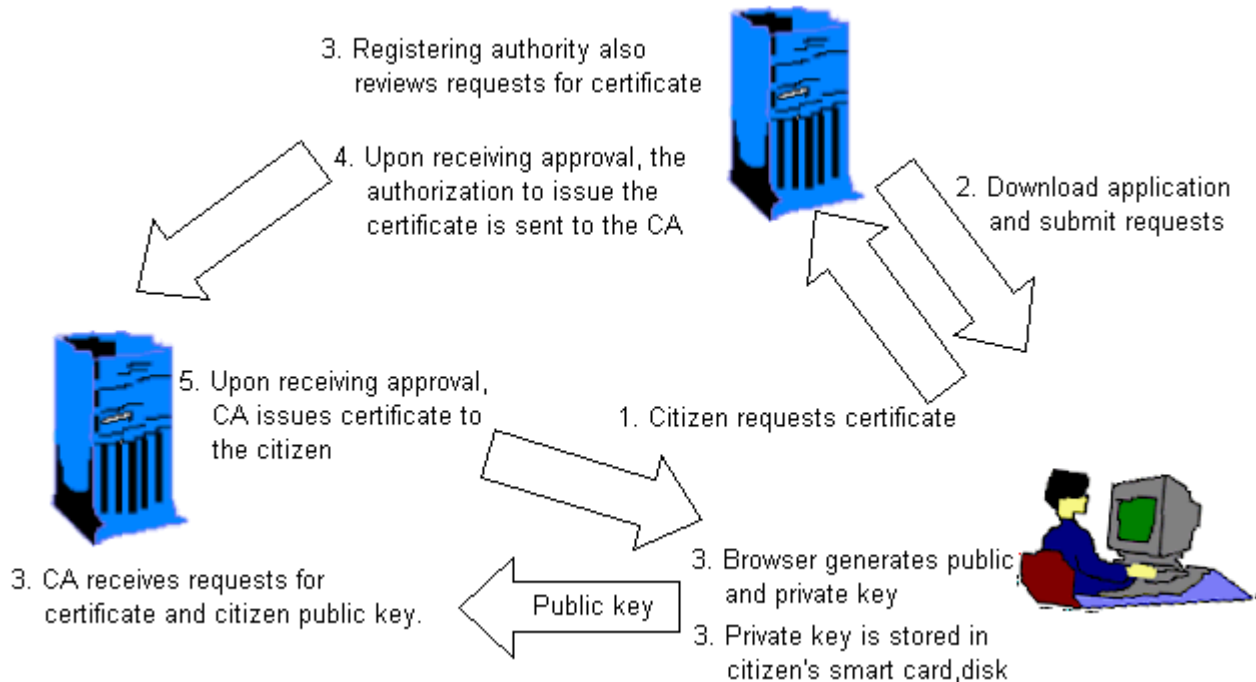
Government provides the infra-structure. Vast amounts of information is stored. Effectiveness depends on collection & use. Users will voluntarily interact with Government via the web only if they are confident of the security integrity. It is vital that the state implement proper security protocols and procedures. Caution & careful planning are the key.

(6) Security Issues

A commonly accepted security practice is to host information on a trusted 3rd party. The website and information are securely protected by firewall technology and site is updated by the agencies remotely. More advanced applications permit retrieval of information from agency database. Sometimes desirable to place sensitive information online but restrict access.

① Digital Certificates

An even more advanced and stringent security procedure calls for the use of digital certificates and signatures. Digital certificate is the electronic equivalent of a driver's license or passport. Digital certificate binds identity to a pair of electronic keys. Keys can be used to encrypt and sign digital information. When implemented, individuals are prompted for their digital certificate when accessing restricted areas. When used properly, digital certificates can provide the ultimate in credibility in allowing access to restricted sensitive agency information.



(7) Accessing Agency Information

The public wants increased functionality. By providing access via the Internet can decrease phone calls and counter traffic. A common application would provide access to an agency database. Web interfaces are built to allow people to conduct their queries.

Requested info is then displayed in requester's browser. This results in numerous savings. Updating of the Agency Information requires a high overhead as it requires a lot of paperwork. Keeping information current & accurate is an issue facing all agencies.

Individuals can now file information & make payments electronically through the web. Agency typing and redundant data entry are reduced or eliminated

(8) Development of standards

Although EDI has been existence for a long time, recent trends are going towards standards for cater for functionality beyond data exchange. These are XML based standards like ebXML. When transactions are to be conducted one such standard that comes from e-business is ebXML. The advantage of this standard is besides supporting transactions it also allows for the description of processes and the type of collaboration between them. It can be used as a standard to describe the infra structure of the government processes.

The ebXML standard comprises the following

Levels	Description
Business reference model	defines a frame of reference for the overall business environment in which the processes occur
Business process identification and discovery	offers an inventory of business process, much like a high-level use cases to note the parties and processes
Business process elaboration	Identifies the individual actors as well as conditions before and after the processes
Business collaboration definition	defines the economic events, which outline the system boundaries and protocols governing the flow of information in the process
Business transaction definition	covers the individual activities taking place in the process by authorized parties within the organizations; while it is important to know who is performing the actions and where they sit in the organization, most of the process definition covers the interactions between, not within, the organizations.
Business information definition	identifies the contents of messages, including such details as required field sizes, data types, descriptions, auditing or tracing requirements, and any additional business context for the reusable semantic components in ebXML

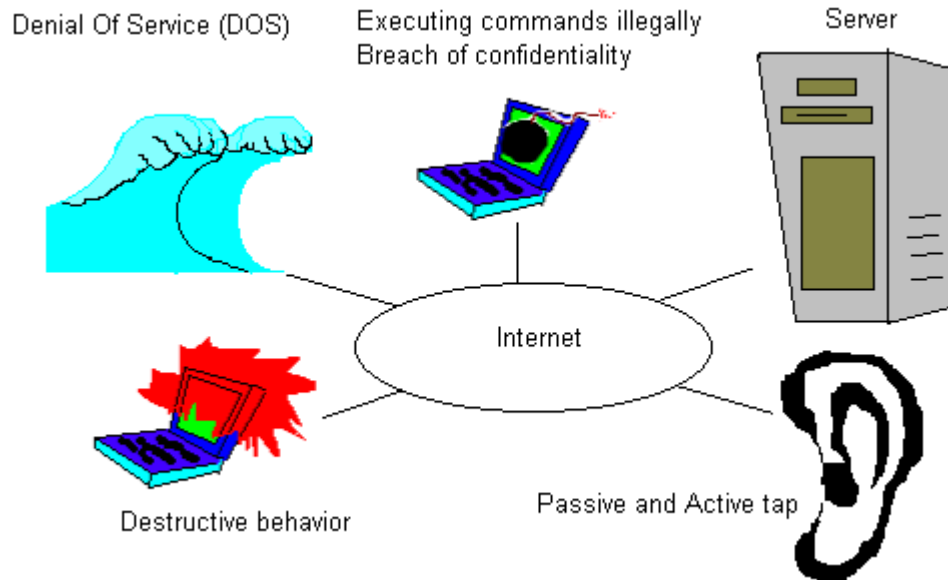
(9) Readiness for e-government

Criteria	Description
Connectivity	Are networks easy and affordable to access and to use. Availability of wireline and wireless communication services, community access centers (free and paid), and networked computers in businesses, schools, and homes. Affordability and reliability of network access, including the cost of service, downtime, and the prevalence of sharing access among individuals.
E-Leadership	<p>Is E-Readiness a national priority?</p> <p>Priority given by government to promoting the development of an e-society on a national level. Extent of demonstrated progress on e-government, including efforts to automate governmental processes.</p> <p>Quality of partnerships between industry leaders and government to improve E-Readiness.</p> <p>Level of effort to promote access for all citizens.</p> <p>Information Security – Can the processing and storage of networked information be trusted?</p> <p>Strength of legal protections and progress in protecting intellectual property rights, especially software. Extent of efforts to protect privacy.</p> <p>Strength and effectiveness of the legal framework to address and prosecute computer crimes, authorize digital signatures, and enable public key infrastructure</p>
Knowledge workers	<p>Are the right people available to support e-business and to build a knowledge-based society?</p> <p>Quality of and participation levels in the education system, with an emphasis on efforts to create and support a knowledge-based society.</p> <p>Culture of local creativity and information sharing within the society.</p> <p>Skills and efficiency of the workforce.</p>
E-Business Climate	<p>How easy is it to do e-business today?</p> <p>Existence of effective competition among communication and information services providers.</p> <p>Transparency and predictability of regulatory implementation, openness of government, rule of law, and general business risk (political stability, financial soundness).</p> <p>Openness to financial and personal participation by foreign investors in ICT businesses.</p> <p>Ability of the financial system to support electronic transactions.</p>

2.2 Security

2.2.1 Internet Security

(1) Types of threats



① Passive Tap

This is listening to the traffic on the network. However no new traffic is introduced.

② Active tap

Traffic is read and new traffic is introduced into the network. Spoofing means pretending to be legitimate service provider to trick the client into revealing the access information. Encryption is used to protect against the passive and active taps.

③ Attacks against IP

a. IP Spoofing

This is where one host claims to have the IP address of another. Since many systems (such as router access control lists) define which packets may and which packets may not pass based on the sender's IP address, this is a useful technique to an attacker: he can send packets to a host, perhaps causing it to take some sort of action.

b. IP Session Hijacking

IP Session Hijacking is an attack whereby a user's session is taken over, being in the control of the attacker. If the user was in the middle of email, the attacker is looking at the email, and then can execute any commands he wishes as the attacked user. The attacked user simply sees his session dropped, and may simply login again, perhaps not even noticing that the attacker is still logged in and doing things.

④ Denial Of Service (DOS) attacks

This means flooding the server with so much traffic that it is overwhelmed.
Actions that can be done to prevent DOS attacks are

- a. Not running your visible-to-the-world servers at a level too close to capacity
- b. Using packet filtering to prevent obviously forged packets from entering into your network address space.
- c. Keeping up-to-date on security-related patches for your hosts' operating systems.

(2) Unauthorized Access

This means access is made by people who are not given these rights.

① Executing Commands Illegally

We must prevent commands from being executed illegally on the system. There are 2 possible types of access.

- a. Normal user access
- b. Administrative access

Normal users should not be given access rights beyond what is needed. The normal users basically require to read the information. The administrative user right allows the attacker to change the configuration of the system.

② Breach of Confidentiality

This refers to the stealing of information from the server. In this case, sufficient access to read the information may be enough for the attacker to gain access to the data.

③ Destructive Behavior

There are two major types of destructive behavior

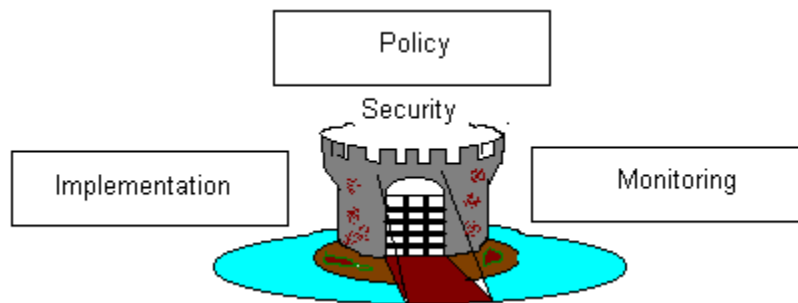
a. Data modification

These may change in the data. The changes may not surface till much later.

b. Destruction of data

These aim to corrupt and delete the data.

(3) Security policy



- ① Develop a strong security policy.
- ② Secure the network.
- ③ Monitor the network and respond to attacks.
- ④ Test existing security safeguards.
- ⑤ Manage and improve corporate security.

The following questions are posed when setting up a security policy

- a. What assets need protecting?
- b. What is the risk to those assets?
- c. What is the impact (in terms of reputation, revenues, profits, research) of a successful break-in?
- d. How much sensitive information is online? What is the impact if this information is damaged or stolen?
- e. Which users have access to those assets?
- f. What do users (and this includes business partners and/or customers) expect in the way of security control procedures and mechanisms?
- g. Are your users mostly accessing assets locally or remotely, or a mixture of both?
- h. Do you need different levels of security for different parts of the organization?
- i. What types of traffic exist on your network?
- j. Are the needs of security consistent with the business/operational needs of the organization?
- k. Is there a strong commitment from management to provide sufficient resources to implement security policies and technologies?
- l. Is there a strong commitment for security awareness training?

⑥ Default should be no access

In many systems, it has been the practice to allow anyone access to anything by default, but to allow system administrators to switch off or restrict access to those things they believe need it. This is the wrong approach; it is easily possible in such a system to forget to restrict access and the default should instead be that nobody has any access until they are explicitly given it.

⑦ Give least privilege possible

When someone is given access to some resource, they should be given the lowest level of privileges which they actually need to do the job they do. Note that privilege should be assessed on task-based need and not based on something like seniority.

A basic rule is that people will complain if they have too little access to something, but they are unlikely to complain if they have too much.

⑧ Check for current authority

For every access to a protected resource, you should check whether an individual has the appropriate privileges. It is not sufficient to rely on the fact that they had access at

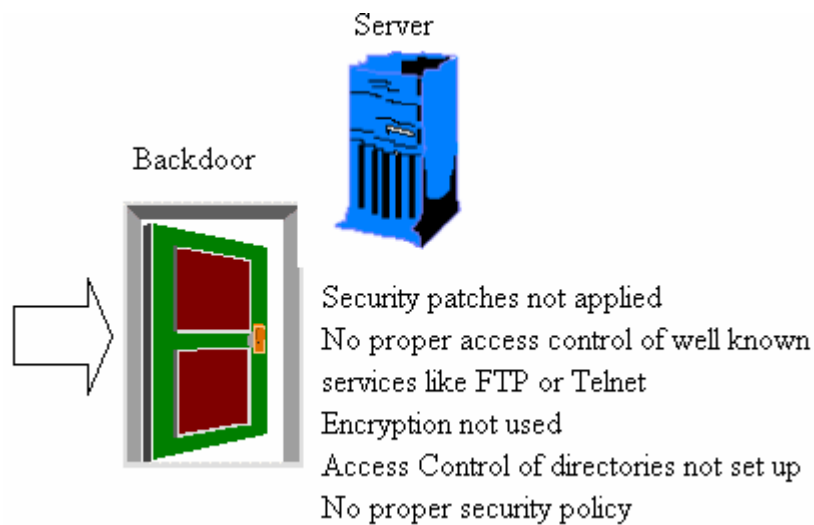
sometime in the past, since their privileges may have been revoked.

(4) Mechanisms for protection

The following can be introduced to enhance security

- ① Use of firewalls
- ② Encryption
- ③ Digital signatures
- ④ Application security
- ⑤ Audit of logs
- a. Loopholes

This is related to the characteristics of the setup. Some services that should be restricted in an exposed server



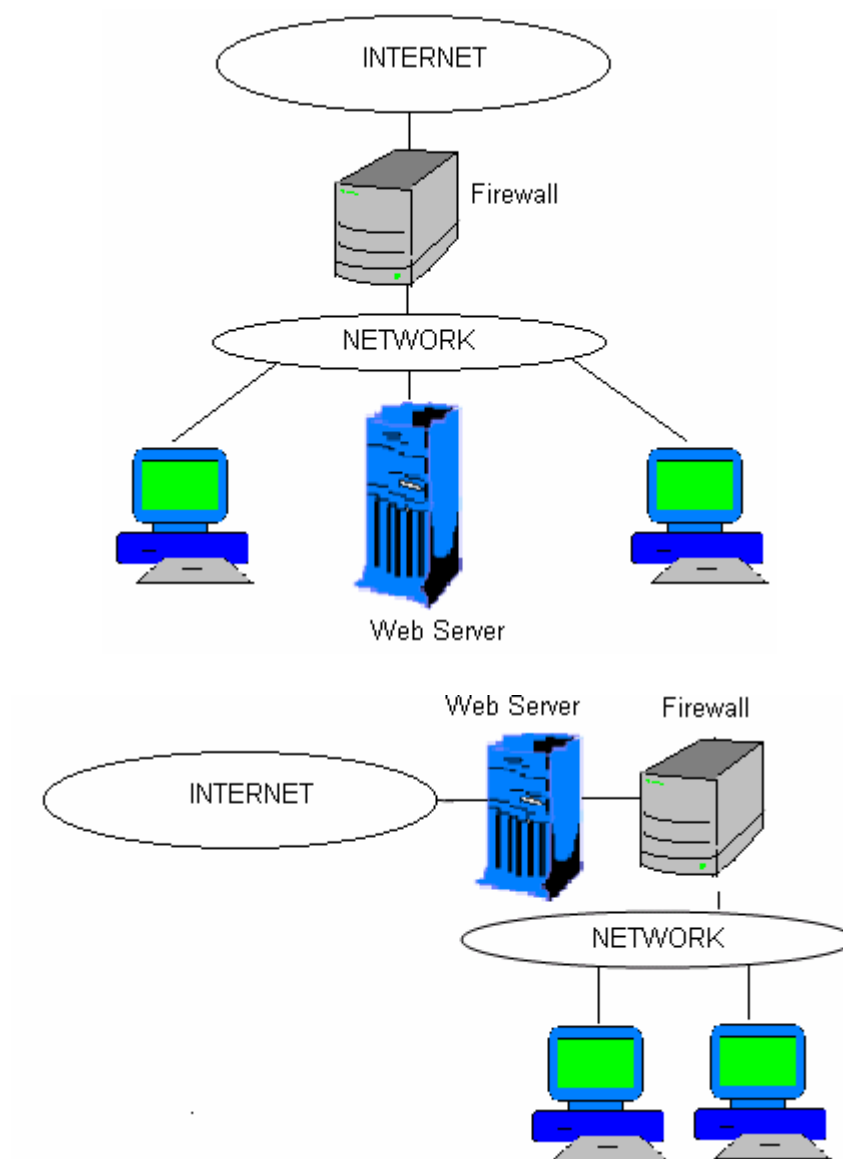
Services	Description
Netstat	This can reveal system configuration and usage patterns
FTP	Anonymous FTP should be executed on a different machine
Telnet	Telnet should be restricted only to administrators
ODBC setting or database location	The database server location should not be installed in the active server

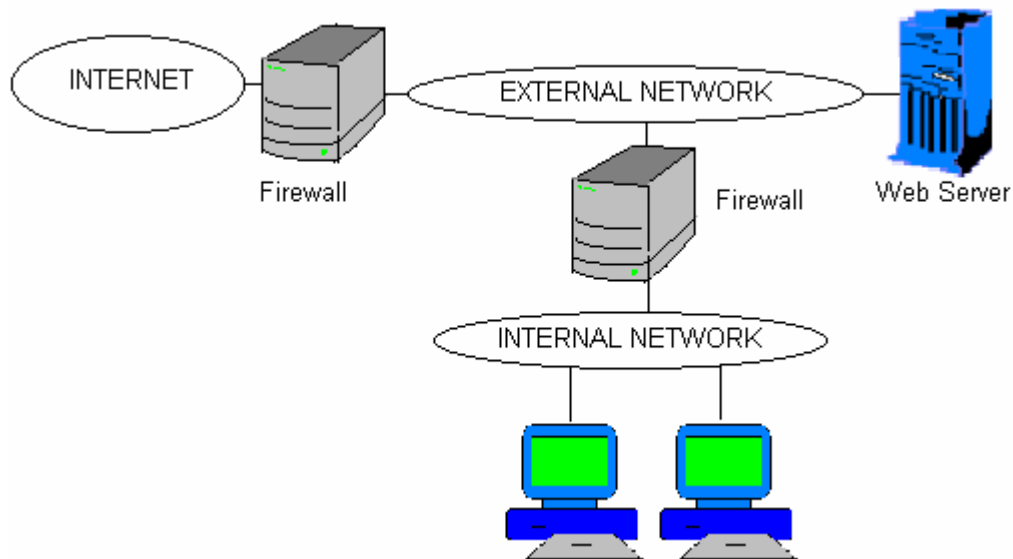
Service patches should be applied and kept current.

32 Chapter 2 Information Processing and Security

b. Firewall

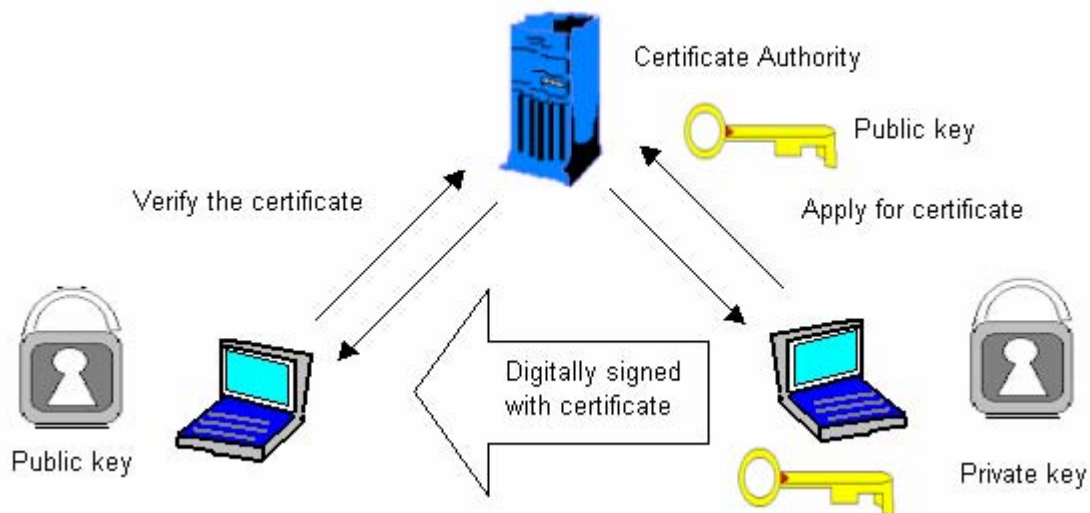
The web server can be secured by using a firewall. The firewall forms a barrier between the network and the internet.





c. Encryption

- Public key algorithms



A separate set of keys is used for encryption and decryption. The encryption key is known as the **public key**. The decryption key is known as the **private key** or secret key. This means the public key can be freely published. Using this public key, a message can be sent securely to the other party. Only the party holding the secret key can decrypt the message.

Public key algorithms are also used for creating digital signatures on the data. The secret key is used to create the digital signature and the public key used to verify it.

- Monitoring

The type of monitoring comes from the analysis of the log information collected by the server. In addition, software may also be introduced to perform some types of monitoring.

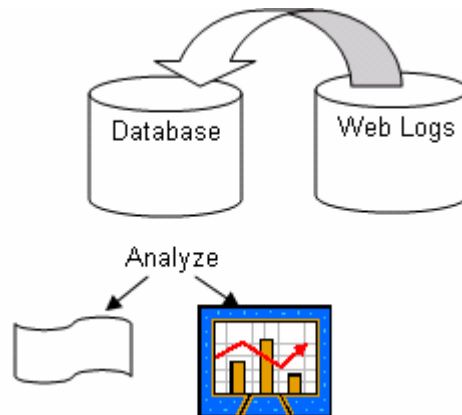
34 Chapter 2 Information Processing and Security

The following monitoring activities are done

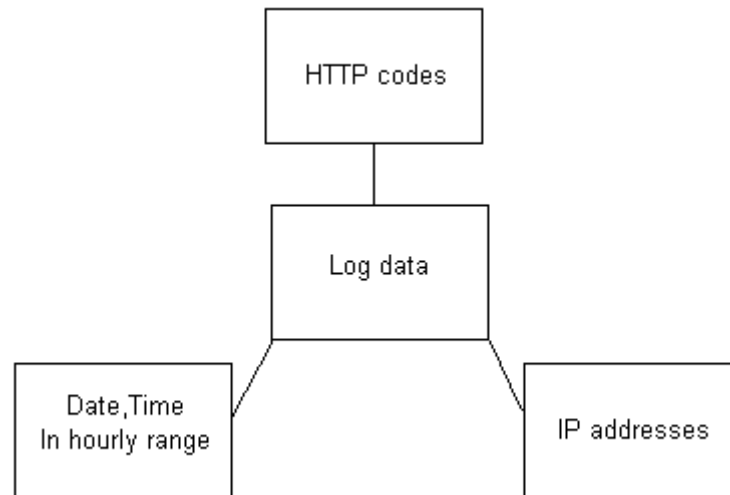
- Web log analysis
- Intrusion detection systems

d. Web log analysis

Various log information is collected for the server. Reports should be generated and analysis done to reveal possible attempts at intrusion. The logs can be stored in the database and analysis can be done.



Example of the criteria used to analyze the web log



The monitoring can be using the following HTTP codes

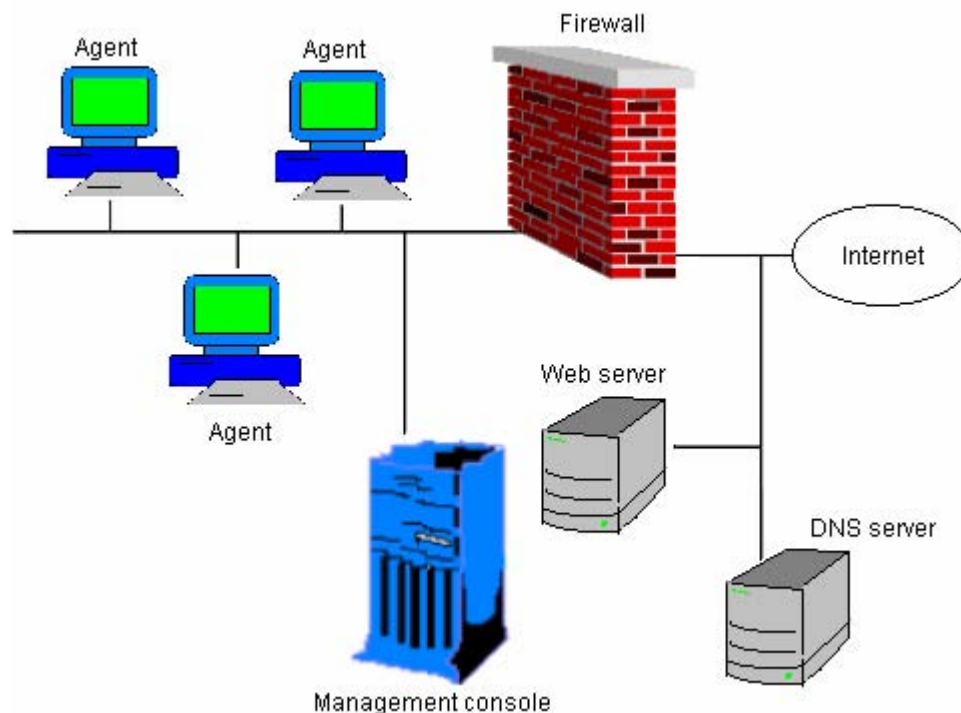
HTTP code	Description
Unauthorized 401	The parameter to this message gives a specification of authorization schemes which are acceptable. The client should retry the request with a suitable Authorization header.
Forbidden 403	The request is for something forbidden. Authorization will not help.
Proxy Authentication Required 407	This code is similar to 401 (Unauthorized), but indicates that the client must first authenticate itself with the proxy.
Method Not Allowed 405	The method specified in the Request-Line is not allowed for the resource identified by the Request-URI.
Service temporarily overloaded 502	The server cannot process the request due to a high load (whether HTTP servicing or other requests). The implication is that this is a temporary condition which maybe alleviated at other times.

e. Intrusion Detection System (IDS)

Host based systems are used to protect systems containing critical information. The agents are loaded at each host. These agents will analyze the audit trails.

There can be defined to track to the desired level of detail. e.g. exactly which user accessed which file. Since the agents have to be installed in each host requiring the tracking, there may not scale well. These agents will compare the activity against a set of rules and trigger an alarm if any of the rules are violated.

- Host based IDS

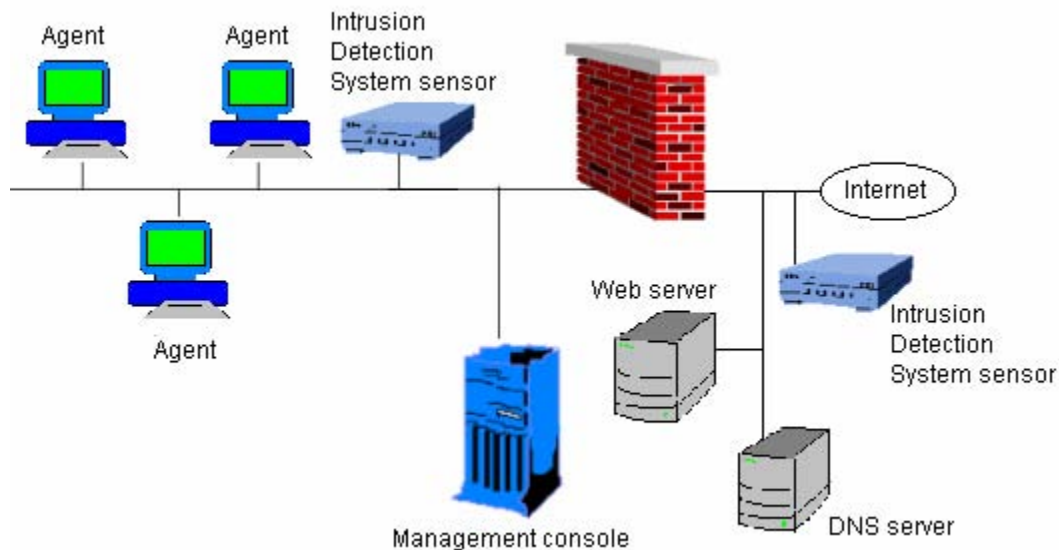


36 Chapter 2 Information Processing and Security

- Network based IDS

Network-based IDS monitor a specific part of the network. There are normally dedicated components comprising a sensor and a management system.

The management system is used to configure the detection and also for display of the alarm information. These sensors perform rules-based or expert system analysis of the traffic using configured parameters. The sensors analyze packet headers to determine



source and destination addresses and type of data being transmitted, and analyze the packet payload to discover information in the data being transmitted. Once the sensor detects misuse, it can perform various security-related actions: log the event, send an alarm to the management console, reset the data connection etc.

- Rules

A profile of the user activity is generated and the detection occurs when any activity falls outside of that range. A user's profile can contain attributes such as files and servers frequently accessed, time spent logged onto the network, location of network access etc. However, it may be difficult to implement as users may change their profile as their assignments require them to access other resources. This is known as Profile-based detection.

Signature-based detection can be compared to a virus checking programs. These signatures are provided by the vendors. The vendor will also provide updates to these signatures.

3 System Development And Operations

Chapter Objectives

Understanding the design of a program by using an object oriented approach.

- ① Understanding the definition of a class and relationships
- ② Understanding how to map the design to a form of implementation
- ③ Understanding the design of Web services by using an object oriented approach

Introduction

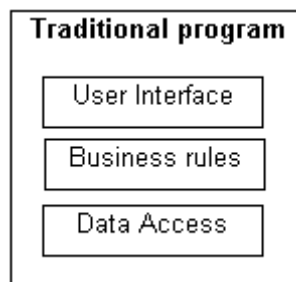
Many software development packages come with some form of objects. The ability to re-use the classes without the need to worry about body allows for loosely coupled systems to be developed easily. In this chapter, the design of a program using the object oriented methodology is explained.

3.1 Program design

3.1.1 Object oriented program design

(1) Organization of an Object Oriented program

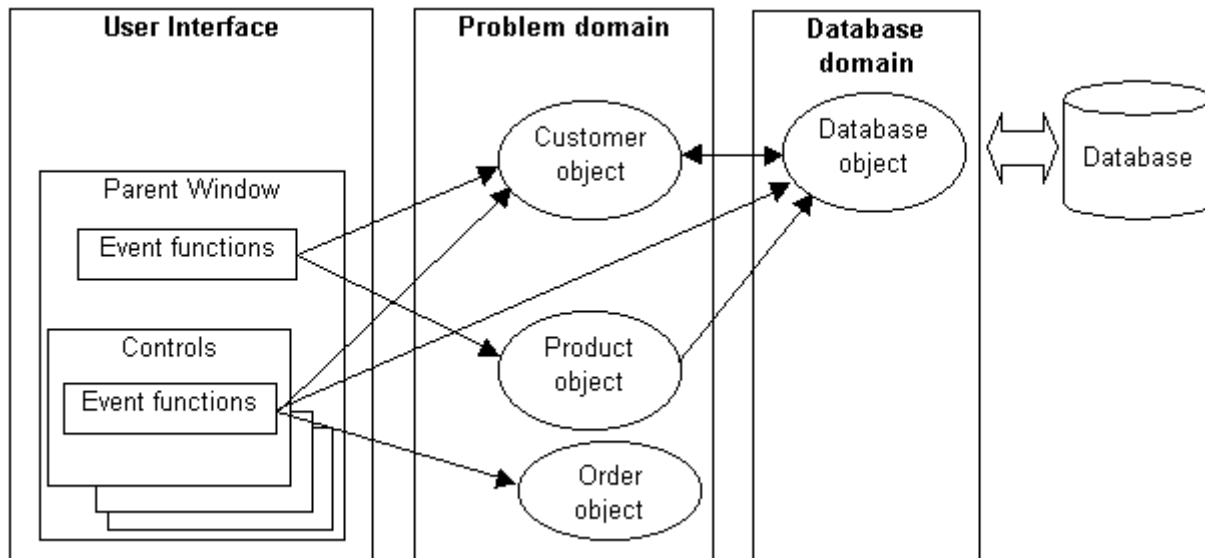
The access of the user interface, reading of the database and business rules are all combined together in a traditional program.



When an object oriented approach is adopted to program design, the individual components have to be separated. Most user interfaces utilize an event driven model. The actions carried out in the real world are translated into a function call. The appropriate event handler in the user interface is modified to process the action. This leads to the following types of classes

- ① User Interface
- ② Problem domain classes
- ③ Database domain class

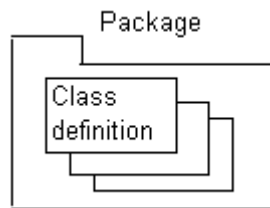
By dividing the program design into these 3 domains , it allows for easier understanding and creation.



This allows a loose coupling between the different parts of the program by breaking up the class definition into these 3 domains. The user interface is on the client to allow the user to interact with the system. The problem domain classes are used to encapsulate the business rules. The database class is used to isolate the location of the database making it transparent to the system. The location of the problem domain classes can be in the server or the client. Since each client uses a different set of values, it is common in practice to allocate the problem domain objects in the client.

(2) Classes and Objects

Objects are instances of classes. You can think of the class as the definition. Classes can be grouped together as a package



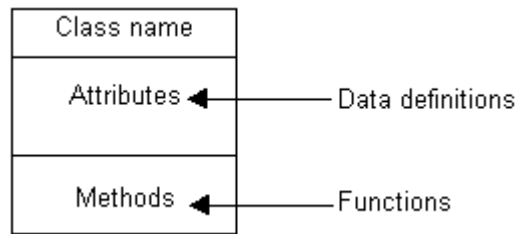
A class comprises 2 parts

① Attributes

This contains the data definition for the class.

② Methods

These become the functions in the class.

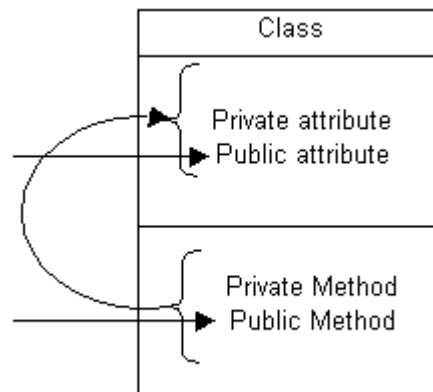


③ Regions of access

There are 3 types of access regions.

a. Public

Public access means the attributes and methods can be accessed externally.

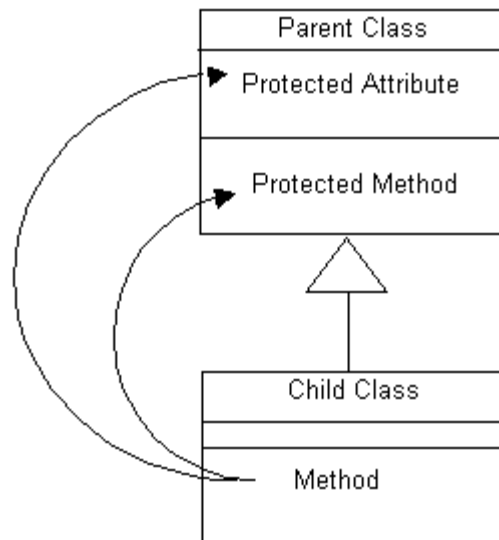


b. Private

This means the attributes or methods can only be accessed by methods within the object. This means even objects of the same class cannot access the other object's private methods. Since the data in each class is managed by the object itself, this means that attributes are usually private and methods are public.

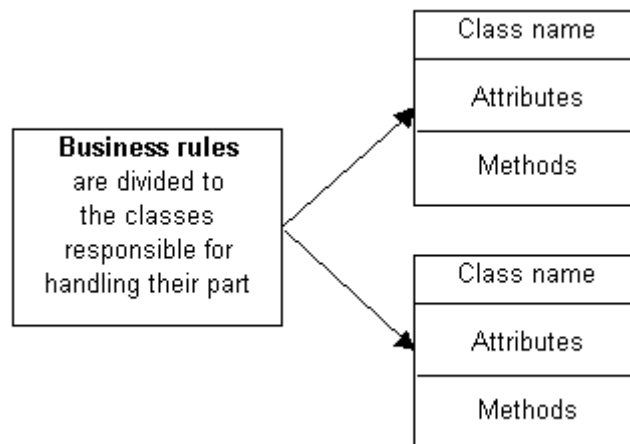
c. Protected

This only applies if inheritance is implemented. Protected access means the descendents of the class can access the attributes and methods.



④ Problem domain classes

The problem domain classes represent the business engine. The business rules are encapsulated in the problem domain classes



(3) Attributes

The attributes represent data that describes the problem domain class. Other types of attributes are

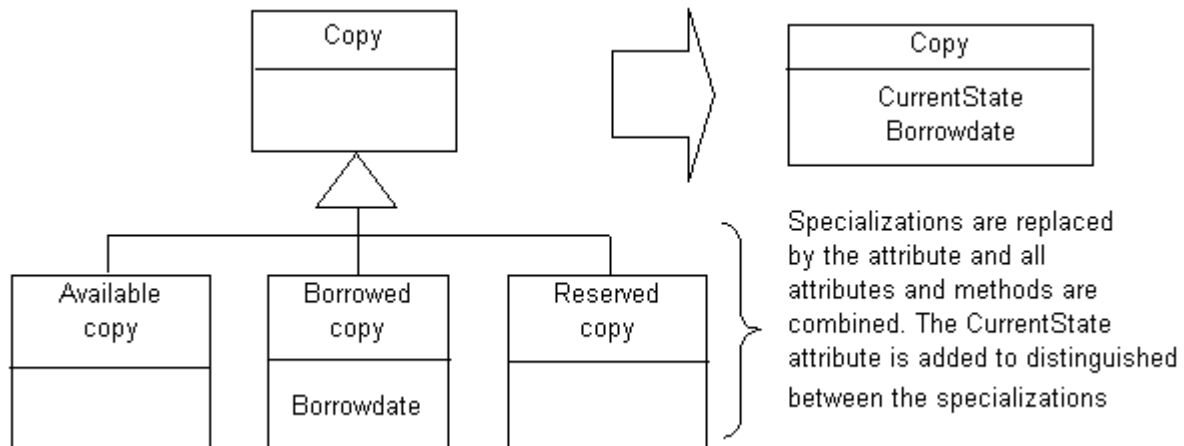
- state attributes

42 Chapter 3 System Development And Operations

- Foreign key relationship attributes
- Cardinality of the object

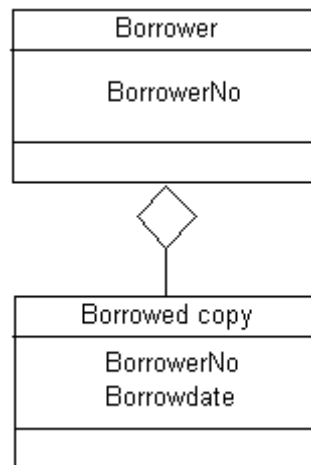
① State attributes

The specialization classes representing the states can be defined by using a state attribute.



② Foreign key relationship attributes

If a class has an association or aggregation with another class, the foreign key appears in the design class.
Example

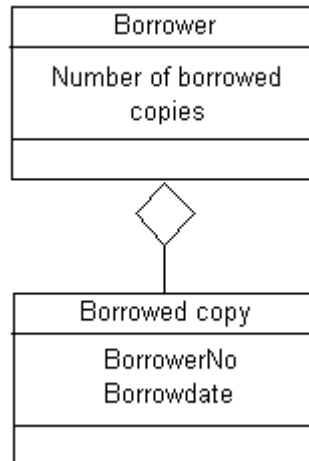


If the BorrowerNo is the object identifier for the borrower, then it also appears as an attribute in the Borrowed copy class.

③ Cardinality of the object

Numeric type attributes are good candidates for determining the limiting states. The cardinality of the object is represented as a attribute in the class.

Example of the borrower and the borrowed copies



An attribute representing the Number of borrowed copies can be added in the borrower class. The total number of borrowed copies allowed per borrower is finite. This attribute can be used to determine if the borrower has exceeded his or her limit.

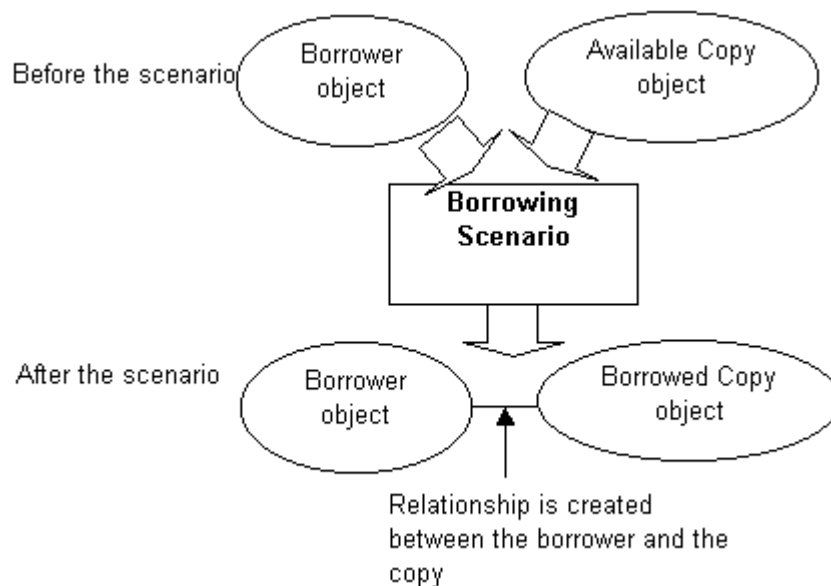
(4) Types of methods in the Problem domain classes

The methods found in the problem domain classes can be classified as

- State verification methods
- State change methods
- Data exchange methods

① State verification methods

Consider a borrowing scenario in a library system. The borrower and copy must satisfy some conditions before they are allowed to participate in this scenario. After the scenario, there is a change in the state of the objects as a relationship is created between the borrower and the copy.



44 Chapter 3 System Development And Operations

State verification methods are used to determine the initial state of the object. The copy objects in the available state are allowed to participate in the borrowing. This means a method like `IsAvailable` returning a `True/False` type output can be used to determine whether the copy object is in the available state.

Example of the state verification method in the `Copy` class (VB example)

```
public Sub IsAvailable( OutReply As Boolean)
```

② State change methods

These methods are used to make the object change its state.

Example in the case of the borrow scenario

The copy object is changed from the state of available to borrowed. This means the ability to identify the associated borrower object has to be passed to the copy object.

Example of the state change method in the `Copy` class

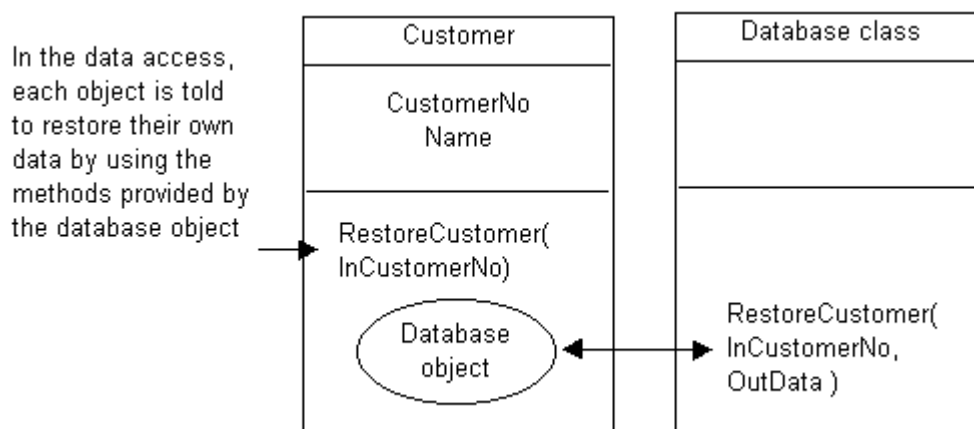
```
public Sub Borrow( InBorrower As Borrower)
```

③ Data exchange methods

These represent the exchange of information between objects. The user interface may require values to be shown to the real world. Instead of creating multiple data exchange methods, an exchange structure can be defined as a parameter. 2 methods can be defined representing data passed in or data returned.

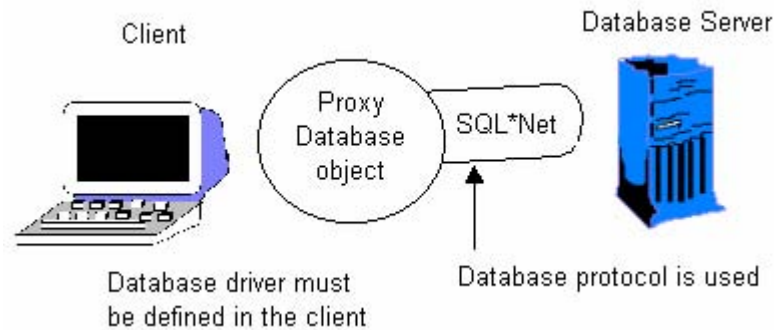
(5) Dependency between objects

Each class has its own area of responsibility. When a specific customer is to be restored, the access criteria e.g. key value is passed into a method in the `Customer` class. The customer class will not read the database directly. The reading and writing of the database is the responsibility of the database class. Instead an object of the database class is created in the body of the `Customer` method. The method offered by the database object is then executed passing the required access criteria and receiving back the data. Unlike traditional programs, the data is used to set the values in the problem domain object i.e. customer.

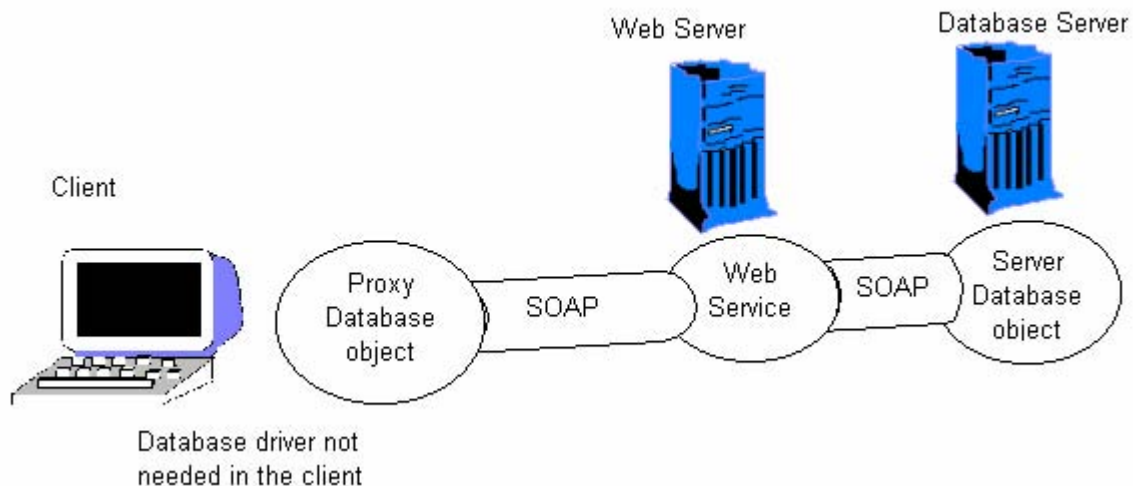


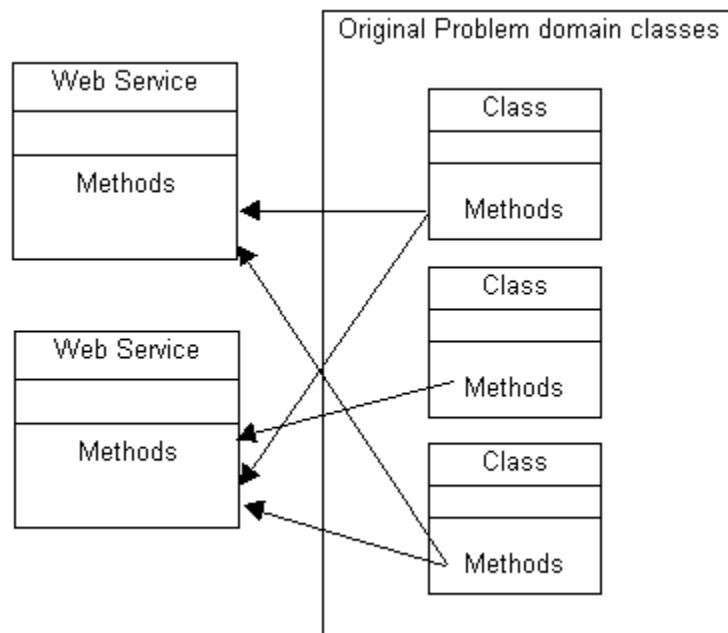
(6) Database class

The database class is used to hide the location of the database server. It is treated like a proxy to the server. In a client server type environment, the database class will actually interface with the database server. The disadvantage is the database client interface has to be installed.



Nowadays, multi-tiered architectures are more common. In a multi-tier architecture, the database object in the client is actually a proxy using some standard protocol like SOAP to access the web services on the server. The web services may then access the database server or another web server to reach the database server. Web services can be thought of as a combination of the required problem domain class methods that are exposed for usage.





4 Network Technology

Chapter Objectives

Understanding the advances in the field of telecommunications

- ① Understanding the new types of WAN technology available and the types of servers
- ② Understanding the concepts behind the wireless LAN and the use of mobile IP and the use of XML
- ③ Understanding the need for security in the internet

Introduction

The recent availability of wireless LAN and broadband services offers a cheap and convenient way to set up a network. In the areas of data exchange, the use of XML based standards is becoming popular. In this chapter, the foundations of broadband and wireless LAN as well as the use of XML is examined.

4.1 Technology advances

4.1.1 Evolution in telecommunications

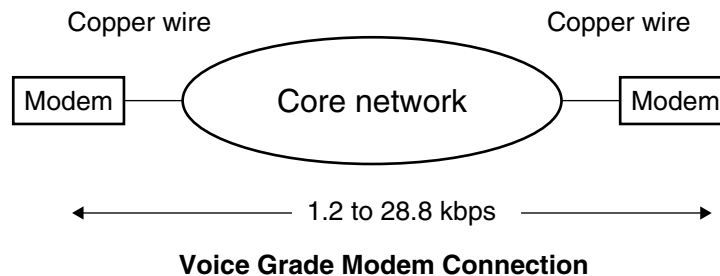
Recent advances in technology have dramatically changed the way of working. The main thrust of this change is the increased availability of network bandwidth (broadband) and the spread of the Internet.

(1) Evolution in telecommunications

The following explains the evolution of the network to support broadband networks.

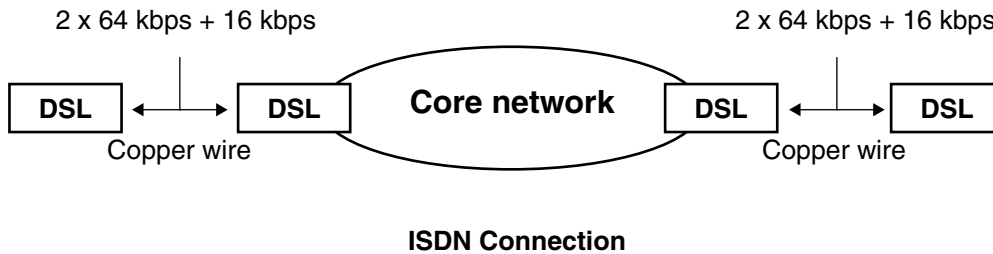
① Voice grade modem

The network handles the signals like voice signals and although very high speeds cannot be achieved, they have the advantage of being easily connectable to the network.

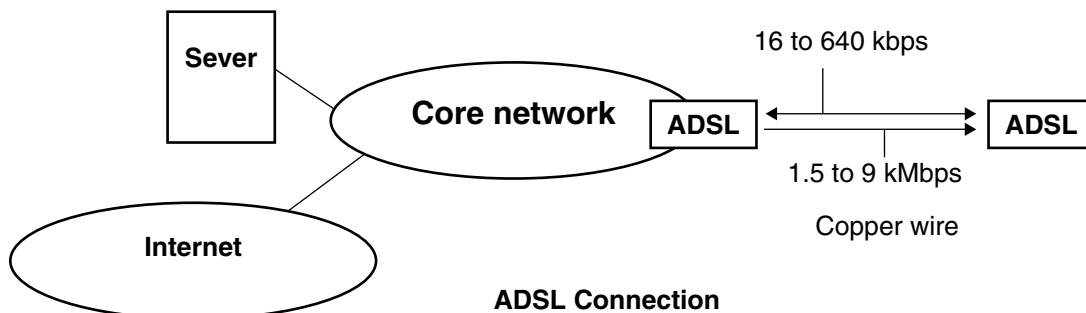


② Digital Subscriber Line

DSL or Digital Subscriber Line means the modem or modem pair. DSL is used for the basic rate ISDN. Data is transmitted in duplex with speeds of 160 kbps. There are two B channels (64 kbps each) and a D channel (16 kbps).



③ ADSL (Asymmetric Digital Subscriber Line)



The data transmission is asymmetric with a higher volume going down to the subscriber. ADSL is associated with broadband. There are 3 basic channels

- high speed down stream
- medium speed duplex
- telephone service

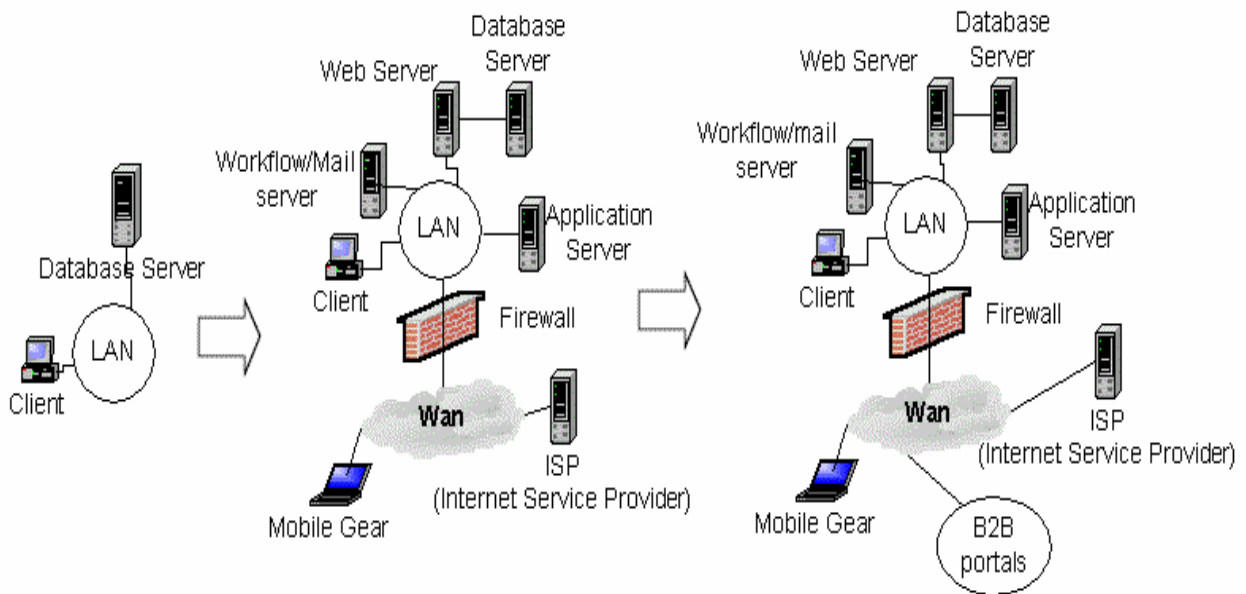
The ADSL modem filters off the telephone channel ensuring it is not interrupted even if ADSL fails. Speeds ranging from 1.5 to 6.1 Mbps for the high speed channel and 16 to 640 kbps for the duplex are available. Each channel may also be sub multiplexed to give lower rate channels. Most applications require a high volume download rather than a high volume upload. Applications like Video on demand, home shopping, Internet access, remote LAN access, multimedia access etc ADSL has a range of downstream speeds depending on distance:

Up to 18,000 feet 1.544 Mbps (T1), 16,000 feet 2.048 Mbps (E1), 12,000 feet 6.312 Mbps (DS2) and 9,000 feet 8.448 Mbps

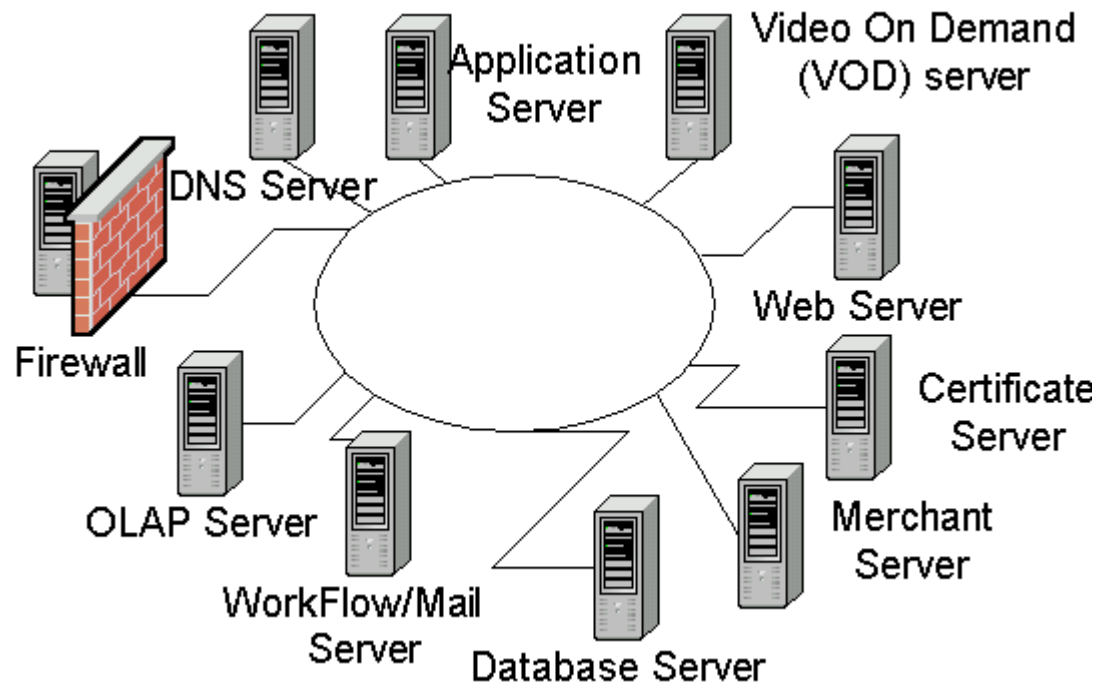
(2) Evolution of the software infrastructure

The **servers** have become specialized in their functionality. This allows the access to be distributed effectively. In addition, the rise of e-commerce and internet as a business medium means the volume of access will be high. This high volume of data has to be analyzed effectively. This means data warehouse or data market applications will become more and more popular. These allow data in the organization to be effectively managed and accessed by management. An added advantage is a corporate data model is now possible as the data market or data warehouse integrates all the information. Applications that utilizes data markets or warehouses are known as **OLAP** (Online Analytical Applications)

The **object based development methodology** will become more popular as many software packages are object based and loosely coupled systems can easily be created and maintained as the details are hidden within the classes. As mentioned previous, broadband applications means that **multimedia** based applications will become more and more prevalent. Another result is the need to have a standard format to allow information to be exchanged easily. Recently, **XML** has gradually become a de facto standard for data exchange.



① Types of servers and their functions



The above shows the different kinds of servers classified based on their application.

a. DNS Server

This is responsible for maintaining the logical address mapping of the network devices.

b. Application Server

This contains the business objects and programs executing the business processes. They often connect to the database server to retrieve or refresh information.

c. Database Server

This platform contains the database. It is used to represent the OLTP database. Data warehouses or data markets should be installed on a separate physical machine.

d. Workflow/Mail server

This server represents the applications which are used by the corporate to facilitate the business process. Workgroup software like Lotus Notes are executed on these platforms. The mail server is normally combined in the same machine.

e. Web Server

This represents the generic machine running the Web Server

f. OLAP server

This is used to manage the data market or the data warehouse. It services queries make to these resources. The data market or data warehouse does not necessary have to be in the same physical machine.

g. Firewall

This represents the host machine that is connected between the internal network and the external internet. It serves to shield the network from external attacks.

h. Video On Demand (VOD) server

This is used to contain the content used for the streaming video. The MPEG 2 or MPEG 4 format is commonly used for the video. The VOD can be used to support the Web based Training programs or used to give more information in the e-commerce application

i. Certificate Server

These are used for the purpose for verifying the validity of the digital certificates issued. They represent normally well known CA (Certificate Authorities) like VeriSign. A company can purchase these certificates from VeriSign and choose to use one of their machines as a certificate server.

j. Merchant Server

These are used to handle payments and orders in e-commerce systems.

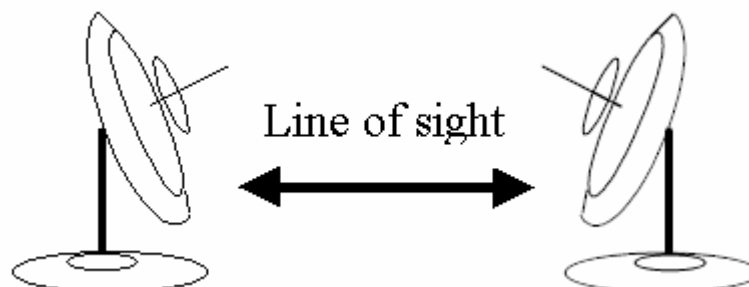
4.1.2 Wireless LAN

An increasing number of LAN users are becoming mobile. These mobile users require that they are connected to the network regardless of where they are Wireless networks can be used in combination with cabled. use one of three transmission techniques: spread spectrum, narrowband microwave, and infrared. The standard used for wireless LAN is 802.11.

(1) Physical interfaces used in the wireless LAN

① Infrared

Infrared systems are cheap and use the same frequencies as fiber optics links. There are not limited by bandwidth. There operate in the light spectrum and do not require any special licensing. However IR based systems cannot penetrate opaque objects. The data rates offered are between 100 Kbps to 4 Mbps



② Microwave

Microwave systems use narrow-band transmission with single frequency modulation. There are set up in the 5.8GHz band. The advantage of using MW systems is higher throughput can be achieved since they do not have the overhead involved with spread spectrum systems. RadioLAN is an example of systems with microwave technology.

③ Radio Frequency

Radio frequency systems use spread spectrum technology. There are two types:

- direct sequence spread spectrum (DSSS)
- frequency hopping spread spectrum (FHSS).

Radio Frequency offers data rates of 1 Mbps to 10 Mbps.

a. Direct Sequence Spread Spectrum (DSSS)

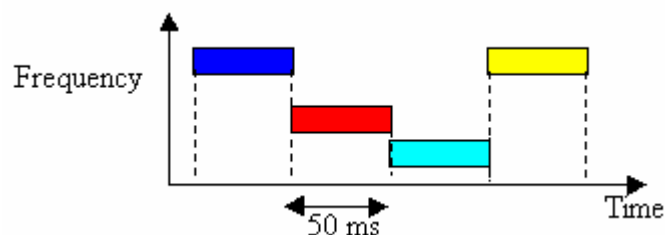
With direct sequence spread spectrum the transmission signal is spread over an allowed band (for example 25MHz). A random binary string is used to modulate the transmitted signal. This random string is called the spreading code. The data bits are mapped to into a pattern of "chips" and mapped back into a bit at the destination. The number of chips that represent a bit is the spreading ratio. The higher the spreading ratio, the more the signal is resistant to interference. The lower the spreading ratio, the more bandwidth is available to the user. The FCC dictates that the spreading ratio must be more than ten.

The IEEE 802.11 standard requires a spreading ratio of eleven. The transmitter and the receiver must be synchronized with the same spreading code.

b. Frequency Hopping Spread Spectrum (FHSS)

Spread spectrum is currently the most widely used transmission technique for wireless LANs. It was initially developed by the military to avoid jamming and eavesdropping of the signals. This is done by spreading the signal over a range of frequencies, that consist of the industrial, scientific, and medical (ISM) bands of the electromagnetic spectrum. The ISM bands include the frequency ranges at 902 MHz to 928 MHz and at 2.4 GHz to 2.484 GHz, which do not require an FCC license. The first type of spread spectrum developed is known as frequency hopping spread spectrum.

It splits the band into many small subchannels (1MHz). The signal then hops from subchannel to subchannel transmitting short bursts of data on each channel for a set period of time, called dwell time. The hopping sequence must be synchronized at the sender and the receiver or information is lost. This technique broadcasts the signal over a seemingly random series of radio frequencies. A receiver, hopping between frequencies in synchronization with the transmitter, receives the message. The message can only be fully received if the series of frequencies is known. Because only the intended receiver knows the transmitter's hopping sequence, only that receiver can successfully receive all of the data. Most vendors develop their own hopping-sequence algorithms, which all but guarantees that two transmitters will not hop to the same frequency at the same time.



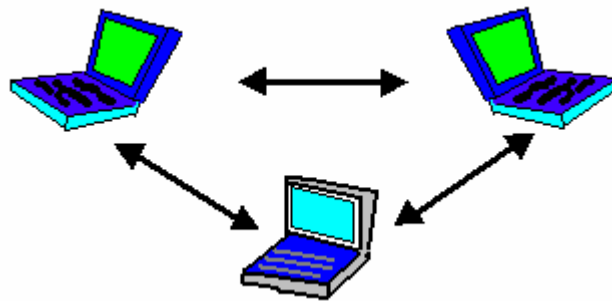
c. Comparison between FHSS and DSSS

Frequency hopping spread spectrum	Direct sequence spread spectrum
It uses only a portion of the bandwidth at any given time	It is more expensive
The implementation of FHSS is simpler than DSSS	It is more resistant to fading
There is a lower cost hardware	It is harder to detect and intercept than FHSS

(2) Architecture of a wireless LAN

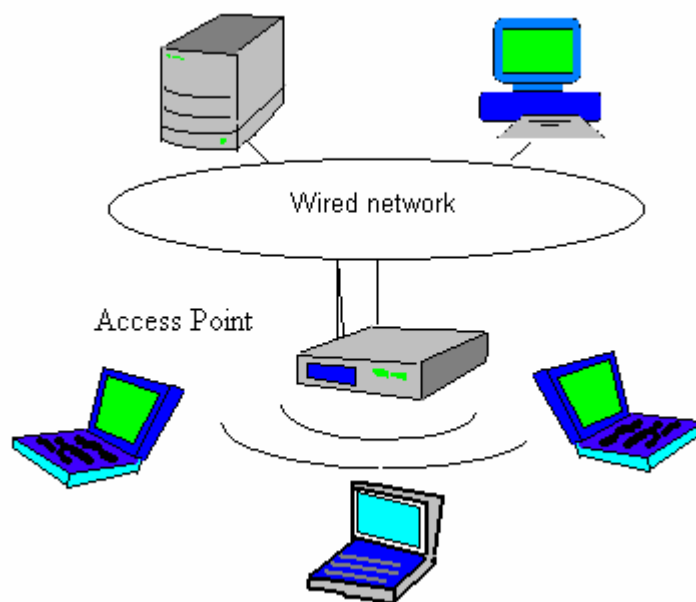
① Basic Service Set

Ad Hoc networks are peer to peer connections without any using any Access Point (AP).

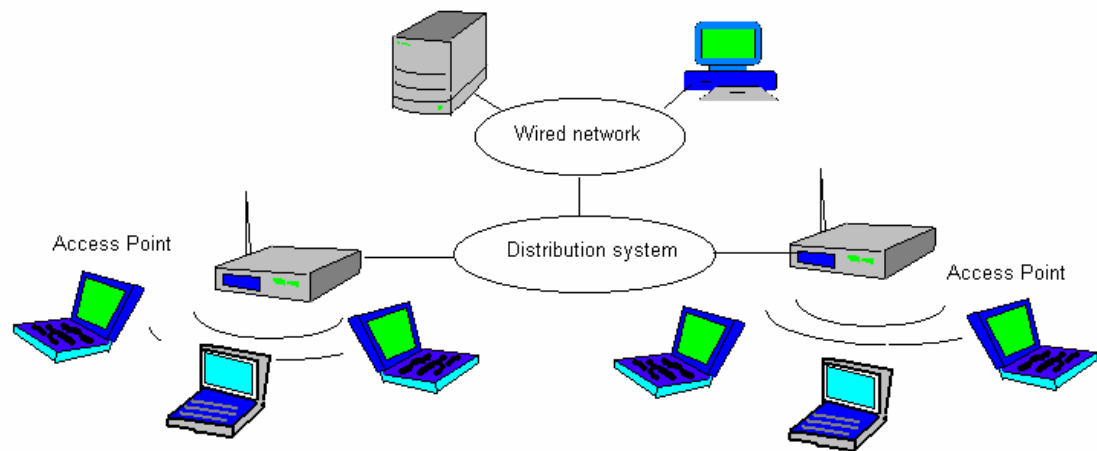


② Infrastructure based

The access point allows the connection between a wireless and wired network.

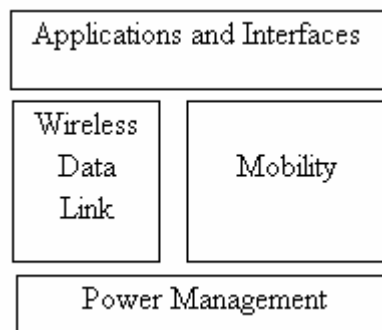


③ Extended Service Set

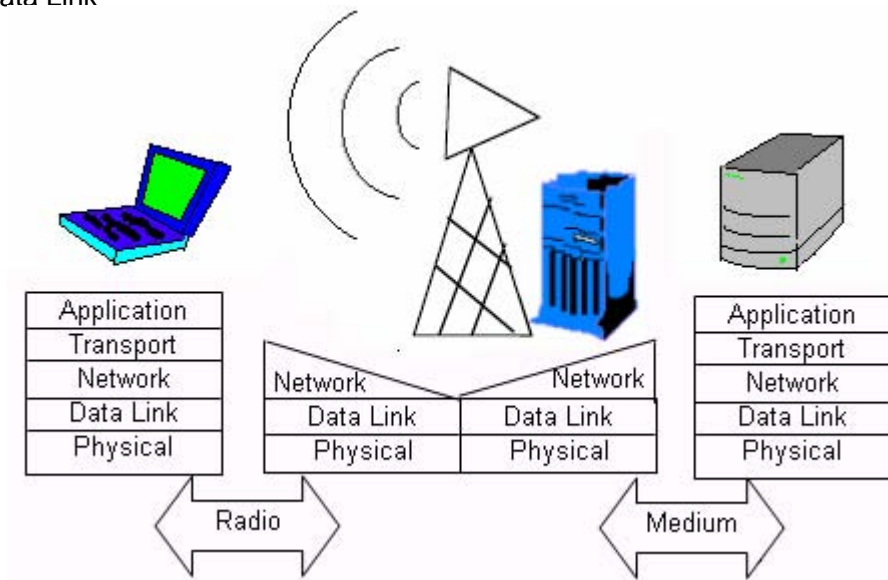


This applies only to infrastructure networks. The distribution may be a backbone using a standard LAN

(3) Layers found in the wireless LAN



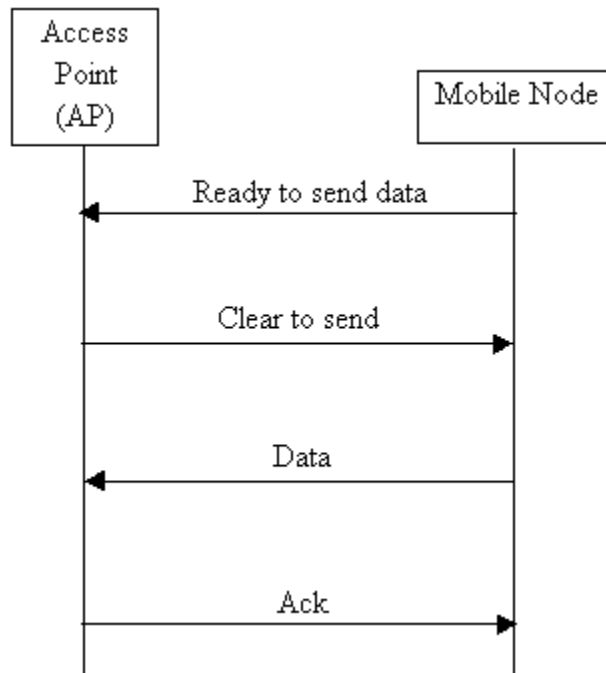
① Wireless Data Link



The standard used is 802.11. There is also a 802.11b standard.

	IEEE802.11b	IEEE802.11
Radio Band	2.4 GHz	2.4 GHz
Modulation Technique	Direct Sequence Spread Spectrum (DSSS)	Direct Sequence Spread Spectrum (DSSS); Frequency Hopping Spread Spectrum (FHSS)
Data Rate Supported	11, 5.5, 2, 1 Mbps - equivalent to Wired Ethernet speed	1,2 Mbps - too slow to support general business requirements

② 4 way hand shake



802.11 uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) is used in this protocol. It listens before you talk. If the medium is busy, the transmitter backs off for a random period. Avoids collision by sending a short message: Ready to send (RTS) RTS contains destination address and duration of message. It tells everyone to backoff for the duration. The destination sends: Clear to send (CTS) It cannot detect collision. Each packet is acked. The MAC level retransmits if not acked.

(4) Security in a Wireless LAN

A new nodes issue a "request for authentication" The Network sends a block of random text. The node encrypts it with network password and returns. Currently, one shared secret key (password) per net. This is known as the Wired Equivalency Privacy (WEP) Algorithm

(5) Mobility

Mobility is possible with the use of mobile IP. This is enhanced with the availability of the IPV6. The characteristics of mobile IP are

- ① It automatically seeks for the IP routers and connects to them
- ② No new IP address or address formats are needed
- ③ Authentication is supported
- ④ Mobile networks are supported

The definition of mobility is the ability to attach from one link to another without the need to change the IP. There are 2 kinds of links

a. Home link

This is the parent network where the node is connected. This means the network prefix is the same as the node's IP address prefix.

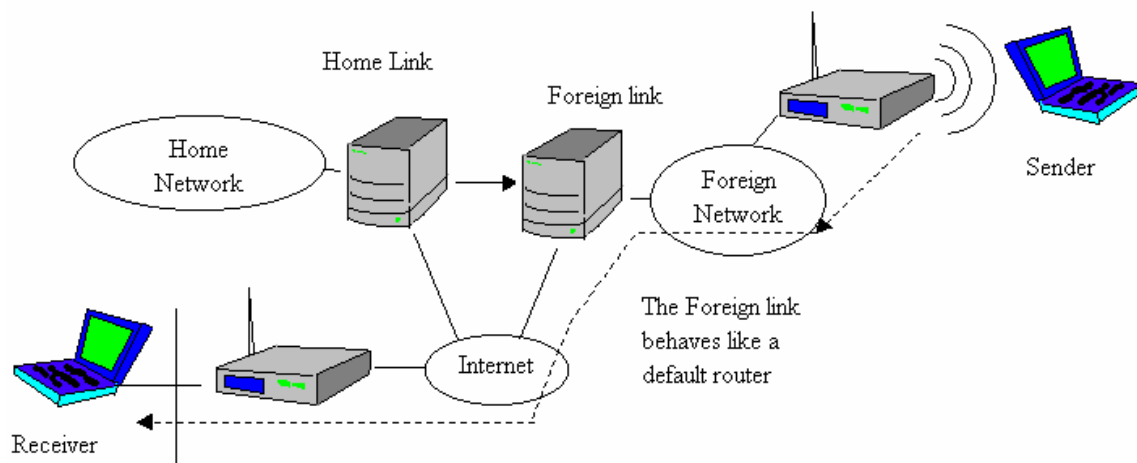
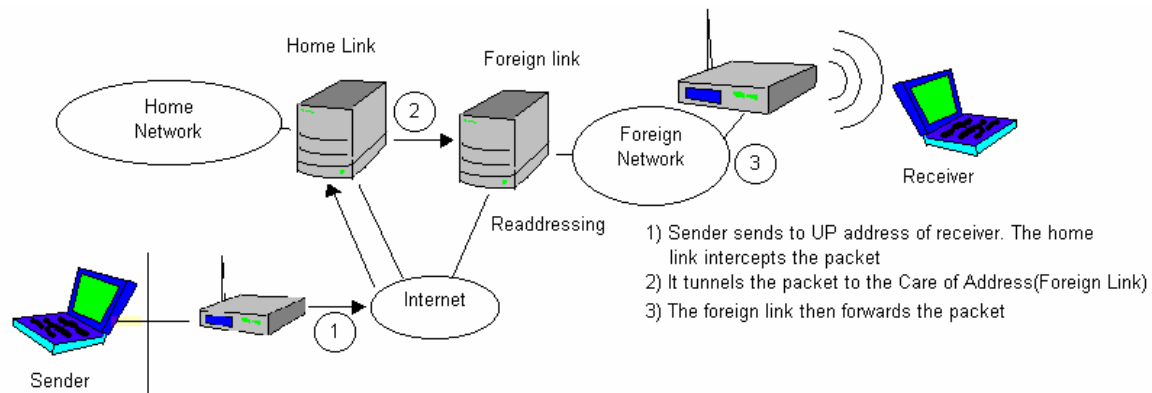
b. Foreign link

This represents a different network from the one originally connected to by the node.

⑤ Mobility model

There are 2 addresses in the mobility management model.

1. One address for locating the host
2. Second address for identifying the endpoint



⑥ Identification of addresses

There are 3 kinds of addresses in the mobility model

a. Home address

This is the known IP for the host

b. Home network (Home link)

The destination network associated with the home address

c. care of address

The IP address used to locate the host

⑦ Roaming node

The source sends the datagram to the home address. The datagram reaches the home network. The datagram has to be re-addressed to reach the foreign link at the care of address. The datagram is then delivered to the care of address. The re-addressing is done at the foreign link to restore original address.

⑧ Power management

A station can be in one of three states:

- a. Transmitter on
- b. Receiver only on
- c. Dozing: Both transmitter and receivers off.

The function of the access point is to buffer traffic for dozing stations. It will also announce which station has buffered frames. The dozing station wakes up to get to listen for traffic. A poll frame is sent if there is any outstanding data for it.

4.2 XML (Extensible Markup Language)

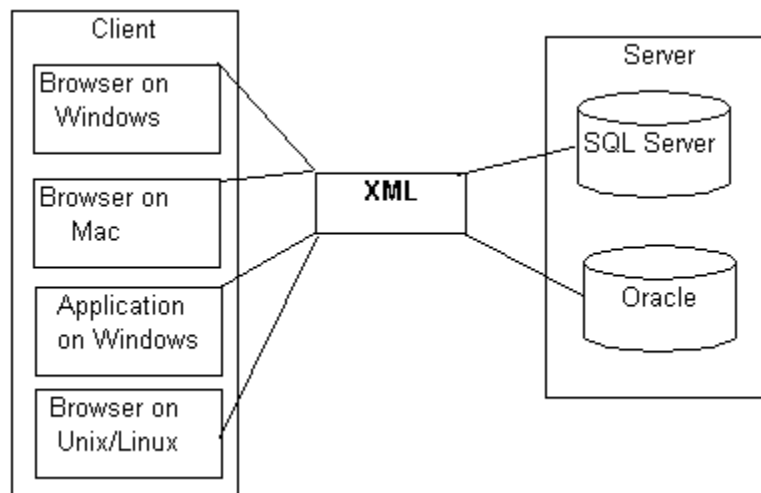
This is the base language used to define many standards like SOAP(Simple Objects Access Protocol) or XHTML.

4.2.1 XML(Extensible Markup Language)

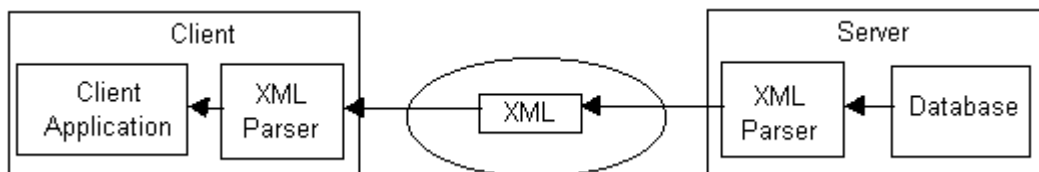
(1) XML (Extensible Markup Language)

XML is a markup language with a set of rules that help to structure the document. It can be used to standardize the exchange format of information.

① Overview of XML



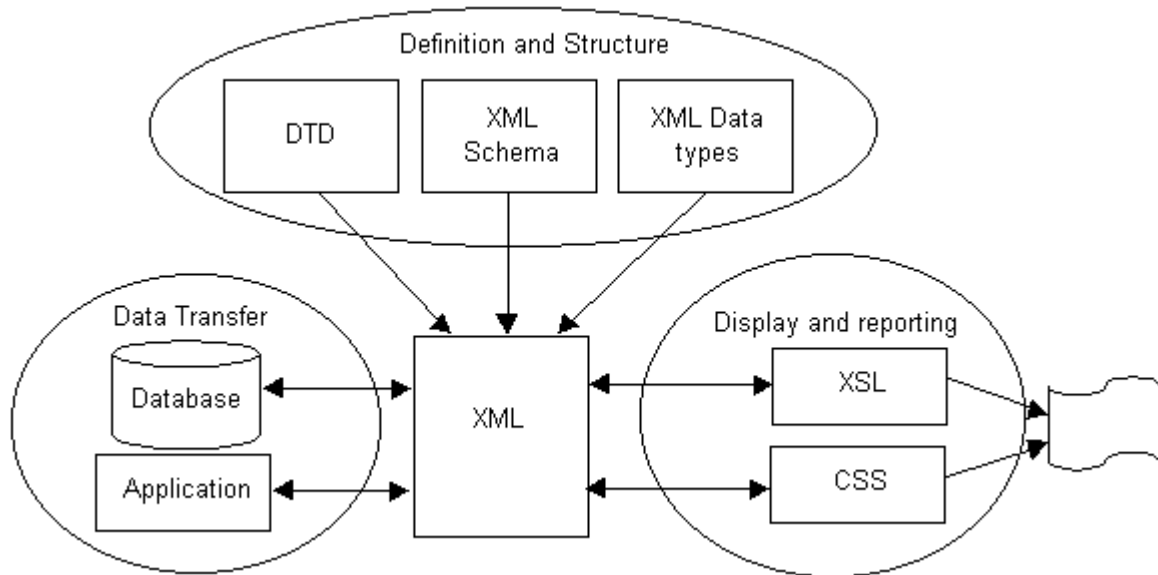
Applications can access the XML document and utilize the information inside. As shown in the diagram, applications on the client and the server can utilize the XML document to update the information and exchange information making it independent of the software.

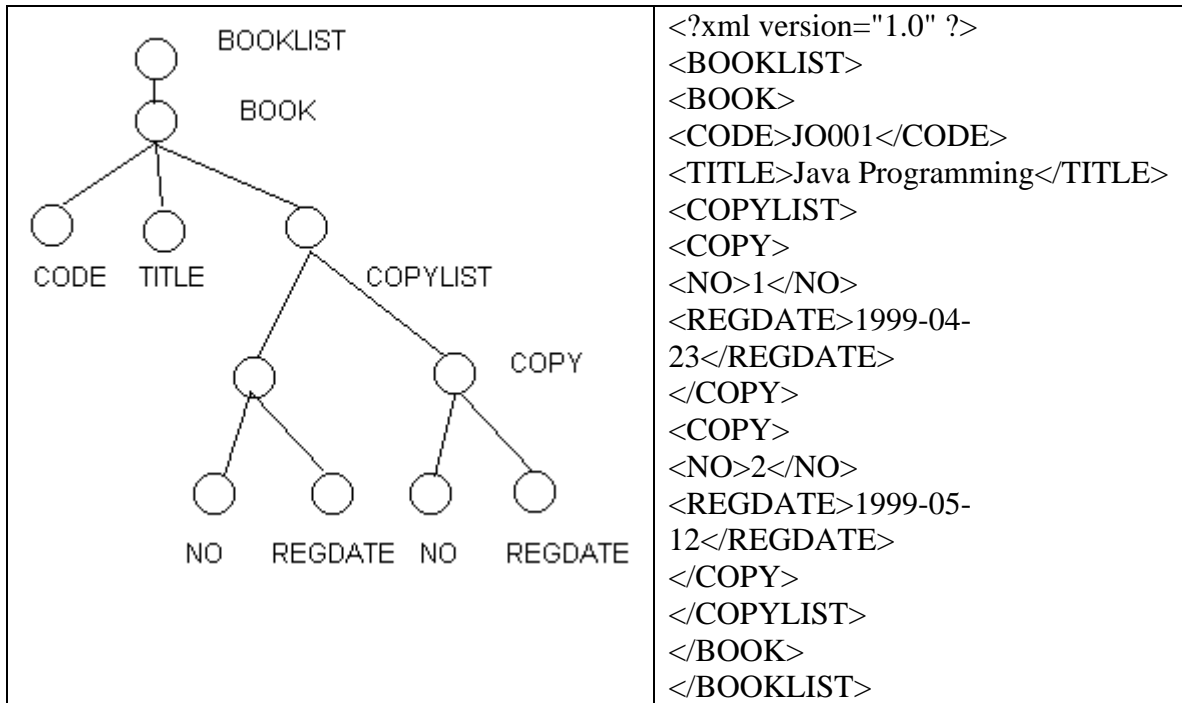


XML parser will parse the XML document i.e. changing it for errors and the application can access the XML document or create XML documents and send it to the parser for checking. XML is a set of rules describing the document,. This means the data is enclosed within the tags like HTML. However, unlike HTML, the tags are not as restricted. In addition, the characteristics of the data enclosed between the tags, e.g. string, repeated etc is defined by using DTD (Document Type Definition) and XML Schemas. XML Schema definition is proposed by Microsoft and is an XML document itself. Formatting of the XML document can be done by creating XSL (Extensible Stylesheet Language) definitions.

There are the following types of definitions.

- a. XML for defining the data in the XML document
- b. DTD for defining the structure of the XML document
- c. Schema for defining the structure of the XML document
- d. XSL for defining the styles to apply





② Well formed documents

This means they is a set of rules that must be satisfied when creating the document. Documents that follow the given rules are known as well formed documents. The rules that apply are

- Start and end tags are required. In the case of empty elements, the definition should be terminated with the "/" slash before the enclosing bracket ">".

Not allowed	Well formed
<pre> <EMPLOYEE> <ID>23001 <NAME>James Born </pre>	<pre> <EMPLOYEE> <ID>23001</ID> <NAME>James Born</NAME> </EMPLOYEE> </pre>
<pre> <EMPLOYEE ID="23001" NAME="James Born" > </pre>	<pre> <EMPLOYEE ID="23001" NAME="James Born" /> </pre>

b. Element definitions cannot overlap.

Not allowed	Well formed
<pre><H1> The following<A>BO0001 </H1> is important. </pre>	<pre><H1> The following<A>BO0001 is important. </H1></pre>

c. Markup characters like "<" , '>', etc cannot be allowed in the parsed content. Instead, the characters are replaced with the equivalent definition.

Markup character	Definition
<	<
>	>
'	'
"	"
;	&

Example

Not allowed	Well formed
<condition> a > 2 </condition>	<condition> a > 2 </condition>

d. You may use only letters or underscores to start the element names and the rest of it may contain either letters, numbers, hyphens, periods and underscores.

③ Valid document

The structure of the XML document can be defined by using DTD or a schema definition. These describe information like the data type of the elements or attributes etc. If the XML document conforms to the constraints, the document is said to be valid.

A script is used for the purpose of written expression. A set of phonetic characters may comprise the script or it be represented by ideographic characters like Chinese. Some languages also have a set of special notations. The following conditions are required to utilize non roman characters in XML.

- A character set for the script
- A font for the character set
- A means to input the characters
- The OS and application must understand this character set.

The XML processor assumes the XML file is coded in UTF-8 unless otherwise stated. The other format that the XML processor recognizes is the raw Unicode format. The character set to use can be defined in the encoding attributes of the starting declaration.

Example

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
```

(1) Unicode character set

Each character is encoded as a 2 byte unsigned number in the range of 0 to 65,535. The Unicode characters 0 to 255 is the same as the Latin-1 characters 0 to 255.

Script	Range	Description
Basic Latin	0 to 127	ASCII . American English
Latin-1 supplement	128 to 255	Characters from ISO 8859 are added

① UTF-8

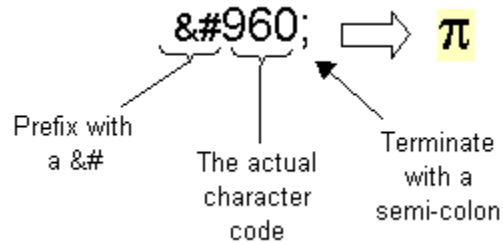
This uses 2 bytes for the coding of each character. UTF-8 uses a single byte for most common characters and is a compressed version of the Unicode. However, it uses 3 bytes for uncommon characters so UTF-8 is good for roman based script but creates very large if the script is non roman. e.g. Japanese

② UCS (Universal Character System)

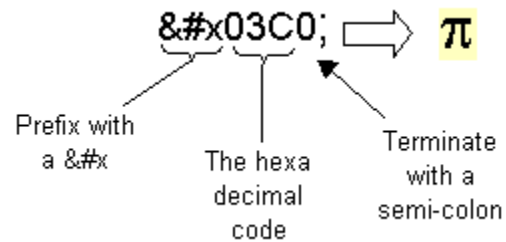
This system also known as ISO 10646 uses 4 bytes per character. Unlike Unicode, that uses 2 bytes. This allows for full character sets to be defined. Pure Unicode is sometimes known as UCS-2. The UTF-16 maps some of the UCS characters into byte strings such that Unicode is unchanged.

③ Using character references

A character reference can be used to insert the character into the XML file. This is done if the text editor cannot support the required character set. The Unicode character reference has the following format



You can also use the hexadecimal notation to encode the character.

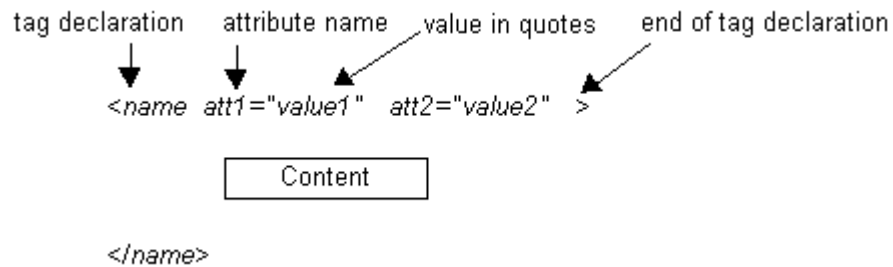


④ Common character sets

Character set name	Language or country	Character set name	Language or country
US-ASCII	English	ISO-10646-UCS-2	Raw Unicode
UTF-8	Compressed Unicode	ISO-10646-UCS-4	Raw UCS
UTF-16	Compressed UCS	ISO-8859-1	Latin-1, Western Europe
ISO-8859-5	ASCII plus Cyrillic	ISO-8859-11	ASCII plus Thai
ISO-2022-JP	Japanese	Shift-JIS	Japanese, Windows
EUC-JP	Japanese Unix	Big5	Chinese, Taiwan
GB2312	Chinese, mainland China		

⑤ Element definition

Elements form the main backbone of the definition. Each element can also comprise of nested elements. Each element can also comprise attributes. Attributes can be defined within the tag.



Example

```
<BOOK>
  <CODE>JO001</CODE>
  <TITLE>Java Programming</TITLE>
</BOOK>
```

Definition with attributes

```
<BOOK CODE="JO001" TITLE="Java Programming" />
```


⑥ Empty elements

In the example for the attributes, all the elements were defined as attributes. This means there are no elements enclosed. When an element has no enclosing tag because there are no elements within, an empty element definition can be defined. Empty elements are definitions without any content but may contain attribute definitions.

end of tag declaration
↓

```
<name att1="value1" att2="value2" />
```

4.2.3 Namespace

The same tag may be used but the context of the meaning may be different. Example CODE may mean the book's code or the borrower's code depending on which element it is defined. In order to distinguish tags with the same name, the namespace concept is used. The xmlns can be defined as an attribute. Each namespace requires a URI (Universal Resource Identifier) to distinguish the namespaces. Currently, it is not possible to connect the URL to the namespace. Instead, qualified namespaces are used.

Example

```
<BORROWER xmlns="http://www.lib.com/bor/" >
```

(1) Qualified namespace

A qualified namespace allows a local identifier to be added in the definition.

Example

```
<BOOKDOC xmlns:forBook="http://www.lib.com/book/"
  xmlns:forMember="http://www.lib.com/bor/">

<forBook:LIST>
<BOOK>
  <forBook:CODE>BO001</CODE>
  <TITLE> Biology</TITLE>
```

A unique name for the namespace is suffix to the xmlns attribute. This namespace is prefixed to the tag name. The tags that do not have the prefix will use the default namespace. The attributes can also be qualified with the namespace.

(2) Definition of the structure of XML document

There are 2 methods used in describing the structure of the XML document.

- ① DTD (Document Type Definition)
- ② Schema

4.2.4 DTD (Document Type Definition)

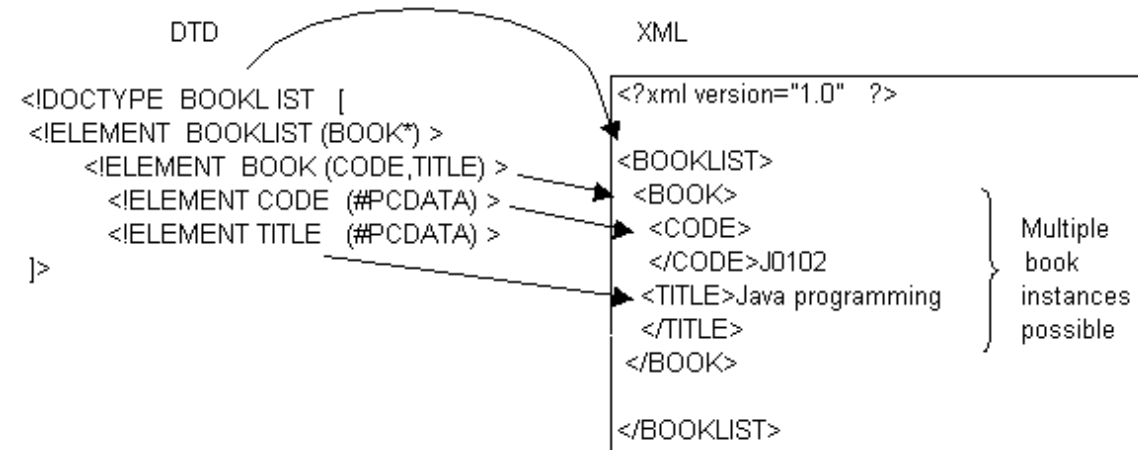
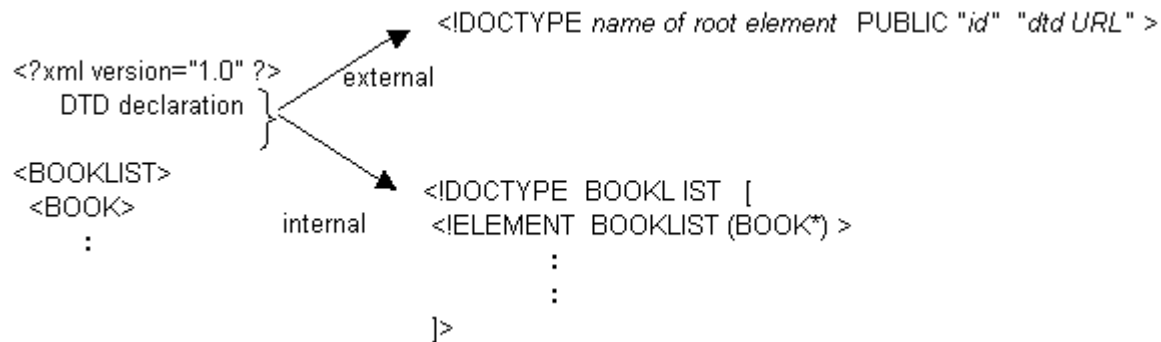
This is based on the SGML. This is used to describe the structure of the XML document. The declaration elements used in DTD are

- DOCTYPE
- ELEMENT
- ATTLIST

DOCTYPE	This is the root of the definition
ELEMENT	The content is defined as elements
ATTLIST	The attributes of the elements are defined

Example

```
<?xml version="1.0" ?>
<!DOCTYPE book [<!ELEMENT book (#PCDATA)>] >
<book>Java programming</book>
```



(1) DOCTYPE

This starts the definition of the declaration.

`<!DOCTYPE root name` The root name corresponds to the name of the root element in the XML document

`<!DOCTYPE BOOKLIST [`

Element definitions

`]>` This terminates the DTD declaration

(2) ELEMENT and ATTLIST

This is used to describe the contents and attributes. `<!ELEMENT tag name type of contents >`

Example `<!ELEMENT BOOKLIST (BOOK*) >`

Type of Contents	Description
ANY	This means all possible elements are allowed
<i>element name repeat</i>	<p>This defines how the tag is repeated or optional</p> <p>BOOK* asterisks means it can appear zero or more times</p> <p>BOOK+ plus means it can appear one or more times</p> <p>BOOK? Asterisks means it can appear zero or once</p> <p>BOOK BORROWER vertical bar is used to separate the elements that can appear</p>
EMPTY	No elements contained
The above apply to all element definitions. The next section applies to the contents and attributes	
#PCDATA	Parsed character. The content cannot contain the reserved characters e.g. <
CDATA	Character data
Modifiers	#REQUIRED value is mandatory #FIXED values are fixed and a default is provided. #IMPLIED value is optional

Example of the use of the required

`<!ELEMENT BOOKTYPE (hardcover | paperback | none) >`

`<!ATTLIST BOOKTYPE DESIGN (color | blackwhite) "color" >`

`<!ATTLIST BOOKTYPE PUBLISHER CDATA #IMPLIED >`

The element BOOKTYPE is defined with both content and 2 attributes. A list of values can be defined to restrict the values allowed. A default value can be defined. The values are defined without quotes.

(3) Entity

Entity references are used to replace the text in the XML document.

```
<!ENTITY mail "yy@tt.com" >
```

In the XML document, the entity reference prefixed with a & can be defined.

```
<EMAIL>&mail;</EMAIL>
```

(4) Comments

Comments can be added to the DTD definition by enclosing it as shown

```
<!-- Comments -->
```

4.2.5 Schema definition

(1) Schema definition

Schema is proposed by Microsoft. It follows the XML rules. The structure of a schema is shown below

```

<Schema ..... >

  <AttributeType>
    <description>.....</description>
    <attribute>.....</attribute>
    <datatype>.....</>
  </AttributeType>

  <ElementType name="BOOKLIST" content="e/Only"
    models="closed" >
    :
    <description>.....</description>
    <AttributeType>
      <description>.....</description>
      <attribute>.....</attribute>
      <datatype>.....</>
    </AttributeType>

    <group ..... >
      :
      <element ..... />
      <attribute ..... />
    </group>
  </ElementType>
</Schema>

```

① Schema

The schema is the root of the definition.

Name	name of the schema
Xmlns	namespace identifier or URI (Uniform resource Identifier)
xmlns:prefix	namespace URI for the datatype attributes. Usually dt is used.

② ElementType

This defines the type of element

Content	empty " no content textOnly only text can be contained eltOnly Elements only and no text mixed Elements and text can be contained
dt:type	The data type of the content
model	open undefined content allowed closed only contents defined by schema is allowed to show
name	The name of the element
order	one only one definition of the element allowed seq The order as defined in the definition many any order allowed

③ element

This defines the instances of the element types

Type	The name of the element type
MinOccurs	0 optional 1 once or more (default)
MaxOccurs	1 only once (default) * any number

④ AttributeType

This defines the type of attributes.

Dt:type	Primitive data types
Dt:values	set of values for enumerated type
Default	default value mutually exclusive with required="yes"
Name	attribute name
Model	open undefined content allowed closed only contents defined by schema is allowed to show
Required	yes or no whether it is mandatory

⑤ Attribute

The instance of the attributeType

Default	default value mutually exclusive with required="yes"
Required	yes or no whether it is mandatory
Type	name of the attributeType

⑥ description

This element is used to describe the element

⑦ datatype

The datatype element describes the characteristics of the data type.

Dt:max	Maximum value inclusive
Dt:maxExclusive	Maximum value exclusive meaning values less than
dt:maxLength	This is usually applied to strings
dt:min	Minimum value inclusive
dt:minExclusive	Maximum value exclusive meaning values greater than
dt:type	primitive data type
dt:values	list of values for enumeration

a. Values used in dt:type

string	characters
number	numeric
Int	integers (whole numbers)
Float	floating point
Boolean	1 true 0 false
Date	yyyy-mm-dd format
Enumeration	enumerated type

⑧ group

This is used to organize the elements

Order	one only one definition of the element allowed seq The order as defined in the definition many any order allowed
MinOccurs	0 optional 1 once or more (default)
MaxOccurs	1 only once (default) * any number

⑨ Example

```
<Schema name="BOOKSchema"
xmlns="urn:schema-microsoft-com:xml-data"
xmlns="urn:schema-microsoft-com:data-types" >
```

```
<ElementType name="BOOKLIST" constant="eltOnly"
models="closed">
<description>List of books</description>
<element type="BOOK" minOccurs="1" maxOccurs="*" />
</ElementType>
```

```
<ElementType name="BOOKLIST" constant="textOnly"
models="closed">
<description>List of books</description>
<group order="seq">
<element type="CODE" />
<element type="TITLE" />
</group>
</ElementType>
```

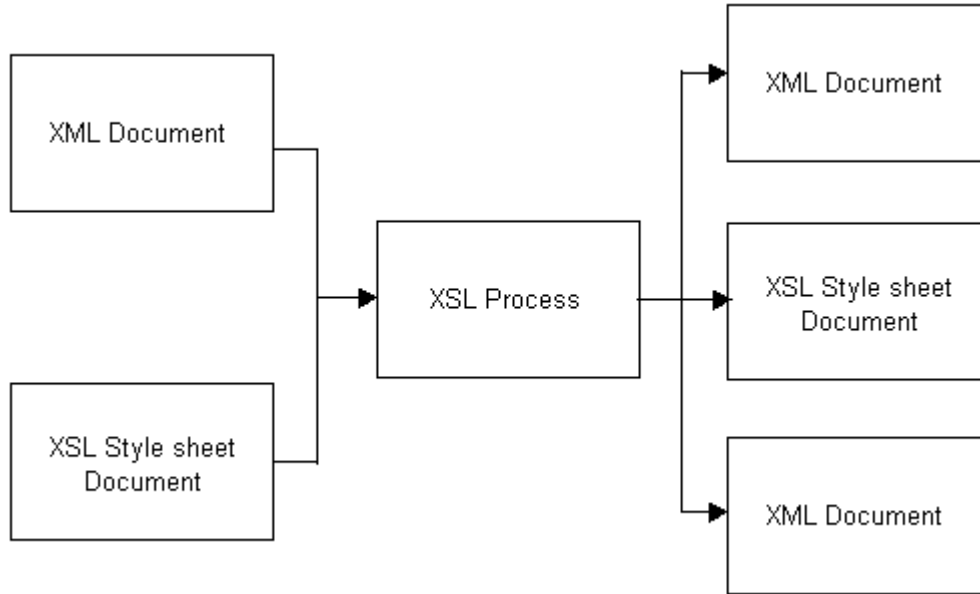
```
<ElementType name="CODE" constant="textOnly"
models="closed" dttype="string" >
</ElementType>
<ElementType name="CODE" constant="textOnly"
models="closed" dttype="string" >
</ElementType>
</Schema>
```

```
<?xml version="1.0" ?>
<BOOKLIST
xmlns="x-schema:BOOKSchema:xml" >
<BOOK>
<CODE>J0102
</CODE>
<TITLE>Java programming
</TITLE>
</BOOK>
</BOOKLIST>
```

Multiple book instances possible

4.2.6 XSL(Extensible Stylesheet Language)

(1) XSL (Extensible Stylesheet Language)



The aim of XSL is to format the XML document for display. It transforms the XML document into another format e.g. HTML for display. It also has a filtering feature represented by templates to specify target nodes with a set of style instructions. It follows the XML format of coding.

Example

```

<?xml version="1.0" ?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/TR/WD-xsl" >
  <xsl:template match="/" >
    <xsl:value-of />
  </xsl:template>
</xsl:stylesheet>
  
```

The template is defined and the processing instructions defined within the template. The processing instruction is known as the action. The template will match the elements and apply the actions to those elements in the XML document that match the search.

```

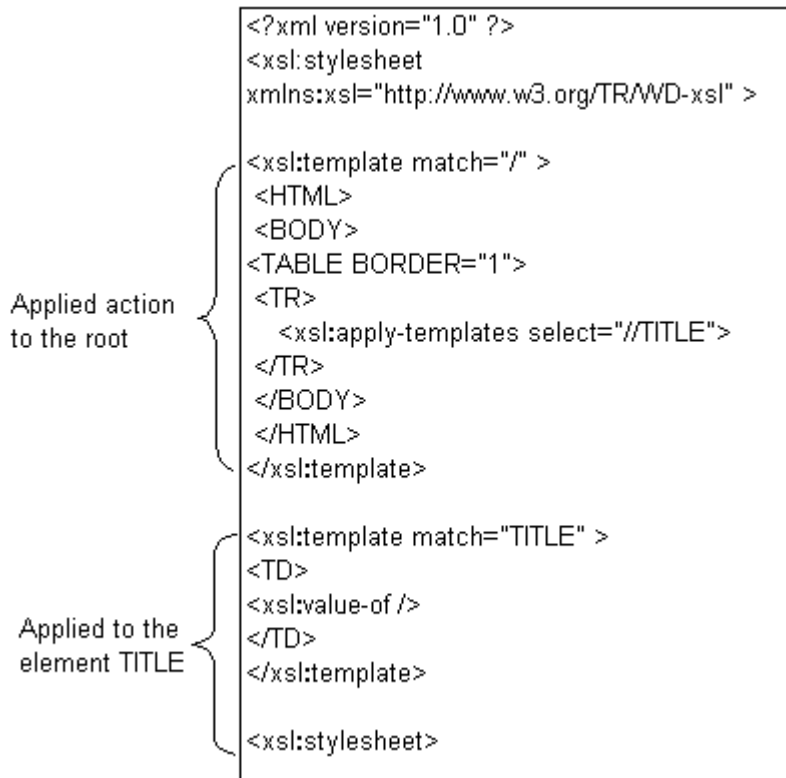
<xsl:template match="/" > <--Pattern to match "/" represents the root
  <xsl:value-of />          <-- action part insert the text value of the matching node
  
```

(2) Using the stylesheet in the XML document

A xml-stylesheet declaration is added and the name of the xsl document is defined in the href parameter.

Example

```
<?xml version="1.0" ?>
<?xml-stylesheet type="text/xml" href="bookstyle.xml" ?>
```



(3) Basic XSL elements

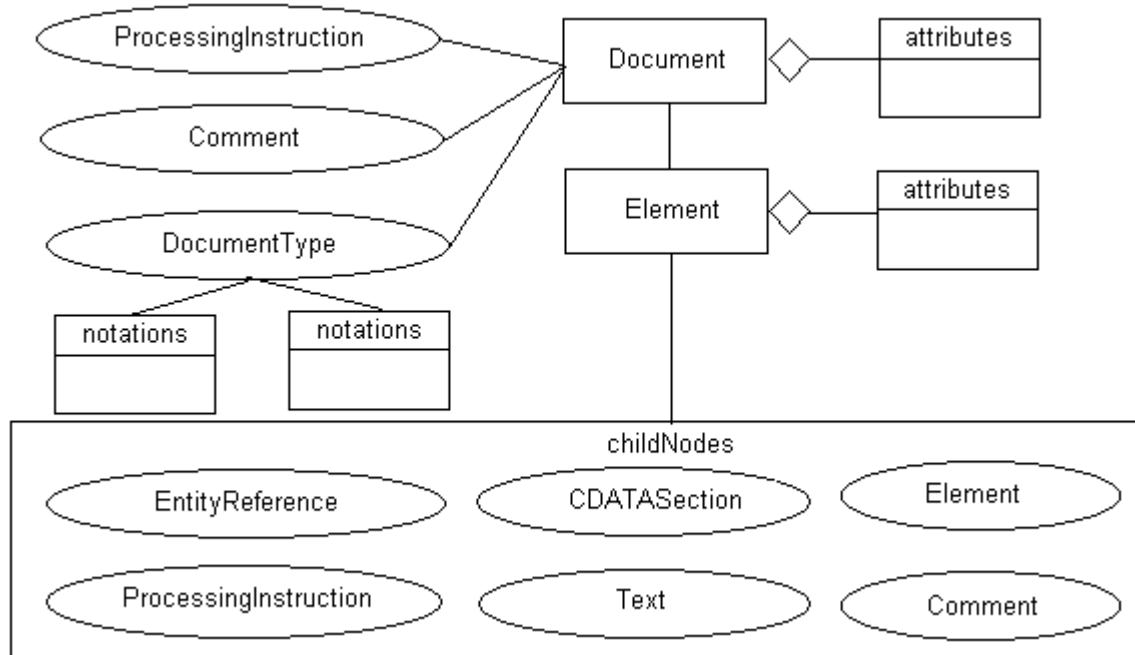
<xsl:stylesheet ... > </xsl:stylesheet>	Root of the XSL sheet <xsl:stylesheet xmlns:xsl=" http://www.w3.org/TR/WD-xsl " > The xmlns value must be the above for it to work in IE
<xsl:template > </xsl:template>	This defines a single template <xsl:template match=" <i>pattern</i> " >
<xsl:apply-templates />	This defines the actual use and which elements to apply to <<xsl:apply-templates select=" <i>pattern</i> " order-by=" <i>pattern list</i> " />
<xsl:value-of />	The value of the node is returned <xsl:value-of select=" <i>pattern</i> " />
<xsl:copy />	This copies the node with all the tags

① Pattern

Pattern	Meaning	Pattern	Meaning
/	root	*	Non root elements
book/copy	<copy> elements that are children of <book>	book/@title	title attributes attached to the book
//	Root and all nodes below		
//*	All elements below the root		

4.2.7 XML Document Object Model

This shows the XML Document object model found in Internet Explorer 5 and above. The XML Document can be created and used in the applications.



(1) Node object

The node object is the major object used and each element is represented as a node. There are 2 collections

① childNodes

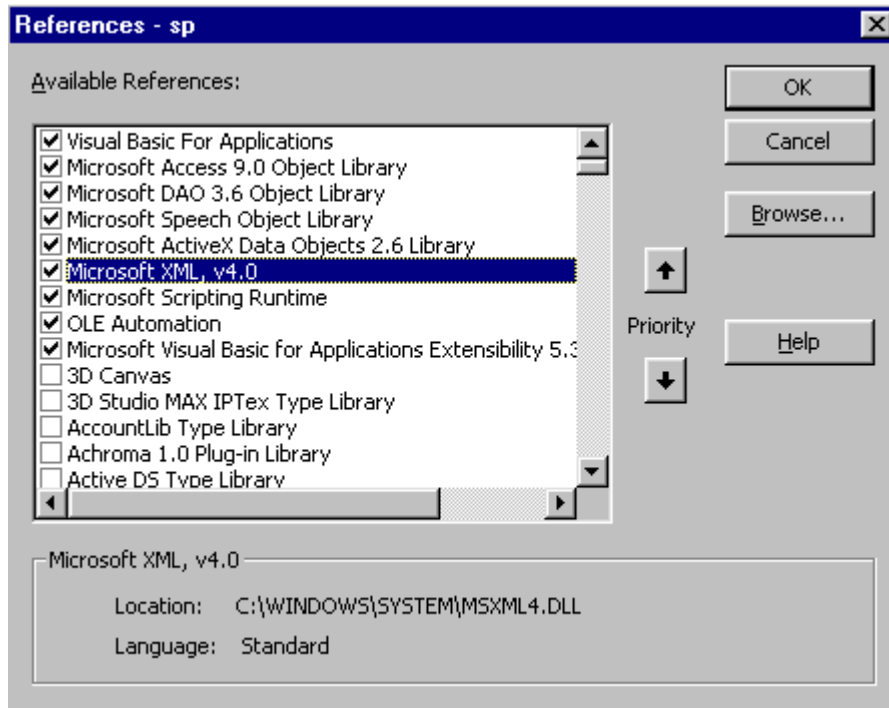
representing any child nodes within the current node

② attributes

representing any attribute definition in the current node

Example of using the Document object in VB

The reference to the XML object has to be added



```
Dim DOMOBJ As DOMObject ' Declare a Document reference
```

```
Set DOMOBJ = New DOMDocument ' Create the Document object
```

```
DOMOBJ.Load "d:\viet\book.xml" ' load an XML file
```

Example of the properties if the current node is the one given

```
<Title>Java Programming</TITLE>
```

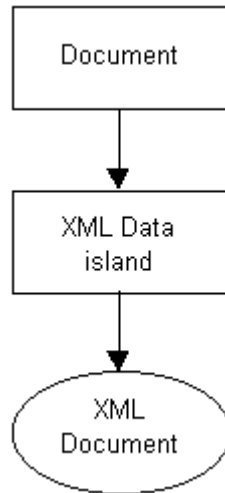
The text property returns the contents i.e. Java Programming

The nodeName returns the element name i.e. TITLE

The xml property gives the XML definition i.e. <Title>Java Programming</TITLE>

4.2.8 Using XML in the HTML document

The XML tag can be used in the HTML definition to definition an XML document in the HTML file.



Example

Data island contents defined directly	XML document referred from HTML
<pre> <HTML> : <XML ID="objBOOK" > <BOOK> <CODE>JO001</CODE> <TITLE>Java Programming</TITLE> </BOOK> </XML> </HTML> </pre>	<pre> <HTML> : <XML ID="objBOOK" SRC="book.xml" </XML> </HTML> </pre>

Client scripting can be added in the HTML document and the RDS (Remote Data Service) object can be used to access the information in the data island. The elements exposed can then be bound to specific HTML elements to display the contents. The ID attribute must be unique as this is used to identify the data island when it is bound to the other HTML tags.

Example

```
<SPAN DATASRC="#objBook" DATAFLD="CODE" ></SPAN>
```

The DATASRC attribute will point to the data island. A # must be prefixed to the name of the Data island ID when it is used. The DATAFLD attribute represents which element's contents is to be connected to the HTML tag.

Example of a table definition bound to the XML island

```
<TABLE DATASRC="#objBOOK" DATAPAGESIZE=5 >
```

The DATAPAGESIZE defines the maximum number of records to display at one time.

Example of the HTML

HTML	XML document
<pre><HTML> <BODY> <XML ID="OBJBOOK" SRC="BOOK.XML" DATAPAGESIZE=10></XML> <TABLE BORDER DATASRC="#OBJBOOK" > <TR><TD VALIGN="TOP"></TD> <TD VALIGN="TOP"></TD> <TD><TABLE BORDER DATASRC="#OBJBOOK" DATAFLD="COPY" DATAPAGESIZE=10 > <TR><TD></TD> <TD> </TD> </TR> </TABLE> </TD> </TR> </TABLE> </BODY> </HTML></pre>	<pre><?xml version="1.0" ?> <BOOKLIST> <BOOK> <CODE>JO001</CODE> <TITLE>Java Programming</TITLE> <COPY> <NO>1</NO> <REGDATE>1999-04-23</REGDATE> </COPY> <COPY> <NO>2</NO> <REGDATE>1999-05-12</REGDATE> </COPY> </BOOK> </BOOKLIST></pre>

Notice that the nested elements have to be put at the same level i.e.

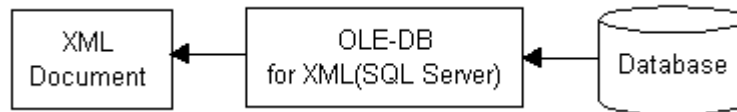
Original with the <COPYLIST> element	Allow for data binding in the HTML
<pre> <COPYLIST> <COPY> <NO>1</NO> <REGDATE>1999-04-23</REGDATE> </COPY> <COPY> <NO>2</NO> <REGDATE>1999-05-12</REGDATE> </COPY> </COPYLIST> </pre>	<pre> <COPY> <NO>1</NO> <REGDATE>1999-04-23</REGDATE> </COPY> <COPY> <NO>2</NO> <REGDATE>1999-05-12</REGDATE> </COPY> </pre>

Output

JO001	Java Programming	1	1999-04-23
		2	1999-05-12

4.2.9 SQL Server OLE-DB XML Interface

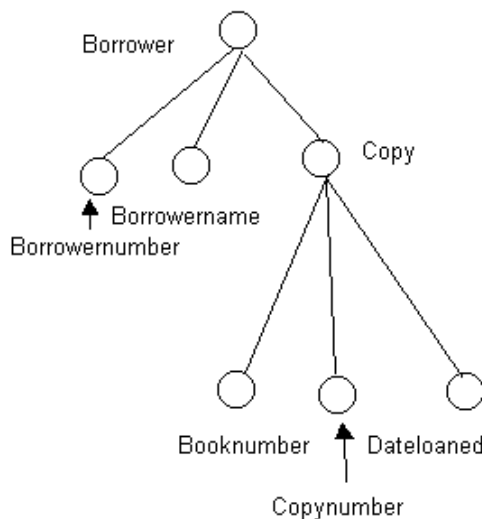
This is an interface provided by Microsoft that allows even SQL Server 7.0 databases to be read and the result returned as an XML document. SQL2000 recognizes these modifiers directly.



The SQL SELECT has the extra modifiers added to return the result as an XML document.

SELECT statement FOR XML { RAW | AUTO | EXPLICIT }

FOR XML RAW	Each row of data is created with the columns as attributes with the table name prefixed and the element name row e.g. <row BOOK_NO="CP0002" BOOK_TYPE="Computing " TITLE="ELEMENTARY FTAM AND MOTIS/MHS "/>
FOR XML AUTO	Auto is like raw except the table name becomes the element name e.g. <BOOK BOOK_NO="CP0002" BOOK_TYPE="Computing " TITLE="ELEMENTARY FTAM AND MOTIS/MHS "/>
FOR XML EXPLICIT	The user has to map the tree structure required. This is normally used with a JOIN statement to relate multiple tables.



```

SELECT 1 as tag,
  NULL as parent,
  Borrower.Borrower_no as [Borrower!1!Borrowernumber],
  Borrower.name as [Borrower!1!Borrowername],
  Null as [Copy!2!Booknumber],
  copy.copy_no as [Copy!2!Copynumber],
  Null as [Copy!2!Dateloaned]
FROM borrower,copy
WHERE copy.date_loaned is Not null AND
copy.Borrower_no = Borrower.Borrower_no
UNION ALL
SELECT 2 as tag,
  1 as parent,
  NULL,
  NULL,
  copy.Book_no,
  copy.copy_no,
  copy.date_loaned
FROM copy,book
WHERE copy.date_loaned is Not null AND
copy.Borrower_no = Borrower.Borrower_no
FOR XML EXPLICIT
  
```

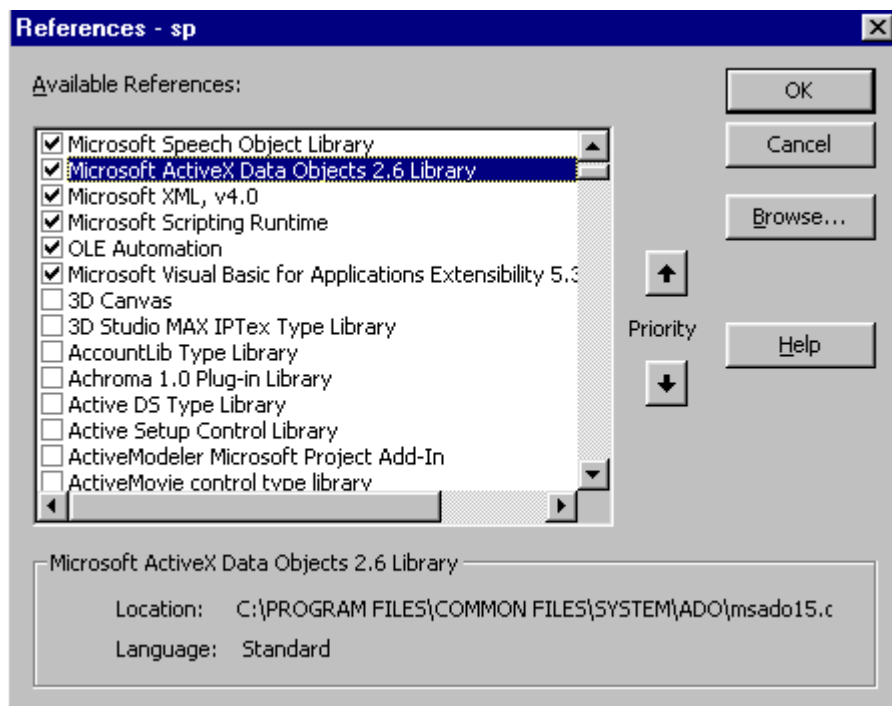
Tag	Unique number referring to the level
Parent	Parent's level
[element name!level!nested element name]	This defines the elements and the grouping it belongs to each section is separated by an exclamation mark and the whole definition is enclosed within square brackets

Example

[Borrower!1!BorrowerName]

The parent element is Borrower at the topmost level i.e. 1 and the element contained is BorrowerName

You need to have the Microsoft ActiveX Data Objects 2.6 reference or above



Sample code to read the database and output the result as an XML document

```

Sub xmlmodes()
Dim i As Integer
Dim wstr As String
Dim sqlstr As String
Dim conobj As New ADODB.Connection
Dim comd As New ADODB.Command
Dim STREAMOBJ As ADODB.Stream

```

```
Set STREAMOBJ = New ADODB.Stream
Set comd = New ADODB.Command
```

```
constr = "PROVIDER=SQLXMOLEDB.2.0;DATA PROVIDER=SQLOLEDB;DATA
SOURCE=NITR215;INITIAL CATALOG=Sales;UID=SA"
```

```
,
' OPEN DATABASE FROM CONNECTION OBJECT
,
```

```
conobj.Open constr
,
```

```
' ASSOCIATE CONNECTION OBJECT TO COMMAND OBJECT
,
```

```
comd.ActiveConnection = conobj
,
```

```
' INFORM DB XML STREAM IS USED ON CLIENT SIDE
,
```

```
comd.Properties("ClientSideXML") = "True"
,
```

```
' PREPARE SQL STATEMENT
,
```

```
sqlstr = "SELECT BOOK_NO, BOOK_TYPE, TITLE " & _
" FROM BOOK " & _
" WHERE TITLE LIKE '%ELE%' " & _
" FOR XML RAW "
,
```

```
' Set the SQL statement to the Command object
,
```

```
comd.CommandText = sqlstr
,
```

```
' OPEN STREAM
,
```

```
STREAMOBJ.Open
,
```

```
' ASSOCIATE STREAM OBJECT TO COMMAND OBJECT
,
```

```
comd.Properties("Output Stream").Value = STREAMOBJ
,
```

```
' Define root name
,
```

```
comd.Properties("xml root") = "root"
,
```

```
' LOCATION TO START
,
```

```
comd.Properties("xml root") = "root"
```

```

' EXECUTE THE DATABASE ACCESS
'
comd.Execute , , adExecuteStream

STREAMOBJ.Position = 0
STREAMOBJ.Charset = "ISO-8859-1" ' "utf-8"

'
' Create the Document object
'
Set DOMOBJ = New DOMDocument

'
' Read the XML document created and put it in a string
'
wstr = STREAMOBJ.ReadText()
'
' Load the XML into the document object
'
DOMOBJ.loadxml wstr
'
' Save the XML document
'

STREAMOBJ.SaveToFile "d:\tmp\xmlexplicit3_layer.txt", adSaveCreateNotExist
'
' Close the stream
'
STREAMOBJ.Close

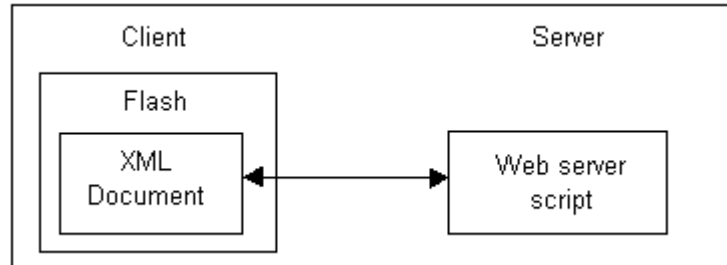
Set comd = Nothing
conobj.Close

Set DOMOBJ = Nothing
End Sub

```

4.2.10 Multimedia Front End (Flash)

(1) Multimedia front end (Flash)



There is an XML object that allows flash5 application to access XML documents or execute a script that returns an XML document.

Example of the asp script that creates an XML document

```

<% response.write "<?xml version='1.0' ?>"

response.write "<BOOKLIST>" & vbCrLf

response.write "<BOOK>" & vbCrLf

response.write "<BOOKNUMBER>BO0001</BOOKNUMBER>" & vbCrLf

response.write "<TITLE>ADVANCE STUDY OF ZOOLOGY</TITLE>" & vbCrLf

response.write "</BOOK>" & vbCrLf

response.write "<BOOK>" & vbCrLf

response.write "<BOOKNUMBER>BO0002</BOOKNUMBER>"

response.write "</BOOK>" & vbCrLf

%>
  
```

Flash script

```

var XMLOBJ;
function myload(success) // function prototype function name (success)
{ // success : true (successful)
  txXML = this.toString();      // display the XML document in the text box
}
//
// Button instance
//
on (release) {
  XMLOBJ = new XML();           // Creation of the XML object
  txXML = "loading"             // message to inform the user it is loading the XML
  XMLOBJ.onLoad = myload;       // function name to handle the downloaded data
  XMLOBJ.load ("/scripts/xml.asp"); // the direct XML file or the script that
  // generates the file
}

```

5 Database Technology

Chapter Objectives

Understanding the management of large databases and the types of applications that utilize such databases.

- ① Understanding the type of applications that require large databases
- ② Understanding the management of these large databases

Introduction

The rapid change in the pace of business and keen competition means that managers need to have not only accurate data but have it available in a time frame where effective action can be taken. In addition, the ability to do forecasting by analyzing the history of transactions becomes very important. This means there is a need to hold large amount of data over a time span in order to perform then analysis. In this chapter, we will look at the type of applications that require large volumes of data and how such databases can be managed.

5.1 Management of large databases

(1) Applications requiring large databases

The increasing volume of information available has led to the need for proper organization and storage. In addition, infra-structure changes like knowledge management, data mining and e-learning has also increase the requirement for large databases.

The following types of applications require large volumes of data

- Knowledge management
- Learning and Content Management systems
- Multimedia
- OLAP (Online Analytical Applications)
- Data mining

① Knowledge management

The recent trend is towards knowledge assets. These knowledge assets however cannot be properly utilized if there is no mechanism available for easy storage and retrieval. The assets come in the form of know how and procedures. This know how is derived from the information needs required to support the processes in the organization. The medium of the knowledge assets can come in a variety of formats

a. text documents

b. Pictures

c. Audio

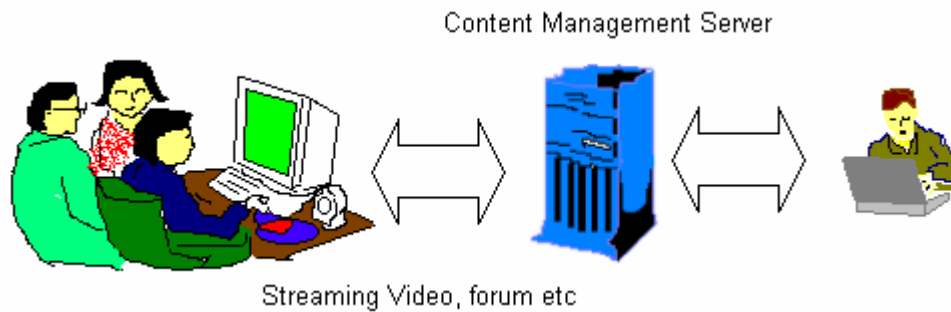
d. Video

These assets require a large volume for storage. Knowledge discovery refers to finding out new knowledge about an application domain using data on the domain usually stored in a database.

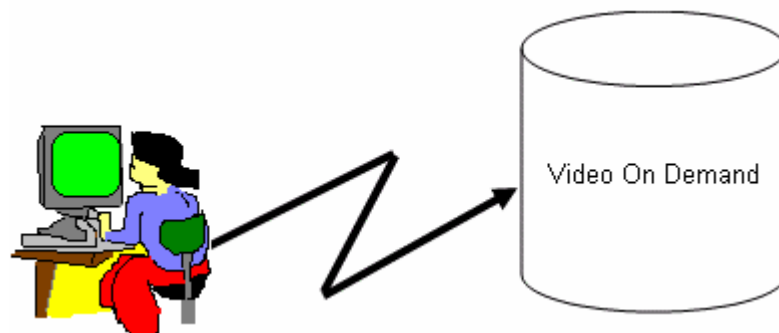
② Learning and Content Management system

There is a trend towards the introduction of e-learning. This means the contents have to be managed and stored. Many e-learning environments involved a heavy use of multimedia. This means there is a requirement for large storage space in the database.

③ Multimedia

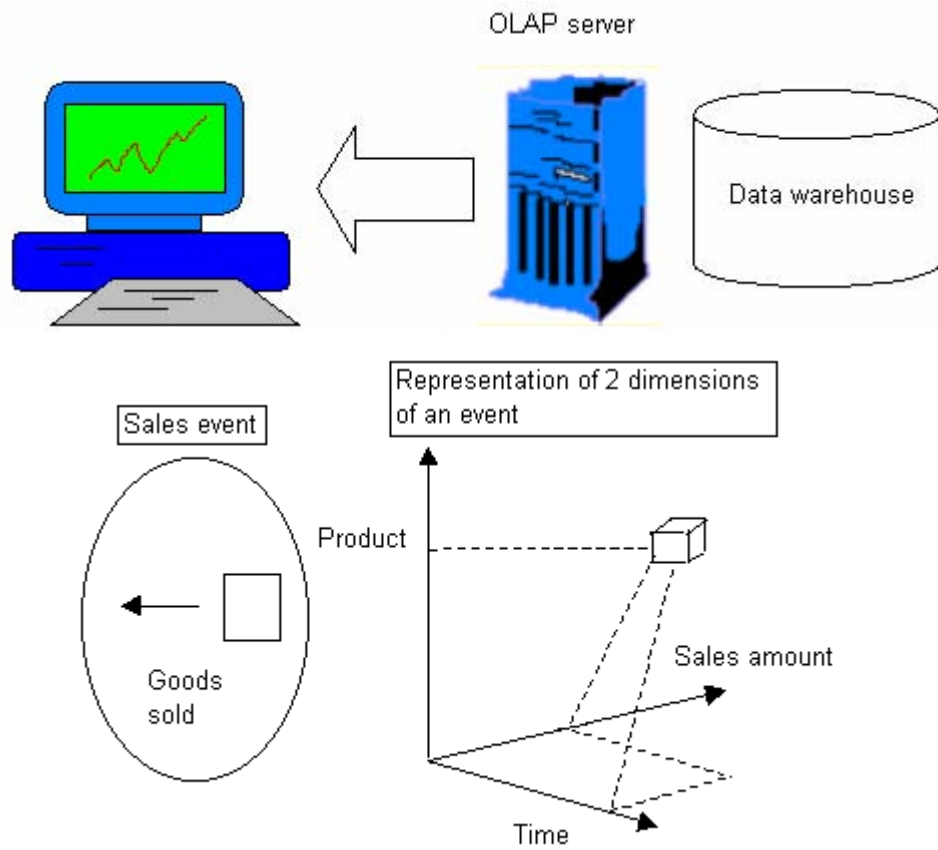


The nature of multimedia contents is the large size compared with text based contents. This applies especially to video contents. With the recent availability of low cost and high bandwidth networks, this means the delivery of these contents become feasible. However, as mentioned earlier, the size of the contents may lead to use of large storage space in the database.

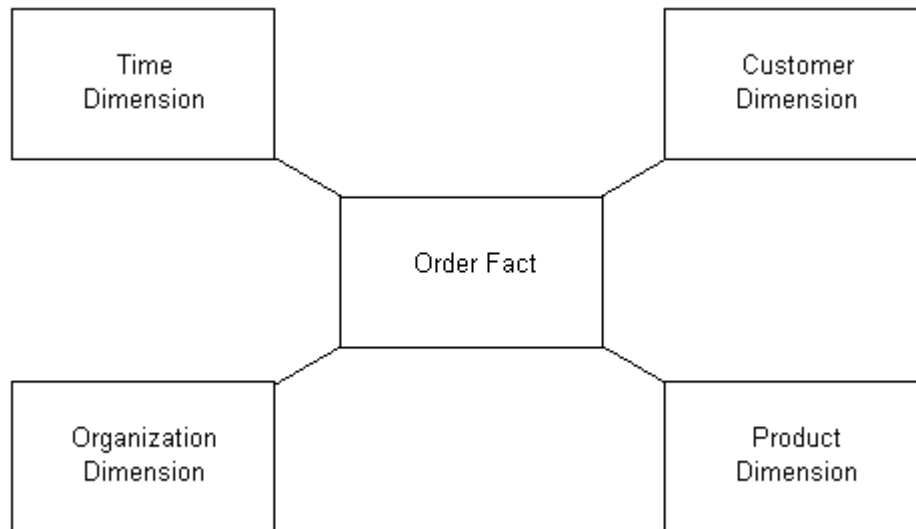


④ OLAP (Online Analytical Applications)

Many organizations are integrating their data into an enterprise data model. The enterprise data model is intended to support the business analyst and management. It offers a flexible snapshot of the existing financial and performance conditions in the organization. Moreover, data mining applications require the structure of a data warehouse to effectively seek for patterns. The data warehouse contains the history of the transactions organized in a format that allows users to navigate easily through the data.



Data warehouses record each occurrence of the process with its associated data. Every instance of the sales process is recorded showing when it occurred, what was sold and how much was sold. Data warehouses are defined as a star schema. The master tables in the OLTP database



become a dimension while the transaction tables represent the measures or fact tables.

Example of a star schema

a. Example of the Data volume in a data warehouse

Using the above star schema as an example, an estimate of the maximum number of records in the fact table is given

Table	Calculation
Time dimension	365 days x 4 processes
Product	Average 20 products per day sold
Customer	Average 10 customers per day
Organization	Average 6 staff servicing
Estimated Order fact records	$365 \times 4 \times 20 \times 10 \times 6 = 1,752,000$ records /year

The order fact is all possible intersections of the values from each dimension table

This means if queries were executed against the atomic fact table, the process time will degrade as the data volume increases.

Instead summary tables are created comprising the summary columns with the pre-calculated measures. Since most of the time the queries are done for a set of grouping conditions

b. Type of SQL used to access the tables

The predicates are used to access the data warehouse. The commonly used predicates are

1. SUM
2. AVG
3. Count

The SQL query has the form

SELECT grouping column, ..., Predicate or expression

FROM dimension table, ..., fact table

WHERE dimension table.warehouse key = fact table . corresponding warehouse key AND
dimension table.warehouse key = fact table . corresponding warehouse key, ..

[AND other filtering conditions]

GROUP BY grouping column, ..

[ORDER BY column,..]

Example

SELECT Year, Month, ProductClass, SUM(Quantity)

FROM timeTable, ProductTable, OrderTable

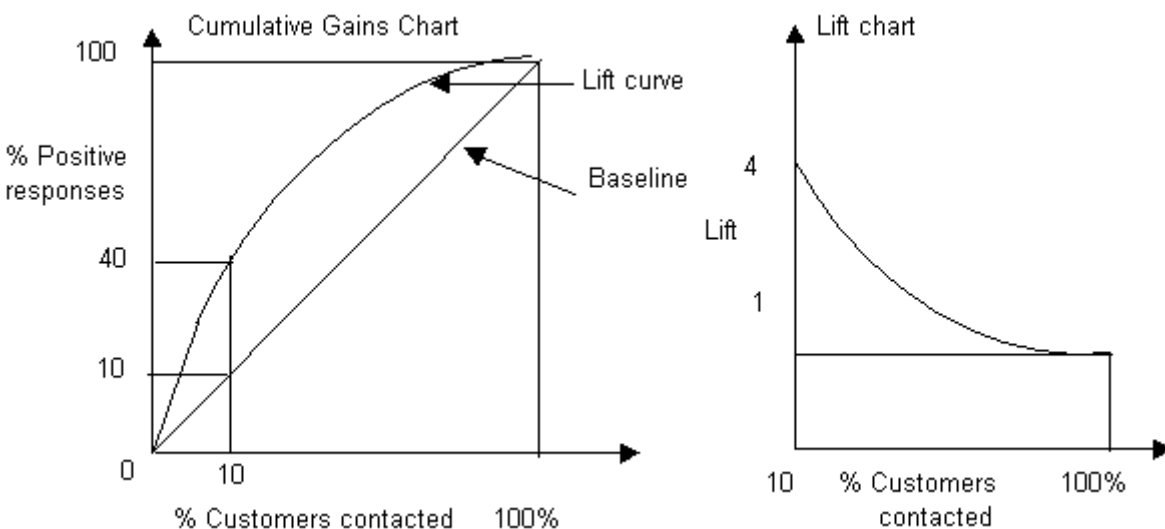
WHERE timeTable.timeKey = OrderTable.timeKey AND

ProductTable.ProductKey = OrderTable.ProductKey

GROUP BY Year, Month, ProductClass

⑤ Data mining

Data mining is defined as an information extraction activity whose goal is to discover hidden facts contained in databases. Data mining cannot be effective if there is insufficient data available. Data mining features are found in many databases. This is used to look for patterns in the historical data collected. These patterns are used to make predictions. This means the data warehouse is used as a source for the data mining algorithms to operate on. The use of the lift chart to show the responses with the use of a model. The baseline represents a 1 : 1 relationship. In the example shown, the lift for 10% customers contacted is $40/10 = 4$.



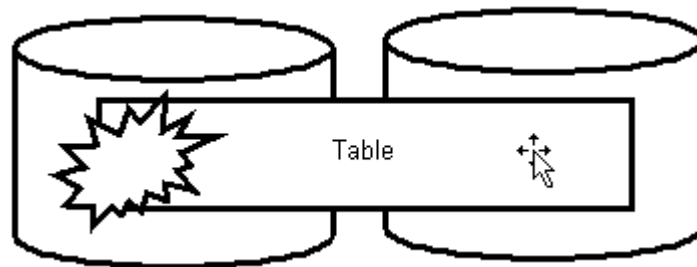
(2) Management of large databases

The following mechanisms are used to manage tables with large volumes of data.

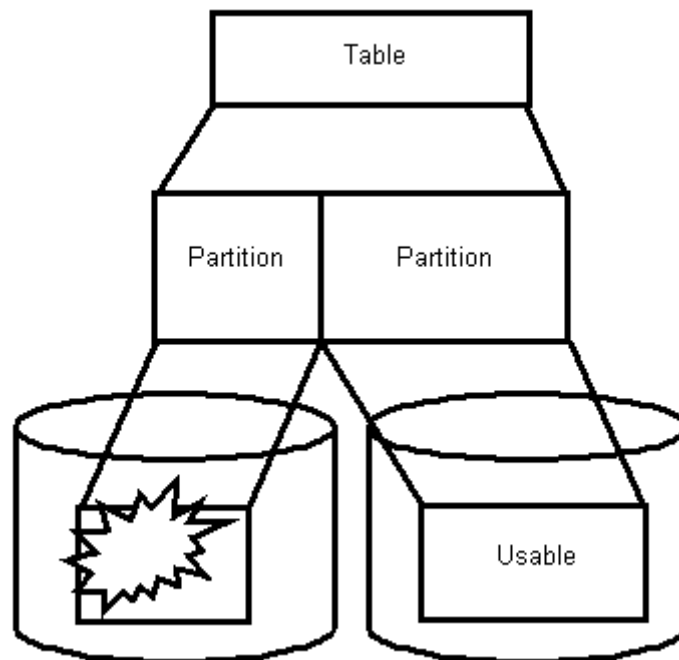
- Partitioning
- Use of bitmap indices
- Creation of summary tables
- Creation of data cubes

① Partitioning

If the table is defined as one logical unit, any physical failure on the device will render the table unusable. This means if one of the devices were to fail, it may cause the whole table to be inaccessible.



The use of partitioning allows for the table to be useable even if one of the disks were to fail. In addition, there is the added benefit of maintenance as partitions can be selectively saved and restored.



There are 2 kinds of partitioning supported

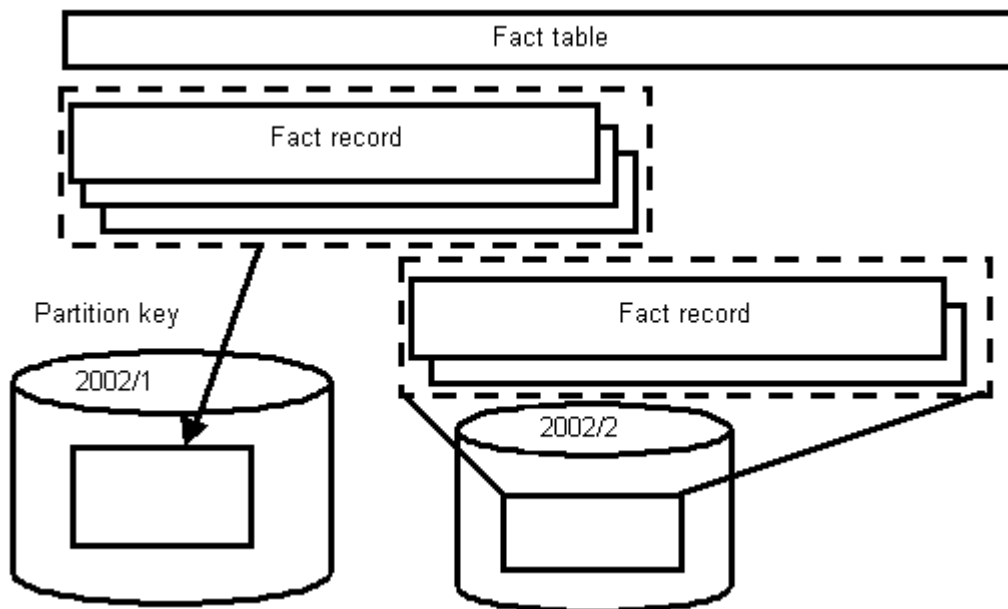
- Vertical
- Horizontal

a. Vertical partitioning

This is to separate a logical record into 2 parts with the binary data usually separate separated from the text data fields. This allows the text fields to be searched without the need for large buffer usage to accommodate the binary data.

b. Horizontal partitioning

This is commonly used in the implementation of data markets or data warehouses. One of the main uses of data warehouses is the analysis of the business. This means comparison is made across periods of time. The partition key is normally a time element.



② Use of bitmap indices

A special kind of index called a bit map index that utilizes bit map patterns instead of a whole byte to represent the index is available in some databases. As the range of summary column values are repeated, bit map indexes are commonly created for summary columns.

③ Creation of summary tables

Summary or aggregate tables contain summarized measurements together with the summary columns. They lose the connection to the original dimension table and the summary columns from these dimension tables are transferred into the summary table.

Sales value in \$	Quantity	Product type	Type of customer	Month	Quarter	Year

④ Creation of data cubes

A relational database can be used to model the enterprise data model. A special kind of structure that is optimized for storage is used for implementing data markets or warehouses. This is known as Multi-dimensional OnLine Analytical Processing database.

(MOLAP) Using only a relational database to implement the data warehouse is known as Relational Analytical Processing database (ROLAP). It is known as HOLAP(Hybrid) if a combination of both is used. The relational database is used to refresh the MOLAP.

The following OLAP databases are available

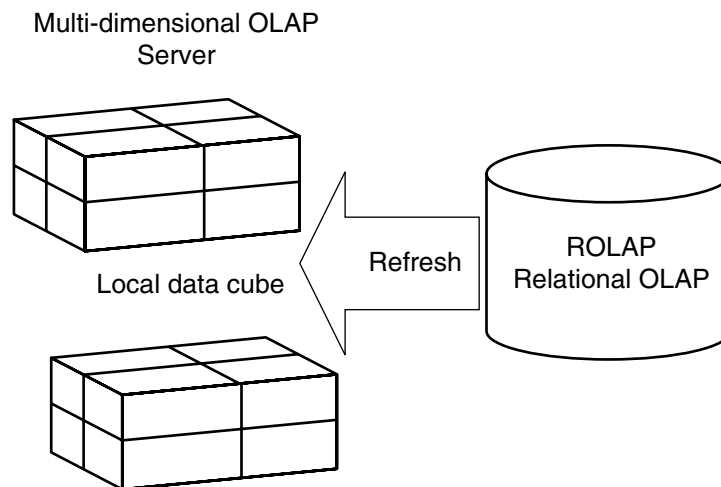
a. ROLAP

b. MOLAP

c. HOLAP

The ROLAP allows insertion and updating to be done easily while MOLAP is optimized for reading and compressed storage.

These data cubes can be directly linked to the spreadsheet like Excel for immediate analysis



Index

[Numerals]	52,53,56,57	G to B	15	Protected access	40
		G to C	15	Prototype	9
802.11	56	[H]		Private access	40
802.11b		Horizontal partition	99	[R]	
[A]				Region of access	40
ADSL	49	[I]		Relationship	42
Attributes	39,41,42	IP session hijacking	29	Readiness for e-government	27
		IP spoofing	29	Rules	36
		Infrared	52	Radio frequency	53
[B]		Intrusion detection system	35	[S]	
Benefits of e-government	24	Infrastructure based	93	Standards in e-learning	20
Basic Service Set	54	IEEE 802.11	69,70,74,75	Schema	96
Business process	16	IEEE 802.11b	74	Storyboard	6,8,9
B to B	14			Standards in e-business	16
B to C	15	[K]		Stylesheet	62,77
		Knowledge Management	93	SCORM	20,21
[C]		keyboard	56	Security policy	30
Cardinality	42	[L]		Security	28
Categories of e-government	22	LCMS	18	Server	50,51
Class	39	LMS	19	State	43,44
CMS	25	Loopholes	31	[T]	
Content aggregation	21			Types of threats	28
		[M]		Telecommunication	48
[D]		Management of large databases	93	[U]	
Database class	45	Multimedia project tea	4	Unicode	66
Data cubes	100	Mechanisms for protection	31	Unauthorized access	29
Data Island	37	Media Production	12		
Data mining	97	Methods	39	[V]	
Denial of Service	29	Microwave	53	Valid document	64
Digital certificates	24	Mobility	58	Vertical Partition	99
Digital Subscriber Line	49	Monitoring	33		
DSSS	53	Multimedia application	3	[W]	
DTD	70			Wireless LAN	52
		[N]		Well formed document	63
[E]		Namespace	69	Web log analysis	34
e-business	14			[X]	
e-Government	21	[O]		XML	61,95
e-learning	17	OLAP	93		
ebXML	16,26				
Encryption	33	[P]			
Extended Service Set	55	Public access	40		
		Partitioning	98		
[F]		Presentation	10		
FHSS	53	Problem domain classes	41		
Firewall	52				
[G]					
G to G	15				