Network Layer

Compiled By :- Er. Nagendra Karn

Network Layer

- The main aim of the **Network Layer** the source-to- destination delivery of a packet across multiple networks (links).
- Whereas the data link layer oversees the delivery of a packet between two systems on the same network (links), the **network layer** ensures that each packet gets from its source to its final destination.
- It also breaks the messages that have to be sent into packets and to assemble incoming packets into messages for higher levels.
- If two systems are attached to the same network, there is usually no need for a **network layer**.
- However, if two systems are attached connecting devices on the different networks (links), so there is
- often needed for the **network layer** to complete the source-to-destination delivery of the message.

Function of Network Layer

- **Routing:** When a packet reaches the router's input link, the router will move the packets to the router's output link. For example, a packet from S1 to R1 must be forwarded to the next router on the path to S2.
- **Logical Addressing:** The data link layer implements the physical addressing and network layer implements the logical addressing. Logical addressing is also used to distinguish between source and destination system. The network layer adds a header to the packet which includes the logical addresses of both the sender and the receiver.
- **Internetworking:** This is the main role of the network layer that it provides the logical connection between different types of networks.
- **Fragmentation:** The fragmentation is a process of breaking the packets into the smallest individual data units that travel through different networks.
- E.g. If I want to access some data from Facebook then I will open my laptop, type URL of Facebook and send an HTTP request to facebook.com for some data. Since the server of Facebook is situated outside my local area network, my request is forwarded to Facebook through the default gateway or router of my institution.

Virtual circuit Vs Datagram subnet

- Computer networks that provide connection-oriented service are called Virtual Circuits while those providing connectionless services are called as Datagram networks.
- For prior knowledge, the Internet which we use is actually based on Datagram network (connectionless) at network level as all packets from a source to a destination do not follow same path.
- **Virtual Circuits:**
  - It is connection-oriented simply meaning that there is a reservation of resources like buffers, CPU, bandwidth etc. for the time in which the newly setup VC is going to be used by a data transfer session.
  - First packet goes and reserves resources for the subsequent packets which as a result follow the same path for the whole connection time.
  - Since all the packets are going to follow the same path, a global header is required only for the first packet of the connection and other packets generally don't require global headers.
  - Since data follows a particular dedicated path, packets reach in order to the destination.
  - Virtual Circuits are highly reliable means of transfer.
  - Since each time a new connection has to be setup with reservation of resources and extra information handling at routers, its simply costly to implement Virtual Circuits.

Datagram Networks

- It is connectionless service. There is no need of reservation of resources as there is no dedicated path for a connection session.
- All packets are free to go to any path on any intermediate router which is decided on the go by dynamically changing routing tables on routers.
- Since every packet is free to choose any path, all packets must be associated with a header with proper information about source and the upper layer data.
- The connectionless property makes data packets reach destination in any order, means they need not reach in the order in which they were sent.
- Datagram networks are not reliable as Virtual Circuits.
- But it is always easy and cost efficient to implement datagram networks as there is no extra headache of reserving resources and making a dedicated each time an application has to communicate.

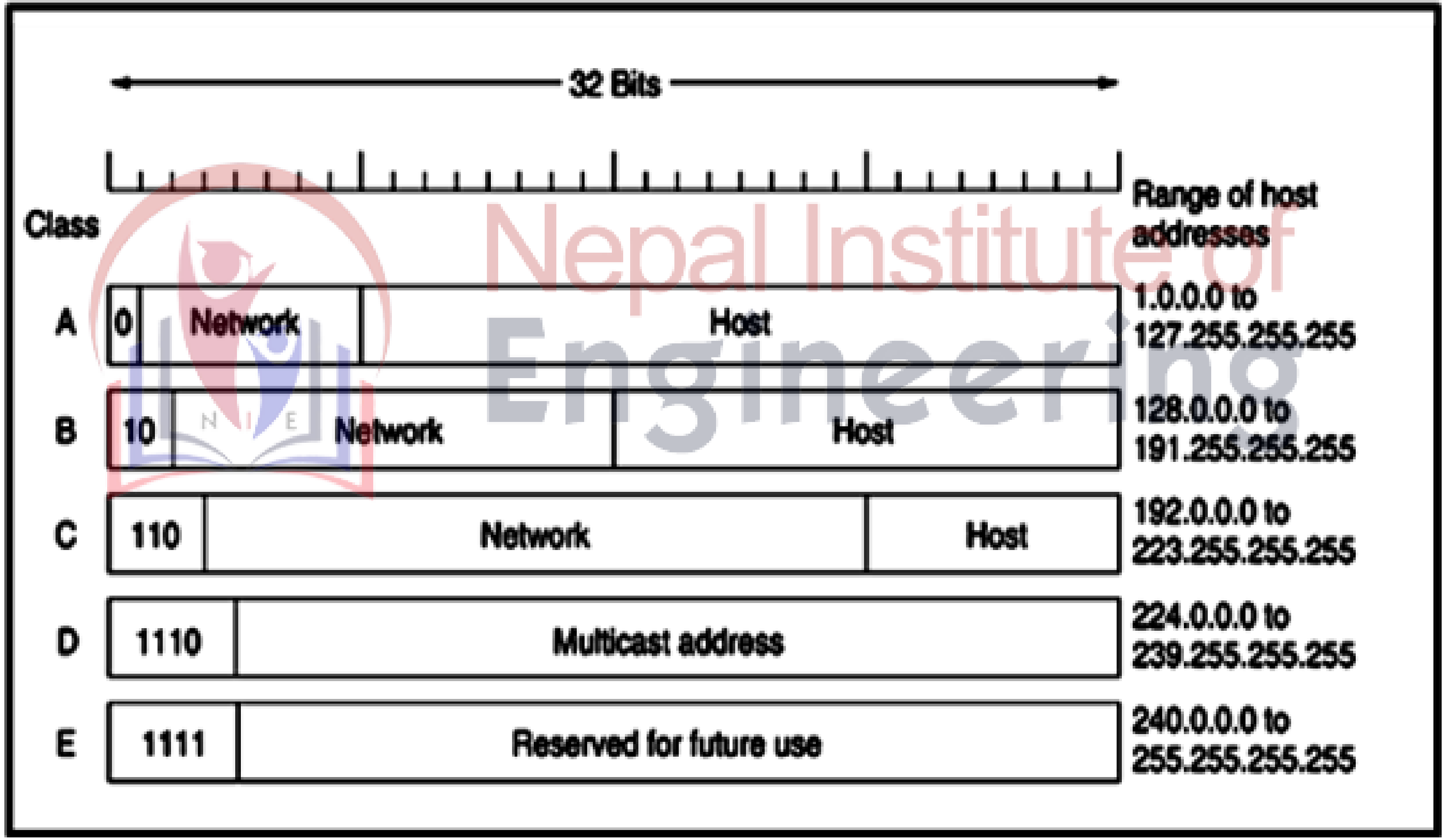
Comparison of Datagram and Virtual Circuit Subnet

	Datagram	Virtual Circuit
Connection Setup	None	Required
Addressing	Packet contains full source and destination address	Packet contains short virtual circuit number identifier.
State Information	None other than router table containing destination network	Each virtual circuit number entered to table on setup, used for routing.
Routing	Packets routed independently.	Route established at setup, all packets follow same route.
Effect of Router failure	Only on packets lost during crash.	All virtual circuits passing through failed router terminated.
Congestion control	Difficult since all packets routed independently router resource requirements can vary.	Simple by pre-allocating enough buffers to each virtual circuit at setup, since maximum number of circuits fixed.

IP Addresses

- An **IPv4** address is a 32-bit address that uniquely and universally defines the connection of a device (e.g., a computer, router) to the internet.
- The 32 binary bits are broken into four octets (1 octets = 8 bits). Each octet is converted to decimal and separted by a period (dot).
- For this reason, an IP address is said to be expressed in dotted decimal format (for example, 192.168.10.1). The value in each octet ranges from 0 to 255 decimal, or 00000000 to 11111111 binary.
- IPv4 addressing at its inception, used the concept of classes. This architecture is called classful addressing.
- In classful addressing, the address space is divided into five classes: A,B,C,D, and E.
- In common usage, the first address is a subset, all binary zero in the host identifier, is reserved for referring to network itself, while the last address, all binary one in the host identifier, is used as a broadcast address for the network; this reduces the number of addresses available for hosts by 2.

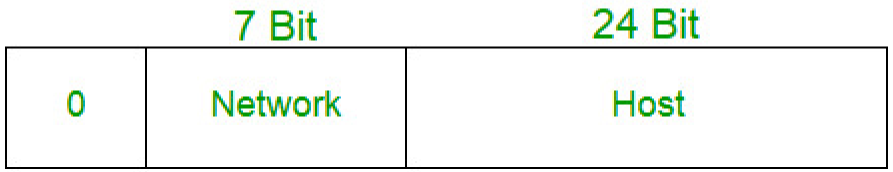
Classful IP Address



Class				Range of host addresses
A	0	Network	Host	1.0.0.0 to 127.255.255.255
B	10	Network	Host	128.0.0.0 to 191.255.255.255
C	110	Network	Host	192.0.0.0 to 223.255.255.255
D	1110	Multicast address		224.0.0.0 to 239.255.255.255
E	1111	Reserved for future use		240.0.0.0 to 255.255.255.255

Class A:

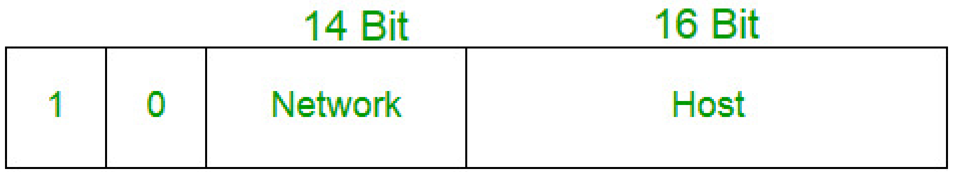
- IP address belonging to class A are assigned to the networks that contain a large number of hosts.
- The network ID is 8 bits long.
- The host ID is 24 bits long.
- The higher order bit of the first octet in class A is always set to 0. The remaining 7 bits in first octet are used to determine network ID.
- The 24 bits of host ID are used to determine the host in any network. The default subnet mask for class A is 255.x.x.x.
- Therefore, class A has a total of:
- 2^7-2= 126 network ID(Here 2 address is subtracted because 0.0.0.0 and 127.x.y.z are special address.)
- 2^24 – 2 = 16,777,214 host ID
- IP addresses belonging to class A ranges from 1.x.x.x – 126.x.x.x



Class A

Class B:

- IP address belonging to class B are assigned to the networks that ranges from medium-sized to large-sized networks.
- The network ID is 16 bits long.
- The host ID is 16 bits long.
- The higher order bits of the first octet of IP addresses of class B are always set to 10.
- The remaining 14 bits are used to determine network ID. The 16 bits of host ID is used to determine the host in any network.
- The default sub-net mask for class B is 255.255.x.x. Class B has a total of:
- 2^14 = 16384 network address
- 2^16 – 2 = 65534 host address
- IP addresses belonging to class B ranges from 128.0.x.x – 191.255.x.x.



14 Bit

16 Bit

1

0

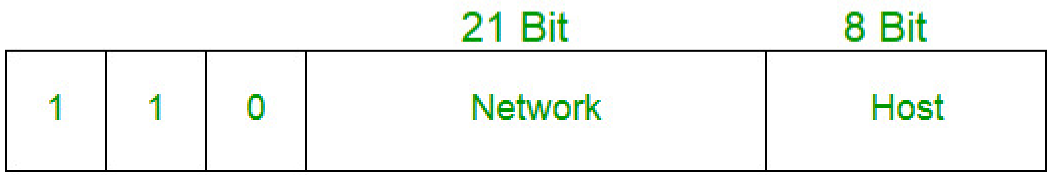
Network

Host

Class B

Class C

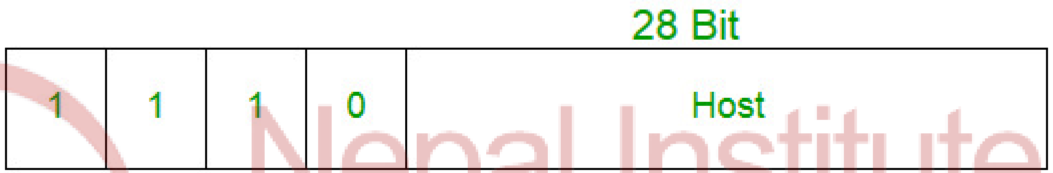
- IP address belonging to class C are assigned to small-sized networks.
  - The network ID is 24 bits long.
  - The host ID is 8 bits long.
- The higher order bits of the first octet of IP addresses of class C are always set to 110.
- The remaining 21 bits are used to determine network ID. The 8 bits of host ID is used to determine the host in any network.
- The default sub-net mask for class C is 255.255.255.x. Class C has a total of:
  - 2^21 = 2097152 network address
  - 2^8 – 2 = 254 host address
- IP addresses belonging to class C ranges from 192.0.0.x – 223.255.255.x.



Class C

Class D

- IP address belonging to class D are reserved for multi-casting. The higher order bits of the first octet of IP addresses belonging to class D are always set to 1110. The remaining bits are for the address that interested hosts recognize.
- Class D does not posses any sub-net mask. IP addresses belonging to class D ranges from 224.0.0.0 – 239.255.255.255.



				28 Bit
1	1	1	0	Host

Class D

Class E

- IP addresses belonging to class E are reserved for experimental and research purposes. IP addresses of class E ranges from 240.0.0.0 – 255.255.255.254. This class doesn't have any sub-net mask. The higher order bits of first octet of class E are always set to 1111.
- **Range of special IP addresses:**
- **127.0.0.0 – 127.0.0.8** : Loop-back addresses

Classless Addressing

- There is no classes hierarchy in the IP address but address is still granted in blocks.
- **Restriction**
- To simplify the handling of addresses, the Internet authorities impose three restrictions on classless address blocks:

1. The addresses in a block must be contiguous, one after another.
2. The number of addresses in a block must be a power of 2 (I, 2, 4, 8, ..)
3. The first address must be evenly divisible by the number of addresses.

Subnet mask

- A subnet mask is a 32 bits number that masks an IP address, and **divides the IP address** into **Network address and Host address.**
- Subnet mask is made by setting Network buts to all “1’s” and setting Host bits to all “0’s”.
- The mask can help us to find the network address and the host address. For example, the mask for a class A address has eight 1’s, which means the first 8 bits of any address in class A define the network address; the next 24 bits define the host address.
- It is used to identify network address of an ip address by preforming a **bitwise AND** operation on the subnet mask.
- The subnet mask is the classical way of representing which bits are part of the network portion of the address verses the host bits of the address. The subnet mask for a /24 network is 255.255.255.0.

Subnetting

- A subnet denotes a range of addresses that can be allocated to hosts, such as 192.168.1.0/24.
- It is a process of dividing large network into the smaller networks know s subnets based on IP address.
- Every computer on network has an IP address that represent its location on network. Two version of IP addresses are available IPv4 and IPv6.
- Example 1:- what is the subnetwork address if the destination address is 200.45.34.56 and the subnet mask is 255.255.240.0?
- Solution :- 11001000 00101101 00100010 00111000

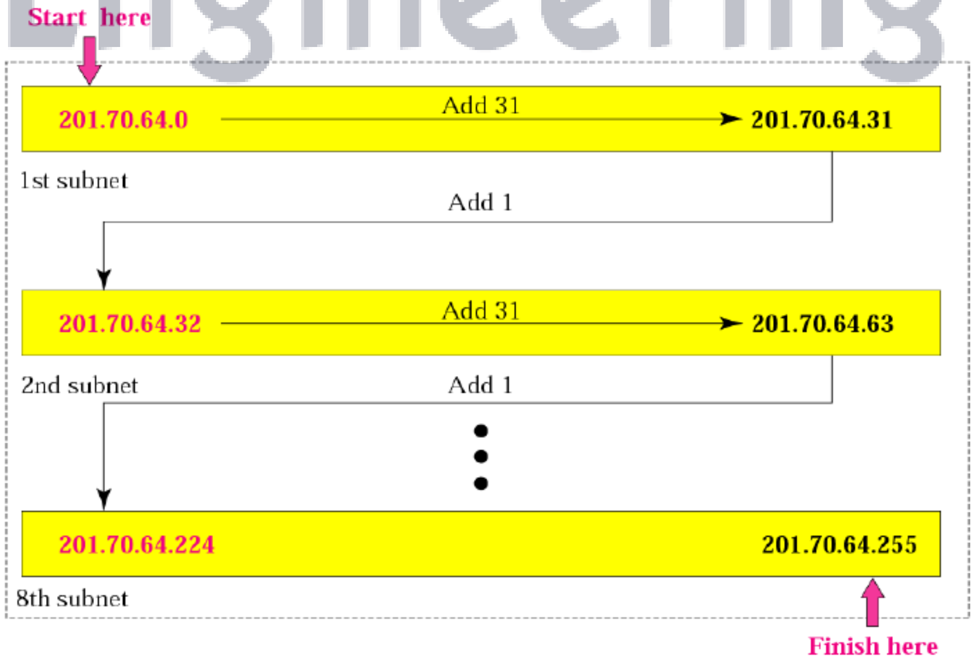
  11111111 11111111 11110000 00000000

  11001000 00101101 00100000 00000000

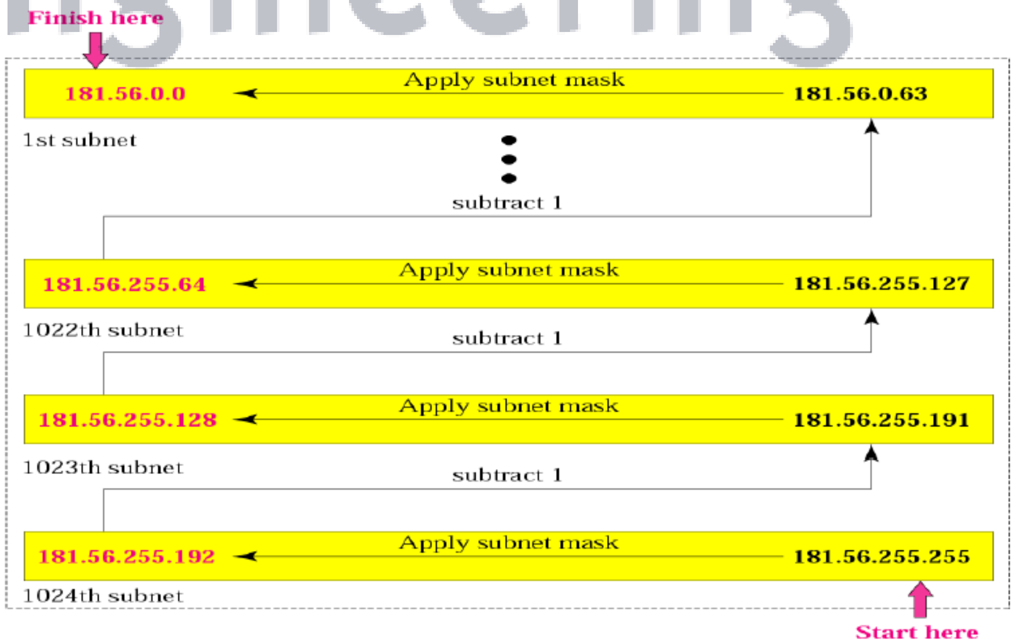
The subnetwork address is 200.45.32.0

Subnetting

- Example 2: A company is granted the site address 201.70.64.0 (class C). The company needs six subnets. Design the subnets.
- Solution :-
- The number of 1 s in the default mask is 24 (class C).
- The company needs six subnets. This number 6 is not a power of 2. The next number that is a power of 2 is 8 (2^3). We need 3 more 1s in the subnet mask. The total number of 1s in the subnet mask is 27 (24+3). The total number of 0 is 5 (32-27).
- The mask is 11111111.11111111.11111111.11100000 OR 255.255.255.224
- The number of subnet is 8. the number of address in each subnet is 2^5 (5 is the number of 0s) or 32



- **Example 3** : A company is granted the site address 181.56.0.0 (class B). The company needs 1000 subnets. Design the subnets.
- Solution :-
- The number of 1s in the default mask is 16 (Class B).
- The company needs 1000 subnets. This number is not a power of 2. the next number that is a power of 2 is 1024 (2^10). We need 10 more 1s in the subnet mask. The total number of 1s in the subnet mask is 26 (16+10).
- The total number of 0s is 6 (32-26).
- The mask is 11111111.11111111.11111111.11000000 OR (255.255.255.192).
- The number of subnets is 1024. the number of addresses in each subnet is 2^6 (6 is the number of 0s or 64.



Classless Interdomain Routing (CIDR)

- A classful addressing is a special case of classless addressing.
- A CIDR IP address looks like a normal IP address except that it ends with a slash followed by a number, called the IP network prefix.
- CIDR addresses reduced the size of routing tables and make more IP addresses available with organization.
- CIDR is based on the variable length subnet masking (VLSM).
- CIDR reduced the problem of wasted address space by providing a new and more flexible way to specify network addresses in routers.
- Example
  - /10 means Network address :10 bits and Host address :22 bits
  - /24 means Network address : 24 buts and Host address :8 bits
  - A CIDR address of 192.168.0.0/24 defines a block of addresses in the range 192.168.0.0 through 192.168.0.255.
  - While 192.168.0.0/20 would define a network 16 times as large from 192.168.0.0 through 192.168.0.255.

Network Address Translation (NAT)

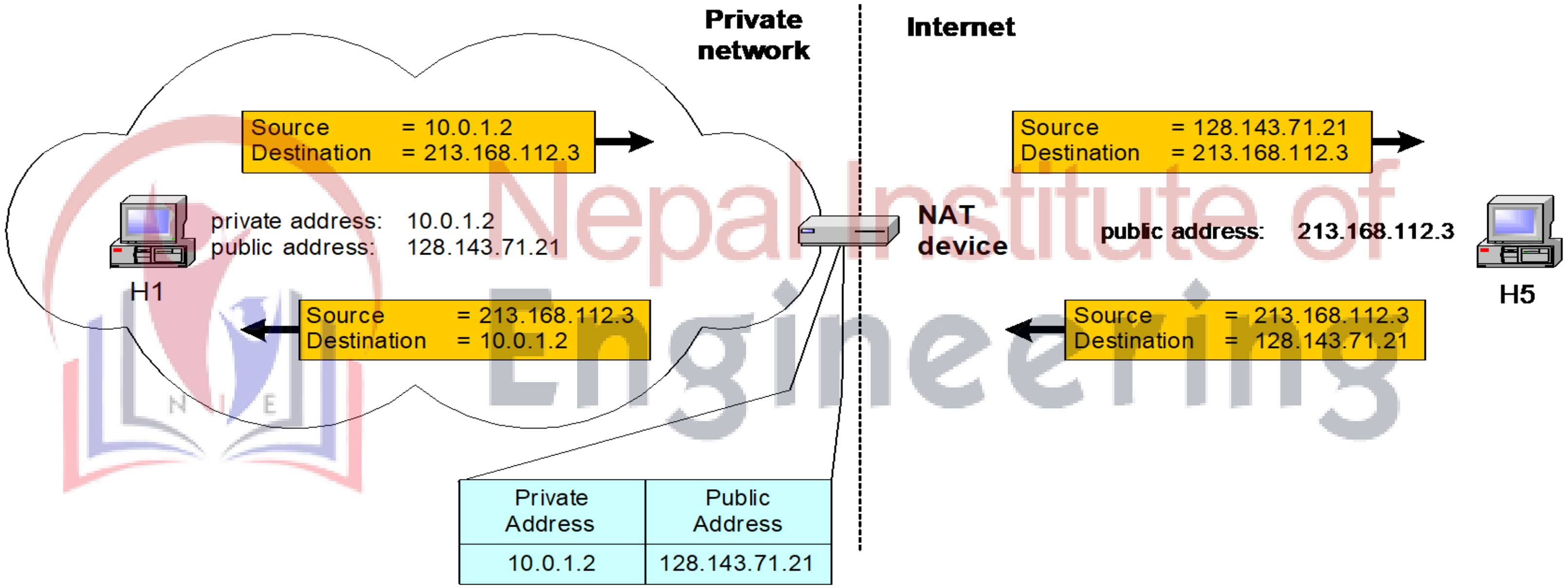
- NAT (Network Address Translation or Network Address Translator) is the translation of an Internet Protocol address (IP address) used within one network to a different IP address known within another network.
- Types of NAT
- Static NAT: A local IP address to one global IP address statically
- Dynamic NAT: A local IP address to any of a rotating pool of global IP addresses that a company may have

Network Address Translation (NAT)

- A short term solution to the problem of the depletion of IP addresses
  - Long term solution is IP v6
  - CIDR (Classless InterDomain Routing) is a possible short term solution
  - NAT is another

- NAT is a way to conserve IP addresses
  - Can be used to hide a number of hosts behind a single IP address
  - Uses private addresses:
    - 10.0.0.0-10.255.255.255,
    - 172.16.0.0-172.32.255.255 or
    - 192.168.0.0-192.168.255.255

Basic Operation of NAT


Private network
Internet
Source = 10.0.1.2
Destination = 213.168.112.3
private address: 10.0.1.2
public address: 128.143.71.21
H1
Source = 213.168.112.3
Destination = 10.0.1.2
NAT device
Source = 128.143.71.21
Destination = 213.168.112.3
public address: 213.168.112.3
H5
Source = 213.168.112.3
Destination = 128.143.71.21

Private Address	Public Address
10.0.1.2	128.143.71.21

- NAT device has address translation table
- One to one address translation

Introduction to IPv6

IPv6 (Internet Protocol version 6)

- IPv4, defines a 32-bit address - 2^{32} (4,294,967,296) IPv4 addresses available
- Major problem with IPv4: eventual depletion of the IP address space with exponential growth in world population.
- Techniques like NAT and CIDR can only buy some more time for IPv4.
- Introduced basically to eliminate the addressing problem

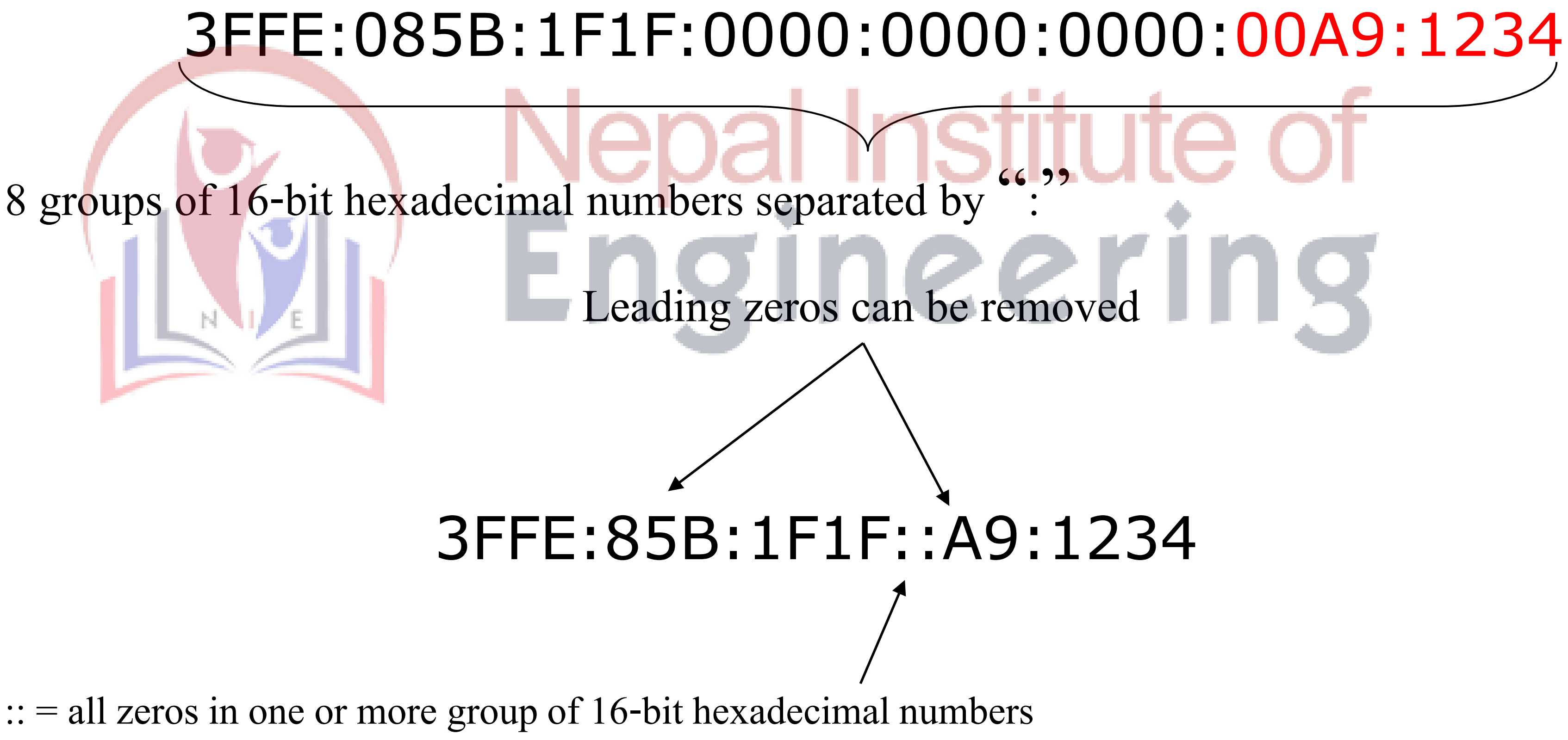
IPv6 (Internet Protocol version 6)

- IPv6 - 128 bits Network Layer Address
- Similar architectural principles as v4 – only bigger
- Also known as IPng(next generation)
- A new version of Internet Protocol
  - Primarily designed to extend address space
  - Enhancements and new features

IPv6- Advantages (Features)

- Very large address space
  - 128-bit address => 2^128 ~ 3 Trillion trillion trillion
  - "Every grain of sand on earth can get Unique IPv6 address"
- Reduce end-to-end delay
  - Processing delay reduces due to fixed header size and no header checksum
- Higher level of security
  - Mandatory Ipsec
- No fragmentation by routers
  - reduces fragmentation / reassembly overhead

128-bit IPv6 Address



3FFE:085B:1F1F:0000:0000:0000:00A9:1234

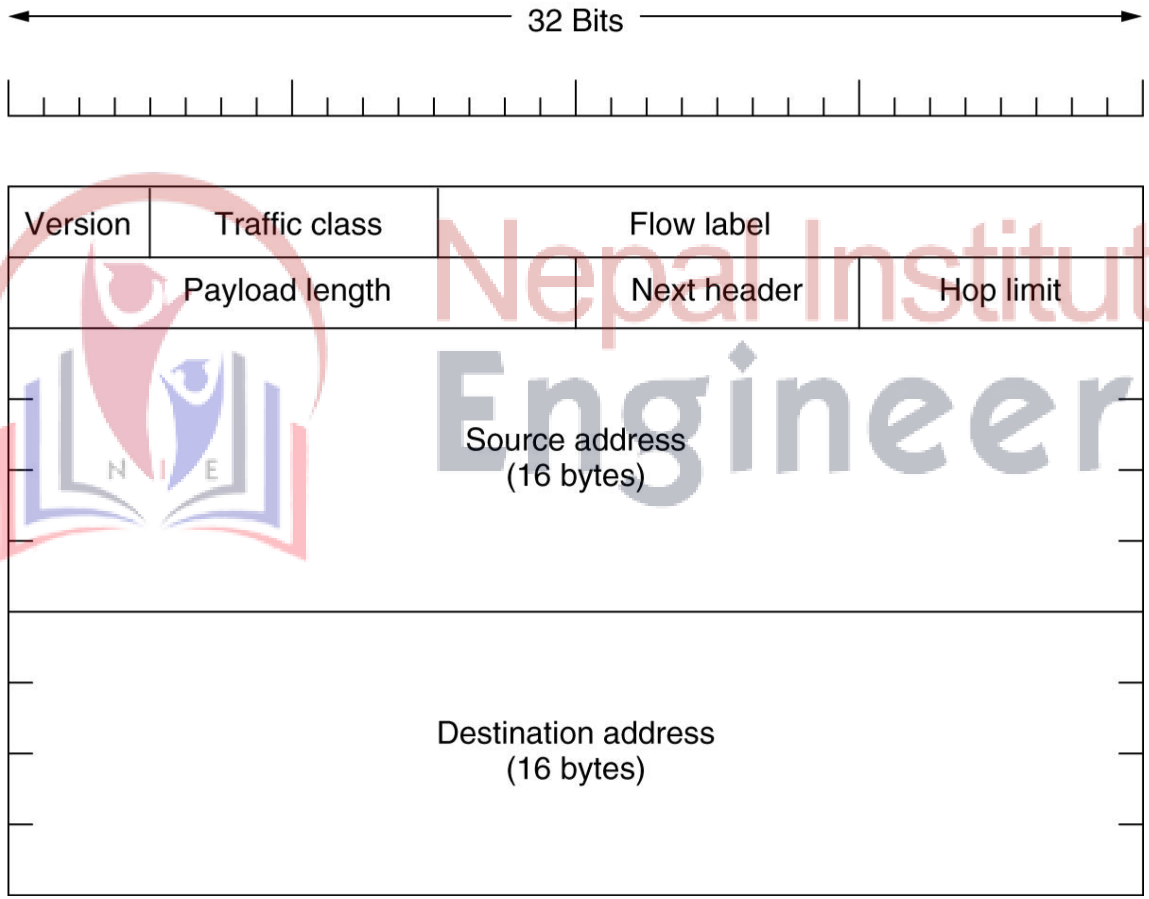
8 groups of 16-bit hexadecimal numbers separated by “:”

Leading zeros can be removed

3FFE:85B:1F1F::A9:1234

:: = all zeros in one or more group of 16-bit hexadecimal numbers

IPv6- Header Format



32 Bits

Version | Traffic class | Flow label

Payload length | Next header | Hop limit

Source address
(16 bytes)

Destination address
(16 bytes)

The IPv6 fixed header (required).

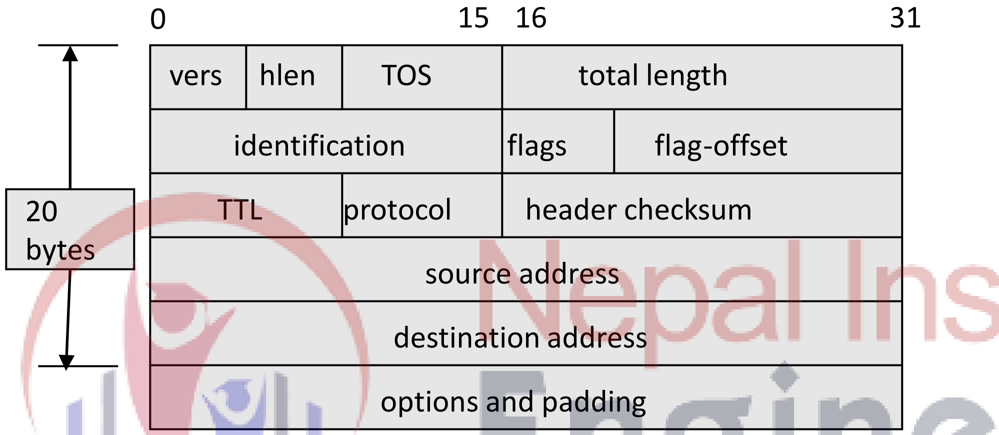
IPv6- Header Format

- **Version** (4 bits) - 4 bits are used to indicate the version of IP and is set to 6
- **Traffic Class** (8 bits) - Same function as the Type of Service field in the IPv4 , distinguish different real-time delivery requirement
- **Flow Label** (20 bits)
  - Identifies a flow and it is intended to enable the router to identify packets that should be treated in a similar way without the need for deep lookups within those packets.
  - Set by the source and should not be changed by routers along the path to destination.
  - Unique & powerful tool to IPv6

IPv6- Header Format

- **Payload Length** (16 bits) – Only the length of the payload (Header length is fixed to 40 bytes)
- **Next Header** (8 bits) - Indicates either the first extension header (if present) or the protocol in the upper layer PDU (such as TCP, UDP, or ICMPv6).
- **Hop Limit** (8 bits) - IPv4 TTL was appropriately renamed Hop Limit because it is a variable that is decremented at each hop, and it does not have a temporal dimension.
- **Source Address** (128 bits) - Stores the IPv6 address of the originating host.
- **Destination Address** (128 bits) - Stores the IPv6 address of the current destination host.

Header comparison



0 15 16 31

vers | hlen | TOS | total length

identification | flags | flag-offset

TTL | protocol | header checksum

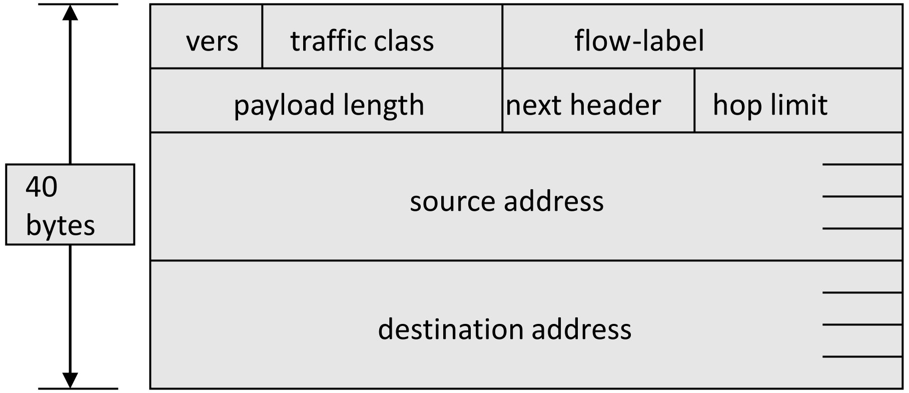
source address

destination address

options and padding

20 bytes

IPv4



vers | traffic class | flow-label

payload length | next header | hop limit

source address

destination address

40 bytes

IPv6

Removed (6)

- ID, flags, flag offset
- TOS, hlen
- header checksum

Changed (3)

- total length => payload
- protocol => next header
- TTL => hop limit

Added (2)

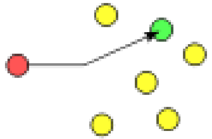
- traffic class
- flow label

Expanded

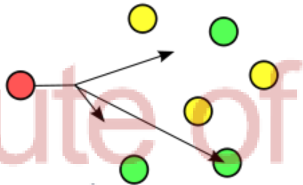
- address 32 to 128 bits

IPv6 Address Types

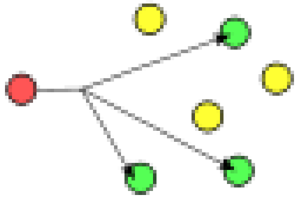
- **Unicast** : One to One Connection



- **Anycast :**
  - One to One of Many Connection
  - Packet sent to an anycast address is delivered to just one of member interfaces, typically the nearest host.



- **Multicast :**
  - One to Many Connection
  - Packet sent to a multicast address is delivered to all interfaces that have joined the corresponding multicast group



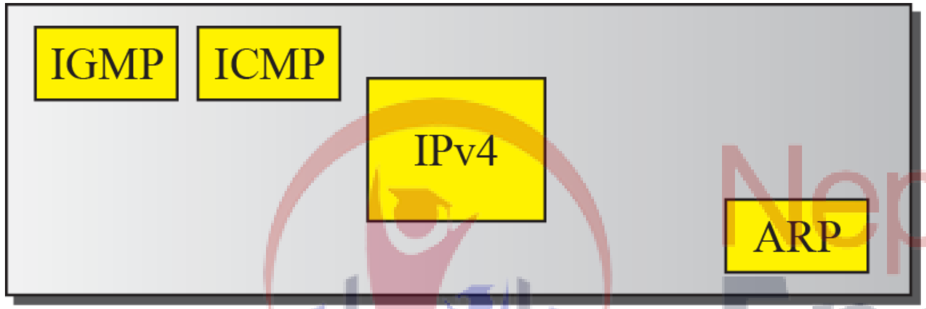
IPv6 Address Types

- No broadcast Address in IPv6
- Unicast , anycast or multicast address are identified by the prefix of the IPv6 address
- Slash notation used to determine prefix(as in IPv4)

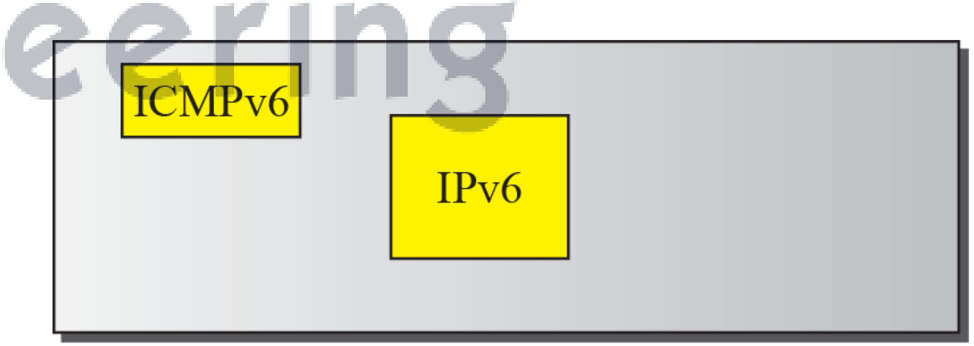
ICMPv6

- Another protocol that has been modified in version 6 of the TCP/IP protocol suite is ICMP.
- This new version, Internet Control Message Protocol version 6 (ICMPv6), follows the same strategy and purposes of version 4.
- ICMPv6, however, is more complicated than ICMPv4:
- some protocols that were independent in version 4 are now part of ICMPv6 and some new messages have been added to make it more useful.

Comparison of Network Layers in Version 4 & 6

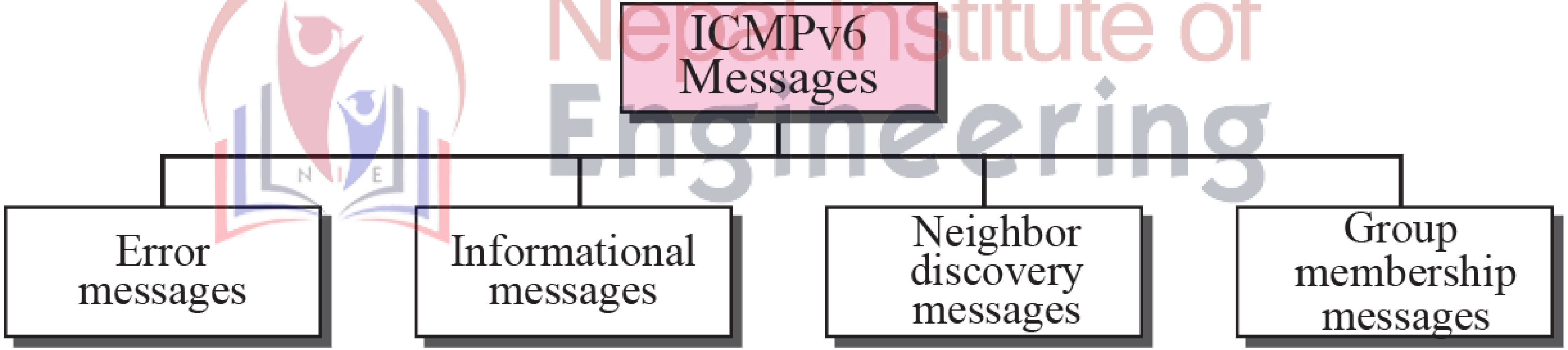


Network layer in version 4



Network layer in version 6

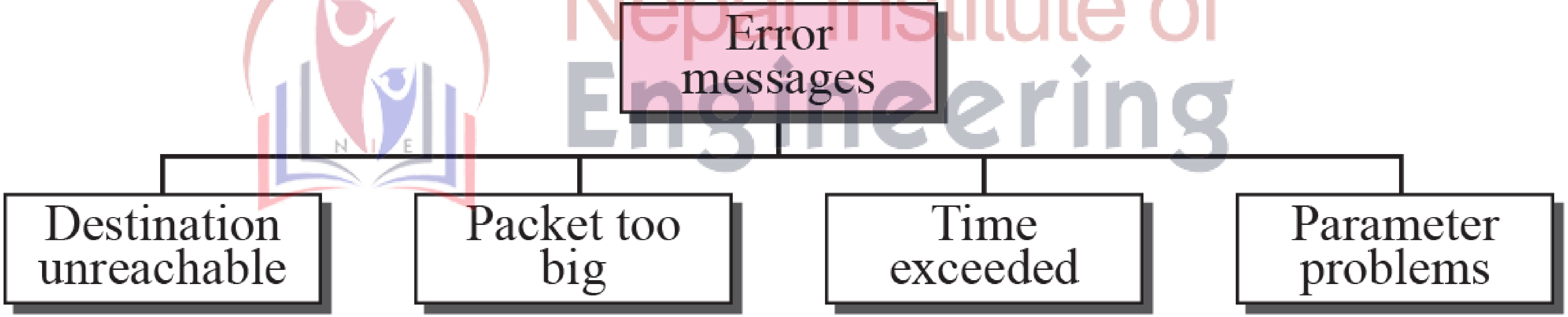
ICMPv6 Messages



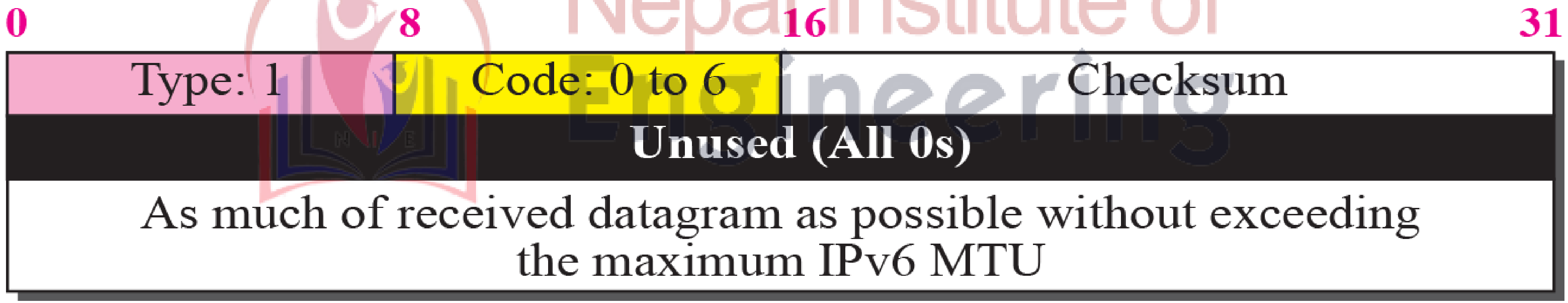
ND Protocol

MLD Protocol

Error Messages



Example: Destination Unreachable



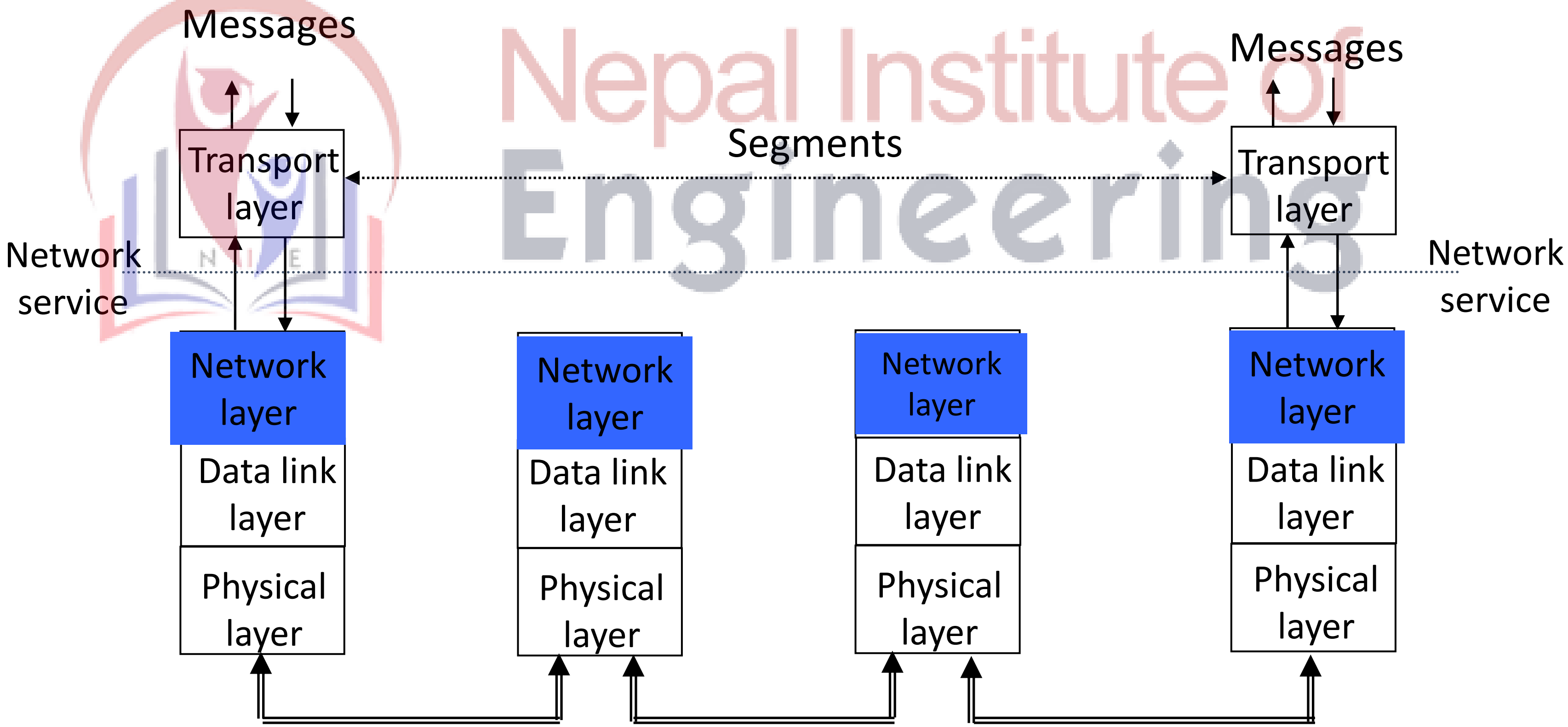
0
8
16
31
Type: 1
Code: 0 to 6
Checksum
Unused (All 0s)
As much of received datagram as possible without exceeding the maximum IPv6 MTU

Information Messages

- ✓Echo-Request Message
- ✓ Echo-Reply Message

Routing

- Transport Layer must be shielded from whatever process going on during routing of a packet from one host to other.



Routing Algorithms and Protocols

- **Routing Algorithms**
  - Shortest Path Algorithm
  - Flooding
  - Optimality Principle
  - Distance Vector Routing Algorithm
  - Link State Routing Algorithm
- **Routing Protocols**
  - RIP(Routing Information Protocol)
  - OSPF(Open Shortest Path First)
  - BGP(Border Gateway Protocol)

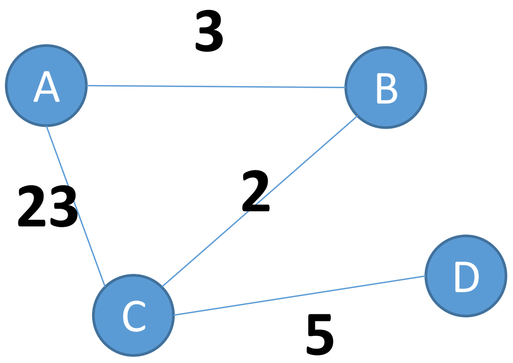
Routing Algorithm

- The routing algorithm is that part of the network layer software responsible for deciding which output line an incoming packet should be transmitted on
- Based on whether algorithm is static or dynamic :
  - Static Routing (Non-Adaptive) Algorithm
    - The choice of the route to use to get from one network to other *is computed in advance, off-line, and downloaded to the routers when the network is* booted
  - Dynamic Routing (Adaptive) Algorithm
    - Change their routing decisions to reflect changes in the topology, and usually the traffic as well. Adaptive algorithms differ in where they get their information (e.g., locally, from adjacent routers, or from all routers)

Distance Vector Algorithm

- Operate by having each router maintain a table (i.e a vector) giving the best known distance to each destination and which line to use to get there.
- These tables are updated by exchanging information with the neighbors.
- Also called Bellman-Ford routing algorithm.
- The router is assumed to know the "distance" to each of its neighbors. If the metric is hops, the distance is just one hop. Metric may be delay time.

Distance Vector



T=0

from A	via A	via B	via C	via D
to A				
to B		3		
to C			23	
to D				

from B	via A	via B	via C	via D
to A	3			
to B				
to C			2	
to D				

from C	via A	via B	via C	via D
to A	23			
to B		2		
to C				
to D				5

from D	via A	via B	via C	via D
to A				
to B				
to C			5	
to D				

T=1

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D			28	

from B	via A	via B	via C	via D
to A	3		25	
to B				
to C	26		2	
to D			7	

from C	via A	via B	via C	via D
to A	23	5		
to B	26	2		
to C				
to D				5

from D	via A	via B	via C	via D
to A			28	
to B			7	
to C			5	
to D				

T=2

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D		10	28	

from B	via A	via B	via C	via D
to A	3		25	
to B				
to C	26		2	
to D	31		7	

from C	via A	via B	via C	via D
to A	23	5		33
to B	26	2		12
to C				
to D	33	9		5

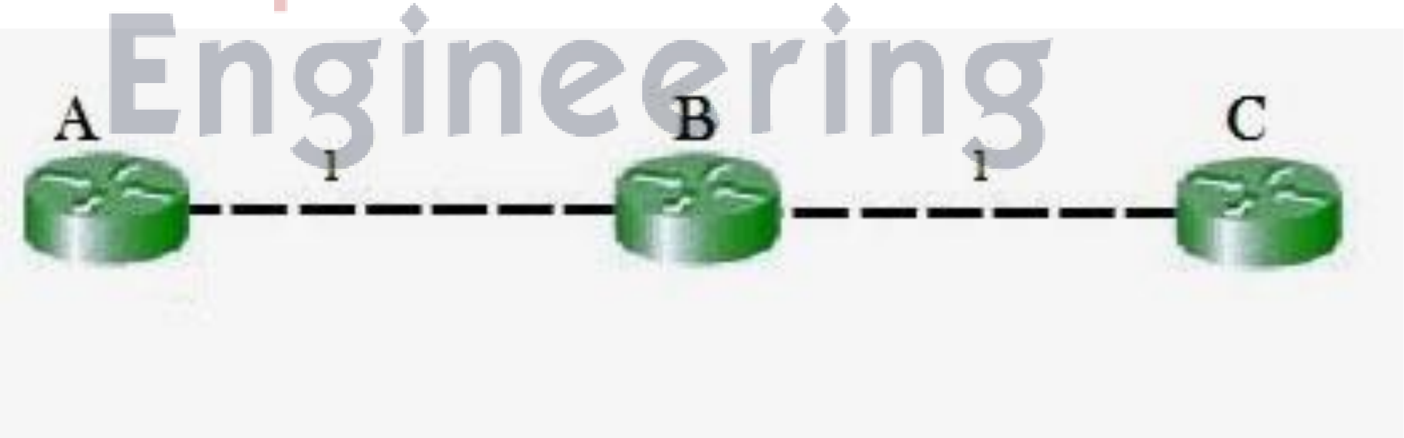
from D	via A	via B	via C	via D
to A			10	
to B			7	
to C			5	
to D				

Drawbacks of Distance vector

- Works properly theoretically but practically it has serious problem
- React quickly to good news does so leisurely to bad news.
- Slowness in converging to correct answer
- Suffer from **count to infinity problem**
- This problem can be solved by split horizon technique

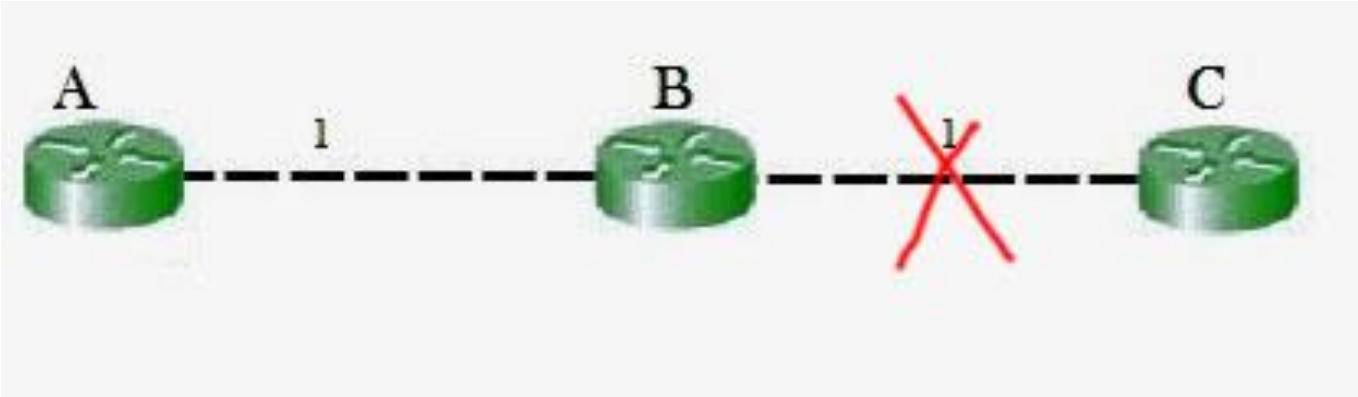
Count to infinity problem

So in this example, the Bellman-Ford algorithm will converge for each router, they will have entries for each other. B will know that it can get to C at a cost of 1, and A will know that it can get to C via B at a cost of 2.



Count to infinity problem

If the link between B and C is disconnected, then B will know that it can no longer get to C via that link and will remove it from it’s table. Before it can send any updates it’s possible that it will receive an update from A which will be advertising that it can get to C at a cost of 2. B can get to A at a cost of 1, so it will update a route to C via A at a cost of 3. A will then receive updates from B later and update its cost to 4. They will then go on feeding each other bad information toward infinity which is called as **Count to Infinity problem**.



Link-state Routing Algorithm(LS)

- Distance Vector do not make good candidate for longer network and
- Convergence and scalability problem for topology changes
- Link state routing is used to overcome those limitations
- LS has full knowledge of network
- Each node maintains the full graph by collecting the updates from all other nodes
- Each node then independently calculates the next best logical *path* from it to every possible destination in the network
- Router's receive topology information from their neighbor router via link state advertisements(LSA)

Link-state Routing Algorithm(LS)

- Use Dijkstra's shortest path first algorithm to determine optimal paths
- Link state protocols don't have to constantly resend their entire LSAs instead they can send small hello LSAs to let their neighbor routers know they are still alive

Each router must do the following:

1.Discover its neighbors, learn their network address.

2.Measure the delay or cost to each of its neighbors.

3.Construct a packet telling all it has just learned.

4.Send this packet to all other routers.

5.Compute the shortest path to every other router.

Dynamic Routing Protocol

Distance Vector	Link State ()
• Entire routing table is sent as an update • Distance vector protocol send periodic update at every 30 or 90 second • Updates are sent to directly connected neighbor only • Routers don't have end to end visibility of entire network. • Suffer from count to infinity problem • Examples: RIP, BGP	• Updates are incremental & entire routing table is not sent as update • Updates are triggered not periodic • Update are sent to entire network & to just directly connected neighbor • Routers have visibility of entire network of that area only. • No routing loops • Convergence is fast because of triggered updates. Examples: OSPF, IS-IS

Hierarchical Routing

- **Problem with above Routing**
    - **Scale :** As no. of routers become large, the overhead involved in computing, storing, and communicating routing information becomes prohibitive.
    - **Administrative autonomy :** An organization should be able to run and administer its network as it wishes, while still being able to connect its network to other networks and hide network's internal organization.
- Both of these problems can be solved by organizing router into autonomous system(AS)

Autonomous System

- Consist of group of routers that are typically under same administrative control, belonging to same company.
- AS Number (ASN) is assigned to each AS to identify them uniquely for routing purpose
- ASNs are 16 bit values, 64512 through 65535 are “private”
- Internet can be viewed as the collection of Interconnected Autonomous Systems (ASes)

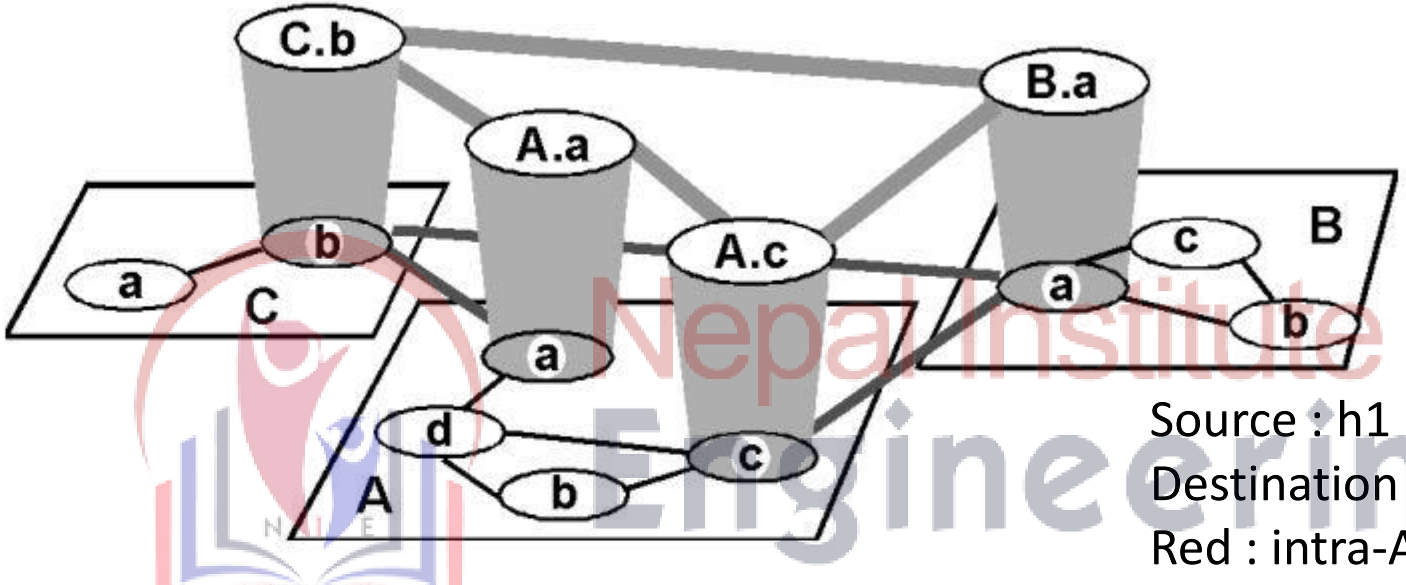
Routing in Internet

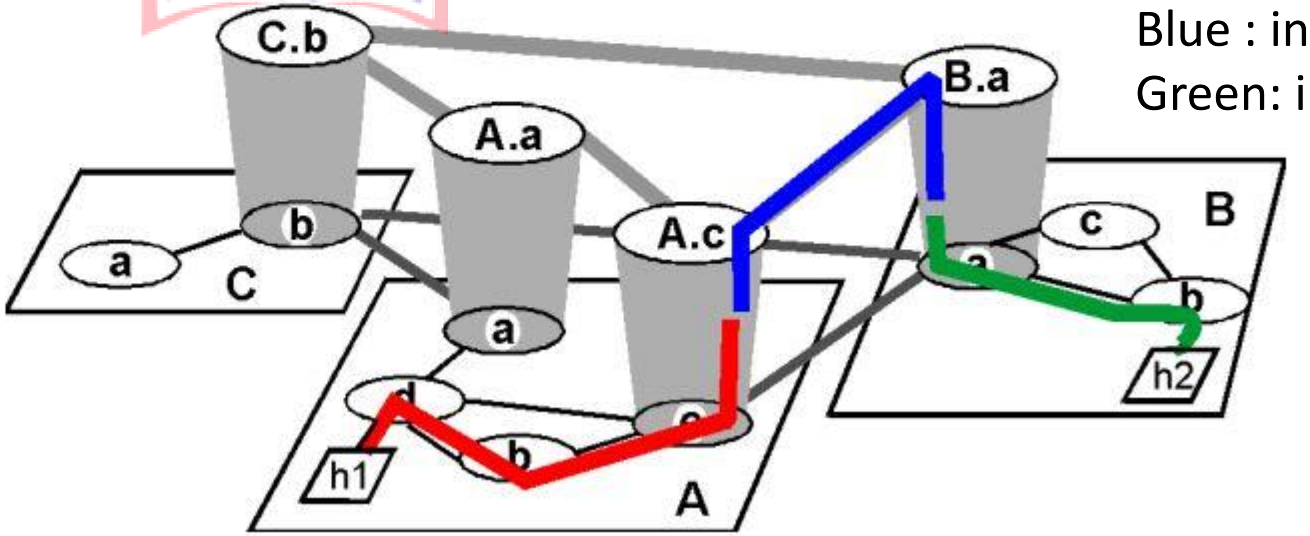
- Routers within same AS run same routing algorithm called **intra-autonomous system routing protocol** or **Interior Gateway Protocol – IGP** (Intra-AS / Intra-Domain)
  - RIP(Routing Information Protocol),
  - OSPF(Open Shortest Path First)
  - IS-IS(Intermediate System-Intermediate System) ,
  - IGRP(Interior Gateway Routing Protocol),
  - EIGRP(Enhanced IGRP)

Routing in Internet

- Routers responsible for routing packets to destination outside the AS are called **gateway router**
- The routing algorithm that gateways use to route among the various AS is known as **inter-autonomous system routing protocol** or **Exterior Gateway Protocol – EGP** (Inter-AS / Inter-Domain)
- Eg. **BGP**(Border Gateway Protocol)

Autonomous System





Source : h1 in AS A

Destination : h2 in AS B

Red : intra-AS routing

Blue : inter-AS routing

Green: intra – AS routing

Intra-AS Routing in the Internet : RIP(Routing Information Protocol)

- One of earliest intra-AS internet routing protocol
- Is distance vector protocol
- Cost Metric = hop count (no. of subnets traversed)
- Max. hop count =15 i.e. RIP is limited to run on small networks
- Routing updates are exchanged between neighbors approximately every 30 seconds using RIP advertisement.

Intra-AS Routing in the Internet : RIP(Routing Information Protocol)

- Upgraded to RIPv2 that uses concept of VLSM and sub-netting
  - RIPv2 carries Subnet Mask and also provides Authentication

Intra-AS Routing in the Internet OSPF(Open Shortest Path First)

- Successor of RIP
- OSPF and its closely related cousin, IS-IS, are typically deployed in upper-tier ISPs whereas RIP is deployed in lower tire ISPs.
- Open indicates routing protocol specification is publicly available
- Link state routing has full knowledge of network.(using link state advertisement)
- Uses flooding of link-state information and Dijkstra least-cost path algorithm

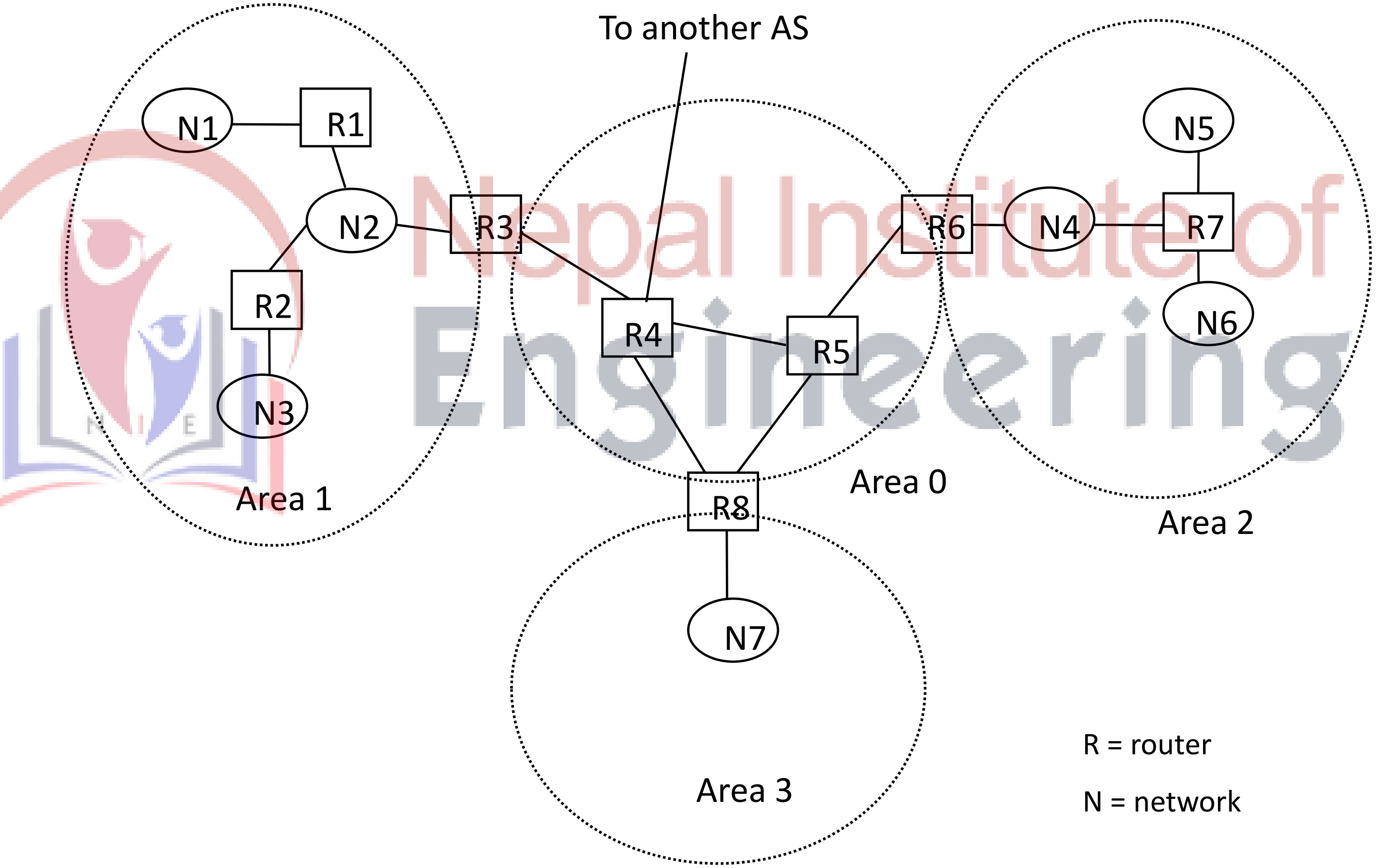
Intra-AS Routing in the Internet OSPF(Open Shortest Path First)

- With OSPF, router broadcast link state information whenever there is a change in link's state
- It also broadcast link's state periodically at least once every 30 minutes
- With OSPF router constructs complete topological map of entire AS
- Then router run DiJkstra's shortest path algorithm
- Link costs are configured by network administrator
- Cost Metric = Link Speed

Intra-AS Routing in the Internet OSPF(Open Shortest Path First)

- No hop limit
- Event Triggered
- Also provide authentication mechanism

OSPF Areas



R = router

N = network

OSPF

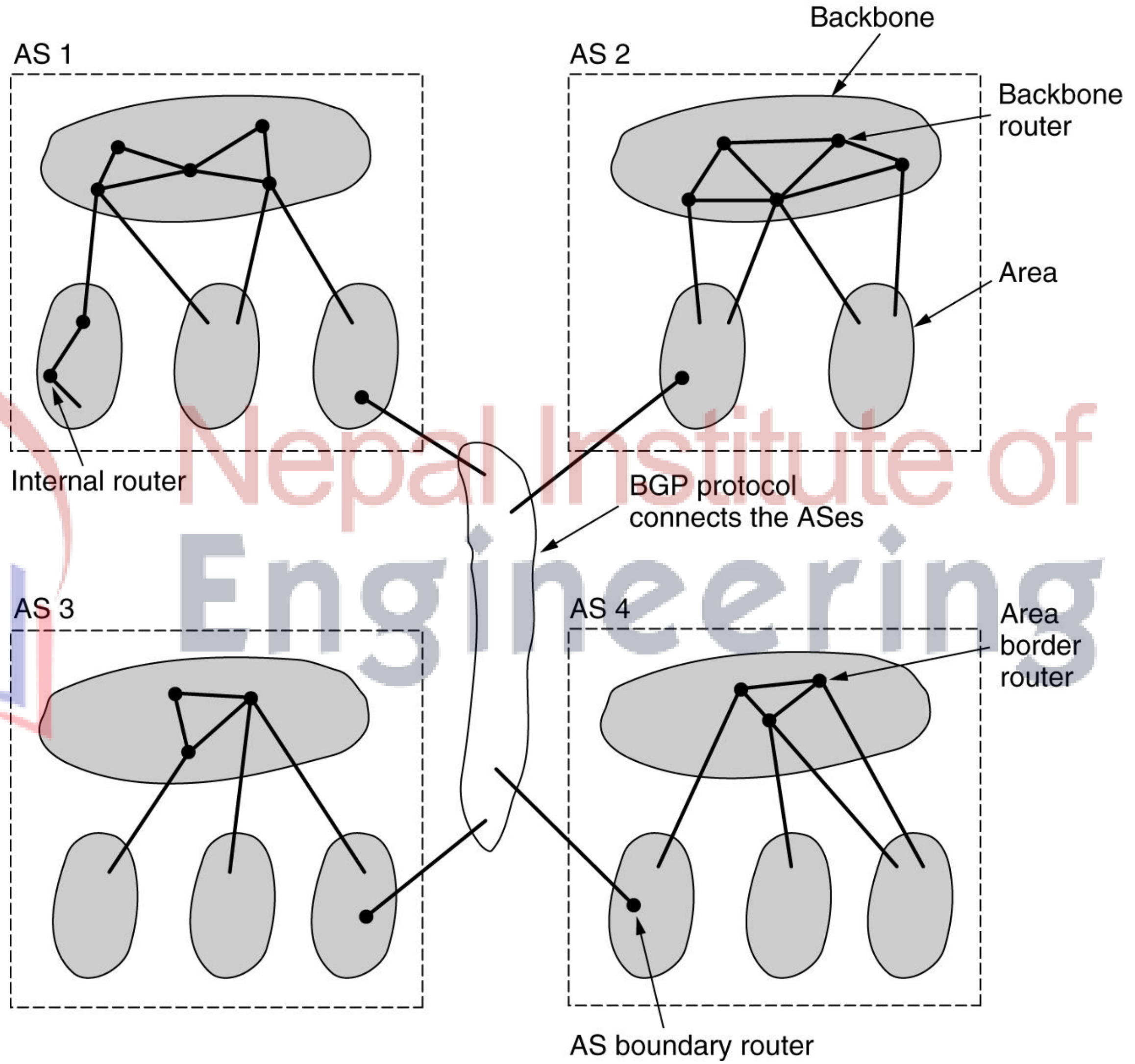


Figure : The relation between ASes, backbones, and areas in OSPF.

Inter-AS Routing in the Internet BGP(Border Gateway Protocol)

- Current version used is BGP-4.
- The Routing Domain of BGP is the entire Internet
- Uses distance vector algorithm
- BGP assumes the Internet is an arbitrary interconnected set of ASes.
- In *interdomain* routing, the goal is to find ANY path to the intended destination that is loop-free.
- The protocols are more concerned with **reachability** than optimality.

Inter-AS Routing in the Internet
BGP (Border Gateway Protocol)

- Types of BGP messages
  - **Open** : Establish a peering session.
  - **Keep Alive** : Handshake at regular intervals.
  - **Notification** : Shuts down a peering session.
  - **Update** : Announcing new routes or withdrawing previously announced routes.

communication mechanisms

- There are three communication mechanisms defined and classified as they related to the Internet
  - **Unicasting**
  - **Multicasting**
  - **Broadcasting**

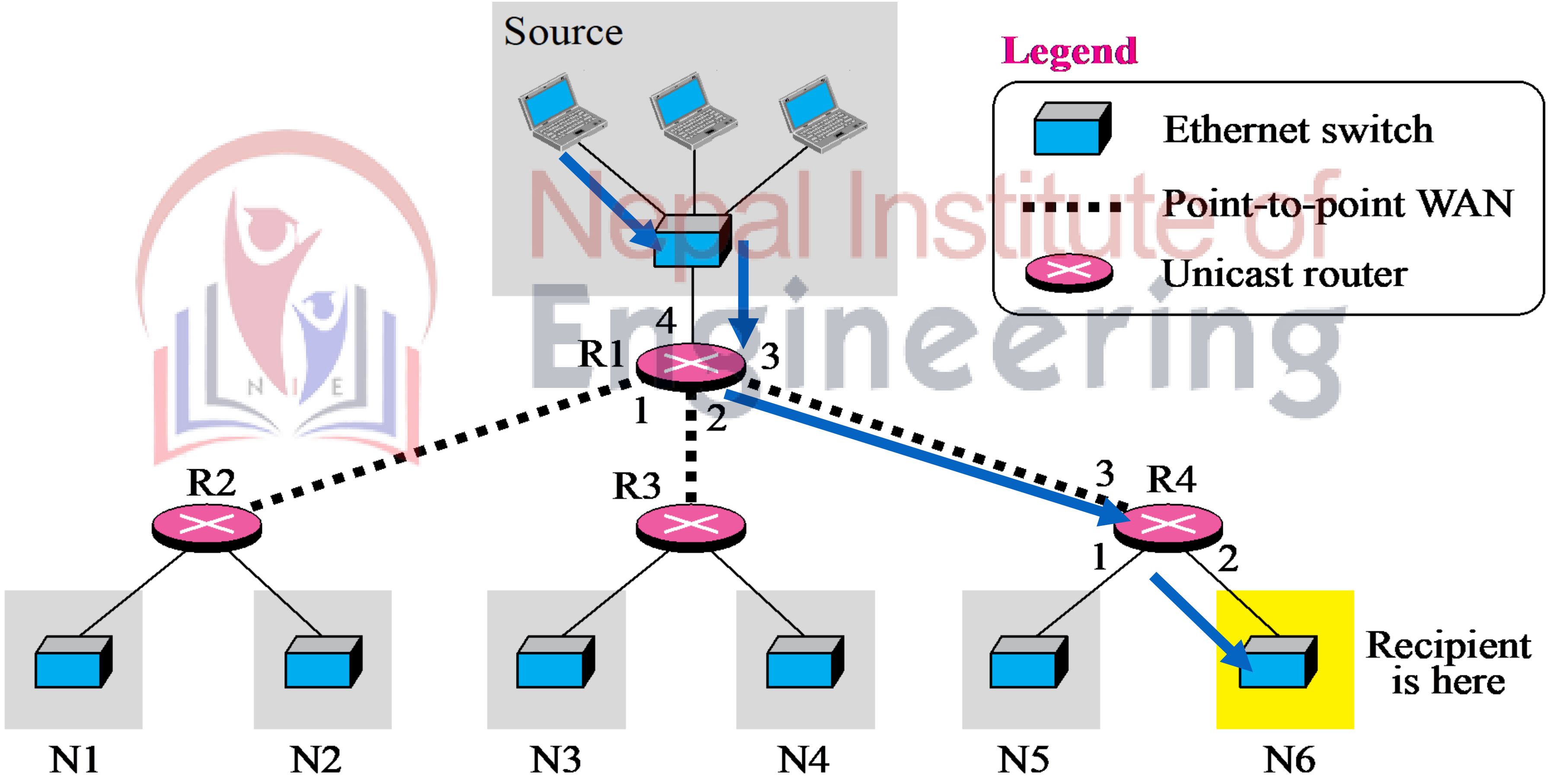
Broadcasting

- In broadcasting communication, the relationship between the source and the destination is one to all
- The Internet does not explicitly support broadcasting because of
  - Traffic
  - Bandwidth

Unicast

- In unicast routing the router forwards the received packet through only one of its interfaces
- In unicast routing, each router in the domain has a table that defines a shortest path tree to possible destinations

Unicast

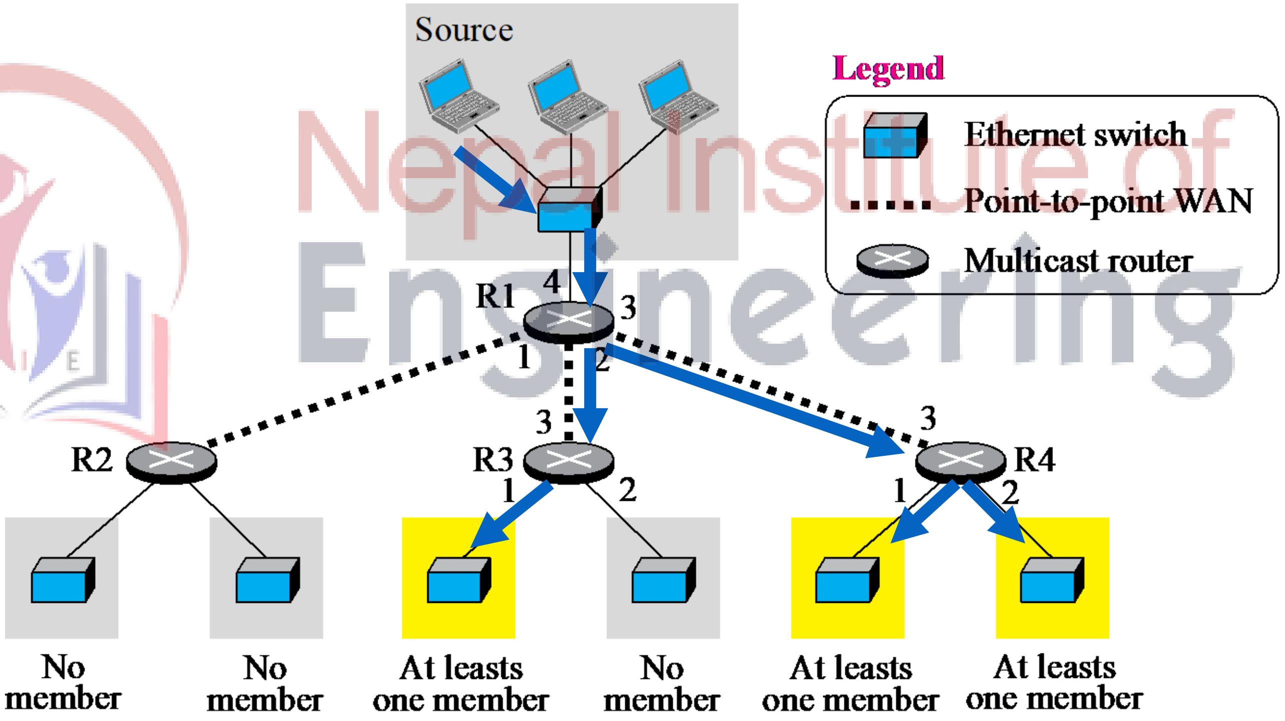


Source
Legend
Ethernet switch
Point-to-point WAN
Unicast router
R1
4
3
1
2
R2
R3
R4
3
1
2
Recipient is here
N1
N2
N3
N4
N5
N6

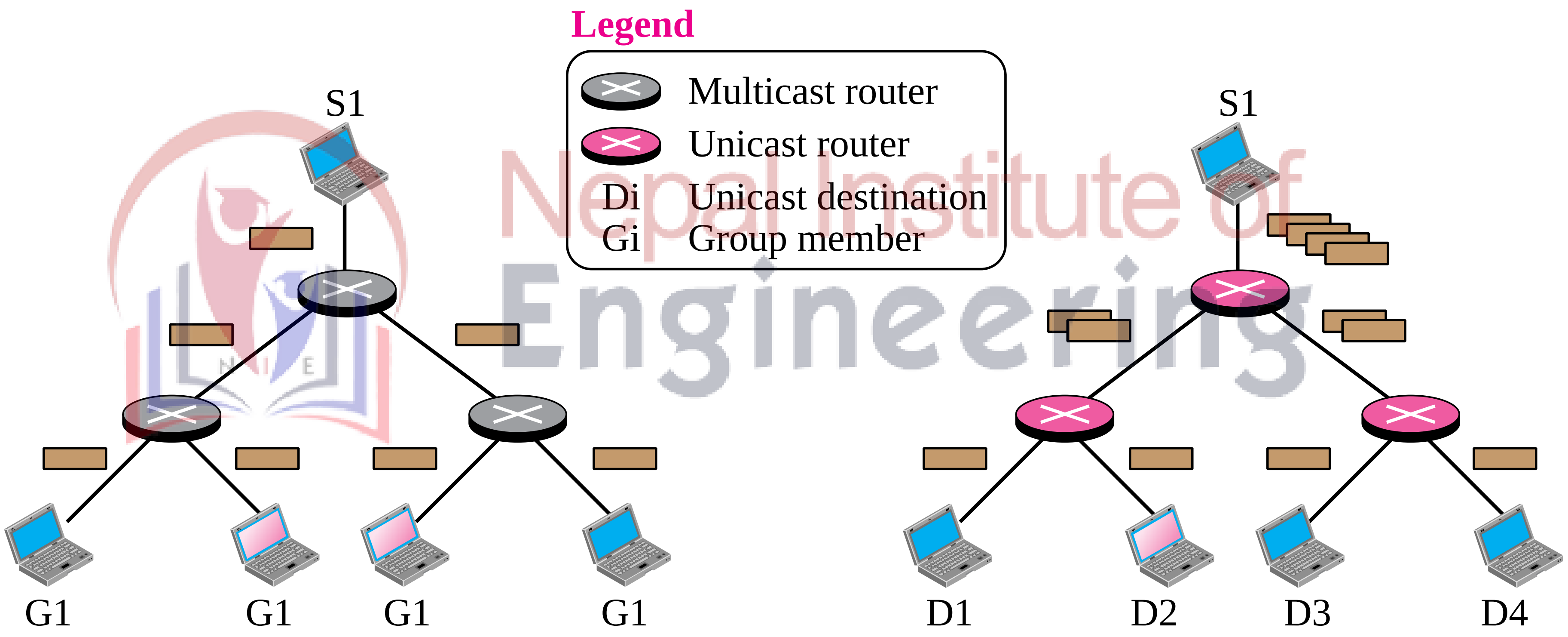
Multicast

- In multicasting, there is one source and a group of destinations
  - The relationship is one to many
  - The router may forward the received packet through several of its interfaces

Multicast


Source
Legend
Ethernet switch
Point-to-point WAN
Multicast router
R1
4
3
1
2
R2
R3
3
1
2
R4
3
1
2
No member
No member
At leasts one member
No member
At leasts one member
At leasts one member

Multicasting versus multiple unicasting



a. Multicasting

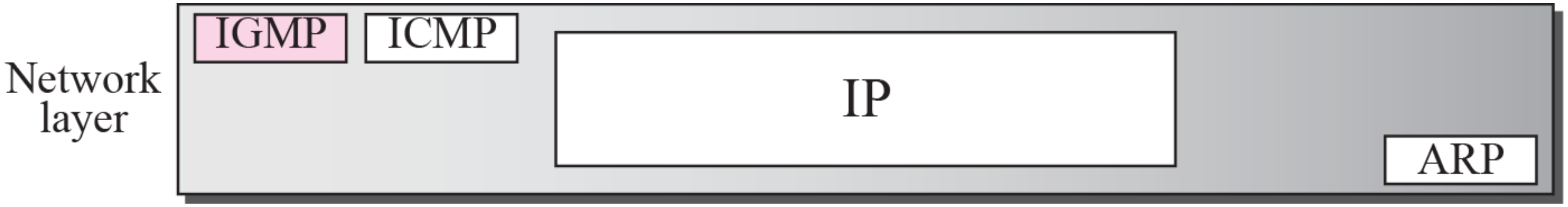
b. Multiple unicasting

Multicast Routing

- In computer networking, multicast is the delivery of a message or information to a group of destination computers simultaneously in a single transmission from the source.
- Copies are automatically created in other network elements, such as routers, but only when the topology of the network requires it.

Internet Group Management Protocol (IGMP)

- In multicast, each multicast router needs to know the list of groups to each interface
  - It means that they need to collect information about members and share it with others
- The Internet Group Management Protocol (IGMP) is responsible for correcting and interpreting information about group members in a network
  - It is one of the protocols designed at the IP layer for this purpose



Applications of Multicast

- Video/audio conference
- IPTV, Video on Demand
- Advertisement, Stock, Distance learning
- Distributed interactive gaming or simulations
- Voice-over-IP
- Synchronizing of distributed database, websites