King Fahd University of Petroleum & Minerals
College of Computer Sciences and Engineering
Computer Engineering Department

COE 444 - Internetwork Design and Management

Spring 2006 (Term 052)

Lecture Notes

Prepared by:
Dr. Mohammed Houssaini Sqalli

Course Topics:

1. *Overview of Computer Networks* **1 week**
Types of computer networks. LANs and WANs. Protocols and protocol families. The OSI reference model. The TCP/IP protocol.

2. *Internetworking* **3 weeks**
Basic terminology. Principles of internetworking. Types of internetworking devices. Repeaters, hubs, bridges, routers, switches and gateways. Transparent and source-routing bridges. Multilayer switches. VLANs. Routing strategies. Addressing.

3. *The Network Development Life Cycle* **1 week**
Network analysis. Network design methodology. Writing of a Request For Proposal (RFP) and quotation analysis. Prototyping/simulation. Implementation.

4. *Enterprise Network Design* **4 weeks**
Enterprise Network Design Model. Backbone design concepts. Network security and firewalls. Structured cabling systems. Case studies.

5. *Topology design and analysis* **3 weeks**
Topology design. Network design algorithms. Terminal assignment. Concentrator location. Traffic flow analysis and performance evaluation. Network reliability.

6. *Network Management* **2 weeks**
Network management standards & models. ISO Functional areas of management. Network management tools and systems. SNMP architecture & operations. Network administration.

7. *Project Presentations* **1 week**
More details will be posted on the course web site about the project.

Chapter 1 Overview of Computer Networks

Topics covered:
Types of computer networks. LANs and WANs. Protocols and protocol families. The OSI reference model. The TCP/IP protocol.

1.1 Terminology

- **Internetwork**: A collection of interconnected networks. It is also called "internet". The worldwide "Internet" is one specific internet. Examples are: LAN-to-LAN, LAN-to-WAN. These networks can be of different types: underlying hardware, protocols used, etc. One known application that runs on top of the Internet is the web which is a distributed system.

- **Design**: a preliminary sketch or outline showing the main features of something to be executed. The arrangement of elements or details in a product or work of art. ***[Merriam-Webster Dictionary]***

- **Network Design**: involves decisions on several issues including topology, architecture, flexibility and other cost and vendor related issues. The final product is a plan to be implemented including cabling, routing strategies, protocols to use.

- **Management:** the act or art of managing: the conducting or supervising of something. ***[Merriam-Webster Dictionary]***

- **Network Management**: The process of controlling a network so as to maximize its efficiency and productivity. ISO model defines 5 functional areas of network management: Configuration Management, Fault Management, Performance Management, Accounting Management, and Security Management.

1.2 Uses of computer networks

A distributed system is a software system built on top of a network.

1.2.1 Business Applications

- Issue of resource sharing
    - Physical resources (e.g., a high-volume networked printer)
    - Information: instant access to relevant documents
- Use of Client/Server model

- Networks provide a communication medium among employees
    - e-mail
    - Writing reports with a group of people located in different areas
    - Videoconferencing to organize meeting between employees at different locations.
    - e-business (e.g., suppliers and customers)
    - e-commerce: to do business with customers over the Internet (buying and selling)

1.2.2 Home Applications

The biggest reason for home computers is the Internet access.

The most popular uses of the Internet for home users are:

- Access to remote information: involves interactions between a person and a remote database. Examples: web, on-line newspapers, on-line digital library (e.g., www.acm.org)

- Person-to-person communication. Examples: email, instant messaging, chat rooms, newsgroups, telephone calls, videophone.

    - Peer-to-peer communication: eliminate the central database (different from the client/server model). Example: Napster.

- Interactive entertainment. Examples: video on demand, game playing.

- Electronic commerce: Examples: Home shopping, access to financial institutions (security issues), B2C (e.g., buying online), B2B (manufacturer orders from suppliers), G2C (Tax forms), C2C (Auctions), P2P (File sharing).

  B: Business, C: Client, G: Government, P: Peer.

1.2.3 Mobile users

- Mobile computers. Examples: notebook computers, personal digital assistants (PDAs)

- Portable office: portable electronic equipment used on the road. Examples: send and receive telephone calls, faxes, email, web, etc.

- Wireless networks: used in taxis, delivery vehicles, military, etc.
    - Fixed wireless
    - Mobile wireless

- Applications:
    - WAP (Wireless Application Protocol) to merge cell phones and PDAs
    - m-commerce
    - Location dependent services (e.g., nearby restaurant requested from network operator)

1.2.4 Social, Political, Ethical, and Moral Issues

- Newsgroups: topics discussed are controversial (may be offensive to some)
- The exchange of information (text, audio, video) may be inappropriate and unacceptable by some people
- Employer check e-mails sent by employees: find out if anything harmful for the company
- Government snoop all incoming & outgoing e-mail: spy on people to find specific information (e.g., illegal activities)
- Cookies on web browsers allow companies to track users' activities in cyberspace. May allow confidential info (e.g., credit card numbers) to leak all over the Internet.
- Anonymous messages
- Information on the Internet can be ill-informed, misleading, or wrong.
- Electronic junk mail (spam)
- Viruses can easily be sent contained in e-mails.

→ Security can solve a lot of these problems. Messages can be encrypted and authenticated (i.e,. right user). This can be costly.

1.3 Types of Computer Networks

Computer networks are frequently classified by:

- Transmission technology: broadcast and point-to-point.
- Geographical area they encompass: LAN, WAN.
- Type of communications path they use and the manner in which data are transmitted across this path: circuit-switched, packet-switched.

1.3.1 Transmission Technology

1.3.1.1 Broadcast Networks

A broadcast network consists of nodes that share a single communications channel. Short messages, e.g., packets, sent by one machine are received by all other nodes connected to the shared channel.

An address field (i.e., destination address) within the packet specifies the intended recipient. Upon receiving a packet, each node checks the address field. If the packet is intended for the receiving node, that node processes the packet; otherwise it ignores it.

- **Analogy:** shouting in a corridor, airport announcement, asking one student a question
- **Broadcasting:** Broadcast systems allow the possibility of addressing a packet to all destinations by using a special code in the address field.
- **Multicasting:** support transmission to a subset of the nodes in a network.

Broadcast networks employ several topologies such as Bus (e.g., Ethernet) and Ring (e.g., Token Ring). Satellite can also be classified as a broadcast system (downlink transmission).

1.3.1.2 Point-to-Point Networks

Point-to-point networks consist of many connections between individual pairs of nodes. For a packet to go from the source to the destination, it may have to visit one or more intermediate nodes. The connection to or from an intermediate node on the path from the source to the destination is called a "hop". One hop implies two directly connected nodes.

Often multiple routes, of different lengths, are possible. There are different routing algorithms to find a good route.

Large networks are usually point-to-point.

Three very common point-to-point topologies are:

- Star
- Loop
  - Complete loop
  - Fully meshed
- Tree

1.3.2 Geographical Area

- PAN (1-10m): computer for one person (e.g., wireless network connecting computer, mouse, keyboard, and printer)
- LAN (10m-10km): room, building, campus
- MAN (10-100km): city (e.g., cableTV)
- WAN (100-1000km): country, continent
- Internet (10000km): planet

1.3.2.1 Local Area Networks

A LAN has the following characteristics:

- LANs are restricted in size (in general < 10km) and have smaller scope (within a single building or campus). This implies that the worst-case transmission time is bounded and known in advance. This makes it possible to use certain kinds of designs and to simplify network management (e.g., configuration, fault, performance, security, and accounting).
- LANs run at speeds ranging from 10 Mbps to 10 Gbps.
- LANs are usually privately-owned networks (i.e., owned by the same organization)

- LANs are usually broadcast systems. They may use a transmission technology consisting of a cable to which all the machines are attached (e.g., Ethernet, Token ring).
- Two of the main topologies used in broadcast LANs are: Bus and Ring.
- Switched LANs and ATM LANs have appeared as well.

1.3.2.2 Wide Area Networks

A WAN has the following characteristics:

- Covers a large geographical area (100-1000 km)
- Rely in part on common carrier circuits (e.g., telephone company, ISP)
- Consists of a number of interconnected switching nodes

1.3.3 Type of Communication Path

1.3.3.1 Circuit switching

A dedicated communications path is established between two stations through the nodes of the network for the duration of a connection.

Example:

Problem:

1.3.3.2 Packet switching

Data are sent out in a sequence of small chunks, called packets. Each packet travels through the network from node to node.

Example:

Problem:

1.3.3.3 Frame relay

Takes advantage of high data rates and low error rates in the modern high-speed telecommunications systems by stripping out most of the overhead involved with error control.

Problem:

1.3.3.4 Asynchronous Transfer Mode (ATM)

ATM is an evolution of frame relay. ATM, in contrast to frame relay, uses fixed-length packets called cells (ATM is sometimes referred to as cell relay). A cell consists of 53 bytes. This reduces the overhead even further. Data rates in the range of 10s and 100s Mbps, and even in the Gbps range.

ATM also extends circuit switching to allow multiple channels with the data rate on each channel dynamically set on demand. ATM can offer constant-data-rate channels.

1.4 Protocols

- A protocol is an agreement between the communicating parties on how communication is to proceed (i.e., set of rules).
- A network architecture is a set of layers and protocols, e.g., OSI reference model and TCP/IP.
- Standards are needed to promote interoperability among vendors.

1.4.1 OSI Reference Model

The OSI (Open Systems Interconnection) reference model is an ISO (International Standards Organization) standard.

The OSI model is an architecture/structure that defines communication tasks and which would:

- serve as a reference model for international standards, and
- facilitate efficient internetworking among systems from different technologies, manufacturers, nationalities, and enterprises.

The OSI model architecture is layered to reduce complexity:

- Each layer offers certain services to the layer immediately above it.
- Each layer shields the higher layer from the details of implementation of how the services are offered.
- Layer N on one station carries on a conversation with layer N on another network station.

Layer functions:

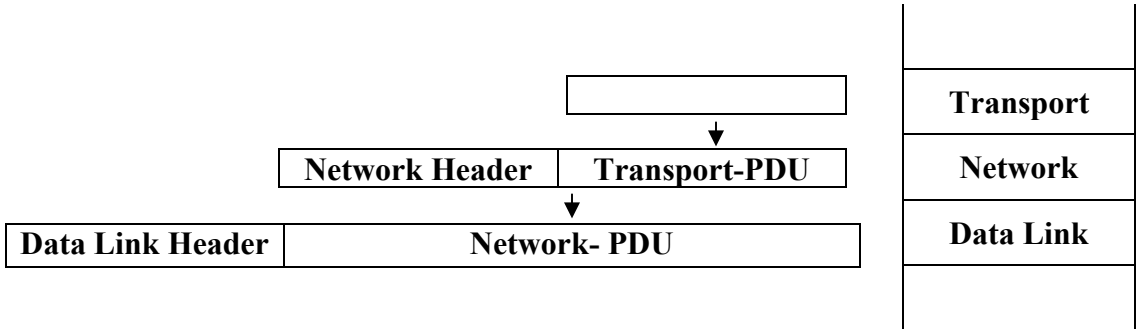
Application	ftp, telnet, email, web
Presentation	Data representation (ASCII, EBCDIC)
Session	Negotiation & connection (e.g., NFS)
Transport	End-to-end delivery → **Segments**
Network	Routing: addresses and best path → **Packets**
Data Link	Access to media → **Frames**
Physical	Binary transmission and cabling → **Bits**

Protocol entities exchange Protocol Data Units (PDUs). Each PDU contains:

- Header: contains control information to be used by the protocol at the peer layer.
- Message: Data from the upper layer.
- Trailer: is defined for some protocols

Header	Message	Trailer

The addition of control information to data is referred to as "Data Encapsulation"



1.4.2 TCP/IP Reference Model

(Referred to as TCP/IP protocol suite)

TCP/IP is a result of protocol research and development conducted on the experimental packet-switched network, ARPANET sponsored by DoD (U.S. Department of Defense).

Application
Transport
Internet
Host-to-Network

1.4.2.1 Transport Layer

Two end-to-end transport protocols have been defined: TCP and UDP.

- TCP (Transmission Control Protocol):
  - is a connection-oriented protocol
  - ensures end-to-end data reliability
  - uses segmentation and reassembly
  - handles flow control

- UDP (User Datagram Protocol):
  - is a connectionless protocol
  - unreliable (no sequencing or flow control)

1.4.2.2 Internet Layer (or Network Layer)

The internet protocol (IP) is used at this layer to provide the routing function across multiple networks.

This protocol is implemented in the end systems and in routers. A router primary function is to relay data from one network to the other.

1.4.2.3 Host-to-Network Layer

The TCP/IP reference model does not define explicitly what happens in this layer except that the host has to connect to the network using some protocol.

LAN protocols occupy the bottom two layers of the OSI reference model: the physical layer and the data link layer.

- The data link layer provides data transport across a physical link. It includes:
  - LLC: Logical Link Control sub-layer
  - MAC: Media Access Control sub-layer
- The physical layer specifies the electrical, mechanical, procedural and functional characteristics of the physical link between end systems.

1.5 References:

Chapter I of "Computer Networks" by Andrew S. Tanenbaum, Fourth Edition

Chapter 2 Internetworking

Topics covered:
Basic terminology. Principles of internetworking. Types of internetworking devices. Repeaters, hubs, bridges, routers, switches and gateways. Transparent and source-routing bridges. Multilayer switches. VLANs. Routing strategies. Addressing.

2.1 Terminology

Internetworking stands for connectivity and communication between two or more networks.

- Internetwork (internet): a collection of communication networks interconnected by bridges, switches and/or routers.

- Intranet: a corporate internet that provides key Internet applications. It is usually isolated and self-contained within an organization.

- End System (ES): a device attached to one of the networks.

- Intermediate System (IS): a device that connects two or more networks (e.g., switch, router). It is called sometimes an IWU (Internetworking Unit) or a relay.

2.2 Principles of Internetworking

2.2.1 Requirements for Internetworking

The overall requirements for an internetworking facility are:

1. Provide a link between networks. At minimum, a physical and link control connection is needed.
2. Provide for the routing and delivery of data between processes on different networks.
3. Provide an accounting service that keeps track of the use of the various networks and routers and maintains status information.
4. Provide the services just listed without requiring modifications to the networking architecture of constituent networks. This means accommodating the following differences:

    - Different addressing schemes: e.g., naming (DNS), DHCP.

    - Different maximum packet size: e.g., segmentation, ATM cells.

    - Different network access mechanisms: e.g., Ethernet, FDDI, ATM.

    - Different timeouts: longer with multiple networks.

- Different error recovery services: some networks will have it, others won't. Internetwork error recovery should be independent of individual networks.

- Different status reporting: how and whether this information can be shared.

- Different routing techniques: may depend on fault detection and congestion control techniques. Coordination is needed.

- Different user access control: authorization for use of the network.

- Connection-oriented vs. connectionless

Some of the above mentioned issues are dealt with in the IWUs.

It may be desirable for an internetwork service not to depend on the characteristics of individual networks.

By fulfilling these requirements, two important problems in interconnecting networks can be addressed:

- Heterogeneity of types of networks

- Scale of internetwork: routing and addressing issues with large growth.

2.2.2 Motivation for Internetworking

- Sharing of computer resources across a number of communications networks

- The use of multiple networks allows for network isolation when needed. This is critical to network performance as failure is contained within one network. Also, a network can be shielded from intrusion (Security).

- Contain the amount of traffic sent between the networks (e.g., Routing domains)

- Network Management that provides centralized support and troubleshooting capabilities in an internetwork.

2.2.3 Components of an Internetwork

- Campus Network: locally connected users in a building or group of buildings. It generally uses LAN technologies.

- Wide Area Networks (WANs): distant campuses connected together usually through connection providers such as a telephone company.

- Remote connections: linking branch offices and mobile users to a corporate campus. They are generally dial-up links or low bandwidth dedicated WAN links.

2.2.4 Routing domains

A routing domain is an administrative entity. Its goal is to establish boundaries for the dissemination of routing information.

- It is also useful for security administration.
- Provides accounting, billing, and revenue services (i.e., Accounting Management).
- Overcome the "flat network" problem by providing a routing hierarchy.

2.3 Internetworking Devices

Devices that interconnect LANs are known as relays and operate at one layer of the OSI model.

There are 5 common types of relays:

-
-
-
-
-

2.3.1 Repeaters (Hubs)

- Overcomes restrictions caused by single segment usage such as number of users, cable length.
- Amplifies or regenerates weak signals.
- Extends cable length.
- Connects LANs of similar type, but may use different media.
- Provides simple connection between adjacent LANs at the expense of increased network congestion.

2.3.2 Bridges

The bridge was designed for interconnection of LANs that use identical protocols at the MAC layer (i.e., layer 2). However, there are bridges capable of mapping between different MAC protocols (e.g., Ethernet and Token Ring).

A bridge main function is forwarding frames from one network to another. A bridge does the following:

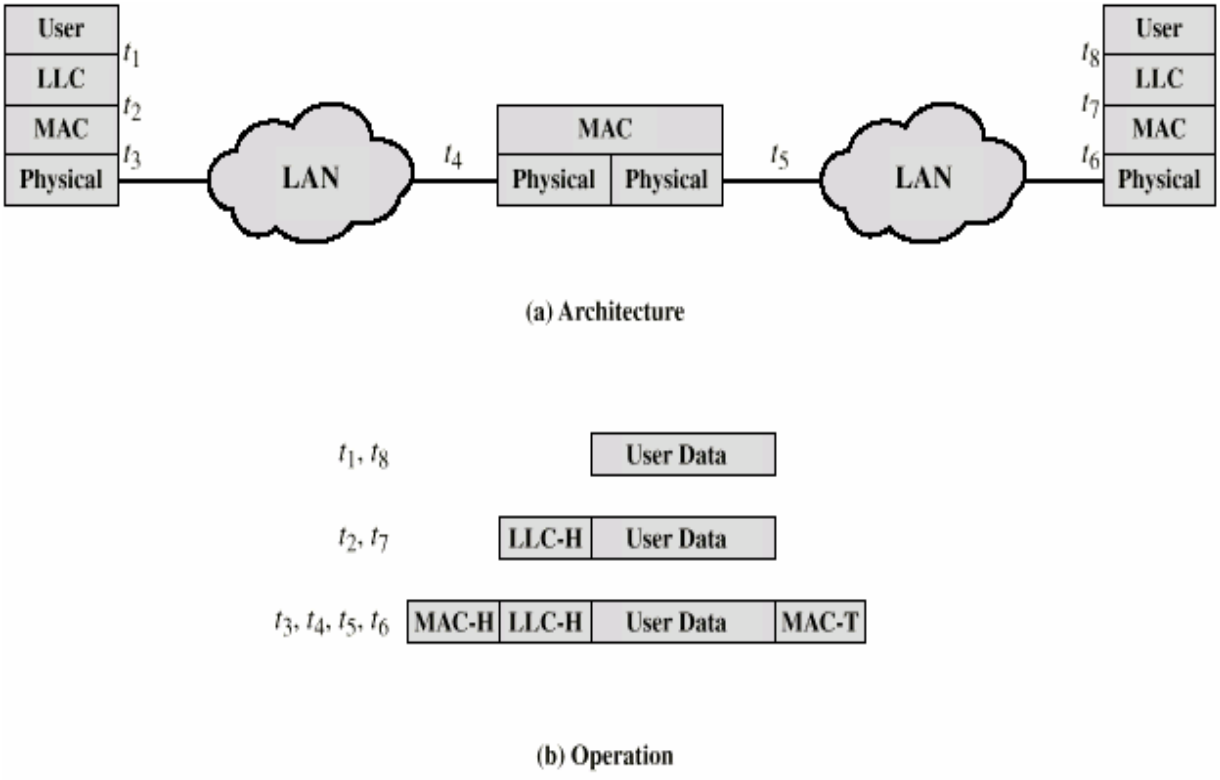


Figure 2-1: Connection of two similar LANs (Stallings)

Characteristics of bridges

- Interconnects two or more LANs (either similar or dissimilar) at the MAC level (e.g., Ethernet and Token Ring)
- Capable of deciding whether or not to forward a frame.
- Creates an extended network and keeps local traffic off.
- Can make minor changes to frame header.
- Does not inspect or modify the network layer packets inside frames.

Reasons for using bridges

- **Reliability**: fault is limited to the network where it happened.
- **Performance**: intra-network traffic stays within one network.
- **Security**: Types of traffic with different security needs are kept on physically separate media.
- **Geography**: LANs may need to be on separate locations.

Bridges have to make a routing decision

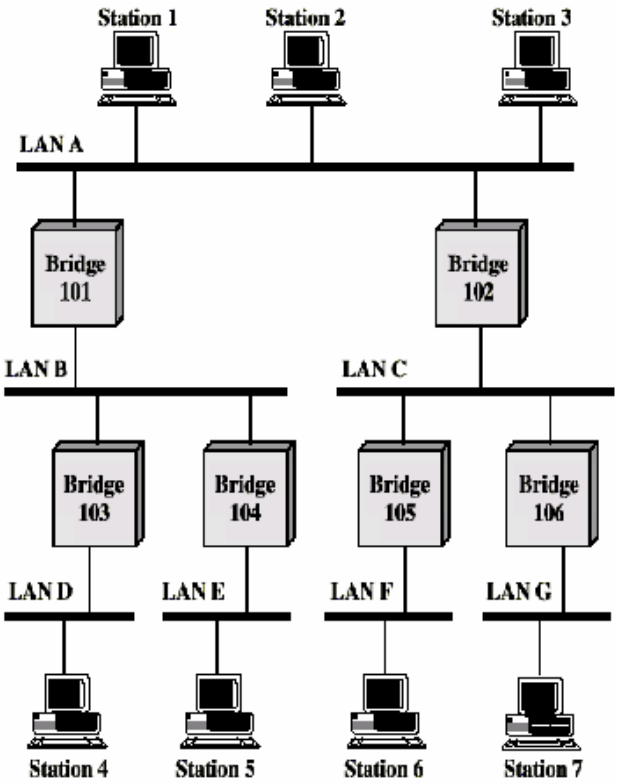


Figure 2-2: Multiple LANs (Stallings)

- S1 transmits a frame on LAN-A intended for S5. B1 and B2 will read the frame. Each one must make a decision of whether or not to retransmit the frame to other LANs. This continues until the frame reaches LAN-E where it is received by S5.

- The routing decision may not always be a simple one. If we add bridge B7 between LAN-A and LAN-E.

- B7 may fail.

-

Many routing strategies are used in bridges:

- Fixed routing
- Spanning tree routing (Transparent bridges)
- Source routing

2.3.2.1 Fixed routing

- A route is selected for each source-destination pair of LANs. If more are available, the one with the least number of hops is selected.
- A central routing matrix is created. It shows the identity of the first bridge on the route.

B1 table

From		From	
Dest	Next	Dest	Next

B2 table

From		From	
Dest	Next	Dest	Next

Advantages:

- Simplicity
- Minimal processing requirements

Disadvantages:

- Bridges can be dynamically added and failures may occur, so tables must change.

2.3.2.2 Spanning Tree Routing (Transparent bridges)

Transparent bridge characteristics:

- It is intended to interconnect LANs that satisfy any of the MAC standards without end stations being aware of its existence (i.e., transparent)

- The routing mechanism is the spanning tree algorithm

The bridge must map the content of the incoming frame into an outbound frame that conforms to the frame format for the outbound LAN, because MAC formats for the various LANs differ.

2.3.2.2.1 Frame Forwarding

A bridge maintains a filtering database. This information can be preloaded into the bridge (i.e., static routing).

2.3.2.2.2 Address Learning

The filtering database can be learned.

2.3.2.2.3 Spanning Tree Algorithm

Address learning is effective with a tree topology (i.e., no closed loop)

- **B1 and B2 know where S2 is:**

  1.

  2.

  3.

  4.

  5.

  6.

  7.

  Problem:

- **B1 and B2 do not know of S2 yet (→ worse problem):**

  1.

  2.

  3.

  4.

  5.

  Problem:

In graph theory: for any connected graph, consisting of nodes and edges connecting pairs of nodes, there is a spanning tree of edges that maintains the connectivity of the graph but contains no closed loops.

Algorithm:

- Each bridge is assigned a unique id

- A special group MAC address is used to send a frame to all bridges

- Each port of a bridge has a unique "port id"

- Each port of a bridge has an associated cost:

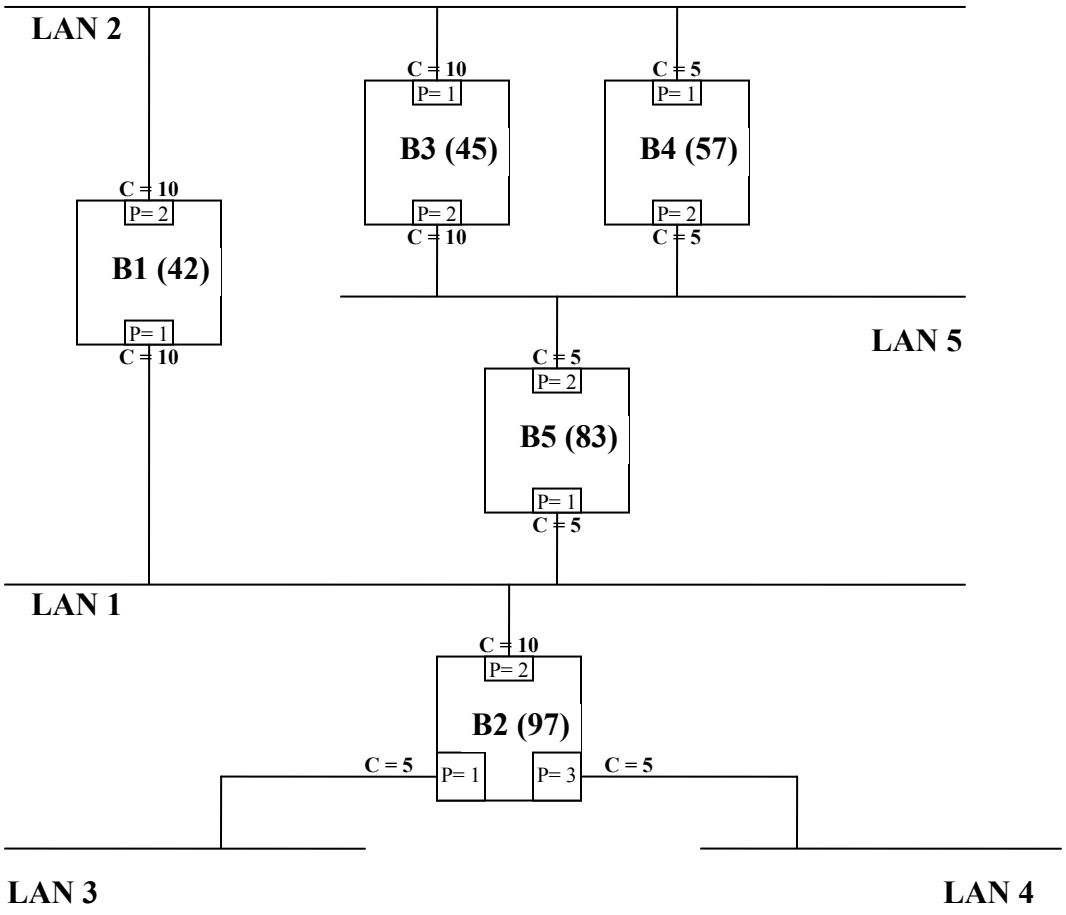
  Default Cost = 1000 Mbps / Data rate (in Mbps)

  Examples: 10 Mbps Ethernet → Cost = 100
  16 Mbps Token Ring → Cost = 62

  A network administrator can change a link's cost.

  For very high-speed LANs, a non-linear relationship between link cost and data rate is defined:

Data rate (in Mbps)	Cost range	Cost recommended value
4	100-1000	250
10	50-600	100
16	40-400	62
100	10-60	19
1000 (1G)	3-10	4
10000 (10G)	1-5	2



The spanning tree is constructed as follows:

1. Determine the root bridge (RB) that is the bridge with the lowest id.

2. Determine the root port (R) on all other bridges. This is the port used for the first hop on the minimum cost path to the root bridge. The lower port number is selected if more than one port exists.

   - The root path cost (rpc) is the cost of the path to the root bridge with minimum cost.

3. Determine the designated port (D) on each LAN. This is the port with the minimum rpc. If more bridges have the same rpc, the one with the highest priority is chosen as designated bridge (i.e., lowest-numbered bridge identifier).

   - The designated bridge is the bridge that provides the minimum cost path to the root bridge.

4. Ports which are neither (R) nor (D) are Blocking (B).

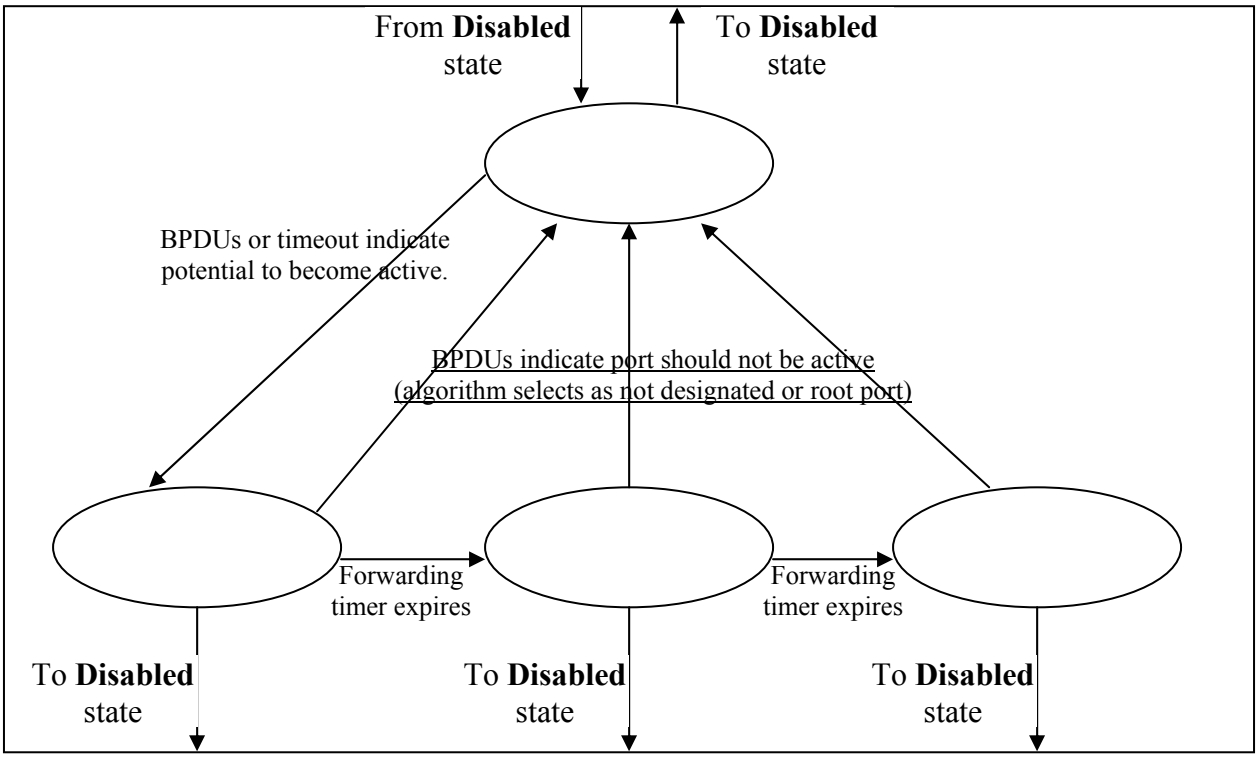
BPDUs (Bridge Protocol Data Units) are used to exchange information between bridges.

- BPDUs are sent by all the bridges each claiming to be the root bridge. B1 is elected as the root bridge.

- All other bridges determine the root port (R) and root path cost (rpc).

- Example: On LAN-5: B3, B4 and B5 send BPDUs claiming to be the designated bridge. B4 and B5 have the lowest RPC. B4 has a higher priority. B4 becomes the designated bridge (D).

2.3.2.2.4 Spanning Tree State Transition Diagram for a bridge port



The following is a table summarizing the actions taken by a bridge in each state.

	Receive BPDUs	**Transmit BPDUs**	**Learn addresses**	**Forward data frames**
Disabled				
Blocking				
Listening				
Learning				
Forwarding				

2.3.2.3 Source Routing Bridges

- Developed by IEEE 802.5 committee

The sending station determines the route to be followed by a frame and includes routing information with this frame.

Each frame includes the type of routing desired:

- Null: no routing desired.

- Nonbroadcast: the frame includes a single route using LANs and bridges.

- All-routes broadcast: the frame will reach each LAN (and the destination station) by all possible routes.

To avoid looping:

- Single-route broadcast: the frame will appear once on each LAN. The frame is forwarded to bridges on the spanning tree with source node as root. The spanning tree is built automatically or manually. The destination receives one copy.

All-routes broadcast and Single-route broadcast types of routing are used to discover route to destination. They are also used for group and all-stations addressing.

Route discovery and selection:

There are three options:

- Manually load information into each station.
    - Problem:

- Stations in the same LAN exchange routing information.
    - Problem:

- Dynamic route discovery procedure by stations

Two approaches are possible:

1.
    - Source station transmits an all-routes request to destination.
    - Destination sends back a nonbroadcast response on each discovered route.
    - Source uses one of these to send subsequent frames.

        - Problems:

2.
    - Source station transmits a single-route request.
    - Destination responds with an all-routes response.
    - Source chooses one for subsequent transmissions.

2.3.2.4 Spanning Tree vs. Source Routing

Characteristics	Transparent bridging	Source route bridging
Transparency		
Topology knowledge		
Frame format		
Frame forwarding		
Bridge mode		
Data Link operation		
Link utilization		
Configuration (LAN numbering, bridge numbering, spanning tree, etc)		
Performance		
Routing		

2.3.2.5 Source Routing Transparent (SRT) Bridges

A key problem is that both (transparent and source routing bridges) are incompatible. To allow the interconnections of LANs using a mixture of transparent and source routing bridges, a new standard was developed by the IEEE 802.5 committee, and that is the Source Routing Transparent (SRT) technique.

2.3.3 Routers

2.3.3.1 Motivation

Bridges do not stop broadcast traffic. This can lead to broadcast storms (e.g., more than 100 non-unicast frames/sec) which can be catastrophic. This can bring the network down.

Some sources of broadcast traffic:

- Address resolution (e.g., ARP, RARP, BOOTP)
- RIP (Routing Information Protocol)
- DHCP (Dynamic Host Configuration Protocol)
- IPX (Internet Packet eXchange) generates broadcast traffic to advertise services and routes
- Netware clients rely on broadcast to find services
- Appletalk: Route discovery protocol

To contain/reduce broadcast traffic, we need to reduce the size of the network (i.e., LAN).

Two approaches are used to do this:

- Use routers to subnet the LAN
- Use VLANs (Virtual LANs)

2.3.3.2 Characteristics

- A router separates traffic of different networks. It does not flood packets.
- Routers route packets at the network layer (layer 3)
- Routers route packets based on the contents of a routing table.
- Routing tables contain a mapping of a destination to a port. They can be static or dynamic.
- Routers "learn" their routing table entries by communicating with their routing peers.
- Routing protocols are used to implement routing (RIP, OSPF, BGP, PNNI)
- Routers perform routing decisions on the basis of the Network ID part of the destination IP address.
- The Host ID part of the destination address is used by the destination router to determine the destination station.

2.3.3.3 IP Addressing

2.3.3.3.1 IP Address Structure

IP address = Network ID + Host ID (32 bits)

Class A:

0	Network ID	Host ID
	7 bits	24 bits

Address range: **1.0.0.1 → 126.255.255.254**
Max. number of networks: **126**
Max. number of hosts: **16,777,214**

Class B:

1 0	Network ID	Host ID
	14 bits	16 bits

Address range: **128.0.0.1 → 191.255.255.254**
Max. number of networks: **16,384**
Max. number of hosts: **65,534**

Class C:

1 1 0	Network ID	Host ID
	21 bits	8 bits

Address range: **192.0.0.1 → 223.255.255.254**
Max. number of networks: **2,097,152**
Max. number of hosts: **254**

Class D:

1 1 1 0	Multicast address
	28 bits

Address range: **224.0.0.0 → 239.255.255.255**

Class E:

1 1 1 1 0	Reserved for future use
	27 bits

Address range: **240.0.0.0 → 247.255.255.255**

Note: The Internet Network Information Center (InterNIC: www.internic.net) assigns IP addresses

Private allocations:

In **RFC 1918**, several IP addresses have been allocated for private addressing. An organization can use these addresses if they are not registered with the Internet. Systems are available that translate private, unregistered addresses to public, registered addresses.

Class A addresses:	10.x.x.x → 10.x.x.x	⇨ 1 network
Class B addresses:	172.16.x.x → 172.31.x.x	⇨ 16 networks
Class C addresses:	192.168.0.x → 192.168.255.x	⇨ 256 networks

2.3.3.3.2 Address Resolution

Address Resolution Protocol (ARP) and the relationship between IP and MAC addresses:

2.3.3.3.3 Subnetting

Subnet Address Structure:

Example of Class B network:

16 bits	8 bits	8 bits
Network ID	Subnet ID	Host ID

Subnet mask: 11111111 11111111 11111111 00000000

1s: identify the network address portion of the IP address.
0s: identify the host address portion of the IP address.

IP routing algorithms are modified to support subnet masks (subnet addressing)

- One problem is how to store, maintain and access many network addresses in one routing table. → The Internet establishes a scheme whereby multiple networks are identified by one address entry in the routing table.

Address aggregation:

Address aggregation is used to reduce the size of the routing tables.

How is subnet mask interpreted?

IP address(Class B)	**128.**	**1.**	**17.**	**1**
Mask	255.	255.	240.	0
IP address (binary)	10000000	00000001	00010001	00000001
Mask (binary)	11111111	11111111	11110000	00000000
Result (Logical AND)	10000000	00000001	00010000	00000000
Logical address	**128.**	**1.**	**16.**	**0**

This subnet address is **128.1.16.0/20** (with 16 bits Network ID, 4 bits Subnet ID, and 12 bits Host ID).

2.3.3.3.4 CIDR - Classless InterDomain Routing ("Supernetting")

- **RFCs: 1518, 1519, 1466, 1447**. (http://www.rfc-editor.org/)

It permits networks to be grouped together logically, and to use one entry in a routing table for multiple class C networks.

2.3.3.4 Key Routing Strategies

2.3.3.4.1 Fixed Routing

A single, permanent route is configured for each source-destination pair of nodes in the network (A least-cost routing algorithm could be used to configure routes). Link costs are based on static variables such as expected traffic or capacity.

Problem:

2.3.3.4.2 Flooding

A packet is sent by a source node to every one of its neighbors and each node retransmits it again to its neighbors (similar to "all-routes broadcast" in source routing bridges). The flooding technique has three properties:

- All possible routes are tried, and there is always a backup route (good for emergency messages)
- One copy of the packet will reach destination by following a minimum-hop route (can be use to setup virtual circuits)
- All nodes are visisted (disseminate information to all nodes)

Problem:

2.3.3.4.3 Random Routing

A node selects only one outgoing path chosen at random for retransmission of an incoming packet.

Problem:

2.3.3.4.4 Adaptive Routing

Routing decisions that are made are updated as conditions on the network change (e.g., failure, congestion). Information about the state of the network must be exchanged.

Problems:

- More complex routing decision.
- Information exchanged is itself a load
- Reaction to changes can be too quick or too slow.

However:

- Adaptive routing can improve performance from the user perspective.
- Adaptive routing can aid in congestion control, because it tends to balance load.

2.3.3.5 Definitions

- Autonomous System (AS):
  - Consists of a group of routers exchanging info via a common routing protocol.
  - A set of routers and networks managed by a single organization.
  - Is connected (i.e., a path exists between any 2 nodes) except in time of failure.

- Interior Router Protocol (IRP, IGP)
  - Passes routing information between routers within an AS (e.g., RIP, OSPF).

- Exterior Router Protocol (ERP/EGP)
  - Passes routing information between routers in different ASes (e.g., BGP)

2.3.3.6 Routing Protocols

2.3.3.6.1 RIP (Routing Information Protocol)

- RFC 1058

RIP is:

- An IRP
- A distance-vector protocol
- A widely used protocol because of its simplicity and ease of use
- Based on the number of intermediate hops to destination
- Based on Bellman-Ford algorithm
- A distributed adaptive algorithm
- Maximum number of hops between a source and destination is 15
- Routing information is sent every 30 seconds to all adjacent routers using broadcast frames.

A distance of **1** means a directly connected network, and a distance of **16** means unreachable network.

Some major problems with RIP are:

- "Count to infinity" and there are several partial solutions to this problem such as "Split Horizon"
- Update of changes in the network is very slow.

2.3.3.6.2 OSPF (Open Shortest Path First)

- RFC 2328

OSPF:

- Is an IRP
- Is a link-state routing protocol
- Is based on Dijkstra's algorithm
- Is a distributive adaptive algorithm
- Routers send link state packets (LSPs) that include information about the cost of each of its links/interfaces
- Relies on two mechanisms:
  - Reliable flooding: the newest information must be flooded to all nodes as quickly as possible, while old information must be removed from the network.
  - Route Calculation: Each node gets a copy of the LSP from all nodes and computes a complete map for the network topology. Then, it decides the best route to each destination.
- Uses flexible routing metrics: distance, delay, cost, etc.
- Allows for scalability
- Uses multiple paths to allow for load balancing
- Supports security measures

2.3.3.6.3 BGP (Border Gateway Protocol)

- RFC 1771 (BGP-4)

- BGP:
  - Is a replacement for EGP (Exterior Gateway Protocol). EGP had limitations that include forcing a tree-like topology onto the network.
  - Provides inter-domain routing.
  - Is more concerned with reachability than optimality.
  - Is the routing protocol employed on the Internet.

- Challenges:
  - Lot of routing information to pass (~90,000 prefixes/routes in BGP routing tables.)
  - Autonomous nature of the domains (different than IRPs). Cost metrics are not the same and don't have the same meaning across ASes.
  - Trust between different providers (e.g., wrong configuration in an AS, competitors, etc.)

- BGP operates with networks with looped topologies.

- It runs on a reliable transport layer protocol (e.g., TCP).
- Each AS is identified by an AS number.
- BGP considers the Internet as a graph of ASes.
- How BGP works:
    - The administrator of each AS picks at least one node to be a "BGP speaker"
    - "BGP speakers" exchange reachability information among ASes.
    - BGP advertises complete paths as an enumerated list of ASes to reach a particular network.
    - Each AS has one or more border gateways.

- BGP prevents the establishment of looping paths (because it uses the complete AS path)
- BGP supports CIDR and address aggregation.
- BGP supports negative advertisement (i.e., withdrawn route) to cancel path(s).
- EBGP: operates between ASes.
- IBGP: is used to tunnel a user's traffic through a transit (pass-through) AS.
- BGP uses policy-based metrics. (RFC 1655: BGP policy-based architecture). Policies include various routing preferences and constraints, such as economic, security, or political considerations. (e.g., preference of internal routes over external routes).

2.3.4 Switches

Switching combines advanced microprocessor technology with the concept of a layer-2 bridge.

Whatever we have said about bridges apply to switches (i.e., a switch is a bridge is a switch).

Sometime the difference between a bridge and a switch is looked at as a marketing distinction rather than a technical one.

A switch has bridge's functionality:

- Learning (generally dynamic)
- Address table (forwarding table) including timers.
- Flooding when destination is unknown.

It can be said that a switch is a high-speed multi-port bridge. A large switch can have more than 100 interfaces.

2.3.4.1 Types of Switches

- **Port switches:** repeaters

- **Switches:** operate at layer 2. They leverage transparent bridging. Typically one port provides a high speed uplink to the backbone.

- **Layer-3 switches (i.e., multilayer switches):** include properties of layer-2 switches and some layer-3 capabilities (i.e., routing capabilities). They use the philosophy of "Switch (bridge) where you can, route where you must".

- **Layer-4 switches:** It does not implement layer-4 functionality, but it prioritizes certain classes of application traffic. Applications are identified using TCP port number.

2.3.4.2 Inside a switch

Switching fabric refers to the hardware and software design of the switch. ASICs (Application Specific Integrated Circuits) and DSPs (Digital Signal Processors) are used to implement switching fabrics.

Two methods of switch operation:

- **"Store-and-forward" switches:**
  - Buffer data.
  - Check for CRC (Cyclic Redundancy Check) errors.
  - Filter out frames

  Problem:

- **"Cut-through" switches:**
  - Frame header is read.
  - Data is switched without being buffered.
  - Only works if both the input and output ports operate at the same data rate.

  Problems:

  Comparison:

Parameters in switches:

- **Backplane speed:** Internal capacity of a switch. It must exceed the summation of all ports capacities, otherwise blocking and frame dropping will occur.
- **Memory:** Used for buffering data. If it is not enough, then frames dropping will occur.

Switch features:

- **Filtering:** Switches, in contrast to traditional bridges, can filter traffic (i.e., forward traffic conditionally) by interpreting the frame beyond the SA (Source Address) and DA (Destination Address). E.g., layer-3 switches.

  Filters can be complex and may result in performance degradation.

- **Forwarding table:** If the size of this table is exceeded constantly, entries are deleted prematurely and lots of flooding of frames will happen.

- **Oversubscription:** where aggregate bandwidth at the leaves exceeds that of the trunk.

2.3.4.3 Layer-3 Switches

They carry the image of switching as high-performance, cost-effective, hardware-based internetworking, together with the feature set associated with network-layer protocols.

(See the internetworking product timeline in table 4.1 of "The Switch Book".)

Operation:

The switch architecture can be optimized for functions that must be performed in real-time, for the majority of packets, known as the **fast path** of the flow.

- Fast path:

A layer-3 switch needs to implement only this fast path in hardware, e.g., implement hardware-based routing for IP.

- Because

Other protocols can be implemented in software.

Exception conditions can also be implemented in software.

The IP fast path:

- Subnet mask represented using 5 bits: used for high-speed routing table lookup operations.

- Packet parsing and validation.

- Routing table lookup.

- Mapping the destination to a local data link address (ARP mapping)

- Update lifetime Control and Checksum

- Fragmentation **is not** usually implemented in the fast path.

2.3.4.4 Virtual Local Area Networks (VLANs)

- VLANs enable the creation of logical groups of network devices across a network.
- Bandwidth Preservation: The broadcast traffic is contained within each VLAN
- LAN Security: VLANs allow for traffic isolation.
- User Mobility: VLANs allow for more flexibility in the positioning of end stations and servers, and reduce the effort of adds, moves, and changes:
    - They can be placed physically anywhere in the building and still remain in the same logical LAN (i.e., VLAN).
    - They can be placed physically in the same location but move to a new logical LAN.
- VLANs are used to partition a flat bridged network using of these techniques:
    - **MAC Address Grouping:** VLAN membership is determined by the device MAC address.
    - **Port Grouping:** A VLAN is a collection of ports across one or more switches. A device attached to one of these ports is a member of this VLAN.

- **Protocol Grouping:** A VLAN group is based on protocol type (e.g., IP) or on network address.

- Some issues with VLANs:

2.3.5 Brouters and Gateways

- **Brouters:** another name for layer-3 switches.

- **Gateways:** more complex as they interface between two dissimilar networks (operates above layer-3). They are necessary when two networks do not share the same network layer protocol.

2.4 References

1. "Data and Computer Communications" by William Stallings, 6th Edition, Prentice Hall, 2000

2. "Computer Networks - A Systems Approach" by Peterson and Davie, 2nd Edition.

3. "Local & Metropolitan Area Networks" by William Stallings, 6th Edition, Prentice Hall, 2000

4. "The Switch Book" by Rich Seifert. John Wiley & Sons Inc., 2000.

5. "Computer Networks" by Andrew S. Tannenbaum, 4th Edition, Prentice Hall, 2003

6. "LAN Technologies Explained" by Philip Miller and Michael Cummins. Digital Press, 2000

Chapter 3 *The Network Development Life Cycle*

Topics covered:
Network analysis. Network design methodology. Writing of a Request For Proposal (RFP) and quotation analysis. Prototyping/simulation. Implementation.

3.1 Introduction

The Network Development Life Cycle (NDLC) depends upon previously completed development processes such as strategic business planning, application development life cycle, etc.

To fulfill strategic business goals, a top-down approach must be taken to the overall information systems development process. (Ref.1. Figure 12.1)

3.2 Information Systems Development: Process & product

- **Process:** used to visualize what should be done at any point of the development cycle.
- **Product:** milestone or deliverable indicating completion of one stage of the development cycle.

There is a need for significant analysis and design, and associated products or deliverables, prior to the commencement of any network analysis and design activities. (Ref.1. Figure 12.2)

3.3 The Network Development Life Cycle

The NDLC is of an ongoing nature. The network design must be dynamic to support any changing requirements.

3.4 Network Analysis and Design Methodology

A network analysis and design methodology is a practical, step-by-step approach to network analysis and design.

3.4.1 Overall Characteristics

- Requirements (business, application, and data) definition is required prior to network design activities.

- Expected compliance with requirements in a Request For Proposal (RFP) by both in-house personnel and outside consultants.

- Activities from various stages often take place simultaneously and backtrack to previous activities is sometimes needed.

- This methodology is an overall guideline to the network development process rather than "cookbook" instructions.

3.4.2 Critical success factors of the NDLC

These factors are best seen as habits or behaviors, rather than discrete events to be scheduled or planned. They include:

- Identification of all potential customers and constituencies:
    - All groups must be consulted.

- Political awareness:
    - Corporate culture: hierarchical, distributed, or open.
    - Backroom politics can play a role in systems design.
    - Find ways to ensure objectivity of the analysis and design process (e.g., measurable goals).

- Buy-in:
    - Reach consensus on the acceptability of results of each stage.
    - Approved results of one stage become the foundation or starting point for the next stage.
    - Makes the final presentation smoother.

- Communication:
    - With all groups.
    - Write memos, communicate with key people in person, etc.

- Detailed project documentation:
    - Prepare agendas
    - Take meeting minutes
    - Action items
    - Use a project binder for all the above

- Process/Product awareness:
    - Stay focused: what is the process/product at each stage?
    - Keep meeting on track: no off-subject discussions.

- Be honest with yourself:
    - Be your own harshest critic (no one else knows the potential weaknesses or areas for improvement in your proposal better than you.
    - Use peer reviews.
    - Not all weaknesses can be corrected (e.g., financial or time constraints).

3.4.3 Overall Guidelines

- Start with a clearly defined problem:
    - Identify affected parties and representatives.
    - Held brainstorming sessions to define problems and requirements of a solution.

- Understand strategic business objectives defined by senior management.

- Collect baseline data from customer groups about the current status of the system and network. This is used to measure eventual impact of the installed network.

- Feasibility studies and buy-in:
    - Feasibility study: problem definition and associated alternatives recommendations for further study.

3.5 Strategic Information System Design

The primary mission of a network is the delivery of the right information at the right time to the right decision-maker in the right place. All these components are determined by the strategic information system design (SISD).

- The SISD process starts with review of strategic business goals articulated by senior management.

- Then, SISD describes the overall characteristics of an information system that fulfills these goals.

- The evaluation criteria associated with these goals is a key product of SISD and must be objective and measurable. This assures the objectivity of the entire network analysis and design phase.

- The importance of these criteria lies in their ability to measure the extent to which the information system designs deliver strategic business goals.

- Identify opportunities for improvement of business processes in areas such as: financial, customer satisfaction, employee retention, etc. Then, identify information required to turn opportunities into reality.

    - If it isn't broken, don't fix it
    - Must have measured how bad the old process was
    - Learn from other's mistakes (related industries with failures)
    - Don't be afraid to admit mistakes (admit them early and make corrections ASAP to minimize the impact)

- Develop specific evaluation criteria: from these opportunities and the information required to turn them into reality

- Prioritization - three pile approach:
    - Priority 1 items: must be implemented
    - Priority 2 items: need to be implemented ASAP (i.e., "work-around" temporarily)
    - Priority 3 items: nice to have (but can live without them)

- Producing the Request For Proposal (RFP):
    - By organizing all the information gathered.
    - All vendors' proposals are measured against RFP requirements.
    - Examine each corporate location: location survey of data and processing requirements.
    - Final RFP preparation. The RFP should include:
        - SISD
        - Corporate location survey results
        - Management abstract:
            - Company profile: number of locations, growth rate, etc.
            - Statement of the problem.
            - Overall system characteristics: vendors can check first if they have the required capabilities to meet requirements.
            - Project phase prioritization: some modules are more critical than others.
            - Proposed project schedule summary.
            - Information requested from vendors.
                - To avoid standard proposals
                - To ensure:
                    - Vendor has significant experience
                    - Vendor has large organization
                    - Vendor is financially solvent

- Percent-to-fit goal:
    - Arbitrary percentage determined by user groups
    - Sets minimum threshold of compliance for vendor proposals to warrant further consideration and invitations for demonstrations (e.g., 50% of priority 1 features are met). This applies to in-house development as well.
    - Objective "score": counting how many features of each priority are present in a proposal.

- Proposal evaluation and the make or buy decision:
    - Invite selected vendors for demonstrations (e.g., Proof of Concept (POC))
    - Buy-in on selected vendors and vendor selection process.
    - Check every feature included in vendor's proposal at the demonstration.
    - Make or buy decision.

- Outsourcing:
    - Hire outside contractors to operate and maintain corporate information systems and networks.

3.6 In-house Network Analysis and Design

A network must be designed to deliver solutions and performance in response to specific, well defined, data, application, and business layer requirements.

- Data traffic analysis:

    - Payload type analysis: e.g., video, voice, and data.

    - Transaction analysis:

        - Examine the source of data, e.g., order entry, pricing lookup
        - Amount of data required to complete each transaction is calculated and documented

        → This influences which type of network to use, e.g., high speed

- Time studies: Analyze when and how these transactions are executed, i.e., counting how often and what time of the day, week, etc. a transaction is executed.
  → This influences bandwidth requirements.

- Traffic volume analysis: Construct a time sensitive traffic volume requirements profile (from transaction analysis and time studies), i.e., average, minimum, maximum bandwidth requirement

- Mission critical analysis: e.g., Electronics funds transfer
  - Requirements: data security, encryption of data transmitted
  - Redundant links may be needed

- Protocol stack analysis: will the network support more than one protocol? What are the bandwidth and network hardware implications?

- Network configuration alternatives (Logical design):

  - Local carriers may be limited in their offering of certain data transmission services → limitation on your design.

  - Capacity: ensure sufficient bandwidth is allocated to handle sudden increase in demand.

  - Reliability: sufficient redundancy is implemented

  - Security

  - Cost (for senior management to decide)

- Network hardware analysis and configuration alternatives (Physical design):

  - Depends on the results of the two previous analysis reports. If these are valid, then networking devices chosen to tie the network together should be valid as well.

- Prepare a comprehensive budget

  - Prevent surprises: required or anticipated facilities upgrade are identified during survey (in RFP preparation)

  - Three cost categories: Acquisition, Operations, and Anticipated growth

- Prepare the final proposal, i.e., RFP response or network design document

3.7 Contents of a Network Design Document

1. Executive Summary: targeted at the managers and key project participants

2. Project Goal: should be business-oriented

3. Project Scope: information on the extend of the project

4. Design Requirements

4.1. Business Goals: how the network will help in providing better products and services.

4.2. Technical goals: Scalability, Availability, Performance, Security, Manageability, Usability, Adaptability, Affordability.

4.3. User Communities and Data Stores: user communities, locations, applications, and data stores (servers and hosts).

4.4. Network Applications: new and existing ones.

5. Current State of the Network: structure and performance of existing network applications

6. Logical Design:

- Network topology
- Addressing and naming models
- Protocols selected for routing, bridging, and switching
- Recommended security mechanisms and products
- Recommended network management architectures, processes, and products

7. Physical Design:

- Features and recommended uses for the technologies and devices selected.
- Pricing for network devices and services.
- Availability of products.

8. Results of Network Design Testing: from prototype or pilot systems implemented

9. Implementation Plan: for installations, outsourcing, informing users, training, measuring design effectiveness, and fallback and future plans

9.1. Project Schedule: at least dates and deliverables for major milestones

10. Project Budget: funds available for purchases, maintenance, support, licenses, training, and staffing

10.1. Return on Investment: how quickly the design will pay for itself

11. Design Document Appendix

3.8 The Network Implementation Process

- Pilot tests: to safely roll out new systems or networks. E.g., deploy/implement the new system on one site, monitor performance, fix problems, and gain experience before deployment on a wider scale.

- Project management:

    - Detailed task lists

    - Manual or using project management software.

- People are important: buy-in at every stage by all affected parties. ***The best designed network will fail miserably without the support of people.***

3.9 Automating the NDLC

- CANE: Computer-Assisted Network Engineering (CANE): Analysis and design software used to model a current network.

- Simulation tools: performance engineering software tools: overall network performance modeled is a result of the effect of a series of mathematical formulas.

    - Ability to predict performance of various networking scenarios (i.e., what-if analysis).
    - Benefits: spot network bottlenecks, test new applications and network configurations before deployment, re-create circumstances, and replicate traffic volume and transaction types.

- Network management tools.

3.10 References

1. "Applied Data Communications - A Business-Oriented Approach" by James E. Goldman, 1998

2. "Top-Down Network Design" by Priscilla Oppenheimer, Cisco Press, 2001

Chapter 4 *Enterprise Network Design*

Topics covered:
Enterprise Network Design Model. Backbone design concepts. Network security and firewalls. Structured cabling systems. Case studies.

Definition

An enterprise network consists of a group of local area networks (LANs) interconnected using wide area networks (WANs). An enterprise network contains a number of internetworking devices (e.g., switches, routers, gateways, etc) and is under the control of one big organization.

4.1 Enterprise Network Design Model (Hierarchical Model)

There really is no "one size fits all" when it comes to network design.

Two design options:

- Design a network infrastructure from the ground up.
- Meld the new technologies into an existing infrastructure.

A model is vital for analyzing large, complex internetworks.

→ Use of guidelines or rules.

Internetworks are generally implemented in a hierarchical manner.

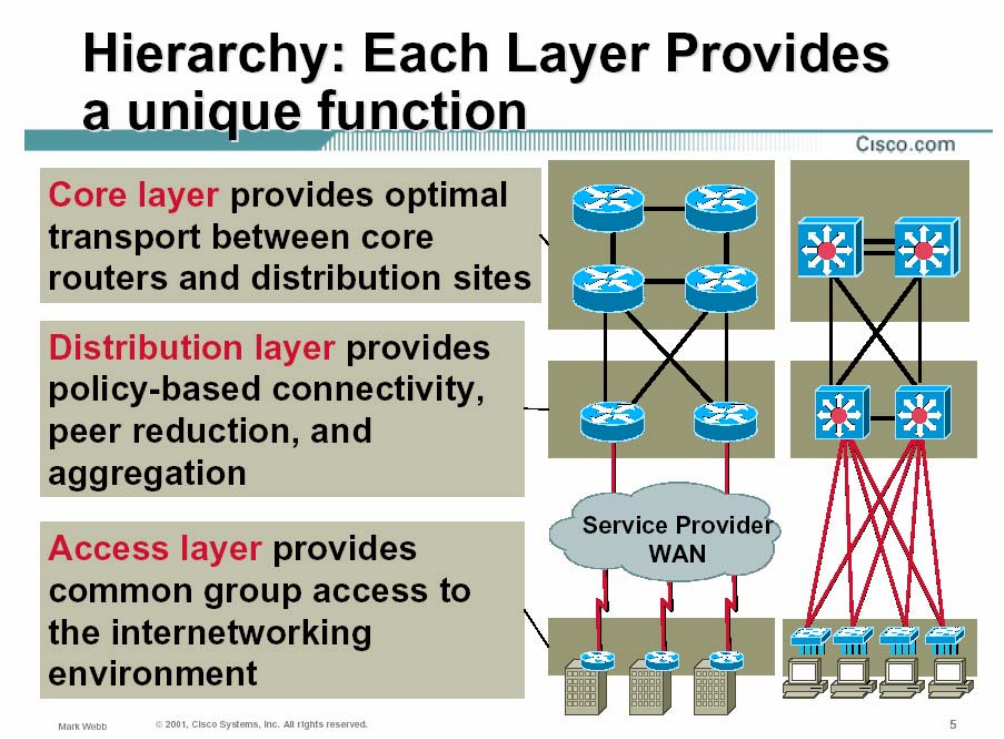
4.1.1 Three-tier hierarchical model

It consists of:

-
-
-

Each level provides a backbone for the level below.

Definition: A backbone is a network whose primary purpose is the interconnection of other networks.



(Slide taken from http://www.cisco.com/warp/public/cc/so/neso/meso/uentd_pg.pdf)

4.1.1.1 Core tier

- Provides optimal wide-area transport between geographically remote sites.
- Connects campus networks in a corporate or enterprise WAN
- Services are typically leased from a telecom service provider
- May use the public Internet as enterprise backbone.
- Focus on redundancy and reliability → Continue to function with circuit outages.
- Need to efficiently use bandwidth because of provider tariffs.
- End Stations should not be put in the core

Design Rule:

4.1.1.2 Distribution tier

- Connects multiple networks (departments) within a campus network environment (one or more buildings).
- Includes campus backbone network, based on FDDI, Fast Ethernet, Gigabit Ethernet, or ATM.
- Acts as a concentrator points for many of its access tier sites.
- Links usually owned and/or controlled by the organization.
- Network policy is often implemented in this tier. E.g., security, firewall, encryption, address translation.
    - Network naming and numbering conventions
    - Network security for access to services (admin privileges, etc)
    - Network security for traffic patterns through definition of path metrics (priority, preference, trust, etc)
    - Address aggregation

Design Rule:

4.1.1.3 Access tier

- Usually a LAN or a group of LANs.
- Typically uses Ethernet, Token Ring, or FDDI.
- Can be divided into two levels (workgroup level & desktop level)
    - Workgroup level: e.g., departmental level
    - Desktop level: where end-user devices are attached.
- Where hosts are attached to the network (e.g., labs)
- Connects workgroups (e.g., marketing, administration)
- Usually within a single building (or single floor)
- Provides logical network segmentation, traffic isolation and distributed environment
- Remote (dialup) users are connected at this tier.

Design Rule:

4.1.2 Benefits of a Hierarchical Design Model

Network designs can be: mesh or hierarchical. In a mesh structure, the network topology is flat.

A hierarchical design model has the following advantages:

1. Scalability

- Design rule: Build hierarchical networks for maximum scalability.

2. Ease of implementation

- Phased approach is more effective due to cost of resources → efficient allocation of resources in each phase of network deployment.

3. Ease of troubleshooting

- Easy to isolate problems in the network
- Use "divide-and-conquer" approach → Temporarily segment the network.
- Does not affect core tier network

4. Predictability

- Makes capacity planning for growth easier.

5. Protocol support

- Mixing new protocol is easier.
- Merger of companies using different protocol is easier

6. Manageability

- Easy to implement network management instrumentation by placing probes at different levels of hierarchy

4.1.3 Variations on the three-tier model

4.1.3.1 One-tier Design – Distributed

- Remote networks connect to a pseudo-core

- Good for small networks with no centralized server location.

Advantage: Faster overall response time between peers, simplicity, and cost effectiveness.

Disadvantage: Loss of centralized management control and higher management cost because of duplicated management functions (Responsibilities such as server backups and network documentation are delegated to the access site).

4.1.3.2 One-tier Design – Hub and Spoke

- Servers are located in central farms.

Advantage: Increased management control (centralized).

Disadvantage: Single points of failure and bandwidth aggregation.

4.1.3.3 Two-tier Design

- A campus backbone interconnects separate buildings
    - VLANs can be used to create separate logical networks (i.e., broadcast domains).

4.1.3.4 Redundant Two-tier Hierarchy

- Core LAN backbone is duplicated for total redundancy → more reliable.

4.1.4 Hierarchical Design Guidelines

- Choose a hierarchical model that best fits your requirements
- Do not always completely mesh all tiers of the network (use the backbone for connections).
    - Core connectivity, however, will generally be meshed for circuit redundancy and network convergence speed.
- Do not place end stations on backbones → Improves the reliability of the backbone.
- Workgroup LANs should keep as much as 80% of their traffic local to the workgroup →

- Use specific features at the appropriate hierarchical level.
- Control the diameter of a hierarchical enterprise network topology (in most cases, 3 major layers are sufficient)

  → Provides low and predictable latency.

  → Helps predict routing paths, traffic flows, & capacity requirements.

  → Makes troubleshooting & network documentation easier.

- Avoid chains at the access layer (e.g., connecting a branch network to another branch, adding a 4^{th} layer)

- Avoid backdoors (i.e., connection between devices in the same layer)

  → Cause unexpected routing problems

  → Make network documentation and troubleshooting more difficult

- Design the access layer first, then the distribution layer, and finally the core layer.

  → Helps, more accurately, perform capacity planning at the distribution and core layers.

4.1.5 Mesh vs. Hierarchical-Mesh Topologies

- In full-mesh topology, every router or switch is connected to every other router or switch. It provides complete redundancy. → Offers good performance.
    - Number of links in this topology is:

- A partial-mesh topology has fewer connections (To reach another router, the network might require traversing intermediate links).

- Mesh networks can be expensive to deploy and maintain
    - Especially full-mesh networks
    - Hard to optimize, troubleshoot, and upgrade
    - Have scalability limits for groups of routers that broadcast routing updates (processing increases)

- A classic hierarchical and redundant enterprise design uses a partial-mesh hierarchy rather than a full mesh.

4.2 Redundant Network Design Topologies

Redundancy:

- Provides network availability by duplicating network links and interconnectivity devices
- Eliminates the possibility of having single point of failure on the network

  Goal:

- Helps you meet the availability goals for users accessing local services (in campus networks)
- Helps you meet overall availability and performance goals (in enterprise networks)
- Adds complexity to the network topology and to network addressing and routing

Note:

4.2.1 Backup Paths

- A backup path:
  - Consists of routers and switches and individual backup links between routers and switches that duplicate devices and links on the primary path.
  - Maintains interconnectivity even when one or more links are down
- Two aspects of the backup path to consider:
  - How much capacity does the backup path support?
  - How quickly will the network begin to use the backup path?

- Use a modeling tool to predict network performance when backup is in use:
    - It can be acceptable that the performance of the backup path is worse than that of the primary path.
- Backup path usually have less capacity than primary path, e.g., a leased line with a backup dial-up line. However, requirements may state that both must provide the same performance → this is expensive → Tradeoff: Cost vs. Reliability.
- Automatic fail-over is necessary for mission-critical applications.
    - Where disruption is not acceptable.
    - If manual reconfiguration is required to switch to a backup path, users will notice disruption.
    - Redundant, partial mesh network design speeds automatic recovery time when a link fails, e.g., spanning tree.
- Backup path must be tested
    - Do not wait for a catastrophe to happen.
- Some backup links are used for load balancing as well as redundancy

  → Advantage:

4.2.2 Load Balancing

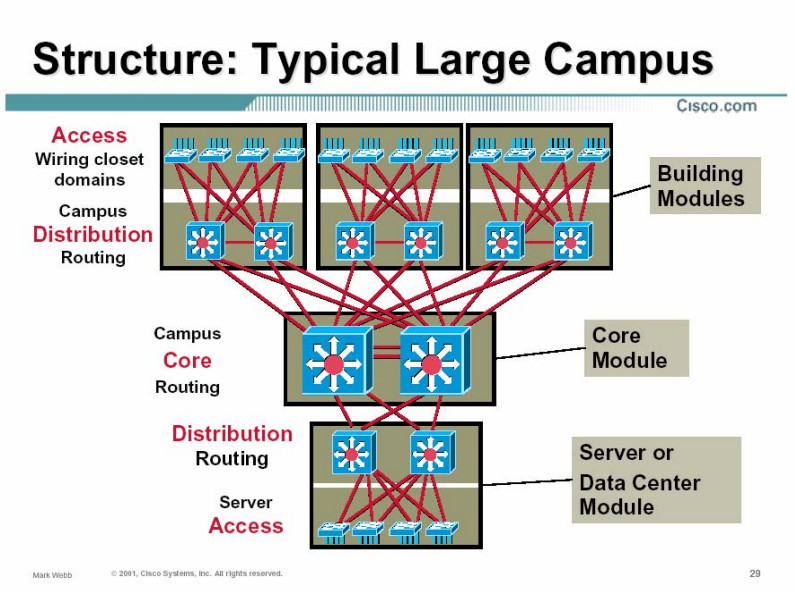
- Redundancy improves performance by supporting load balancing across parallel links.
- Load balancing must be planned and in some cases configured.
    - However, some protocols do not support load balancing by default.
- Some internetworking devices support balancing across multiple parallel paths.

4.3 Designing a Campus Network Design Topology

- Redundant LAN segments
    - Design redundant links between LAN switches.
    - The spanning tree algorithm guarantees that only one path is active between two stations → Good solution for redundancy, but not for load balancing.

- Cisco switches let you implement one spanning tree per VLAN. Redundant links can offer load balancing and fault tolerance.

- Server redundancy
  - Depends on the customer's requirements
  - Services include: file, web, DHCP (Dynamic Host Configuration Protocol), name, database.
  - Use redundant servers when needed (e.g., DHCP). The servers should hold redundant (mirrored) copies of the DHCP database. DHCP servers can be placed at either the access or distribution layer.



4.4 Designing an Enterprise Network Design Topology

- Enterprise network design topology should meet a customer's goals for availability and performance by featuring:
    - Redundant LAN and WAN segments in the intranet
    - Multiple paths to extranets and the Internet.
        - Extranet: an internal internetwork that is accessible by outside parties, e.g., suppliers, resellers, etc.
- Redundant WAN segments
    - Usually uses a hierarchical partial-mesh topology.
    - **Circuit diversity**: physical circuit routing of backup WAN links and primary WAN links should be different than each other.
    - **Multihoming** the Internet connection: provides an enterprise network more than one entry into the Internet (i.e., redundancy and fault tolerance)
        - **Definition:** ***Multihoming*** - provide more than one connection for a system to access and offer network services
            - Example: A server is multihomed if it has more than one network-layer address.
        - **Options for multihoming the Internet connection** (i.e., the enterprise network is multihomed to the Internet)

- **Virtual Private Networks (VPNs):** enables the use of a public network, such as the Internet, to provide a secure connection among sites on the organization's internetwork.
    - A public network is used as a backbone for the enterprise network.
    - Links remote offices together.
    - Inexpensive compared to private leased lines.
    - Private data is encrypted for routing through the public network
    - No permanent link is required.
    - Can use Dial-on-demand routing (DDR).

4.5 Secure Network Design Topologies

- Planning for physical security: protection from unauthorized access, theft, vandalism, and natural disasters (e.g., floods, fires, storms, and earthquakes)

  → Not an aspect of logical network design, but it has an impact on it.

- Meeting security goals with firewall topologies:

  - **Definition:** ***Firewall*** – a system or combination of systems that enforces a boundary between two or more networks (according to the National Computer Security Association (NCSA)).

  - A firewall can be:
    - a router with access control lists (ACLs),
    - a dedicated hardware box (e.g., PIX), or
    - a software running on a PC or UNIX system.

  - A firewall should be placed in the network so that all traffic from outside the protected network must pass through the firewall.

  - A firewall is especially important at the boundary between the enterprise network and the Internet.

  - A basic firewall topology is simply a router with:
    - a WAN connection to the Internet,
    - a LAN connection to the enterprise network, and
    - a software that has security feature.

  - Larger companies use a dedicated firewall in addition to a router between the Internet and the enterprise network.

  - Firewall topology can include a public LAN that hosts Web, FTP, DNS, and SMTP servers (for customers who need to publish public data).
    - This public LAN is referred to as: ***demilitarized or free-trade zone***.

4.6 Backbone Design

There are two types of backbone design:

- Distributed backbones
- Collapsed backbones

4.6.1 Distributed Backbones

4.6.1.1 Distributed Backbones in Buildings (Figure 5-3 in [1])

- Each floor's router is directly connected to a centralized backbone.
- The backbone is typically and FDDI ring. This provides maximum fault tolerance.
- Generally, do not contain a single point of failure
- Requires extra input and output ports for each component
  - Faults quickly corrected by isolation process
  - High cost
- **Drawbacks:**
  - Multiple IP network numbers → difficult to add, move, or change users.
  - More expensive
  - Migration to switching not easy.
  - Less-flexible approach to wiring a building.

4.6.1.2 Distributed Backbones on the Campus (Figure 5-4 in [1])

- More resource-efficient solution than in a building.
- **Drawback:**
    - Lack of flexibility in connecting to other buildings on the campus.

4.6.2 Collapsed Backbones (Figures 5-5 and 5-6 in [1])

- Has a single concentration point connecting all floors.
- All floor-to-floor connectivity passes through the backbone component.
- Problem isolation is simple, while finding problem's root cause is difficult.
- More flexible and cost-effective approach to wiring a building.
- Changes can be easily made.
- Can be extended to accommodate VLANs.
    - VLANs in a building
        - More flexibility in positioning of end stations and servers.
    - VLANs accross a campus
        - One switch acts as the backbone for the entire campus.
        - Assign stations to VLANs such that only 20% of their traffic is destined to other VLANs.

- Single point of failure (Router)

  → **Solution:** Router with HSRP (Hot Standby Router Protocol).

  - **HSRP:** Provides a way for an IP workstation to keep communicating on an internetwork even it its default router becomes unavailable.
  - HSRP works by creating a phantom router with its own IP and MAC addresses.
    - Each workstation uses the phantom as its default router.
    - When a workstation broadcasts an ARP frame to find its default router, the active HSRP router responds with the phantom's MAC address.
    - If the active HSRP router goes offline, a standby router takes over as active router.
    - HSRP routers on a LAN communicate to designate an active and standby router.

4.7 Structured cabling Systems (SCS)

4.7.1 SCS Principles

- Studies have shown that more than 50% of all network disruptions are related to cabling.
- IBM & AT&T developed generic cabling systems based on STP cables and UTP cables, respectively.

- SCS objectives:
    - Use a single common cable type that supports many applications
    - Remain cost effective (i.e., minimum additional equipment required)
    - Based on a "flood wiring" approach.
        - To minimize the impact of moves, additions, and changes → minimize the ongoing cost of ownership
    - Ability to support any given application
    - Reliability of the system.
- SCS topology:
    - Based on "star" topology in a tree-like fashion.
    - Distribution point provides the administration (patching) points for the system
    - All systems must comprise at least the horizontal distribution level
    - At each distribution point, application specific equipment (e.g., computer systems, repeaters, switches, etc) are patched into the system for user connectivity.

- SCS standards:

    - EIA/TIA-568 standard: "Commercial Building Telecommunications Wiring Standard" (1990)
        - Included the use of both 10Base2 and 10Base5 media (i.e., coaxial)
        - EIA: Electronics Industry Association (www.eia.org)
        - TIA: Telecommunications Industry Association (www.tiaonline.org)

    - ANSI/EIA/TIA-568-A standard: provides ideal design platform
        - Different media is possible
        - SCS terminology in ANSI is different from ISO's
            - ANSI: American National Standards Institute
            - ISO: International Standards Organization

4.7.2 Areas within a SCS System

- SCS comprises 3 cabling subsystems:
    - Horizontal cabling subsystem.
    - Building backbone subsystem.
    - Campus backbone subsystem.

- The work area cabling is also necessary but outside the scope of SCS standards.

4.7.2.1 Horizontal Cabling Subsystem

- Includes:
    - Horizontal distribution cables
    - Connecting hardware
    - Cross-connect patching at the Floor Distributor (FD)

- Media options: UTP, STP, optical fiber (multimode)

- Horizontal distribution cable should be continuous wherever possible

    - However, a single transition point can be included between FD and TO, e.g.:
        - Transition between 2 types of cabling
        - Use of 25 pair cables, then 4 pair cable to the TO.

    → Source of increased crosstalk.

- Maximum distance is 90 meters from FD to TO.

    - Maximum 5 meters for patch cords and work area fly leads.

    → Maximum 5+90+5=100 meters between equipment and end-user.
    (100 meters is the maximum transmission distance for high speed data over twisted pair)

4.7.2.2 Building Backbone Cabling Subsystem

- Includes:
  - Building backbone cable
  - Termination hardware
  - Cross-connect patching at the Building Distributor (BD)

- Media options: UTP, STP, optical fiber (multimode and single mode)

- Transition points not allowed.

- Maximum distance is 500 meters.

- Maximum 20 meters for patch cords length.

4.7.2.3 Campus Backbone Cabling Subsystem

- Includes:
  - Campus backbone cable
  - Mechanical termination of backbone cable
  - Cross-connect patching within the Campus Distributor (CD)

- Media options: mainly optical fiber (for longer distances and electrical isolation)

- Maximum distance is 1500 meters.

  - Added to the building backbone maximum distance (i.e., 500 meters)

  → 2 km, that is the maximum supported distance for high speed data over multimode fiber optic.

- Maximum 20 meters for patch cords length.

4.7.3 Application Classes

- Applications must be taken in consideration when designing an SCS

Example: Ethernet maximum transmission limit is 100 meters over UTP cables.
→ Ethernet will not run over UTP backbone of 500 meters.
→ Backbone should be reduced to 100 meters or media changed to fiber optic

- Mapping of LAN applications onto SCS:
  - Ethernet is the easiest to map onto SCS because standards were written for twisted pair media.

4.7.4 SCS Patching

4.7.4.1 Inter-Connect (Direct) Patching

- Convenient when port presentation is the same on equipment and patch panel (e.g., RJ-45)

- Requires fewer connections.
  → Minimizes the amount of crosstalk

- **Problem:** Patch cables can become a tangled mess with cables going in all directions (bad presentation).

4.7.4.2 Cross-Connect (Indirect) Patching

- Involves the addition of extra patch panels where equipments are permanently terminated.

- Much neater cable presentation.

- **Problem:** More crosstalk on the link mainly if the full bandwidth of the cabling is being pushed to its limits over the maximum 90 meters distribution distance.

  → This can be avoided with good installation practices.

4.7.5 Design Guidelines

- Start from the edge (i.e., work area) and work back to the center.

  Example: sizing of work areas, media type of horizontal cabling, location of FD, cabling pathways, etc.

- System administration must start at the planning stage to maximize the potential of the cabling system.

  Example: outlet identification, numbering/naming, etc.

4.7.5.1 Work Area

- 1000 m^2 of floor space is the maximum area to be supported from one FD.
- Work area sizing: 2 m^2 → 10 m^2 (However, this is site specific decision)
- Number of TOs per work area: at least 2 (1 copper, and 1 fiber or copper).
- Design work areas to form logical "zones"
  - Some buildings may require multiple FDs to service all locations on a floor.
- Overlaying, or interleaving the cabling (multiple pathways are used) can provide a high degree of resilience when a pathway is damaged.
  - Much higher installation and material cost.

4.7.5.2 Distributor Layout

- Type of patch panel, FD housing (cabinet, rack, etc), sizing, location, etc.

- Each cabinet contains a proportion of all elements (e.g., some horizontal cabling, some equipment, etc.)

  Advantage: If a cabinet is lost (e.g., due to power failure), a proportion of users will not be affected.

4.7.5.3 The Backbone

- May need to provide two backbones: one for voice and one for data.
- Resilience: implement multiple backbones via multiple risers and have additional capacity in each backbone.

4.8 References

1. "Cisco Internetwork Design" edited by Matthew H. Birkner. Cisco Systems, 2000

2. "Top-Down Network Design" by Priscilla Oppenheimer, Cisco Press, 2001

3. http://www.cisco.com/warp/public/cc/so/neso/meso/uentd_pg.pdf

4. "The Switch Book" by Rich Seifert. John Wiley & Sons Inc., 2000.

5. "LAN Technologies Explained" by Philip Miller and Michael Cummins. Digital Press, 2000

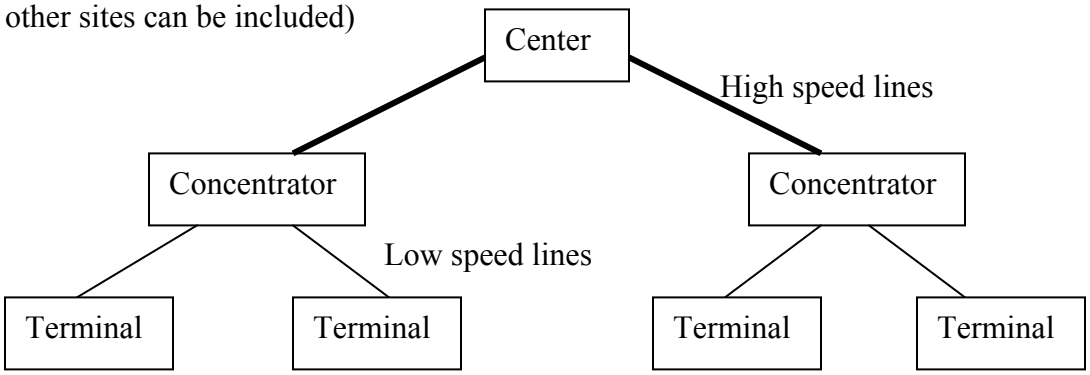
Chapter 5 *Topology design and analysis*

Topics covered:
Topology design. Network design algorithms. Terminal assignment. Concentrator location. Traffic flow analysis and performance evaluation. Network reliability.

5.1 Topology design

5.1.1 Centralized Network design

- **Centralized network:** is where all communication is to and from a single central site.
- The "central site" is capable of making routing decisions.
  → Tree topology provides only one path through the center (For reliability, lines between other sites can be included)



- Three different problems:
  - **Multipoint line topology:** selection of links connecting terminals to concentrators or directly to the center.
  - **Terminal assignment:** association of terminals with specific concentrators.
  - **Concentrator location:** deciding where to place concentrators, and whether or not to use them at all.

5.1.2 Finding Trees in Graphs

- Used to design and analyze networks.
- Connect a number of nodes to a central node:
  - **Node:** Hub, Switch, Router, etc.
  - **Central node:** backbone

- A tree is a graph with no loops, with only one path between any pair of nodes.
- Trees are minimal networks: provide connectivity without any unnecessary additional links:
    - Minimally reliable and robust
    - Networks are more highly connected (but design starts with a tree)

5.1.2.1 Tree Traversals

- Visit all nodes in a tree: edges are traversed twice.
- First, identify a node as the root
- Assume the tree is directed (outward from the root)
- Two algorithms:
    - BFS (Breadth First Search):
        - Nodes closest to root are visited first
        - Implemented using a queue (FIFO)
    - DFS (Depth First Search):
        - Visits an unvisited neighbor of the node just visited.
        - Implemented using a stack (LIFO)
- Both traversals (BFS and DFS) can be preorder traversals (i.e., visit nodes then successors) or post-order traversals (i.e., successors visited first).
- Traversal is generalized to undirected graphs by keeping track of which nodes were visited, and not visiting them again.
- In a BFS or DFS traversal, edges visited form a tree (if the graph is connected) or a forest (if the graph is not connected).

5.1.2.2 Minimum Spanning Trees (MSTs)

- Use DFS to find a spanning tree in a graph, if one exists
  → Arbitrary tree
- Useful to find the "best" tree
  → Minimum Spanning Tree (e.g., minimum total length. Where length is: distance, cost, function(delay), function(reliability), etc.)

- If the graph is not connected → minimum spanning forest
  - For ***n*** nodes, ***c*** components, and ***e*** edges, we have: ***n = c + e***
  - For a tree, ***c = 1***.

- DFS will not, in general, find the spanning tree with minimum total cost.

5.1.2.2.1 The Greedy Algorithm

- At each stage, select the shortest edge possible.

- May not find a feasible solution when one exists.

- Efficient and simple to implement → widely used.

- Basis of other more complex and effective algorithms.

- In the case of MST, the greedy algorithm guarantees both optimality and reasonable computational complexity.

  - Start with empty solution ***s***
  - While elements exist
    - Find ***e***, the best element not yet considered
    - If adding ***e*** to ***s*** is feasible, add it; if not, discard it.

5.1.2.2.2 Kruskal's Algorithm

- A greedy algorithm for finding MSTs.

- Sort the edges, shortest first and then include all edges which do not form cycles with the edges previously selected.

- n: number of nodes

- **<u>Algorithm:</u>**

  1. Sort all edges in ascending order (least cost first)

  2. Select among edges not yet selected, the one with the least cost.

  3. Add it if it does not create a cycle.

  4. If the number of edges selected < ***n-1***, go to step (2), otherwise exit (tree completed)

- **<u>Complexity:</u>**

  O(m log m), m = number of edges

5.1.2.2.3 Prim's Algorithm

- A greedy algorithm for finding MSTs.
- Advantageous if the network is dense.
- Well suited to parallel implementation.
- **Algorithm:**

  1. Start with one node (root node) in the tree
  2. Find node ***i***, not in the tree, which is the nearest to the tree.
  3. Add node ***i*** to the tree and edge ***e*** connecting ***i*** to the tree.

- **Complexity:**

  $O(n^2)$

5.1.2.2.4 Comparison of the Complexity of Kruskal's and Prim's Algorithms

- If the network is dense → $m \sim O(n^2)$ → Prim's algorithm is faster
- If the network is not dense → $m \sim O(n)$ → Kruskal' algorithm is faster

5.1.3 Constrained/Capacitated MST (CMST)

- The algorithms presented in the previous subsections are called "unconstrained MST algorithms"
  - No constraint on flow of information
  - No constraint on the number of ports at each node.
- For the unconstrained spanning tree problem, all these algorithms produce a minimum cost spanning tree.
- **CMST Problem:** Given a central node N_0 and a set of other nodes (N_1, N_2, …, N_n), as et of weights (W_1, W_2, …, W_n) for each node, the capacity of a link, W_{max}, and a cost matrix C_{ij} = Cost(i,j), find a set of trees T_1, T_2, …, T_k such that each N_i belongs to exactly one T_j and each T_j contains N_0.

- **Objective:** Find a tree of minimum cost and which satisfies a number of constraints such as:
  - Flow over a link
  - Number of ports

- **Example:**
  - Assume we are allowed to use one type of links only that can accommodate a maximum of 5 units of flow per unit time.

  - Assume that the flow generated from each node to the central node (N_1) is as follows: $f_1=0$, $f_2=2$, $f_3=3$, $f_4=2$, $f_5=1$ (in units/time_unit).

  - Effect of constraint violation:
    - As a result, a queue will build up since node 3 can service only 5 units/time_unit. If node 3 does not have a large queue to accommodate all coming units, some units will be lost. So, these units are retransmitted, which may cause the network to collapse.

- The CSMT problem is NP-hard (i.e., cannot be solved in polynomial time)
  → Resort to heuristics (approximate algorithms)

- These heuristics will attempt to find a good feasible solution, not necessarily the best, that:
  - Minimizes the cost
  - Satisfies all the constraints

- Well-known heuristics:
  - Kruskal
  - Prim
  - Esau-Williams

5.1.3.1 Kruskal's Algorithm for CMST

Algorithm:

1. Sort all edges in ascending order, $e \leftarrow 0$.
2. Select edge with minimum cost (from edges not yet selected)
3. If it satisfies constraints (i.e., no cycles and no violation of flows on links)
   - Then: add it to the tree, $e \leftarrow e + 1$
   - Else: go to step (**2**)
4. If ($e = n - 1$) then exit, else go to step (**2**)

Example:

Given a network with five nodes, labelled **1** to **5**, and characterized by the following cost matrix:

	1	**2**	**3**	**4**	**5**
1	-	3	3	5	10
2	3	-	6	4	8
3	3	6	-	3	5
4	5	4	3	-	7
5	10	8	5	7	-

Node 1 is the central backbone node.

$f_{max}=5$, $f_1=0$, $f_2=2$, $f_3=3$, $f_4=2$, $f_5=1$.

5.1.3.2 Prim's Algorithm for CMST

Algorithm:

1. Start with one node (root node) in the tree.

2. Find node ***i***, not in the tree, which is the nearest to the tree

3. Add node ***i*** to the tree and edge ***e*** connecting ***i*** to the tree if it satisfies constraints (i.e., no violation of flows on links)

Example:

5.1.3.3 Esau-Williams Algorithm for CMST

Node **1** is the central node.

t_{ij}: is the tradeoff of connecting i to j or i directly to the root.

- If ($t_{ij} < 0$) → better to connect i to j
- If ($t_{ij} \geq 0$) → better to connect i directly to the root

Algorithm:

1. Compute $t_{ij} = c_{ij} - c_{i1}$ for all $i, j \neq 1$, and $i \neq j$

2. Select the link (m,n) such that: $t_{mn} = \min(t_{ij})$

3. If $t_{mn} < 0$, then go to step (**4**)
Else (i.e., $t_{mn} \geq 0$ for all m,n), connect to node 1 all nodes not connected yet, and **exit**.

4. Verify constraints (e.g., does not exceed the maximum weight)
- If satisfied go to step (**5**)
- Else: $t_{mn} = \infty$ and $t_{nm} = \infty$, go to step (**2**)

5. Add link (m, n), remove link (m, 1) and update t_{ij} to indicate that **m** is now connected to **n**.
→ $t_{mn} = \infty$ and $t_{nm} = \infty$
→ if $t_{mj} \neq \infty$, $t_{mj} = c_{mj} - \min(c_{k1})$ [$k \in C_m$, where C_i = component containing node **i**]

6. Go to step (**2**)

Example:

Given a network with five nodes, labelled **1** to **5**, and characterized by the following cost matrix:

	1	**2**	**3**	**4**	**5**
1	-	3	3	5	10
2	3	-	6	4	8
3	3	6	-	3	5
4	5	4	3	-	7
5	10	8	5	7	-

$W_{max}=5$, $W_1=0$, $W_2=2$, $W_3=3$, $W_4=2$, $W_5=1$.

5.1.4 Terminal Assignment

5.1.4.1 Problem Statement

- **Terminal Assignment:** Association of terminals with specific concentrators.

Given:

T terminals (stations) i = 1, 2, …, T

C Concentrators (hubs/switches) j = 1, 2, …, C

$\mathbf{C_{ij}}$: cost of connecting terminal i to concentrator j

$\mathbf{W_j}$: capacity of concentrator j

Assume that terminal i requires W_i units of a concentrator capacity.

Assume that the cost of all concentrators is the same.

- $x_{ij} = 1$; if terminal i is assigned to concentrator j.
- $x_{ij} = 0$; otherwise.

Objective:

5.1.4.2 Augmenting path algorithm

Based on the following observations:

1. Ideally, every terminal is assigned to the nearest concentrator.

2. Terminals on concentrators that are full are moved only to make room for another terminal that would cause a higher overall cost if assigned to another concentrator.

3. An optimal partial solution with ***k+1*** terminals can be found by finding the least expensive way of adding the $\mathbf{(k+1)^{th}}$ terminal to the ***k*** terminal solution.

Assignment problem:

Given a cost matrix:

- One column per concentrator
- One row per terminal

Assume that:

- Weight of each terminal is 1 (i.e., each terminal consumes exactly one unit of concentrator capacity)
- A concentrator has a capacity of W terminals (e.g., number of ports)

A feasible solution exists iff $\boldsymbol{T \leq W * C}$

Algorithm:

1. Initially, try to associate each terminal to its nearest concentrator

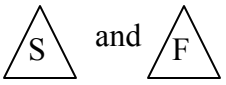
2. If successful in assigning all terminals without violating capacity constraints, then stop (i.e., an optimal solution is found)

3. Else,
   - **Repeat**
     i. Build a compressed auxiliary graph
     ii. Find an optimal augmentation
   - **Until** all terminals are assigned

Building a compressed auxiliary graph:

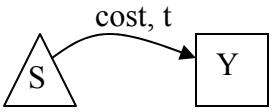
U: set of unassociated terminals

T(Y): set of terminals associated with Y

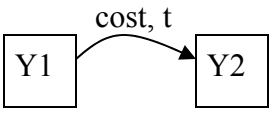


are the start and finish of all augmenting paths

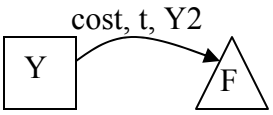
represents a fully loaded concentrator



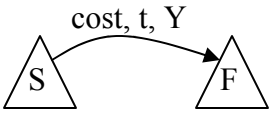
Assign t to a fully loaded concentrator Y. (t Є U)
cost = c(tY) = min c(xY) for x Є U



Move t from a fully loaded concentrator Y1 to another fully loaded concentrator Y2. (t Є T(Y1))
cost = c(tY2) – c(tY1)= min (c(xY2) – c(xY1)) for x Є t(Y1)



Move t from Y to a concentrator Y2 with spare capacity. (t Є T(Y))
cost = c(tY2) – c(tY)= min (c(xY2) – c(xY)) for x Є t(Y)



Assign t to a concentrator Y with spare capacity. (t Є U)
cost = c(tY) = min c(xY) for x Є U

Example:

5.2 Traffic Flow Analysis and Performance Evaluation

5.2.1 Traffic Flow Analysis Objective

- Estimate:
  - Delay
  - Utilization of resources (links)

- Traffic flow across a network depends on:
  - Topology
  - Routing
  - Traffic workload (from all traffic sources)

- Desirable topology and routing are associated with:
  - Low delays
  - Reasonable link utilization (no bottlenecks)

- Assumptions:
  - Topology is fixed and stable
  - Links and routers are 100% reliable
  - Processing time at the routers is negligible
  - Capacity of all links is given $C = [C_i]$ (in bps [bits per second])
  - Traffic workload is given $\Gamma = [\gamma_{jk}]$ (in pps [packets per second])
  - Routing is given $R = [r_{jk}]$
  - Average packet size is $1/\mu$ bits.

5.2.2 Queuing Analysis

Projections of performance are made on the basis of either:

- The existing load information, or
- The estimated load for the new environment.

Approaches that could be used:

- Do an after-the-fact analysis based on actual values

- Make a simple projection from existing to expected environment

- Develop an analytic model based on queuing theory

- Program and run a simulation tool

5.2.2.1 Queuing Models

- The notation **X/Y/N** is used for queuing models.
  - X = distribution of the interarrival times
  - Y = distribution of service times
  - N = number of servers

- The most common distributions are:
  - G = general independent arrivals or service times
  - M = negative exponential distribution
  - D = deterministic arrivals or fixed length service

- Example: **M/M/1**

- **Single-server queues**

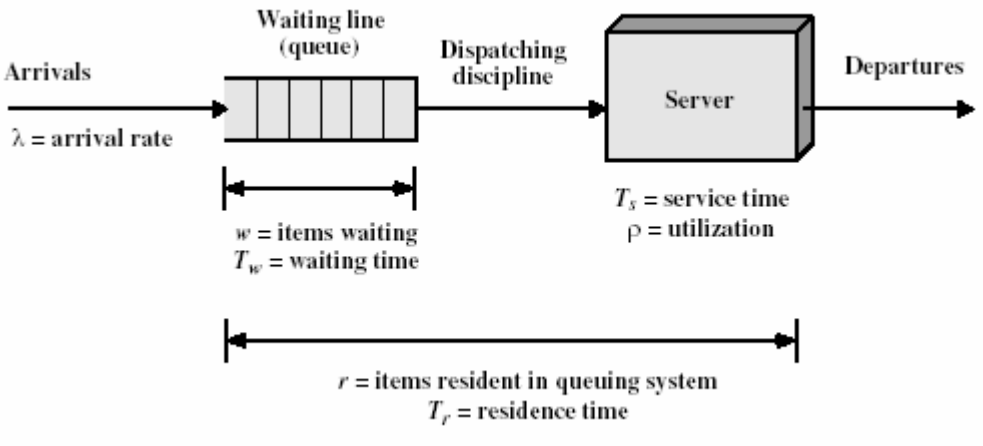


Figure 5.1: Queuing System Structure and Parameters for Single-Server Queue (Taken from "Queuing Analysis" by William Stallings)

Queue parameters:

λ = arrival rate; mean number of arrivals per second

T_s = mean service time for each arrival; amount of time being served, not counting time waiting in the queue

ρ = utilization; fraction of time facility (server or servers) is busy

r = mean number of items in system, waiting and being served (residence time)

T_r = mean time an item spends in system (residence time)

w = mean number of items waiting to be served

T_w = mean waiting time (including items that have to wait and items with waiting time = 0)

Basic Queuing relationship:

- $\rho = \lambda * T_s$
- $r = w + \rho$
- $\lambda_{max} = 1/T_s$
- $r = \lambda * T_r$ **(Little's formula)**
- $w = \lambda * T_w$ (Little's formula)
- $T_r = T_w + T_s$
- $r = \rho/(1-\rho)$

- **Multiserver queue**

  N = number of servers

  ρ = utilization of each server

  $N\rho$ = utilization of all servers (= $\lambda * T_s$)

5.2.2.2 M/M/1 Queues – Application to Networks

- Each link is seen as a service station servicing packets.

  λ_i = arrival rate (in pps); mean number of packets that arrive to link i in one second.

  μC_i = average service rate (in pps); mean number of packets that will get out of the link i in one second. (= $1/T_s$)

- Utilization of link i is:

  ρ_i =

- Stability condition of a network is:

- The external workload offered to the network is:

$$\gamma =$$

Where:

γ = total workload in packets per second
γ_{jk} = workload between source *j* and destination *k*
N = total number of sources and destinations

- The internal workload on link *i* is:

$$\lambda_i =$$

Where:

γ_{jk} = workload between source *j* and destination *k*
Π_{jk} = path followed by packets to go from source *j* and destination *k*

- The total internal workload is:

$$\lambda =$$

Where:

λ = total load on all of the links in the network
λ_i = load on link *i*
L = total number of links

- The average length for all paths is given by:

$$\textbf{E[number of links in a path]} = \lambda/\gamma$$

- The average number of items waiting and being served for link i is:

$$\mathbf{r_i} =$$

- The number of packets waiting and being served in the network can be expressed as:

$$\gamma * T =$$

Where:

T = average delay experienced by a packet through the network.

$$T =$$

- T_{ri} is the residence time at each queue. If we assume that each queue can be treated as an independent M/M/1 model (Jackson's Theorem), then:

$$T_{ri} =$$

Where: T_{si} is the service time for link i

$$T_{si} =$$

Where:

- C_i = data rate on the link (in bps)
- $M = 1/\mu$ = average packet length in bits

Example:

5.3 Network Reliability

5.3.1 Introduction

- A network model is a set of facilities. A facility could be a device or a link.
- A network must contain some slack to allow it to function even if some of its facilities have failed.
- Any network facility is either:
  - Working (**p**)
  - Failing (q = **1-p**)
- **MTBF:** Mean Time Between Failures (**f**).
- **MTTR:** Mean Time To Repair (**r**)

- For any facility i, we'll know from measurements of f_i and r_i:

  P_i = Prob [facility i is working] =

  Therefore:

- We assume that all facilities are independent:

  P(ij) = Prob[facility i and facility j are working] =

  P(i|j) = Prob[facility i or facility j is working] =

- Simplest measure of network reliability:

  P_c(G) = Prob[Network is connected]

  Where: **c** stands for the connectivity of the network, and
  G stands for the graph representing the network

$P_c(G)$ = Prob[All nodes are working and there is a spanning tree of working links]

$P_c(G)$ =

- Since enumerating all trees in G requires an exponential amount of effort, **$P_c(G)$** is very difficult (if not impossible) to compute.

  → We seek simpler measures of network reliability.

5.3.2 Reliability of Tree Networks

- A typical enterprise/campus network includes trees:

- Given a tree T:

  $P_c(T)$ = Prob[A tree network, T, being connected]
  = Prob[All components (nodes and links) are working]

  $P_c(T)$ =

- **$P_c(T)$** can also be computed recursively:

  $P_c(T)$ =

  Where: **T-i** is the tree T without node i, and
  j is the link between node i and the rest of the tree

- Given a particular tree with root r:

  $P_c(i)$ = Prob[node i can communicate with root r]

  $P_c(i)$ =

  Where: **j** is the link between nodes i and k, and
  k is the predecessor of node i

  $P_c(r)$ =

- The expected number of nodes communicating with the root r is:

  E(r) =

- This expression can be computed efficiently for any node as follows:

  E(i) = the expected number of nodes communicating with the node i

  E(i) =

- If node i is a leaf, then:

  E(i) =

Example:

- The expected number of node pairs communicating through the root r is:

EPR(r) =

Example:

5.4 References

1. "Telecommunications Network Design Algorithms" by Aaron Kershenbaum, 1993

2. "Queuing Analysis" by William Stalling, 2000

Chapter 6 Network Management

Topics covered:
Network management standards & models. ISO Functional areas of management. Network management tools and systems. SNMP architecture & operations. Network administration.

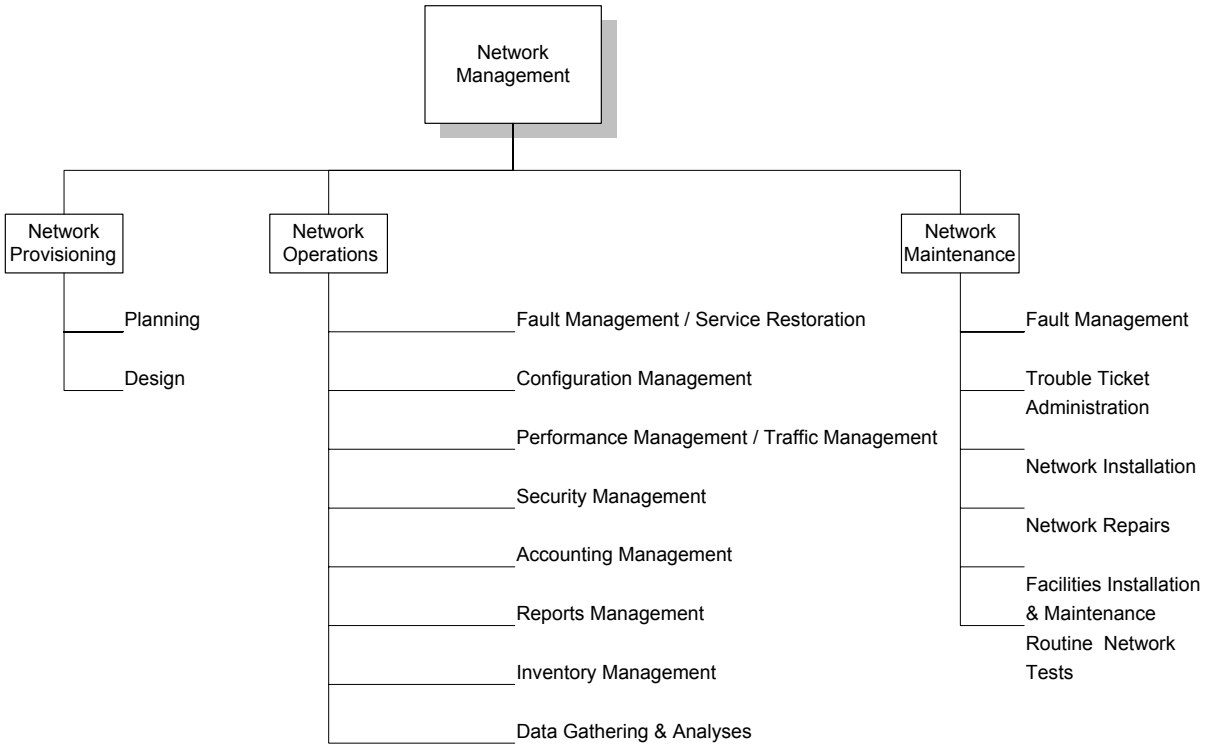
Note: Most of the information in this chapter is taken from [1], and accompanying slides that are © Mani Subramanian 2000

6.1 Introduction

- Network Management is the management of the network resources comprising nodes (e.g., hubs, switches, routers) and links (e.g., connectivity between two nodes).

- System Management is the management of systems and system resources in the network.

- Network Management can also be defined as OAM&P (Operations, Administration, Maintenance, and Provisioning) of network and services.



I Network Management Functional Groupings

- Common Network Problems

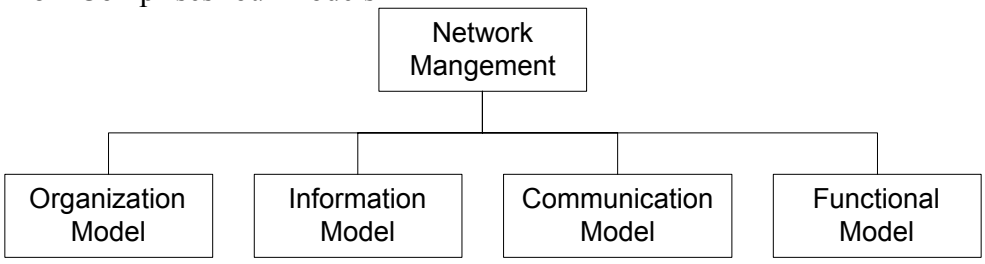
  - Loss of connectivity
  - Duplicate IP address
  - Intermittent problems
  - Network configuration issues
  - Non-problems
  - Performance problems

6.2 Network Management Standards

- NM Standards:
  - OSI/CMIP: Common Management Information Protocol
  - SNMP/Internet: Simple Network Management Protocol (IETF)
  - TMN: Telecommunications Management Network (ITU-T)
  - IEEE standards
  - Web-based Management
- SNMP is the most widely used
- SNMP and CMIP:
  - Use polling methodology → additional load on the network
  - Requires dedicated workstations for the NMS (Network Management System)

6.3 Network Management Model

- OSI Network Management Architecture and Model
  - Most superior of all models
  - Comprises four models



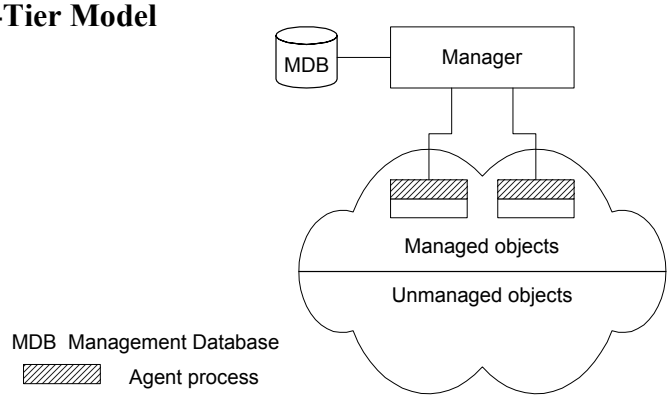
OSI Network Management Model

- SNMP Network Management Architecture and Model:
    - Not defined explicitly.
    - The first 3 models are similar to the OSI models.
    - Addresses the functional model in terms of operations, administration, and security.

6.3.1 Organization Model

- Describes components of network management and their relationship
- Defines the terms: object, agent and manager
- Manager
    - Manages the managed elements
    - Sends requests to agents
    - Monitors alarms
    - Houses applications
    - Provides user interface
- Agent
    - Gathers information from objects
    - Configures parameters of objects
    - Responds to managers' requests
    - Generates alarms and sends them to managers
- Managed object
    - Network element that is managed
    - Houses management agent
    - All objects are either managed or unmanaged

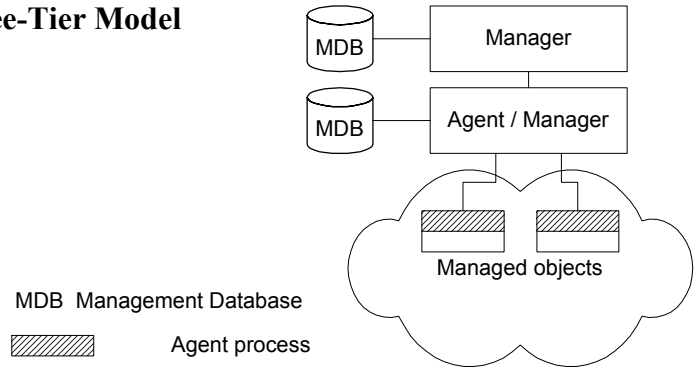
❖ **Two-Tier Model**



Two-Tier Network Mangement Organization Model

- Agent built into network element
  - Example: Managed hub, managed router
- A manager can manage multiple elements
  - Example: Switched hub, ATM switch
- MDB is a physical database
- Unmanaged objects are network elements that are not managed - both physical (unmanaged hub) and logical (passive elements)

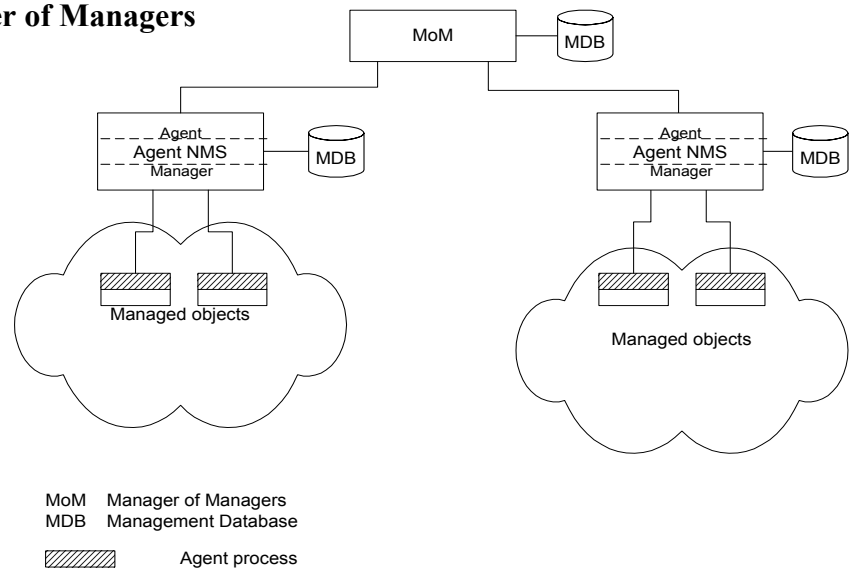
❖ Three-Tier Model



Three-Tier Network Mangement Organization Model

- Middle layer plays the dual role
  - Agent to the top-level manager
  - Manager to the managed objects
- Example of middle level: Remote monitoring probe/agent (RMON)

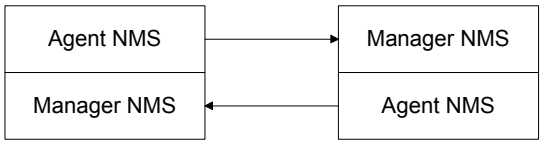
❖ Manager of Managers



Network Mangement Organization Model with MoM

- Agent NMS manages the domain
- MoM presents integrated view of domains
- Domain may be geographical, administrative, vendor-specific products, etc.

❖ **Peer NMSs**



Dual Role of Management Process

- Dual role of both NMSs
- Network management system acts as peers
- Notice that the manager and agent functions are processes and not systems

6.3.2 Information Model

- Concerned with the structure and the storage of information. Similar to information stored in the library (e.g., ISBN)
- Specifies the information base to describe managed objects and their relationships
- The **Structure of Management Information (SMI)** defines for a managed object:
  - Syntax
  - Semantics
  - plus additional information such as status

  Example

  sysDescr: { system 1 }

Syntax:	OCTET STRING
Definition:	"A textual description of the entity."
Access:	read-only
Status:	mandatory

- **Management Information Base (MIB)**
  - Information base contains information about objects
  - Organized by grouping of related objects
  - Defines relationship between objects
  - It is NOT a physical database. It is a ***virtual* database** that is compiled into management module

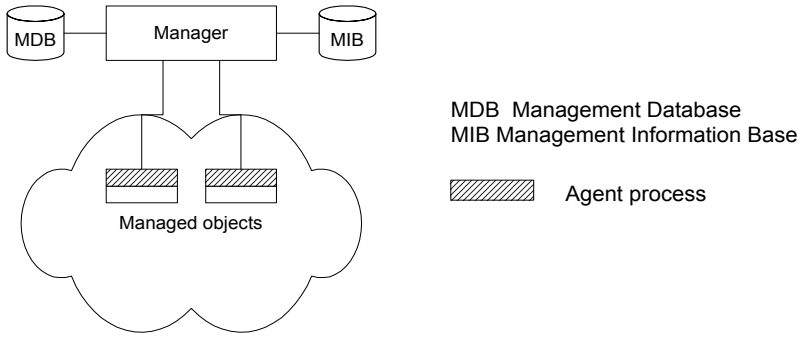
- **MIB View and Access of an Object**
  - A managed object has many attributes - its information base
  - There are several operations that can be performed on the objects
  - A user (manager) can view and perform only certain operations on the object by invoking the management agent
  - The view of the object attributes that the agent perceives is the **MIB view**
  - The operation that a user can perform is the MIB access

- **Management Data Base / Information Base**

  - Distinction between MDB and MIB
    - MDB physical database; e.g., Oracle, Sybase
    - MIB virtual database; schema compiled into management software

  - An NMS can automatically discover a managed object, such as a hub, when added to the network

  - The NMS can identify the new object as hub only after the MIB schema of the hub is compiled into NMS software



Network Configuration with Data and Information Base

- **Managed Objects can be:**

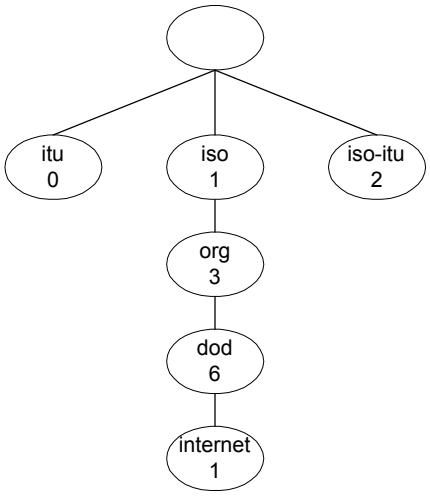
  - Network elements (hardware, system): hubs, bridges, routers, transmission facilities

  - Software (non-physical): programs, algorithms

  - Administrative information: contact person, name of group of objects (IP group)

6.3.2.1 Management Information Tree (MIT)

- Managed objects are uniquely defined by a tree structure specified by the OSI model.



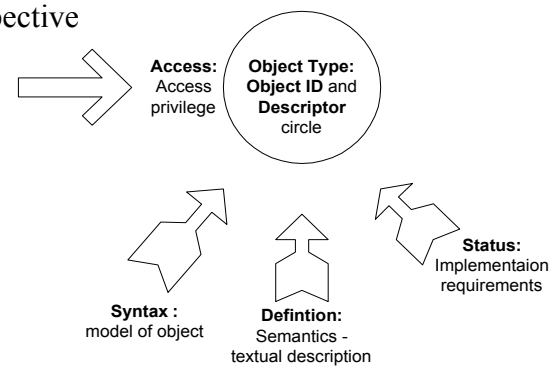
OSI Management Information Tree

- Each node is a managed object (e.g., the Internet is designated as 1.3.6.1)

  → All Internet-managed objects start with 1.3.6.1

6.3.2.2 Managed Objects

- Internet Perspective



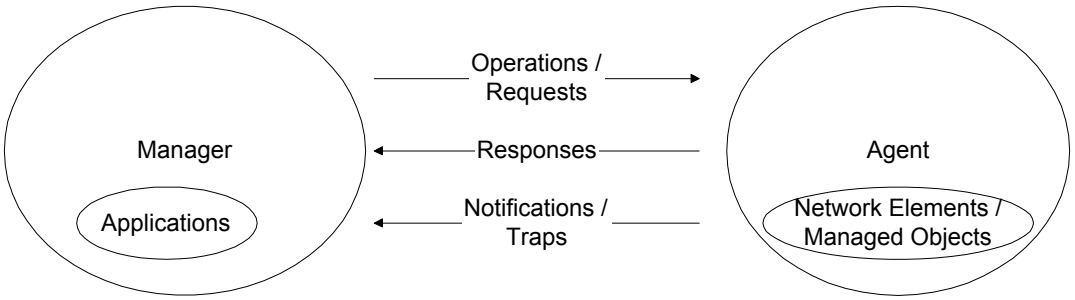
a) Internet Perspective

- Internet specifications for the object "Packet Counter"

Characteristics	Example
Object type	PktCounter
Syntax	Counter
Access	Read-only
Status	Mandatory
Description	Counts number of packets

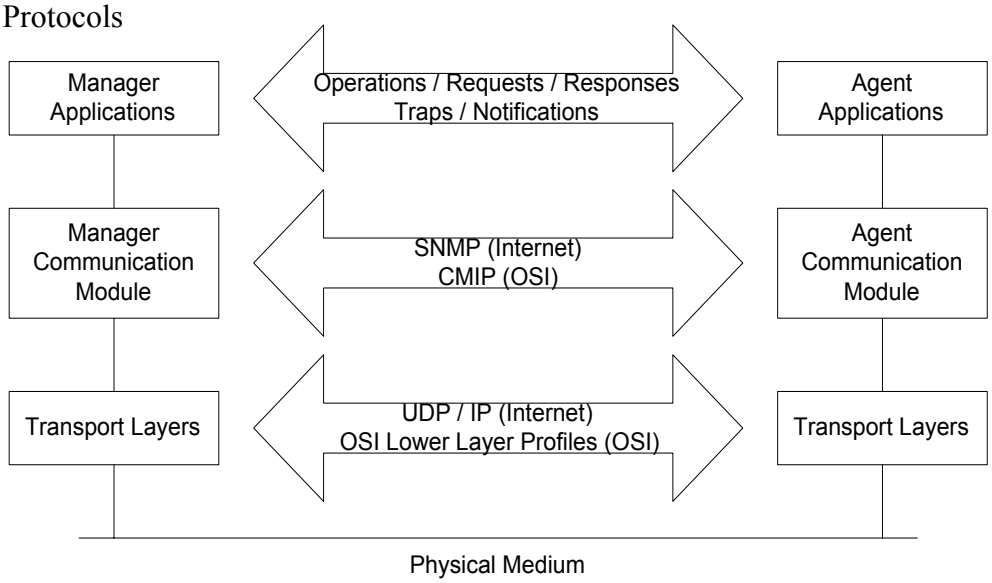
6.3.3 Communication Model

- Addresses the way information is exchanged between systems (agents/managers)



Management Message Communication Model

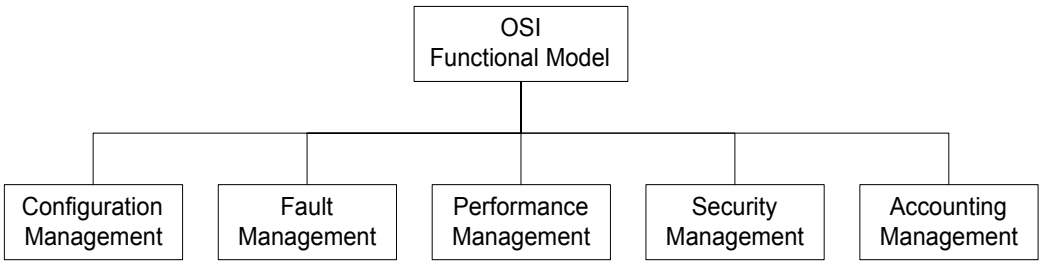
- Transfer Protocols



Management Communication Transfer Protocols

6.3.4 Functional Model

- Addresses the user-oriented applications
- Formally specified in the OSI model as follows:



6.3.4.1 Configuration Management (CM)

- Basic functionality
  - Set and change network configuration and component parameters
  - Set up alarm thresholds

- Network Provisioning
  - Provisioning of network resources: design, installation and maintenance

- Inventory Management
  - Equipment
  - Facilities
  - Database Considerations

- Network Topology
  - Manual
  - Auto-discovery by NMS using
    - Broadcast *ping*
    - ARP table in devices
  - Views
    - Physical
    - Logical (e.g., VLANs)

6.3.4.2 Fault Management (FM)

- Summary
  - Detection and isolation of failures in network
  - Trouble ticket administration

- Fault is a failure of a network component

- Results in loss of connectivity

- Fault management involves a 5-step process:
  - Fault detection (trouble ticket generated)
    - Polling
    - Traps: *linkDown, egpNeighborLoss*
  - Fault location
    - Detect all components failed and trace down the tree topology to the source
    - Fault isolation by network and SNMP tools
    - Use artificial intelligence / correlation techniques
  - Restoration of service (has higher priority)
  - Identification of root cause of the problem
  - Problem resolution (trouble ticket closed)

6.3.4.3 Performance Management (PM)

- Monitor performance of network
- Tools (e.g., analyzers)
- Performance Metrics
  - Macro-level: throughput, response time, availability, reliability
  - Micro-level: bandwidth, utilization, error rate, peak load, average load
- Data Monitoring and Problem Isolation
  - Normal behavior
  - Abnormal behavior (e.g., excessive collisions, high packet loss, etc)
  - Manual and automatic clearing of alarms
- Performance Statistics
  - Traffic statistics
  - Error statistics
  - Used in
    - QoS tracking
    - Performance tuning
    - Validation of SLA (Service Level Agreement)
    - Trend analysis
    - Facility planning
    - Functional accounting

6.3.4.4 Security Management (SM)

- Security threats
  - Modification of information
  - Masquerade
  - Message stream modification
  - Disclosure
- Secure communication
  - Integrity protection
  - Authentication validation
- Policies and Procedures
- Resources to prevent security breaches
  - Firewalls (e.g., packet filtering using a TCP/UDP port address)
  - Cryptography (encryption)
  - Authentication (e.g., data integrity & data origin)
  - Authorization (e.g., read, read-write, no-access)

6.3.4.5 Accounting Management (AM)

- Functional accounting of network usage
- Least developed
- Usage of resources
- Identification of hidden cost of IT usage

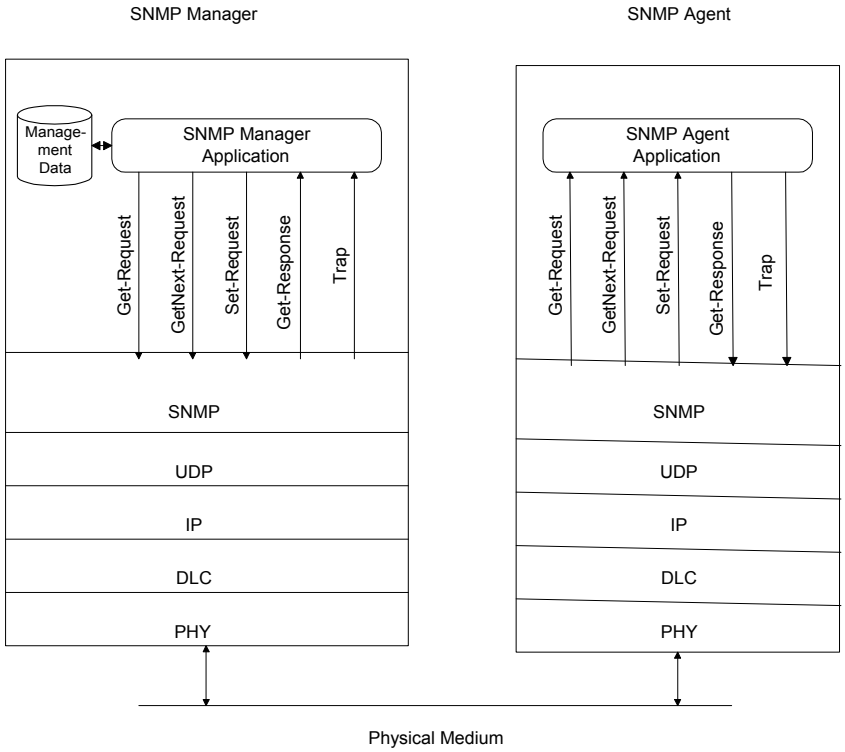
6.4 SNMPv1: Communication Model

An SNMP-based Network Management System consists of 3 main elements:

-
-
-

6.4.1 SNMP Architecture

- Five SNMP messages, three from manager and two from agent.

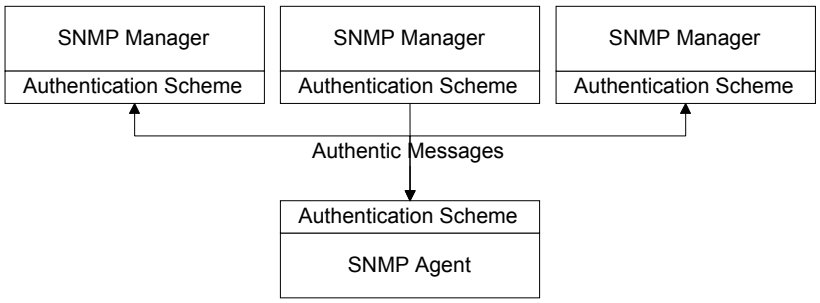


SNMP Network Management Architecture

6.4.2 Administrative Model

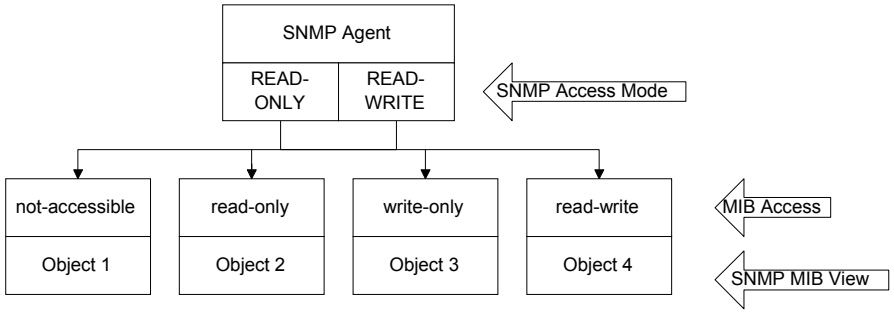
- Based on community profile and policy
- SNMP Entities:
    - SNMP application entities
      - Reside in management stations and network elements
      - Manager and agent
    - SNMP protocol entities
      - Communication processes (PDU handlers)
      - Peer processes that support application entities

6.4.2.1 SNMP Community



- Security in SNMPv1 is community-based
- Authentication scheme is a filter module in manager and agent (e.g., common community name)
- Community: Pairing of two application entities
- Community name: String of octets
- Two applications in the same community communicate with each other
- Application could have multiple community names
- Communication is not secured in SNMPv1 - no encryption

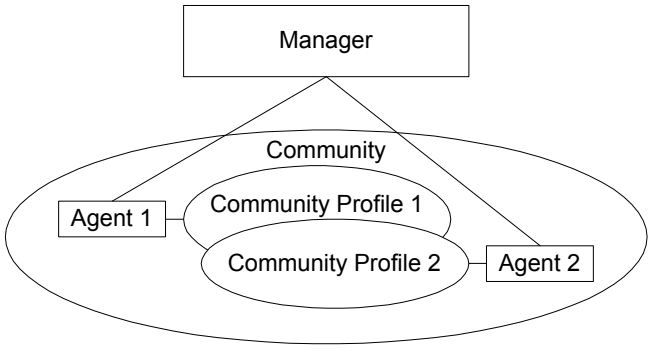
6.4.2.2 SNMP Community Profile



- SNMP MIB view
  - An agent is programmed to view only a subset of managed objects of a network element
- SNMP access mode
  - Each community name is assigned an access mode: read-only and read-write
- Community profile: SNMP MIB view + SNMP access mode
- Operations on an object determined by community profile and the access mode of the managed object
- Total of four access privileges
- Some objects, such as table and table entry are non-accessible
- Most objects available for the public community are read-only.

6.4.2.3 Access Policy

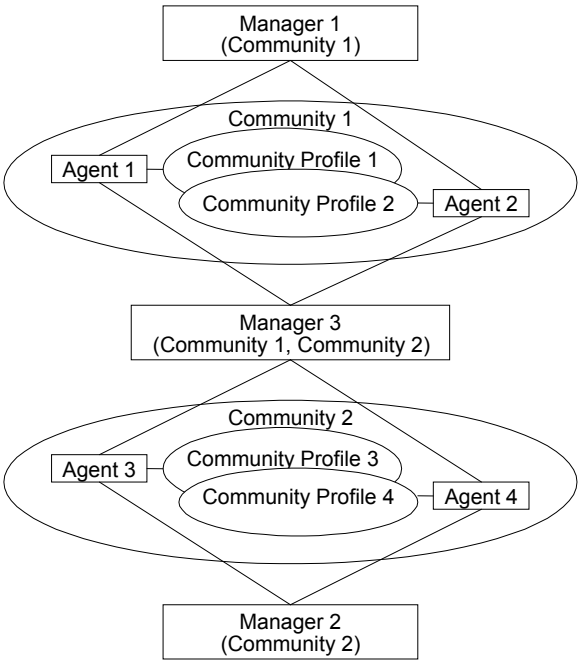
- The SNMP access policy defines the administrative model
- SNMP community paired with SNMP community profile is SNMP access policy



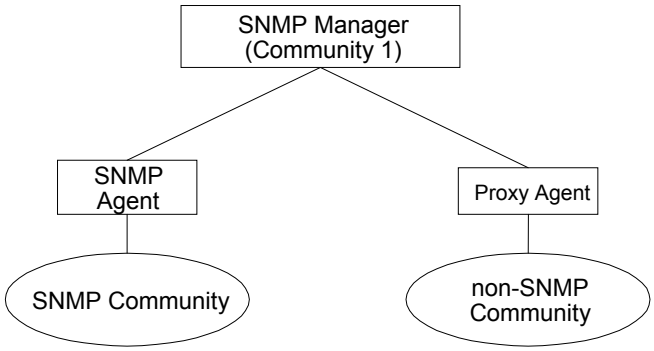
- Manager manages Community 1 and 2 network components via Agents 1 and 2
- Agent 1 has only view of Community Profile 1, e.g. Cisco components
- Agent 2 has only view of Community Profile 2, e.g. 3Com components
- Manager has total view of both Cisco and 3Com components

6.4.2.4 Generalized Administration Model

- Manager 1 manages community 1, manager 2 community 2, and manager 3 (MoM) both communities 1 and 2



6.4.2.5 Proxy Access Policy

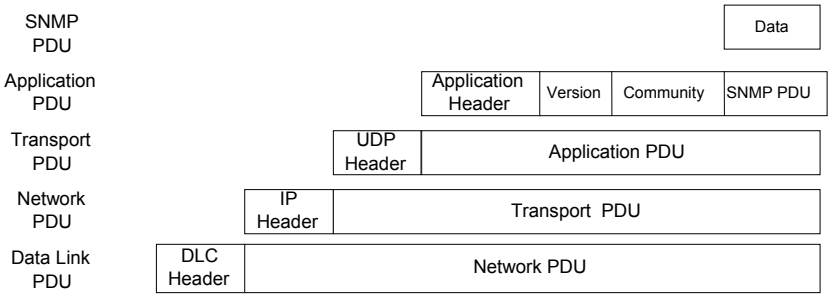


- Proxy agent enables non-SNMP community elements to be managed by an SNMP manager.

- An SNMP MIB is created to handle the non-SNMP objects

6.4.3 SNMP Protocol Specifications

6.4.3.1 Protocol Entities



Encapsulated SNMP Message

- Protocol entities support application entities
- Communication between remote peer processes
- Message consists of
  - Version identifier
  - Community name
  - Protocol Data Unit
- Message encapsulated and transmitted

6.4.3.2 Get and Set PDU

PDU Type	RequestID	Error Status	Error Index	VarBind 1 name	VarBind 1 value	…	VarBind n name	VarBind n value

- VarBindList: multiple instances of VarBind pairs
- PDU Type:

get-request	[0]
get-next-request	[1]
get-response	[2]
set-request	[3]
trap	[4]

- Error in Response

```
ErrorStatus ::=
INTEGER {
        noError(0),
        tooBig(1),
        noSuchName(2),
        badValue(3),
        readOnly(4),
        genErr(5) }
```

- Error Index: No. of VarBind where the first error occurred

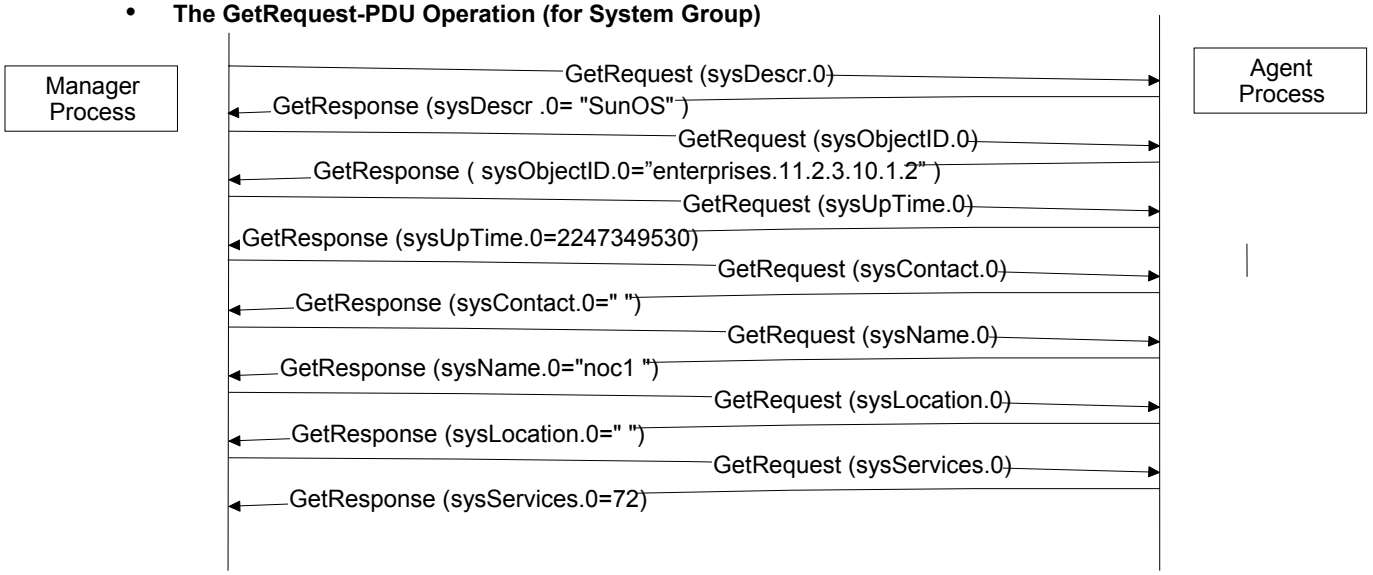
6.4.3.3 Trap PDU

PDU Type	Enterprise	Agent Address	Generic Trap Type	Specific Trap Type	Timestamp	VarBind 1 name	VarBind 1 value	...	VarBind n name	VarBind n value

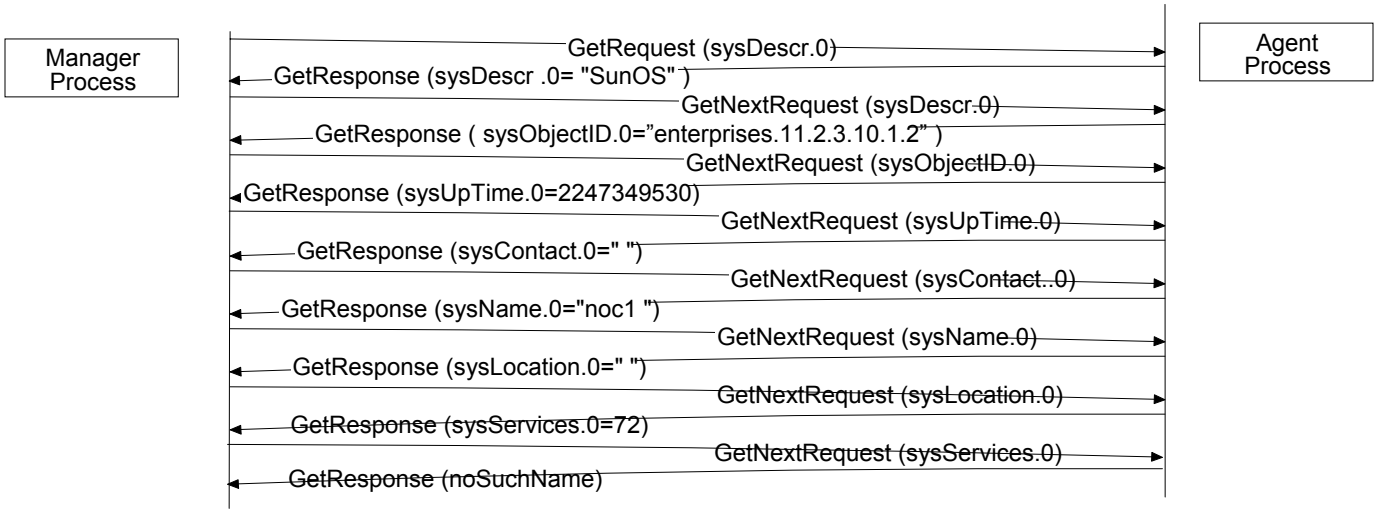
- Enterprise and agent address pertain to the system generating the trap
- Seven generic traps specified by enumerated INTEGER
- The enterprise-specific trap is used by the private organizations to define their device-specific traps. If the Generic Trap type value is 6, the trap is enterprise specific and is defined in a private MIB.
- Timestamp indicates elapsed time since last re-initialization

6.4.4 SNMP Operations

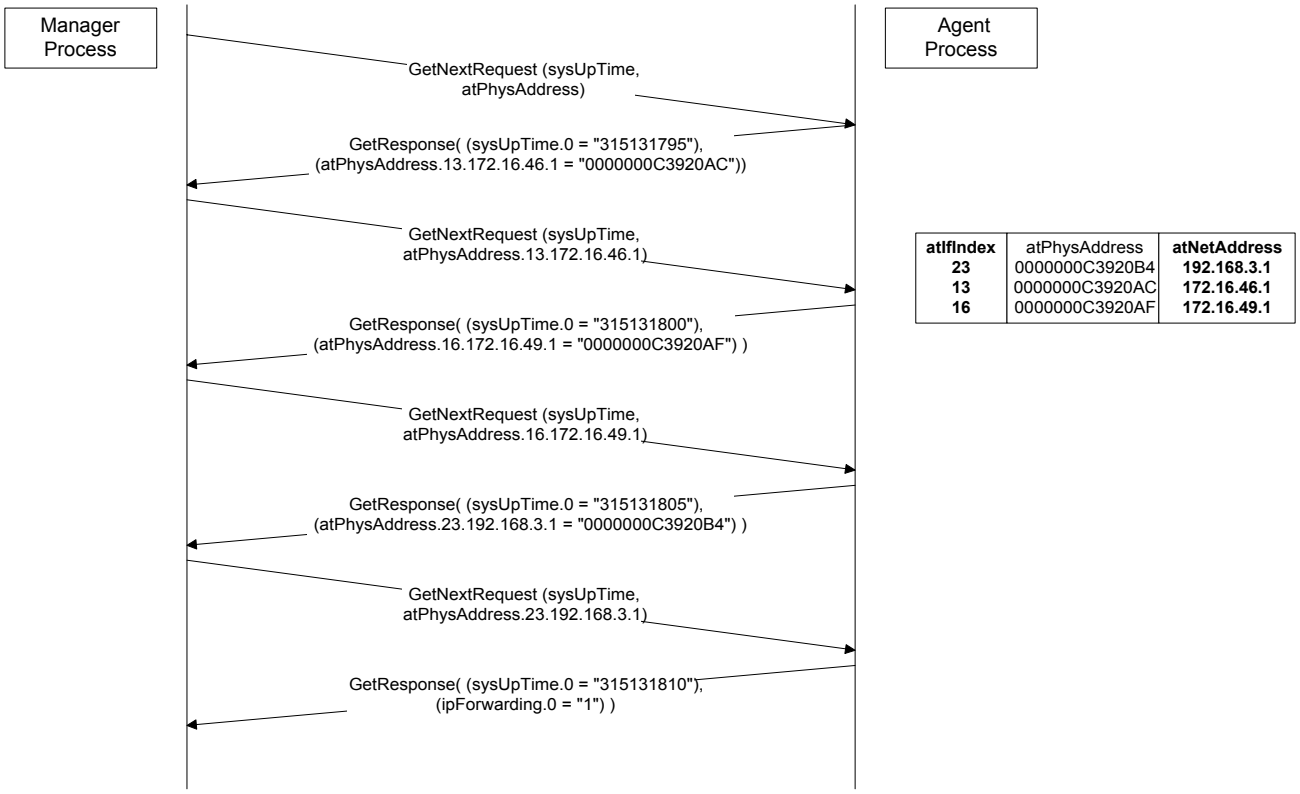
- **The GetRequest-PDU Operation (for System Group)**



- **The GetNextRequest-PDU Operation (for System Group)**



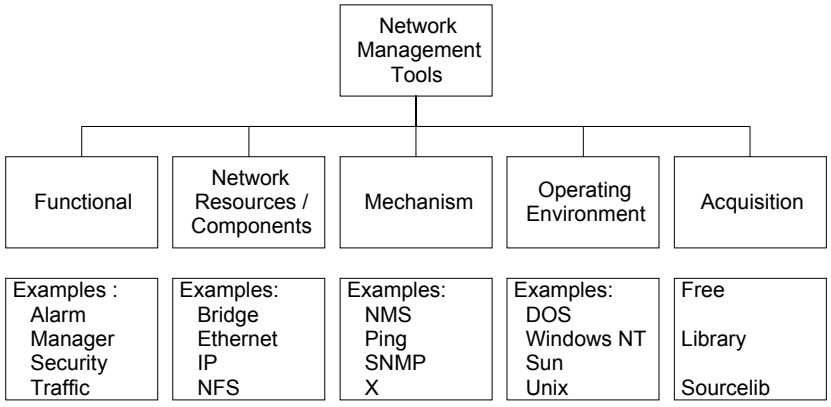
- **GetNextRequest with Indices (faster method)**
    - Uses Lexicographic Order to traverse the MIB subtree
    - A GetNextRequest Example with Indices



GetNextRequest Example with Indices

6.5 SNMP Tools & Systems

6.5.1 Tools Catalog



NOC Tool Categories (RFC 1470)

6.5.2 Network Software Tools

- Status monitoring tools
- Traffic monitoring tools
- Route monitoring tools

6.5.2.1 Network Status Monitoring Tools

NAME	OPERATING SYSTEM	DESCRIPTION
Ifconfig	UNIX	Obtains and configures networking interface parameters and status
ping	UNIX Windows	Checks the status of node / host
nslookup	UNIX Windows NT	Looks up DNS for name / IP address translation
dig	UNIX	Queries DNS server
host	UNIX	Displays information on Internet hosts / domains

6.5.2.2 Network Traffic Monitoring Tools

Name	Operating System	Description
ping	UNIX Windows	Used for measuring roundtrip packet loss
bing	UNIX	Measures point-to-point bandwidth of a link
etherfind	UNIX	Inspects Ethernet packets
snoop	UNIX	Captures and inspects network packets
tcpdump	UNIX	Dumps traffic on a network
getethers	UNIX	Acquires all host addresses of an Ethernet LAN segment
iptrace	UNIX	Measures performance of gateways

6.5.2.3 Network Routing Tools

Name	Operating System	Description
netstat	UNIX	Displays the contents of various network related data structures
arp rarp	UNIX, Windows 95/x/00NT	Displays and modifies the Internet-to Ethernet address translation tables
traceroute tracert	UNIX Windows	Traces route to a destination with routing delays

6.5.3 SNMP MIB Tools

- SNMP MIB Browsers
- SNMP command-line tools

6.5.3.1 SNMP MIB Browsers

- User friendly tools
- May have a GUI
- Specify hostname or IP address & request information on a specific MIB object, MIB group or entire MIB
- Response returns object id(s) and value(s)

6.5.3.2 SNMP Command-Line Tools

- snmpget
- snmpgetnext
- snmpset
- snmptrap
- snmpwalk
- snmpnetstat

6.6 References

1. "Network Management - Principles and Practice" by Mani Subramanian, 2000

2. "TCP/IP Illustrated, Volume 1 - The protocols" by Richard Stevens