Congestion Control

- Congestion control refers to techniques and mechanisms that can either prevent congestion before it happens or remove congestion after it has happened.
- In general, we can divide congestion control mechanisms into two broad categories: **open-loop congestion control** (prevention) and **closed-loop congestion control** (removal).

Open-Loop Congestion Control:

- In open-loop congestion control, policies are applied to prevent congestion before it happens. In these mechanisms, congestion control is handled by either the source or the destination.

Retransmission Policy:

- Retransmission is sometimes unavoidable. If the sender feels that a sent packet is lost or corrupted, the packet needs to be retransmitted.
- Retransmission in general may increase congestion in the network.
- The retransmission policy and the retransmission timers must be designed to optimize efficiency and at the same time prevent congestion.

Window Policy:

- The type of window at the sender may also affect congestion.
- The Selective Repeat window is better than the Go-Back-N window for congestion control. In the Go-Back-N window, when the timer for a packet times out, several packets may beresent.
- This duplication may make the congestion worse. The Selective Repeat window, on the other hand, tries to send the specific packets that have been lost or corrupted.

Acknowledgment Policy:

- The acknowledgment policy imposed by the receiver may also affect congestion.

Discarding Policy:

- A good discarding policy by the routers may prevent congestion and at the same time may not harm the integrity of the transmission.

Admission Policy:

- An admission policy, which is a quality-of-service mechanism can also prevent congestion in virtual-circuitnetworks.

Closed-Loop Congestion Control:

- Closed-loop congestion control mechanisms try to alleviate congestion after it happens.

Backpressure:

- The technique of backpressure refers to a congestion control mechanism in which a congested node stops receiving data from the immediate upstream node or nodes.
- This may cause the upstream node or nodes to become congested, and they, in turn, reject data from their upstream node or nodes, and so on.

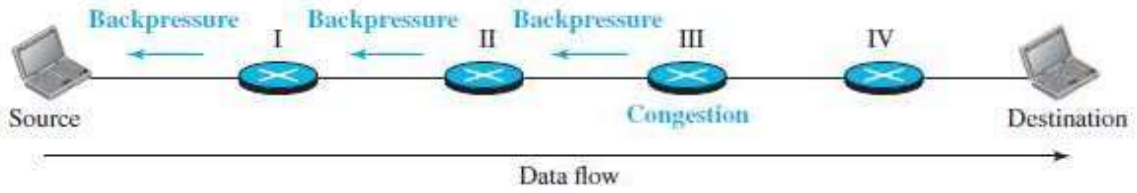


Fig: Backpressure method for alleviating congestion.

Choke Packet:

- A **choke packet** is a packet sent by a node to the source to inform it of congestion. Note the difference between the backpressure and choke-packet methods.
- In backpressure, the warning is from one node to its upstream node, although the warning may eventually reach the source station.
- In the choke-packet method, the warning is from the router, which has encountered congestion, directly to the source station.

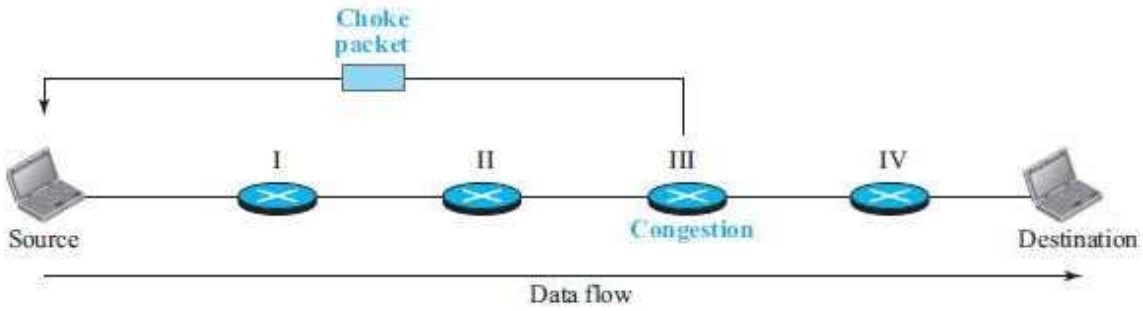


Fig: Choke packet.

Implicit Signaling:

- In implicit signaling, there is no communication between the congested node or nodes and the source.

- The source guesses that there is congestion somewhere in the network from other symptoms.

Explicit Signaling:

- The node that experiences congestion can explicitly send a signal to the source or destination.

Dynamic Host Configuration Protocol (DHCP)

- **Dynamic Host Configuration Protocol (DHCP).** DHCP is an application-layer program, using the client-server paradigm, that actually helps TCP/IP at the network layer.
- DHCP has found such widespread use in the Internet that it is often called a plugandplay protocol. In can be used in many situations.

DHCP Message Format:

- DHCP is a client-server protocol in which the client sends a request message and theserver returns a responsemessage.

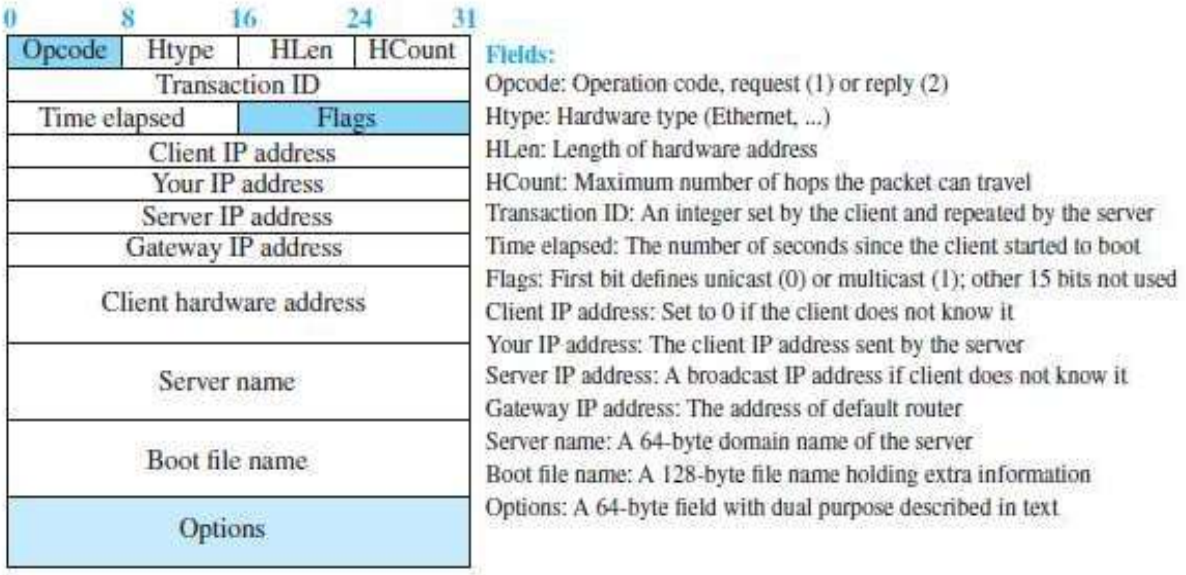


Fig: DHCP message format

- The 64-byte option field has a dual purpose. It can carry either additional informationor some specific vendor information.
- The server uses a number, called a **magiccookie,** in the format of an IP address with the value of 99.130.83.99.When the client finishes reading the message, it looks for this magic cookie.
- If present, the next 60 bytes are options. An option is composed of three fields: a 1-byte tag field, a 1-byte length field, and a variable-length value field.
- There are several tag fields that are mostly used by vendors. If the tag field is 53.

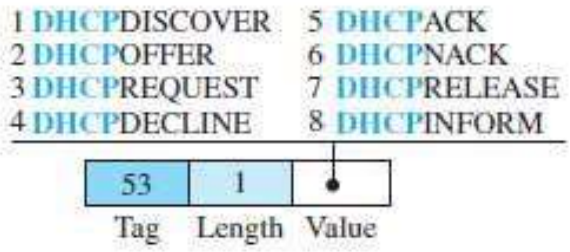


Fig: Option format.

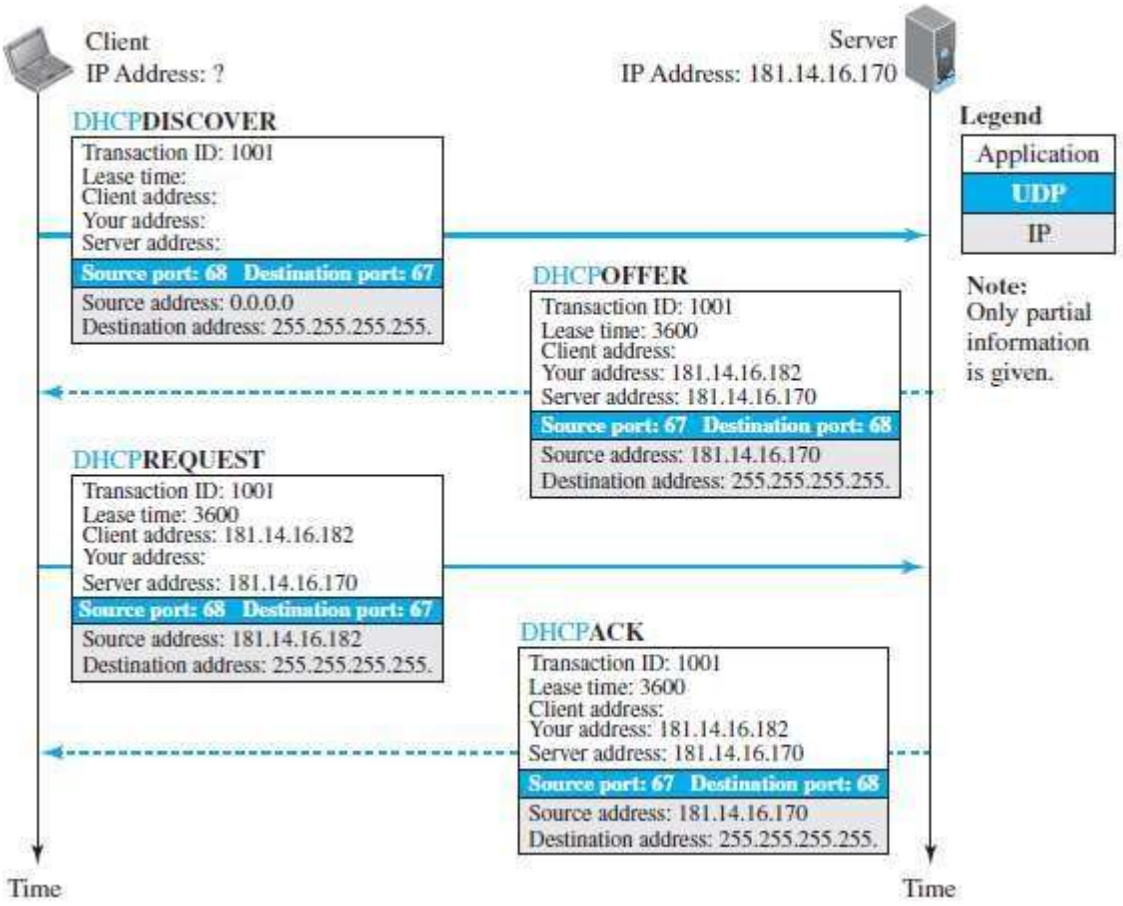


Fig: Operation of DHCP.

1. The joining host creates a DHCPDISCOVER message in which only the transaction-ID field is set to a random number. No other field can be set because the host has no knowledge with which to do so. This message is encapsulated in a UDP user datagram with the source port set to 68 and the destination port set to 67. We will discuss the reason for using two well-known port numbers later. The user datagram is encapsulated in an IP datagram with the source address set to **0.0.0.0** ("thishost") and the destination address set to **255.255.255.255** (broadcast

address).The reason is that the joining host knows neither its own address nor the server address.

2. The DHCP server or servers (if more than one) responds with a DHCPOFFER message in which the your address field defines the offered IP address for the joining host and the server address field includes the IP address of the server.
3. The joining host receives one or more offers and selects the best of them. The joining host then sends a DHCPREQUEST message to the server that has given the best offer.
4. the server sends a DHCPNACK message and the client needs to repeat the process.

Two Well-Known Ports:

- DHCP uses two well-known ports (68 and 67).

Error Control:

- DHCP uses the service of UDP, which is not reliable. To provide error control, DHCP usestwo strategies.
- First, DHCP requires that UDP use the checksum. the use of the checksum in UDP is optional.
- Second, the DHCP client uses timers and a retransmission policy if it does not receive the DHCP reply to a request.

Transition States:

- The operation of the DHCP were very simple. To provide dynamic address allocation, the DHCP client acts as a state machine that performs transitions from one state to another depending on the messages it receives or sends.

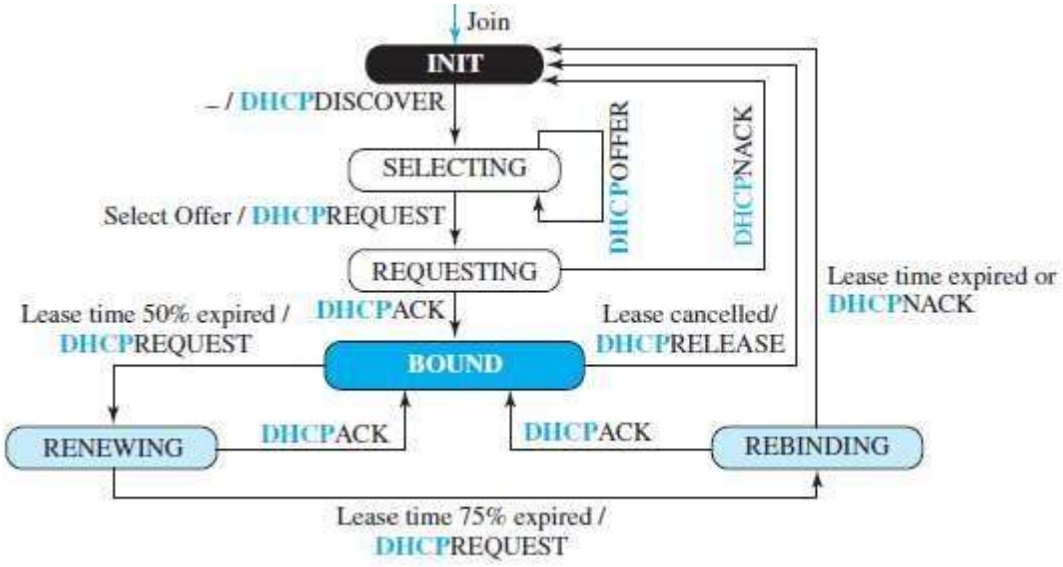


Fig: FSM for the DHCP Client.

- When the DHCP client first starts, it is in the INIT state (initializing state). The client broadcasts a discover message. When it receives an offer, the client goes to the SELECTING state.
- While it is there, it may receive more offers. After it selects an offer, it sends a request message and goes to the REQUESTING state. If an ACK arrives while the client is in this state, it goes to the BOUND state and uses the IP address.
- When the lease is50 percent expired, the client tries to renew it by moving to the RENEWING state. If the server renews the lease, the client moves to the BOUND state again. If the lease is not renewed and the lease time is 75 percent expired, the client moves to the REBINDING state.
- If the server agrees with the lease (ACK message arrives), the client moves to the BOUND state and continues using the IP address; otherwise, the client moves to the INIT state and requests another IP address.
- Note that the client can use the IP address only when it is in the BOUND, RENEWING, or REBINDING state. The above procedure requires that the client uses three timers: renewal timer (set to 50 percent of the lease time), rebinding timer (set to 75 percent of the lease time), and expiration timer (set to the lease time).

Network Address Resolution (NAT):

- The technology allows a site to use a set of private addresses for internal communication and a set of global Internet addresses (at least one) for communication with the rest of the world. The site must have only one connection to the global Internet through a NAT-capable router that runs NAT software.

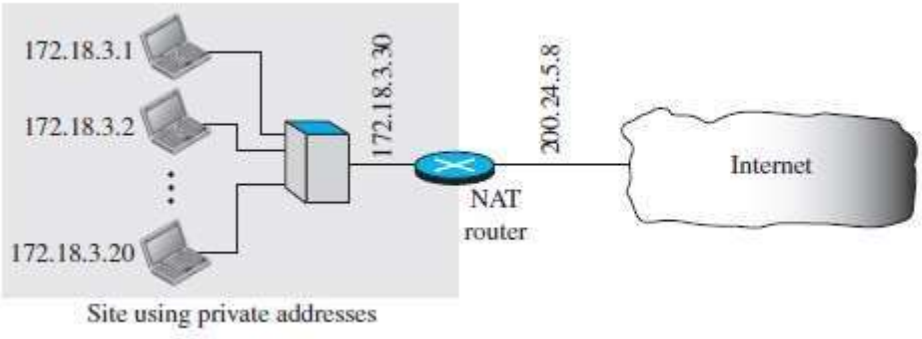


Fig: NAT.

- All of the outgoing packets go through the NAT router, which replaces the source address in the packet with the global NAT address.

- All incoming packets also pass through the NAT router, which replaces the destination address in the packet (the NAT router global address) with the appropriate private address.

Translation Table:

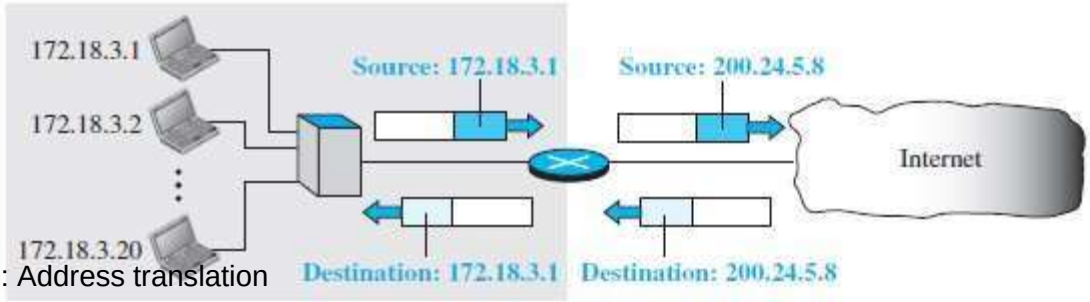


Fig: Address translation

Using One IP Address In this strategy, the router has only one IP address for the private addresses

- In its simplest form, a translation table has only two columns: the private address and the external address (destination address of the packet).
- When the router translates the source address of the outgoing packet, it also makes note of the destination address—where the packet is going.
- When the response comes back from the destination, the router uses the source address of the packet.

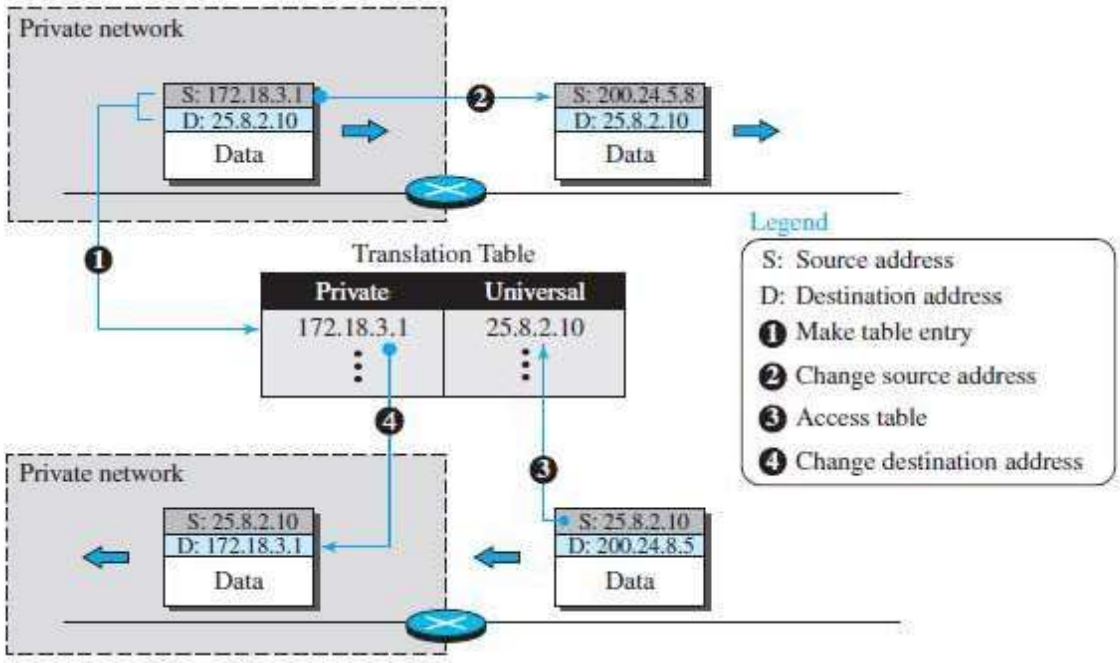


Fig: Translation

Using a Pool of IP Addresses:

- The use of only one global address by the NAT router allows only one private-network host to access a given external host. To remove this restriction, the NAT router can use a pool of global addresses.

Using Both IP Addresses and Port Addresses:

- To allow a many-to-many relationship between private-network hosts and external server programs, we need more information in the translation table.

Table: Five- column translation table

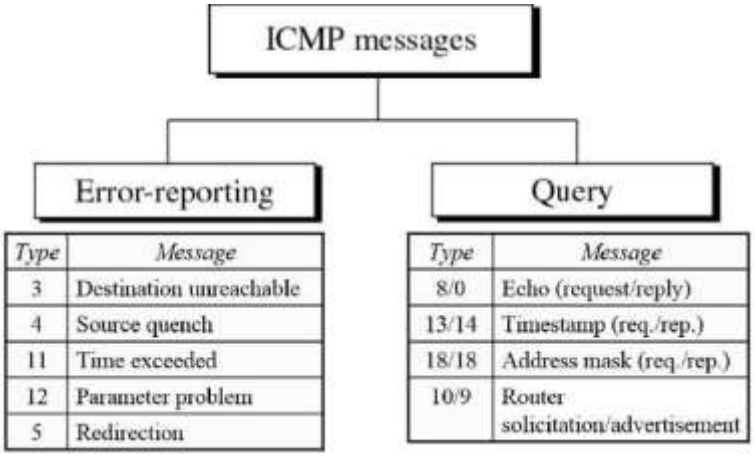
Private Address	Private Port	External Address	External Port	Transport protocol
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
. . .	. . .	. . .	. . .	. . .

ICMP -Internet Control Message Protocol

- ICMP is a network-layer protocol.
- It is a companion to the IP protocol.
- Internet Control Message Protocol (ICMP) defines a collection of error messages that are sent back to the source host whenever a router or host is unable to process an IP datagram successfully.

ICMP MESSAGE TYPES

- ICMP messages are divided into two broad categories: error-reporting messages and query messages.
  - The error-reporting messages report problems that a router or a host (destination) may encounter when it processes an IP packet.
- The query messages help a host or a network manager get specific information from a router or another host.



Type	Message
3	Destination unreachable
4	Source quench
11	Time exceeded
12	Parameter problem
5	Redirection

Type	Message
8/0	Echo (request/reply)
13/14	Timestamp (req./rep.)
18/18	Address mask (req./rep.)
10/9	Router solicitation/advertisement

ICMP Error – Reporting Messages

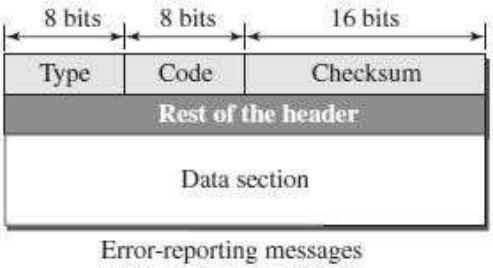
- **Destination Unreachable**—When a router cannot route a datagram, the datagram is discarded and sends a destination unreachable message to source host.
- **Source Quench**—When a router or host discards a datagram due to congestion, it sends a source-quench message to the source host. This message acts as flow

control.

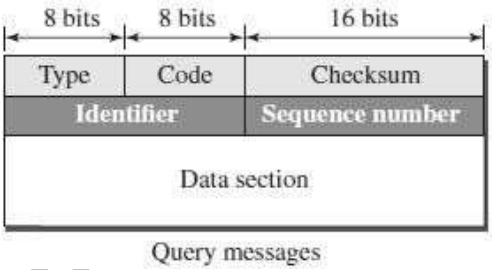
- **Time Exceeded**—Router discards a datagram when TTL field becomes 0 and a time exceeded message is sent to the source host.
- **Parameter Problem**—If a router discovers ambiguous or missing value in any field of the datagram, it discards the datagram and sends parameter problem message to source.
- **Redirection**—Redirect messages are sent by the default router to inform the source host to update its forwarding table when the packet is routed on a wrong path.

ICMP MESSAGE FORMAT

An ICMP message has an 8-byte header and a variable-size data section.



Error-reporting messages



Query messages

Type	Defines the type of the message
Code	Specifies the reason for the particular message type
Checksum	Used for error detection
Rest of the header	Specific for each message type
Data	Used to carry information
Identifier	Used to match the request with the reply
Sequence Number	Sequence Number of the ICMP packet

ICMP DEBUGGING TOOLS

Two tools are used for debugging purpose. They are (1) Ping (2) Traceroute

Ping

- The *ping* program is used to find if a host is alive and responding.
- The source host sends ICMP echo-request messages; the destination, if alive, responds with ICMP echo-reply messages.
- The *ping* program sets the identifier field in the echo–request and echo–reply message and starts the sequence number from 0; this number is incremented by 1 each time a new message is sent.
- The ping program can calculate the round-trip time.
- It inserts the sending time in the data section of the message.
- When the packet arrives, it subtracts the arrival time from the departure time to get the round-trip time (RTT).

$ ping google.com

Traceroute or Tracert

- The *traceroute* program in UNIX or *tracert* in Windows can be used to trace the path of a packet from a source to the destination.
- It can find the IP addresses of all the routers that are visited along the path.
- The program is usually set to check for the maximum of 30 hops (routers) to be visited.
- The number of hops in the Internet is normally less than this.

$ traceroute google.com

IPV4 ADDRESSES:

- An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a host or a router to the Internet.
- The IP address is the address of the connection, not the host or the router, because if the device is moved to another network, the IP address may be changed.

Address Space:

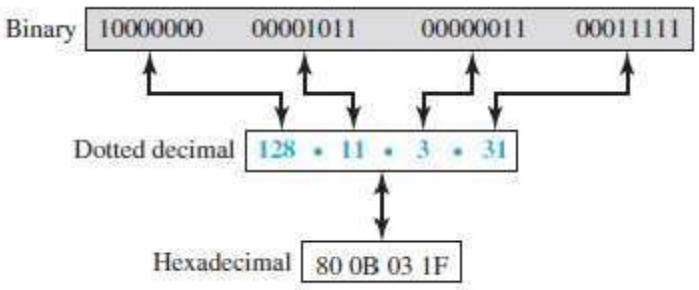
- A protocol like IPv4 that defines addresses has an address space. An **address space** is the total number of addresses used by the protocol.
- If a protocol uses b bits to define an address, the address space is 2^b.

Notation:

- There are three common notations to show an IPv4 address: binary notation (base 2), dotted-decimal notation (base 256), and hexadecimal notation (base 16).
- In binary notation, an IPv4 address is displayed as 32 bits. To make the address more readable, one or more spaces are usually inserted between each octet (8 bits). Each octet is often referred to as a byte.
- To make the IPv4 address more compact and easier to read, it is usually written in decimal form with a decimal point (dot) separating the bytes.
- This format is referred to as dotted-decimal notation. each number in the dotted-decimal notation is between 0 and 255.
- IPv4 address in hexadecimal notation. Each hexadecimal digit is equivalent to four bits. This means that a 32-bit address has 8 hexadecimal digits. This notation is often used in network programming.

Fig: Three different notations in IPV4 addressing

- A 32-bit IPv4 address is also hierarchical, but divided only into two parts.



- The firstpart of the address, called the prefix, defines the network; the second part of the address, called the suffix, defines the node.

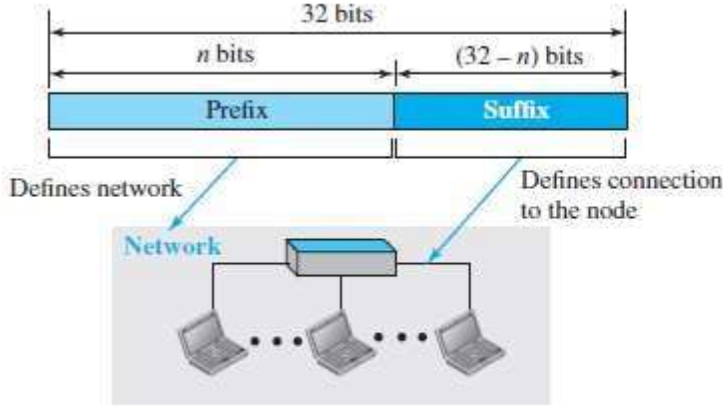
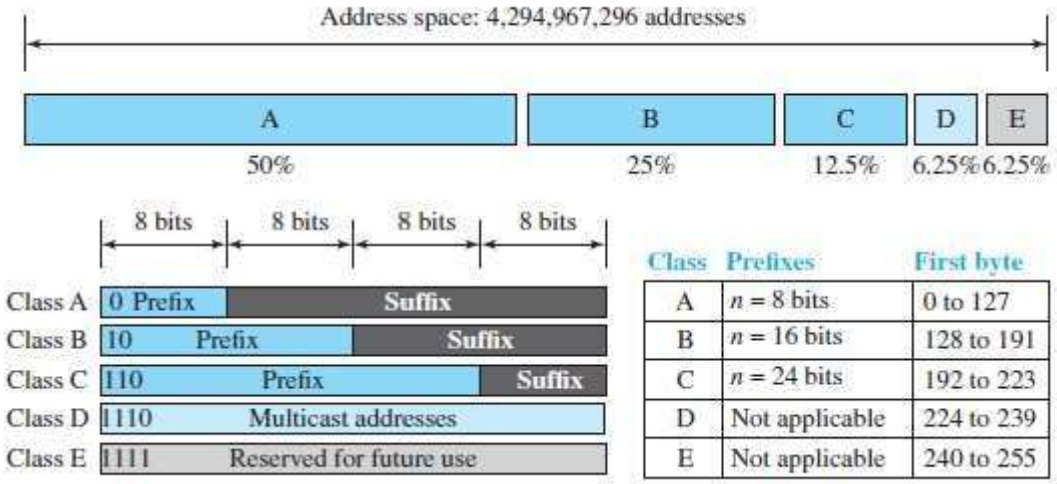


Fig: Hierarchy in addressing.

- IPv4 was first designed as a fixed-length prefix. This scheme, isreferred to as classful addressing. Classless addressing, uses a variable-length network prefix.

Classful Addressing:

- IPv4 address was designed with a fixed-length prefix, butto accommodate both small and large networks, three fixed-length prefixes were designed instead of one (n = 8, n = 16, and n = 24).
- The whole address space was divided into five classes (class A, B, C, D, and E), as shown in Fig This scheme is referred to as **classful addressing.** class A, the network length is 8 bits, the first bit, which is 0, 2^7 = 128.
- In class B, the network length is 16 bits, but since the first two bits, which are $(10)^2$, define the class, we can have only 14 bits as the network identifier. This means there are only 2^{14} = 16,384 networks in the world that can have a class B address. All addresses that start with $(110)^2$ belong to class C.
- In class C, the network length is 24 bits, but since three bits define the class, we can have only 21 bits as the network identifier. This means there are 2^{21} = 2,097,152 networks in the world that can have a class C address.



Class	Prefixes	First byte
A	$n = 8$ bits	0 to 127
B	$n = 16$ bits	128 to 191
C	$n = 24$ bits	192 to 223
D	Not applicable	224 to 239
E	Not applicable	240 to 255

Fig: Occupation of the address space in classful addressing.

- Class D is not divided into prefix and suffix. It is used for multicast addresses. All addresses that start with 1111 in binary belong to class E.
- As in Class D, Class E is not divided into prefix and suffix and is used as reserve.

Address Depletion:

- The reason that classful addressing has become obsolete is address depletion. Since the addresses were not distributed properly, the Internet was faced with the problem of the addresses being rapidly used up, resulting in no more addresses available for organizations and individuals that needed to be connected to the Internet
- Class A can be assigned to only 128 organizations a single network. Class B addresses were designed for midsize organizations, but many of the addresses in this class also remained unused.
- Class C the number of addresses that can be used in each network (256) was so small that most companies.

Subnetting and Supernetting:

- In subnetting, a class A or class B block is divided into several subnets. if a network in class A is divided into four subnets,
- While subnetting was devised to divide a large block into smaller ones, supernetting was devised to combine several class C blocks into a larger block to be attractive to organizations that need more than the 256 addresses available in a class C block.

Advantage of Classful Addressing:

- **A**dvantage: Given an address, we can easily find the class of the address and, since the prefix length for each class is fixed, we can find the prefix length immediately.

Classless Addressing:

- **Classless addressing:** In classless addressing, variable-length blocks are used that belong to no classes.
- We can have a block of 1 address, 2 addresses, 4 addresses, 128 addresses, and so on. In classless addressing, the whole address space is divided into variable length blocks. The prefix in an address defines the block (network); the suffix defines the node (device).

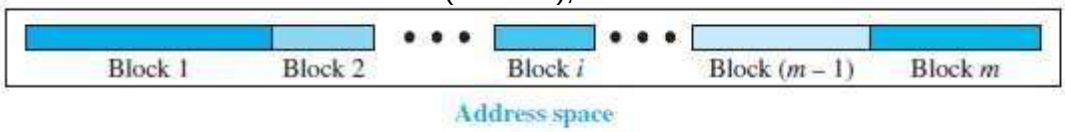


Fig: Variable- length blocks in classless addressing.

Unlike classful addressing, the prefix length in classless addressing is variable.
We can have a prefix length that ranges from 0 to 32.

Prefix Length: Slash Notation:

- The notation is informally referred to as slash notation and formally as **classless interdomain routing** or **CIDR.**

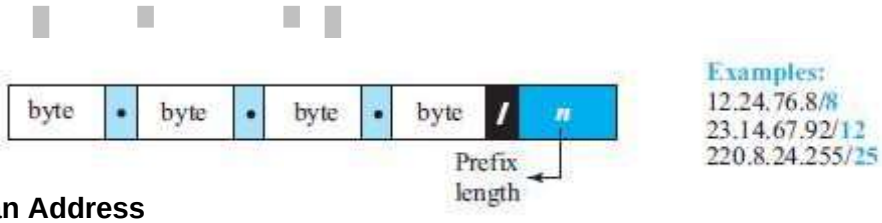


Fig: Slash notation (CIDR)

Extracting Information from an Address

1. The number of addresses in the block is found as $N = 2^{32-n}$.
2. To find the first address, we keep the n leftmost bits and set the (32 - n) rightmost bits all to 0s.
3. To find the last address, we keep the n leftmost bits and set the (32 - n) rightmost bits all to 1s.

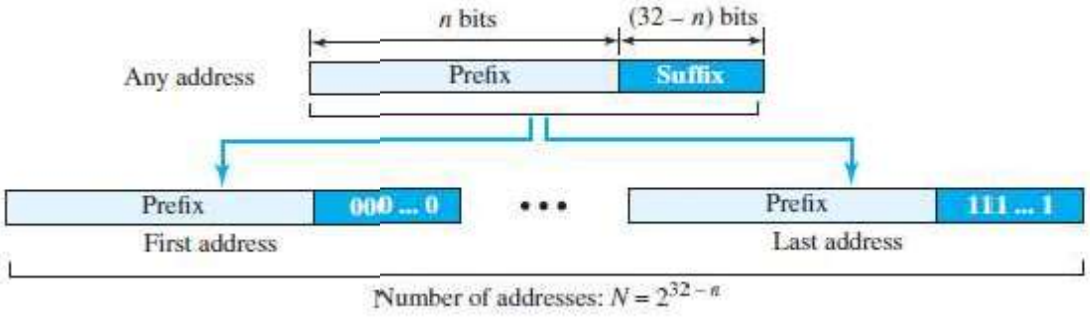


Fig: Information extraction in classes addressing.

Address Mask:

- Another way to find the first and last addresses in the block is to use the address mask.
- The address mask is a 32-bit number in which the n leftmost bits are set to 1s and the rest of the bits (32 − n) are set to 0s.
- The reason for defining a mask in this way is that it can be used by a computer program to extract the information in a block, using the three bit-wise operations NOT, AND, and OR.

1. The number of addresses in the block N = NOT(mask) + 1.
2. The first address in the block = (Any address in the block) AND(mask).
3. The last address in the block = (Any address in the block) OR[(NOT(mask)].

Network Address Resolution (NAT)

- The technology allows a site to use a set of private addresses for internal communication and a set of global Internet addresses (at least one) for communication with the rest of the world. The site must have only one connection to the global Internet through a NAT-capable router that runs NAT software.

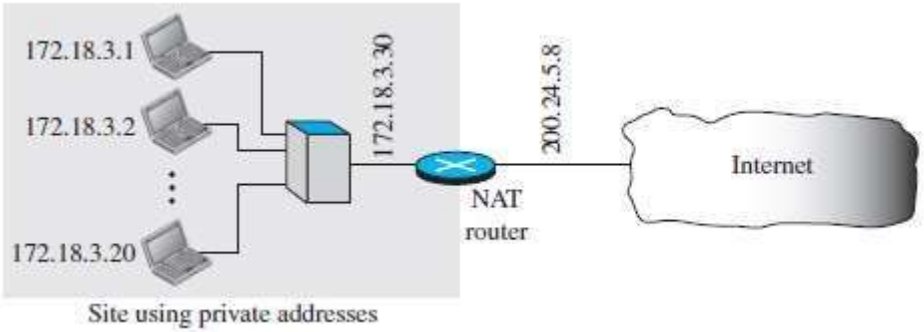


Fig: NAT.

- All of the outgoing packets go through the NAT router, which replaces the source address in the packet with the global NAT address.
- All incoming packets also pass through the NAT router, which replaces the destination address in the packet (the NAT router global address) with the appropriate private address.

Translation Table:

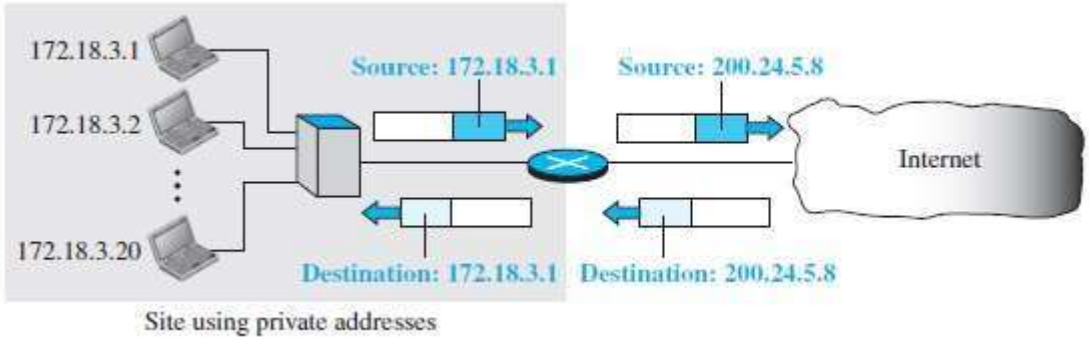


Fig: Address translation

Using One IP Address

- In its simplest form, a translation table has only two columns: the private address and the external address (destination address of the packet).
- When the router translates the source address of the outgoing packet, it also makes note of the destination address—where the packet is going.

- When the response comes back from the destination, the router uses the source address of the packet.

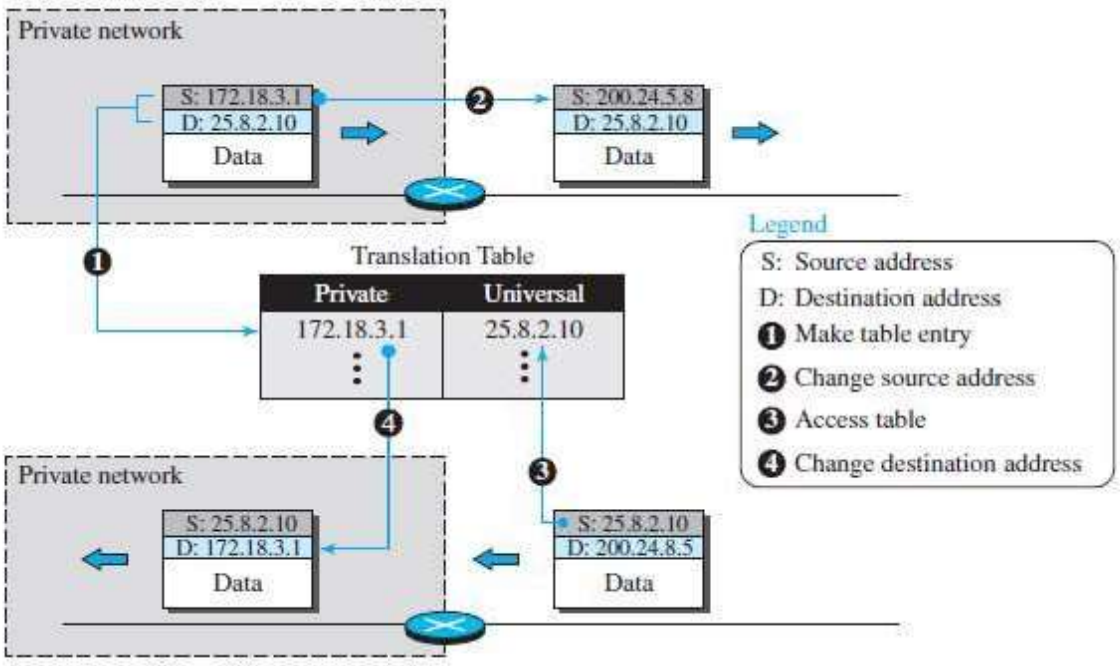


Fig: Translation

Using a Pool of IP Addresses:

- The use of only one global address by the NAT router allows only one private-network host to access a given external host. To remove this restriction, the NAT router can use a pool of global addresses.

Using Both IP Addresses and Port Addresses:

- To allow a many-to-many relationship between private-network hosts and external server programs, we need more information in the translation table.

Table: Five- column translation table

Private Address	Private Port	External Address	External Port	Transport protocol
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

FORWARDING OF IP PACKETS:

Forwarding means to place the packet in its route to its destination.

Forwarding Based on Destination Address:

- This is a traditional approach, forwarding requires a host or a router to have a forwarding table.
- When a host has a packet to send or when a router has received a packet to be forwarded, it looks at this table to find the next hop to deliver the packet to.

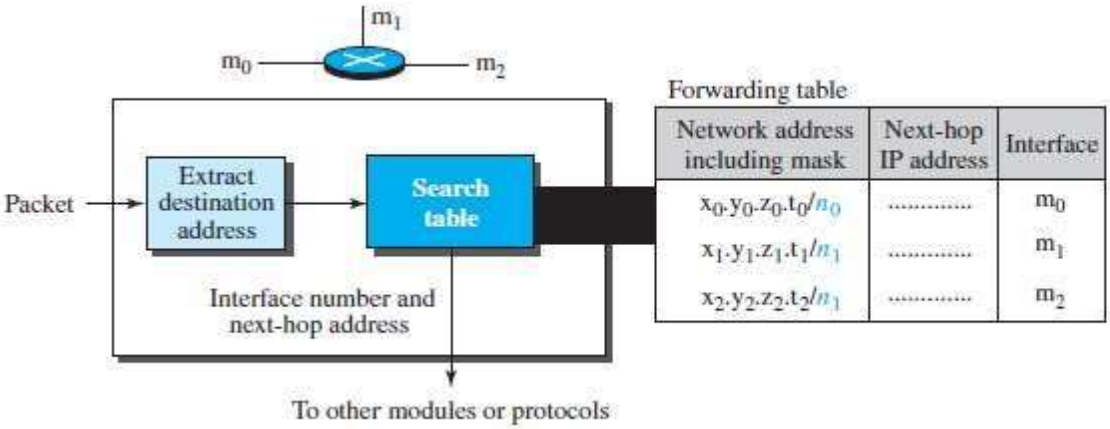


Fig: Simplified forwarding module in classless address.

NETWORK-LAYER SERVICES

Packetizing

The first duty of the network layer is definitely **packetizing:** encapsulating the payload (data received from upper layer) in a network-layer packet at the source and decapsulating the payload from the network-layer packet at the destination.

Routing and Forwarding

Other duties of the network layer, which are as important as the first, are routing and forwarding, which are directly related to each other.

Routing

The network layer is responsible for routing the packet from its source to the destination.

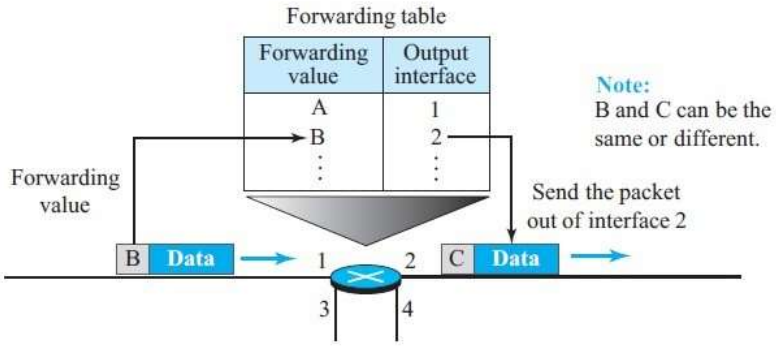
A physical network is a combination of networks (LANs and WANs) and routers that connect them. This means that there is more than one route from the source to thedestination. The network layer is responsible for finding the best one among these possible routes.

Forwarding

Forwarding can be defined as the action applied by each router when a packet arrives at one of its interfaces. The decision-making table a router normally uses for applying this action is sometimes called the *forwarding table* and sometimes the *routing table*. When a router receives a packet from one of its attached networks, it needs to forward the packet to another attached network

To make this decision, the router uses a piece of information in the packet header, which can be the destination address or a label, to find the corresponding output interface number in the forwarding table.

Figure 18.2 *Forwarding process*



Other Services

Error Control

The designers of the network layer, however, have added a checksum field to the datagram to control any corruption in the header, but not in the whole datagram. This checksum may prevent any changes or corruptions in the header of the datagram.

We need to mention that although the network layer in the Internet does not directly provide error control, the Internet uses an auxiliary protocol, ICMP, that provides some kind of error control if the datagram is discarded or has some unknown information in the header.

Flow Control

Flow control regulates the amount of data a source can send without overwhelming the receiver. If the upper layer at the source computer produces data faster than the upper layer at the destination computer can consume it, the receiver will be overwhelmed with data. To control the flow of data, the receiver needs to send some feedback to the
sender to inform the latter that it is overwhelmed with data.

NETWORK- LAYER PERFORMANCE

The performance of a network can be measured in terms of *delay, throughput,* and *packet loss.*

Delay:

The delays in a network can be divided into four types: transmission delay, propagation delay,
processing delay, and queuing delay.

Transmission Delay:

$Delay_{tr}$= (Packet length) / (Transmission rate).

Propagation Delay:

Propagation delay is the time it takes for a bit to travel from point A to point B in the transmission media. propagation delay is

$Delay_{pg}$= (Distance) / (Propagation speed).

Processing Delay

The processing delay is the time required for a router or a destination host to receive a packet from its input port, remove the header, perform an error detection procedure, and deliver the packet to the output port (in the case of a router) or deliver the packet to the upper-layer protocol (in the case of the destination host).

$Delay_{pr}$ = Time required to process a packet in a router or a destination host Queuing Delay

- Queuing delay can normally happen in a router.
- The queuing delay for a packet in a router is measured as the time a packet waits in the input queue and output queue of a router.
  $Delay_{qu}$ = The time a packet waits in input and output queues in a router.

Total Delay

Total delay = (n + 1) ($Delay_{tr}$ + $Delay_{pg}$ + $Delay_{pr}$) + (n) ($Delay_{qu}$)

Throughput:

- Throughput at any point in a network is defined as the number of bits passing throughthe point in a second, which is actually the transmission rate of data at that point.
- Throughput = minimum {TR1, TR2, . . . TRn}.

Packet Loss:

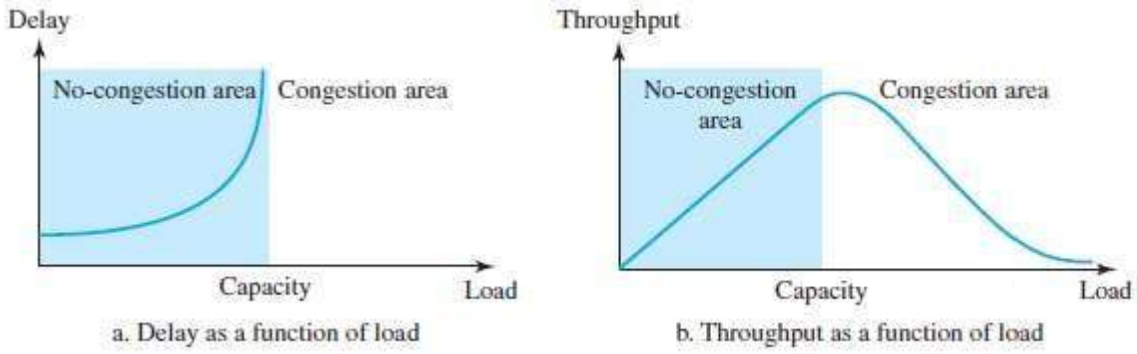


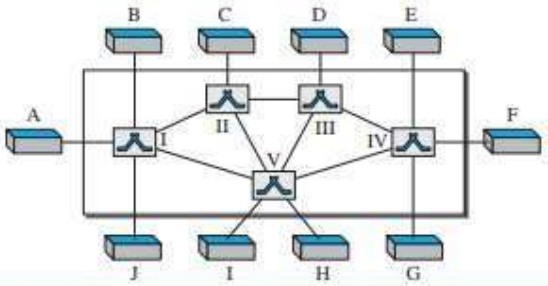
Fig: Pocket delay and throughput as functions of load.

Switching

INTRODUCTION

A switched network consists of a series of interlinked nodes, called switches. Switches are devices capable of creating temporary connections between two or more devices linked to the

Figure 8.1 *Switched network*

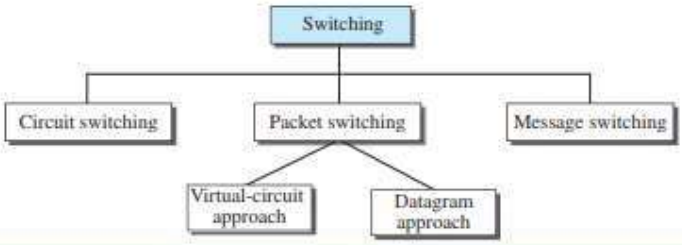


- The above figure is a switching network, the end systems (communicating devices) are labeled A, B, C, D, and so on, and the switches are labeled I, II, III, IV, and V. Each switch is connected to multiple links.

Three Methods of Switching

- Traditionally, three methods of switching have been discussed: circuit switching, packet switching, and messageswitching.
- Packet switching can further be divided into two subcategories—virtualcircuit approach and datagram approach
  - In data communications, we need to send messages from one end system to another. If the message is going to pass through a packet-switched network, it needs to be divided into packets of fixed or variable size. The size of the packet is determined by the network and the governing protocol.

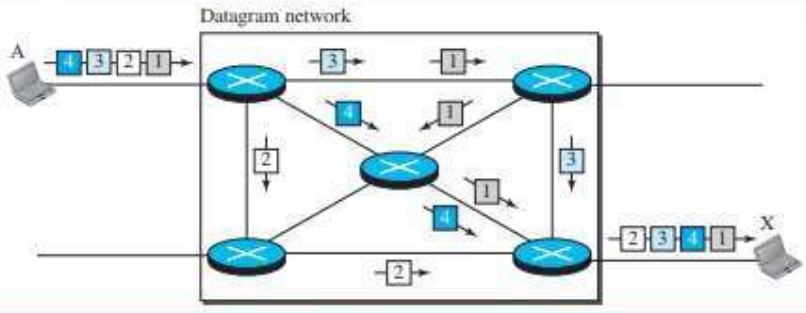
Figure 8.2 *Taxonomy of switched networks*



PACKETSWITCHING

- In a packet-switched network, there is no resource reservation; resources are allocated on demand.
- We can have two types of packet-switched networks: datagram networks and virtual circuit networks.
- **Datagram Networks**
  - In a datagram network, each packet is treated independently of all others. Even if a packet is part of a multipacket transmission, the network treats it as though it existed alone. Packets in this approach are referred to as datagrams.
  - Datagram switching is normally done at the network layer.
  - The datagram networks are sometimes referred to as connectionless networks. The term connectionless here means that the switch (packet switch) does not keep information about the connection state. There are no setup or teardown phases. Each packet is treated the same by a switch regardless of its source or destination.

Figure 8.7 *A datagram network with four switches (routers)*



- A switch in a datagram network uses a routing table that is based on the destination address.
- The routing tables are dynamic and are updated periodically. The destination addresses and the corresponding forwarding output ports are recorded in the tables.
- Every packet in a datagram network carries a header that contains, among other information, the destination address of the packet.
- The destination address in the header of a packet in a datagram network remains the same during the entire journey of the packet.
- **Efficiency** :The efficiency of a datagram network is better than that of a circuit-switched network; resources are allocated only when there are packets to be transferred. If a source sends a packet and there is a delay of a few minutes before another packet can be sent, the resources can be reallocated during these minutes for other packets from other sources.
- **Delay**: There may be greater delay in a datagram network than in a virtual-circuit network. Although there are no setup and teardown phases, each packet may experience a wait at a switch before it is forwarded.

Unicast Routing – Link State Routing

Unicast–Unicastmeansthetransmissionfromasinglesendertoasinglereceiver.Itisapointtopoint communication between sender and receiver. There are various unicast protocols such as TCP, HTTP, etc.

- TCPisthemostcommonlyusedunicastprotocol.Itisaconnectionorientedprotocolthat relay on acknowledgement from the receiver side.
- HTTPstandsforHyperTextTransferProtocol.Itisanobjectorientedprotocol for communication.

There are three major protocols for unicast routing:

1. Distance VectorRouting
2. Link StateRouting
3. Path-Vector Routing

Link State Routing –

Link state routing is the second family of routing protocols. While distance vector routers use a distributed algorithm to compute their routing tables, link-state routing uses link-state routers to exchangemessagesthatalloweachroutertolearntheentirenetworktopology.Basedonthis learned topology, each router is then able to compute its routing table by using a shortest path computation.

Features of link state routing protocols –

- Link state packet – A small packet that contains routing information.
- Link state database – A collection information gathered from link state packet.
- Shortest path first algorithm (Dijkstra algorithm) – A calculation performed on the database results into shortest path
- Routingtable–Alistofknownpathsand

interfaces. Calculation of shortest path –

Tofindshortestpath,eachnodeneedtorunthefamousDijkstraalgorithm.Thisfamousalgorithm uses the following steps:

- Step-1: The node is taken and chosen as a root node of the tree, this creates the tree with a single node, and now set the total cost of each node to some value based on the information in Link State Database
- Step-2: Now the node selects one node, among all the nodes not in the tree like structure, which is nearest to the root, and adds this to the tree. The shape of the tree gets changed .
- Step-3: After this node is added to the tree, the cost of all the nodes not in the tree needs to be updated because the paths may have been changed.
- Step-4: The node repeats the Step 2. and Step 3. until all the nodes are added in the tree

Link State protocols in comparison to Distance Vector protocols have:

1. It requires large amount of memory.
2. Shortest path computations require many CPU circles.
3. If network use the little bandwidth ; it quickly reacts to topology changes
4. All items in the database must be sent to neighbors to form link state packets.
5. All neighbors must be trusted in the topology.
6. Authentication mechanisms can be used to avoid undesired adjacency and problems.
7. No split horizon techniques are possible in the link state routing.

Open shortest path first (OSPF) routing protocol –

- Open Shortest Path First (OSPF) is a unicast routing protocol developed by working group of the Internet Engineering Task Force (IETF).
- It is a intra domain routing protocol.
- It is an open source protocol.
- It is similar to Routing Information Protocol (RIP)
- OSPF is a classless routing protocol, which means that in its updates, it includes the subnet of each route it knows about, thus, enabling variable-length subnet masks. With variable-length subnet masks, an IP network can be broken into many subnets of various sizes. This provides network administrators with extra network-configuration flexibility. These updates are multicasts at specific addresses (224.0.0.5 and 224.0.0.6).
- OSPF is implemented as a program in the network layer using the services provided by the Internet Protocol

- IP datagram that carries the messages from OSPF sets the value of protocol field to 89
- OSPF is based on the SPF algorithm, which sometimes is referred to as the Dijkstra algorithm
- OSPF has two versions – version 1 and version 2. Version 2 is used mostly

OSPF Messages – OSPF is a very complex protocol. It uses five different types of messages. These are as follows:

10. Hello message (Type 1) – It is used by the routers to introduce itself to the other routers.
11. Database description message (Type 2) – It is normally send in response to the Hello message.
12. Link-state request message (Type 3) – It is used by the routers that need information about specific Link-State packet.
13. Link-state update message (Type 4) – It is the main OSPF message for building Link-State Database.
14. Link-state acknowledgement message (Type 5) – It is used to create reliability in the OSPF protocol.