IP Datagram Forwarding

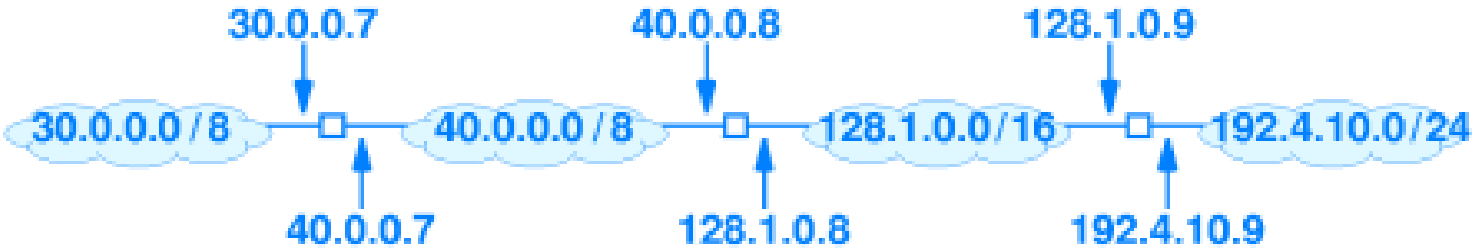
IP Datagram Forwarding

- Forwarding –delivering to the next hop
  - – place the packet in its route to its destination
- Based on destination address
- Based on labels
- When IP is used as a connectionless protocol, forwarding is based on the destination address of the IP datagram;
- when the IP is used as a connection-oriented protocol, forwarding is based on the label attached to an IP datagram.

Types of Forwarding

- *Direct forwarding* takes place where the destination IPv4 address is on a network attached to the router or host. Direct Forwarding: they can choose among many interfaces
- *Indirect forwarding* takes place where the destination IPv4 address is not on a network connected to the router or host, and the datagram therefore has to be forwarded to a router further along the path through the internetwork.
- Indirect Forwarding: is based on routing tables

IP Datagram Forwarding



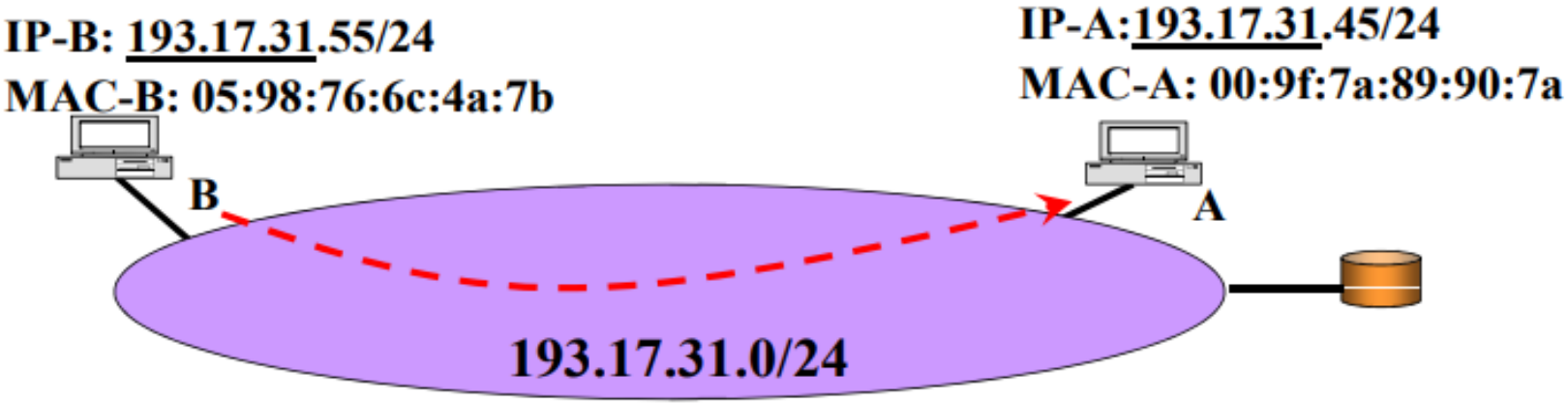
Routing Table at the Middle Router

Destination	Mask	Next Hop
30.0.0.0	255.0.0.0	40.0.0.7
40.0.0.0	255.0.0.0	deliver direct
128.1.0.0	255.255.0.0	deliver direct
192.4.10.0	255.255.255.0	128.1.0.9

192.4.10.3

Host Direct Forwarding

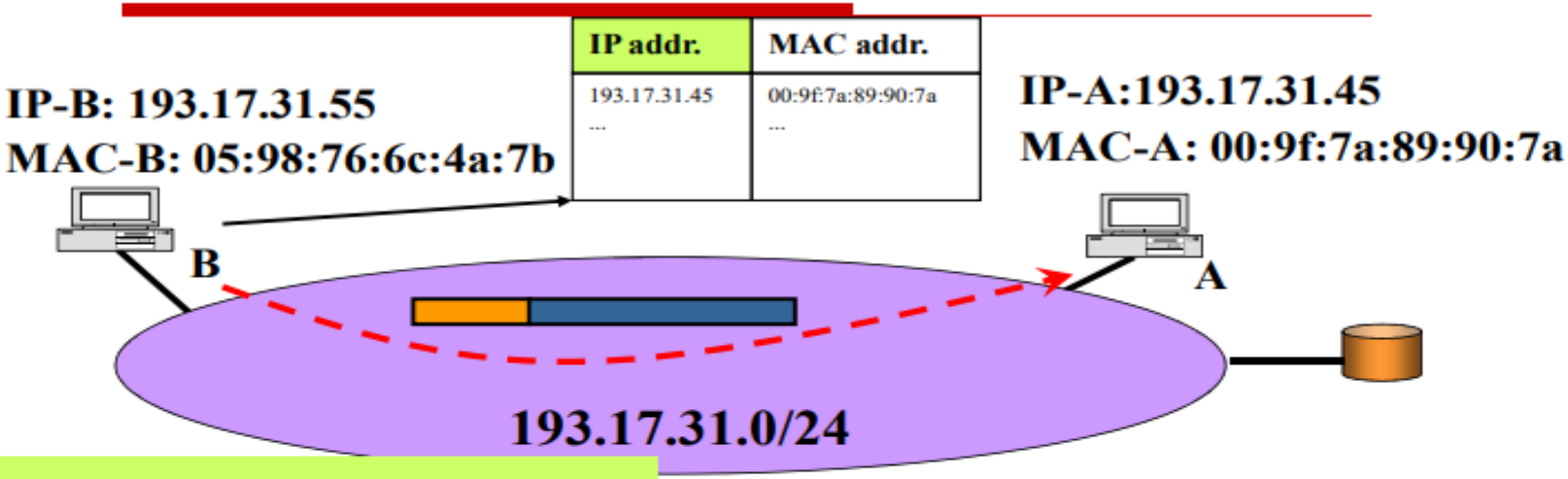
- LAN coincides with IP subnet



1. Host B needs to send a IP packet to host A

2. B knows its own IP address (IP-B) and knows that A is on the same subnet (by comparing the NetIDs)

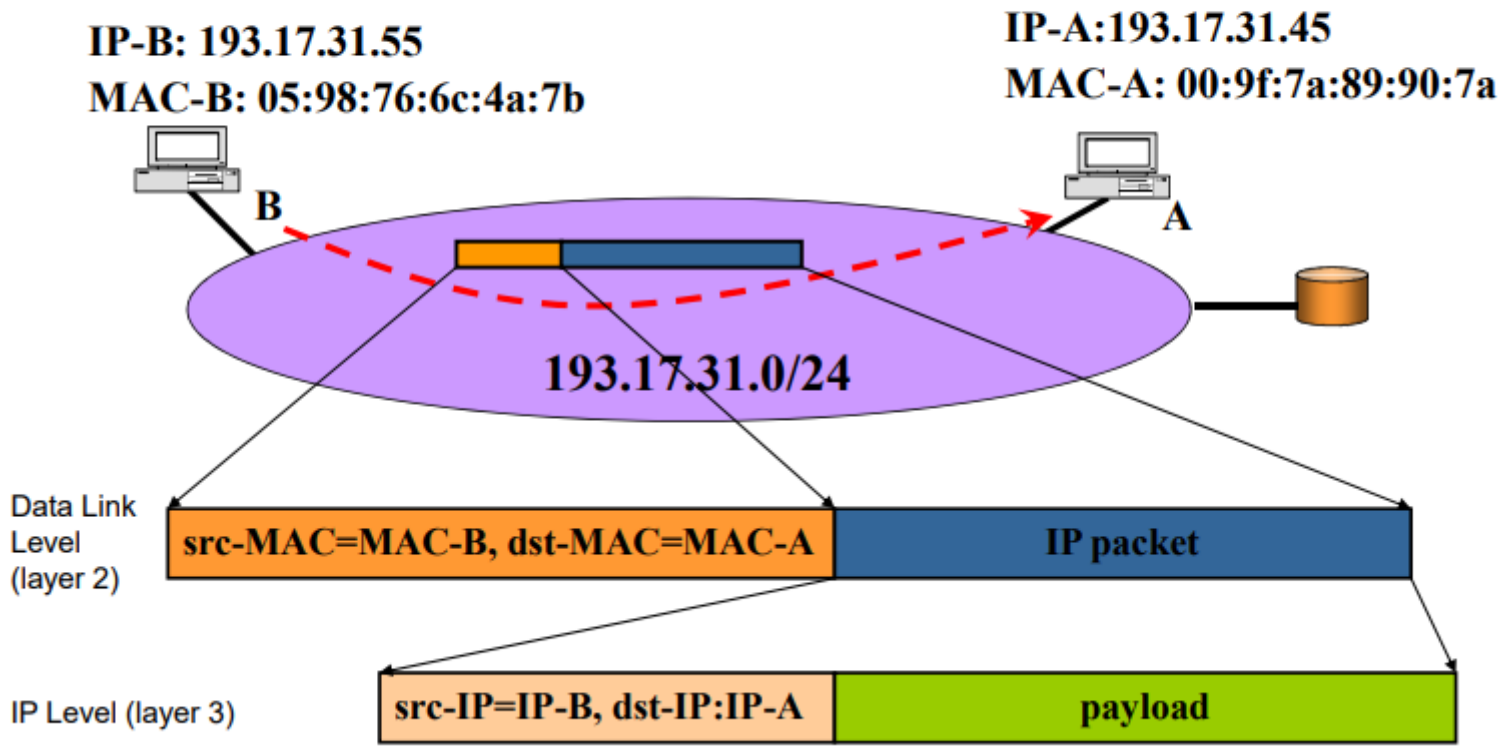
Host Direct Forwarding



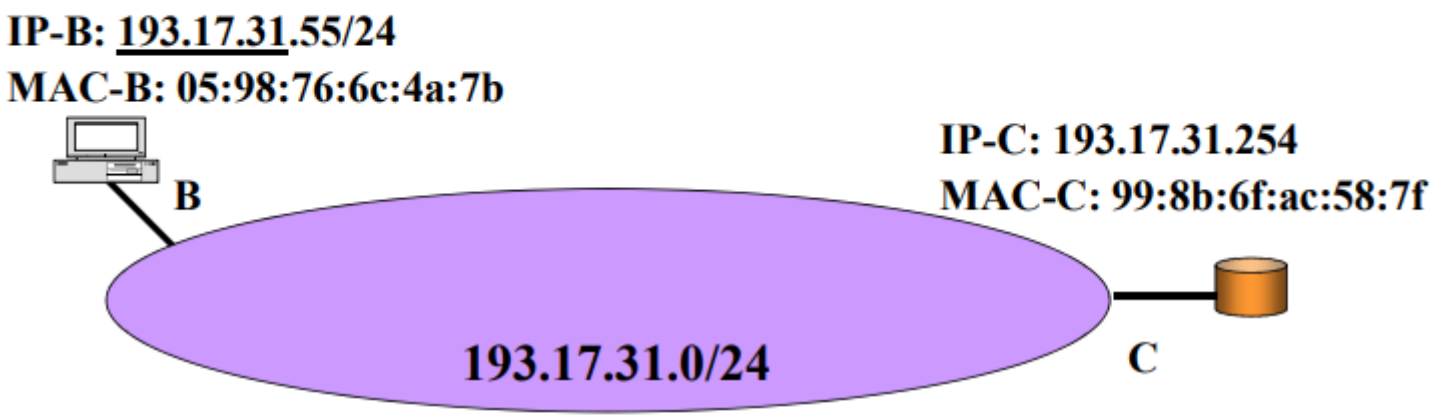
3. B searches a table for the physical address corresponding to the IP destination address IP-A (ARP Table)

4. The IP layer of B passes down the packet to the lower layer (Ethernet …) which is responsible of the forwarding (destination MAC-A)

Host Direct Forwarding



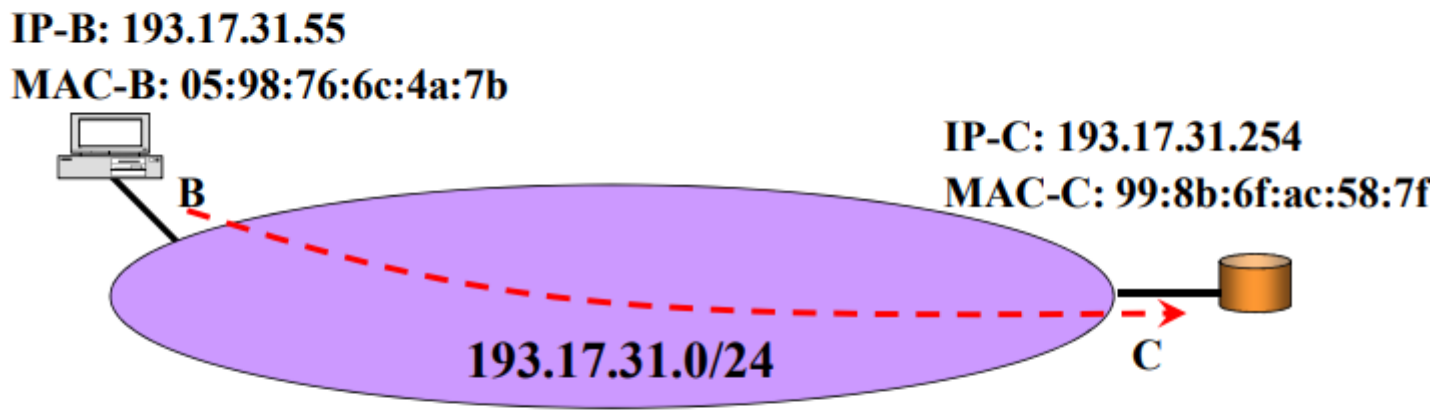
Host Indirect Forwarding



1. Host B needs to send a IP packet to destination *IP-D=131.17.23.4*

2. B knows its own IP address (IP-B) and knows that D is NOT on the same subnet (by comparing the NetIDs)

Host Indirect Forwarding

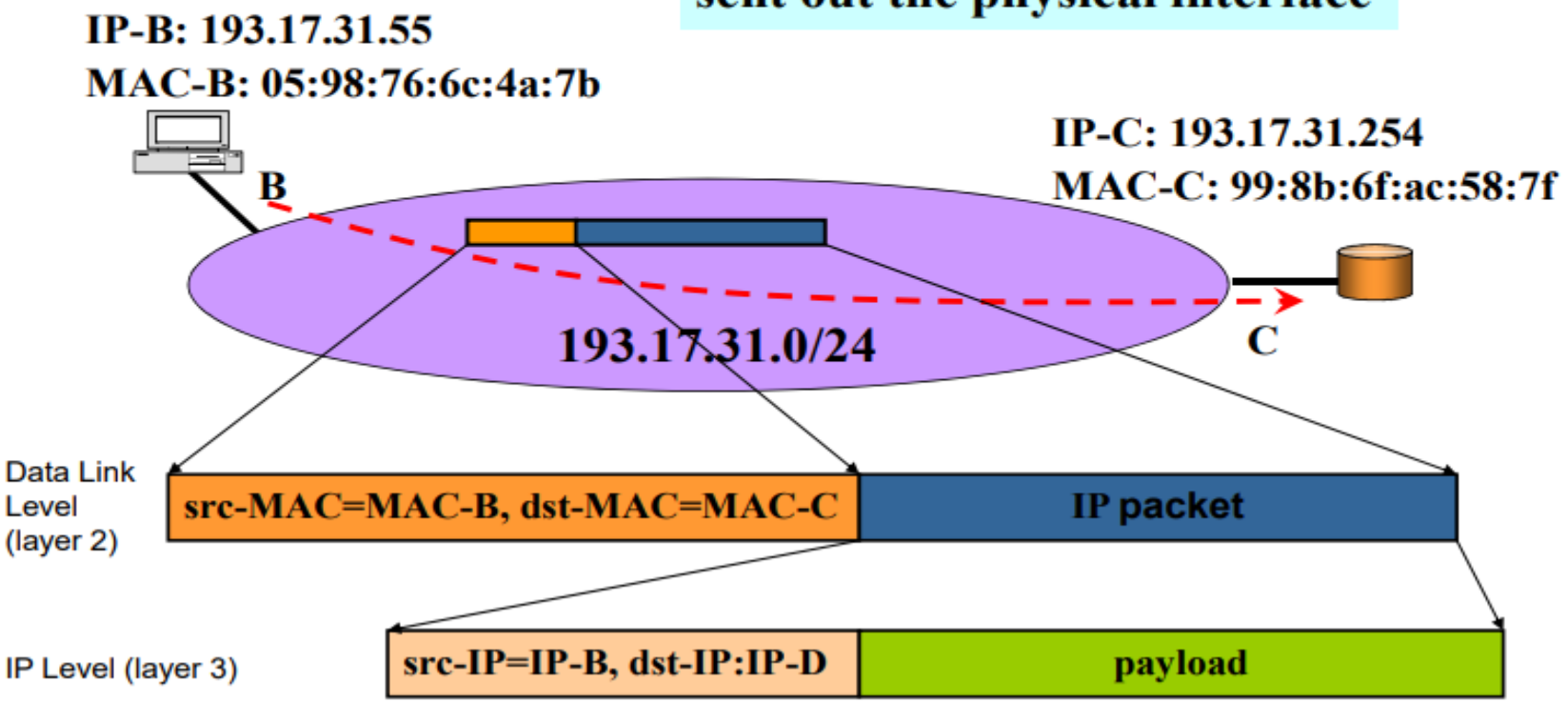


3. B needs to rely upon a router (usually just one default router is configured for a host)

4. B gathers the MAC address of the default router and passes downwards the packet

Host Indirect Forwarding

5. The frame is built up and sent out the physical interface



(i) Based on destination address

- DESTINATION BASED: driven by the destination address
- NEXT HOP ROUTING: for each destination in the routing table, only the next relaying node is reported.

IP Datagram Forwarding

- **The mask field** is used to extract the network prefix of the IP address during lookup.
- Given a destination IP address D, for each entry in the routing table, the routing software computes the Boolean AND of the mask in an entry and D to extract the network prefix of D. If the extracted network prefix of D matches with that of the Destination network prefix in the entry, then an address match has occurred; the packet can be forwarded through the next-hop listed in the entry.

- Example: Consider the network topology in the previous slide. Consider a datagram destined for address **192.4.10.3** arrives at the center router.
- The routing software computes a Boolean AND of 192.4.10.3 with each of 255.0.0.0, 255.255.0.0, 255.255.255.0 to find the matching destination network prefix.
- The routing software succeeds in the fourth entry as 192.4.10.3 & 255.255.255.0 = 192.4.10.0, which is the “Destination” network prefix in that entry. Hence, the datagram is forwarded to the next-hop router whose IP address is 128.1.0.9.

IP Datagram Forwarding Example

- Suppose a router has built up the routing table shown in the following table. The router can deliver packets directly over interfaces 0 and 1, or it can forward packets to routers R1, R2 or R3.
- Determine what the router does with a packet addressed to each of the following destinations:
  - (a) 148.110.17.131
  - (b) 148.110.16.18
  - (c) 212.40.163.90
  - (d) 212.40.163.17

Network Prefix/ Destination Network	Subnet Mask	Next Hop
148.110.16.0	255.255.255.128	Interface 0
148.110.16.128	255.255.255.128	Interface 1
148.110.17.0	255.255.255.128	R1
212.40.163.0	255.255.255.192	R2
<default>		R3

IP Datagram Forwarding Solution

- (a) Consider the destination IP address 148.110.17.131
- Taking a bit-wise AND with the subnet mask 255.255.255.128:

  148. 110. 17. 1 0 0 0 0 0 1 1

  255. 255. 255. 1 0 0 0 0 0 0 0

  ----

  148. 110. 17. 1 0 0 0 0 0 0 0

  148. 110. 17. 128 - it does not match with any of the corresponding network prefixes (148.110.16.0, 148.110.16.128 or 148.110.17.0) for the subnet mask 255.255.255.128 in the routing table.
- Taking a bit-wise AND with the subnet mask 255.255.255.192:

  148. 110. 17. 1 1 0 0 0 0 1 1

  255. 255. 255. 1 1 0 0 0 0 0 0

  ----

  148. 110. 17. 1 1 0 0 0 0 0 0

  148. 110. 17. 192 - it does not match with the network prefix 212.40.163.0.
- Since, there is no match with any of the routing table entries with their masks, we have to forward the packet to the default router, R3.

IP Datagram Forwarding Solution

- (b) Consider the destination IP address 148.110.16.18
- Taking a bit-wise AND with the subnet mask 255.255.255.128:
  - **148. 110. 16. 0 0 0 1 0 0 1 0**
  - **255. 255. 255. 1 0 0 0 0 0 0 0**
  - **----------**
  - **148. 110. 16. 0 0 0 0 0 0 0 0**
  - **148. 110. 16. 0 - it matches with the entry for which the next hop is interface 0.**
- (c) Consider the destination IP address 212.40.163.90
- Taking a bit-wise AND with the subnet mask 255.255.255.192:
  - **212. 40. 163. 0 1 0 1 1 0 1 0**
  - **255. 255. 255. 1 1 0 0 0 0 0 0**
  - **----------**
  - **212. 40. 163. 0 1 0 0 0 0 0 0**
  - **212. 40. 163. 64 - it does not match with the entry 212.40.163.0. Hence, send the packet to the default next hop router, R3.**

IP Datagram Forwarding Solution

- (d) Consider the destination IP address 212.40.163.17
- Taking a bit-wise AND with the subnet mask 255.255.255.192:
  - **212. 40. 163. 0 0 0 1 0 0 0 1**
  - **255. 255. 255. 1 1 0 0 0 0 0 0**
  - **----------**
  - **212. 40. 163. 0 0 0 0 0 0 0 0**
  - **212. 40. 163. 0 - it matches with the entry for which the next hop is R2.**

Routing Table: example 3

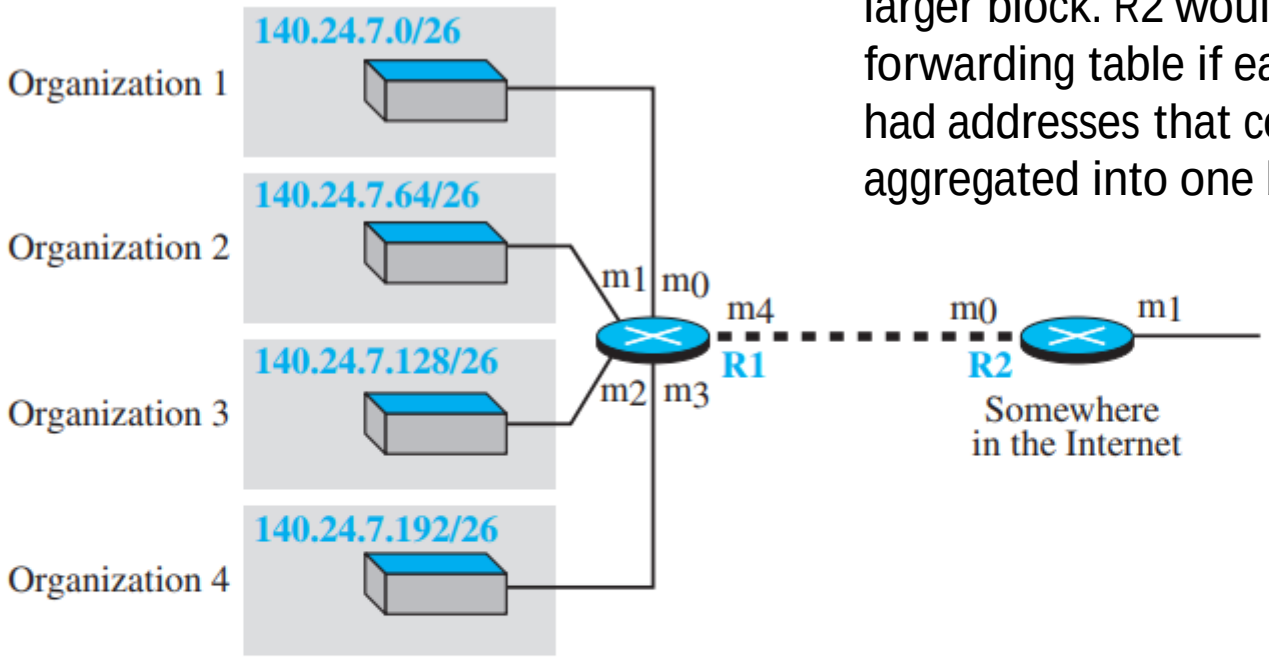
network	netmask	first hop	
131.175.15.0	255.255.255.0	131.175.21.1	x
131.175.16.0	255.255.255.0	131.175.21.2	x
131.175.17.0	255.255.255.0	131.175.21.3	x
131.180.23.0	255.255.255.0	131.175.21.4	x
131.180.18.0	255.255.255.0	131.175.21.4	x
131.180.21.0	255.255.255.0	131.175.21.4	OK
131.180.0.0	255.255.0.0	131.175.21.5	OK
0.0.0.0	0.0.0.0	131.175.12.254	OK

131.180.21.78

longest mask matching. This principle states that the forwarding table is sorted from the longest mask to the shortest mask. In other words, if there are three masks, /27, /26, and /24, the mask /27 must be the first entry and /24 must be the last.

Address Aggregation

Address aggregation because the blocks of addresses for four organizations are aggregated into one larger block. R2 would have a longer forwarding table if each organization had addresses that could not be aggregated into one block



Forwarding table for R1

Network address/mask	Next-hop address	Interface
140.24.7.0/26	----------	m0
140.24.7.64/26	----------	m1
140.24.7.128/26	----------	m2
140.24.7.192/26	----------	m3
0.0.0.0/0	address of R2	m4

Forwarding table for R2

Network address/mask	Next-hop address	Interface
140.24.7.0/24	----------	m0
0.0.0.0/0	default router	m1

IP Datagram Forwarding (Alternate Approach)

Network Prefix/ Destination Network	Subnet Mask	Next Hop
148.110.16.0	255.255.255.128	Interface 0
148.110.16.128	255.255.255.128	Interface 1
148.110.17.0	255.255.255.128	R1
212.40.163.0	255.255.255.192	R2
<default>		R3

- For each of the above combinations of network prefix and subnet mask, determine the range of valid IP addresses.
- If the *test* IP addresses fall within the range, then the packet is sent through the appropriate next hop.

148.110.16.0 255.255.255.128
We have 25 bits for the network + subnet part and 7 bits for the host part
148.110.16.0 0 0 0 0 0 0 0
The range of unicast IP addresses are: **148.110.16.1 to 148.110.16.126**

148.110.16.128 255.255.255.128
We have 25 bits for the network + subnet part and 7 bits for the host part
148.110.16.1 0 0 0 0 0 0 0
The range of unicast IP addresses are: **148.110.16.129 to 148.110.16.254**

148.110.17.0 255.255.255.128
We have 25 bits for the network + subnet part and 7 bits for the host part
148.110.17.0 0 0 0 0 0 0 0
The range of unicast IP addresses are: **148.110.17.1 to 148.110.17.126**

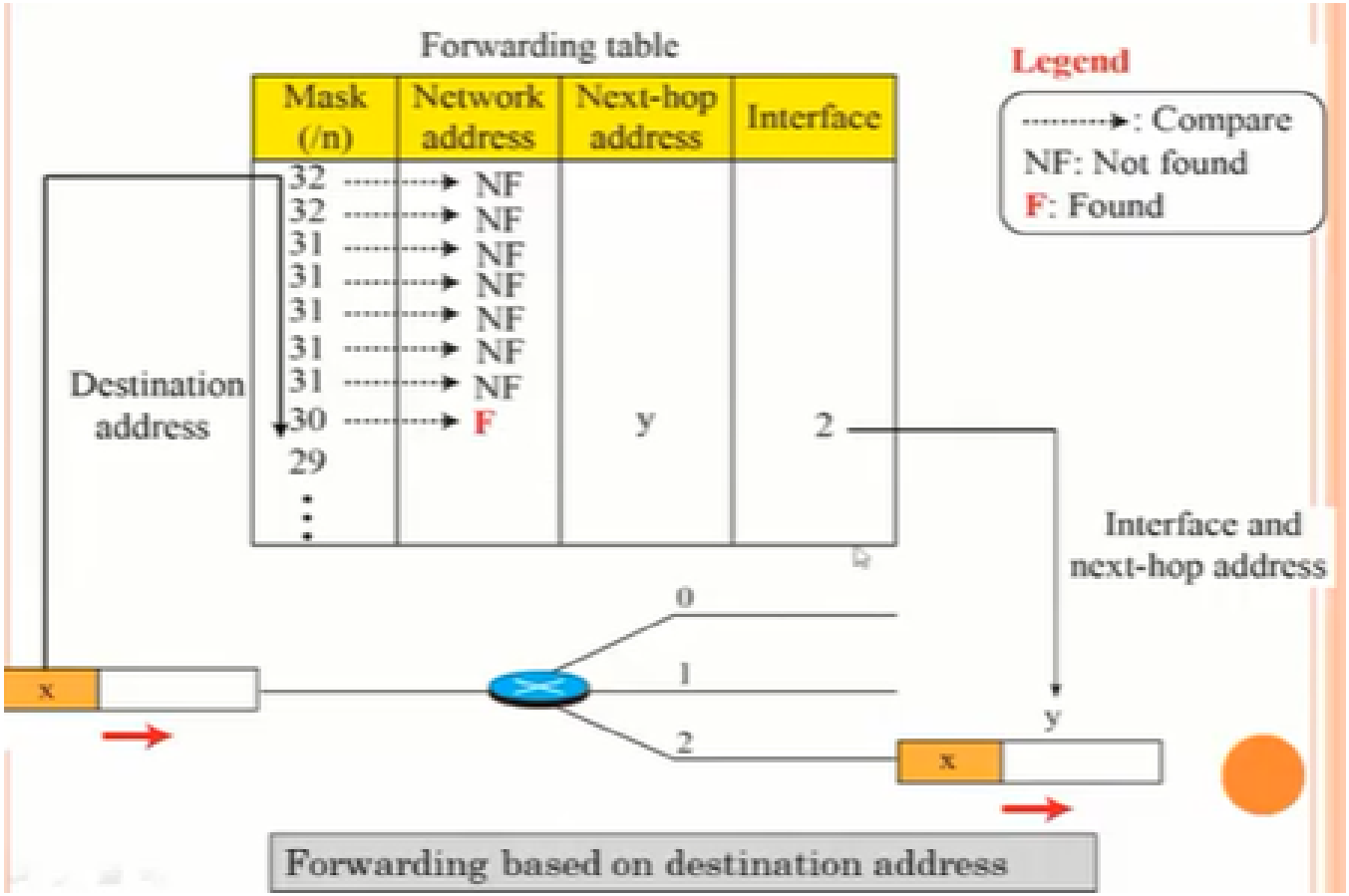
212.40.163.0 255.255.255.192
We have 26 bits for the network + subnet part and 6 bits for the host part
212.40.163.0 0 0 0 0 0 0 0
The range of unicast IP addresses are: **212.40.163.1 to 212.40.163.62**

- (a) 148.110.17.131 – Default R3
- (b) 148.110.16.18 – Interface 0
- (c) 212.40.163.90 - Default R3
- (d) 212.40.163.17 – R2

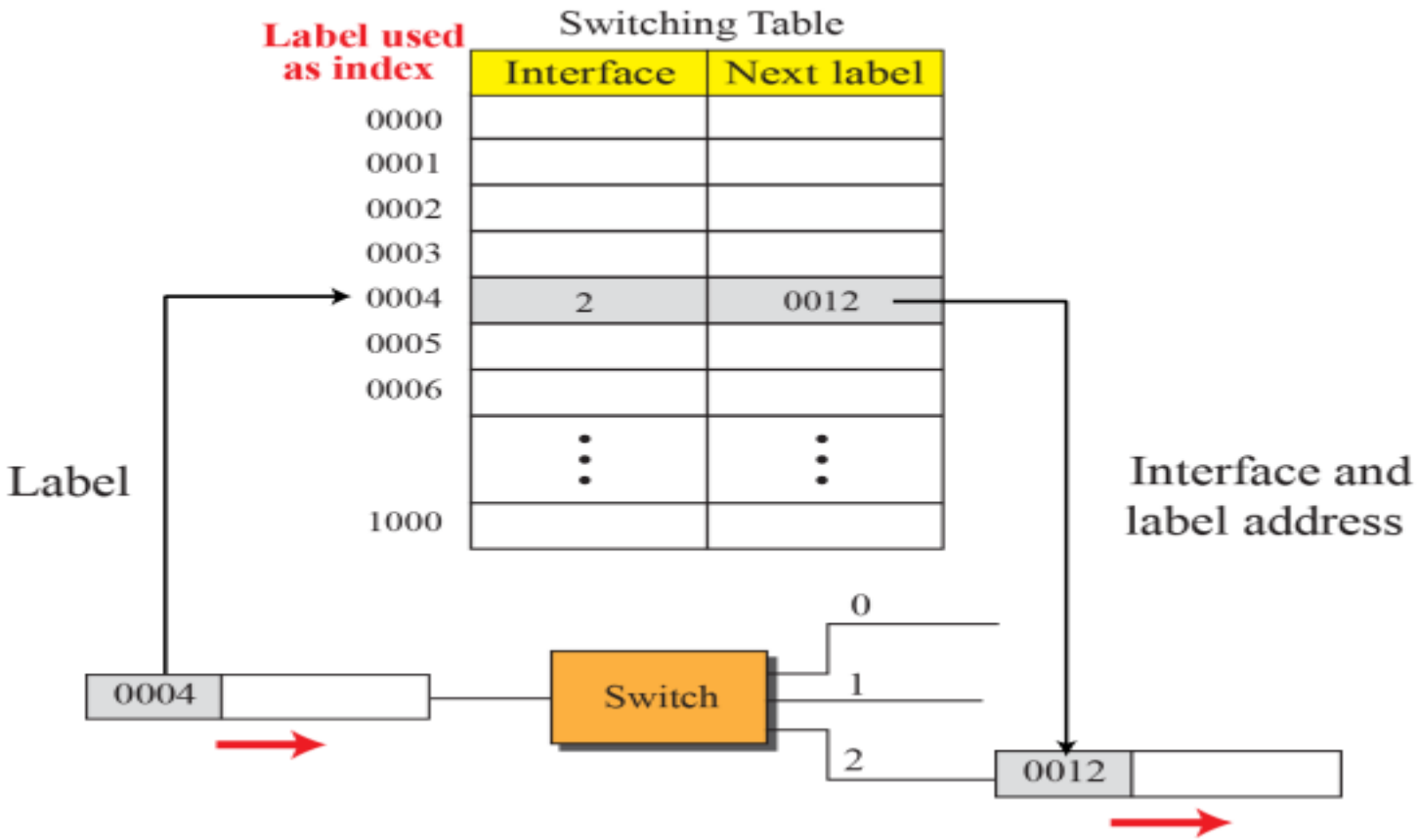
(ii) Based on label

- In the 1980s, an effort started to somehow change IP to behave like a connection-oriented protocol in which the routing is replaced by switching.
- In a connection-oriented network (virtual-circuit approach), a switch forwards a packet based on the label attached to the packet. Routing is normally based on searching the contents of a table; switching can be done by accessing a table using an index.
- Label map to a switch,each label will have an index
- In other words, routing involves searching; switching involves accessing.

Forwarding based on destination



Forwarding based on label



6.4 IP Auxiliary Protocols and Technologies

Address Resolution Protocol (ARP)

- Motivation for Address Translation
- IP datagrams contain IP addresses, but the physical network interface hardware cannot understand it. So, if the datagram has to be sent to a host or next hop router in a given physical network, it has to be encapsulated in a frame and addressed to the hardware address (link-level address) of the host/router interface on that network.

- ARP cache/table: The protocol enables each host on a network to build a table of mappings between IP addresses and link-level addresses.
- The entries in the table are discarded if not refreshed.

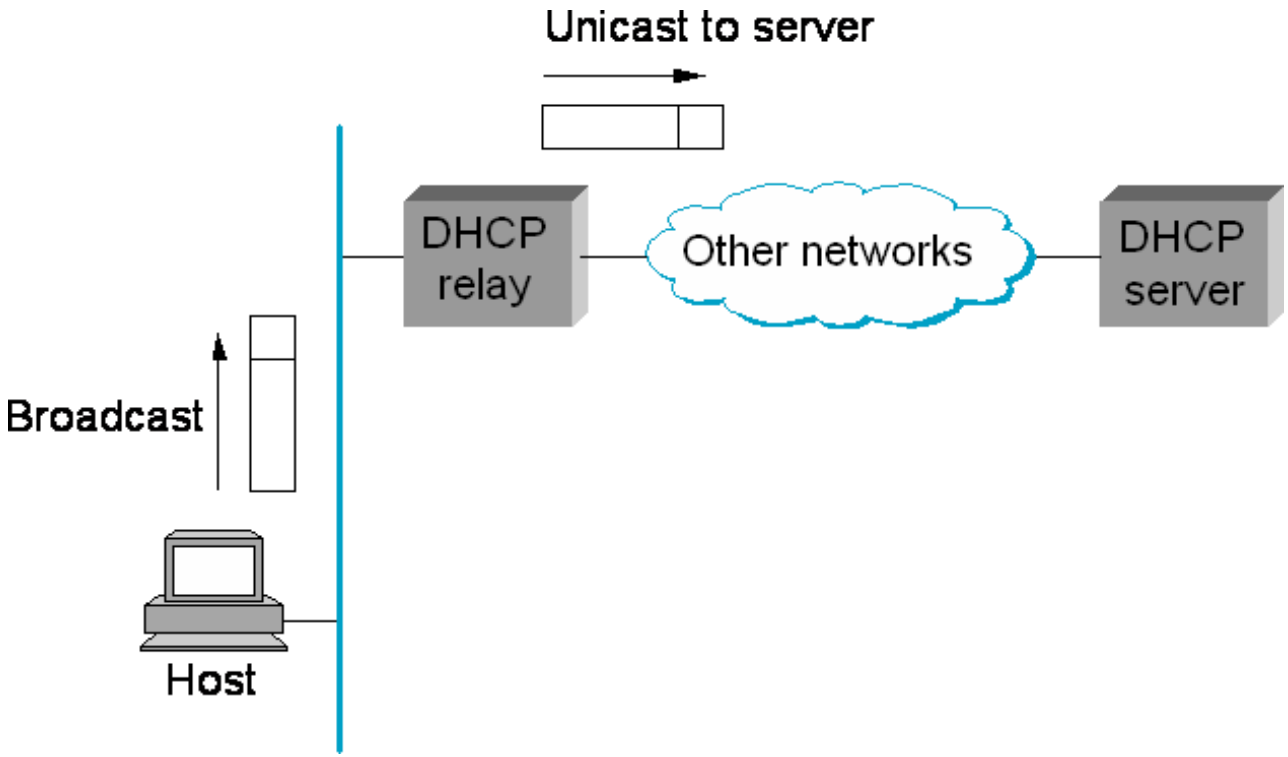
Address Resolution Protocol (ARP)

- If a host wants to send an IP datagram to a host (or router) that is known to be on the same network (i.e., the sending and receiving host have the same IP network number), the host first checks its ARP table for a mapping of the receiver's IP address to link-level address.
- If no mapping is found, the host broadcasts a ARP query (that contains the IP address in question) onto the network.
- Each host receives the query and checks to see if it matches its IP address. If it does match, the host sends a response message that contains its link-level address back to the originator of the query.
- The target host of the ARP query also creates an entry in its ARP table and stores the mapping between the IP address and link-level address of the query's sender. All the other hosts discard the ARP query.
- The ARP query also includes the IP address and link-level address of its sender. This information is read by all the hosts receiving the broadcast. Hosts add an entry in the ARP table to create a mapping between the IP address and link-level address of the sender (or refresh the mapping using the query, if an entry already exists).

Dynamic Host Configuration Protocol (DHCP)

- IP addresses need to be reconfigurable as the network part of the IP address of a host has to be at least changed depending on the network to which the host is attached.
- When a host connects to a network, it should be able to get assigned the IP address corresponding to that particular network. The protocol that handles this problem is called the Dynamic Host Configuration Protocol (DHCP)
- DHCP relies on the existence of a DHCP server that provides the configuration information to hosts. The DHCP server would maintain a pool of available IP addresses from which it will handover one address to a requesting host. The network administrator has to only allocate a range of IP addresses (all with the same IP network number) to each network.
- A host that boots up or connects to a network needs to be able to contact the DHCP server.

Dynamic Host Configuration Protocol (DHCP)



Dynamic Host Configuration Protocol (DHCP)

- The host broadcasts a DHCPDISCOVER message (with target IP address 255.255.255.255) to all the hosts in the network.
- Only the DHCP server or the DHCP relay agent pick up the DHCP message. If the message is picked up by a DHCP server, it assigns an IP address to the requesting host and notifies it through a DHCPREPLY message.
- If the DHCPDISCOVER message is picked up by a DHCP relay agent, it unicasts the message to the DHCP server, which assikns the address and notifies the requesting host through the relay agent.
- An IP address is assigned on a “lease basis” and the requesting client has to periodically refresh it. If not refreshed, the IP address is returned to the available pool of addresses and is free to be assigned to another host.

Internet Control Message Protocol (ICMP)

- IP is configured with a companion protocol called ICMP that 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.
- Some of the ICMP error messages are:
- Destination host unreachable
- Reassembly process failed
- TTL reached 0
- Checksum failed
- Cannot fragment

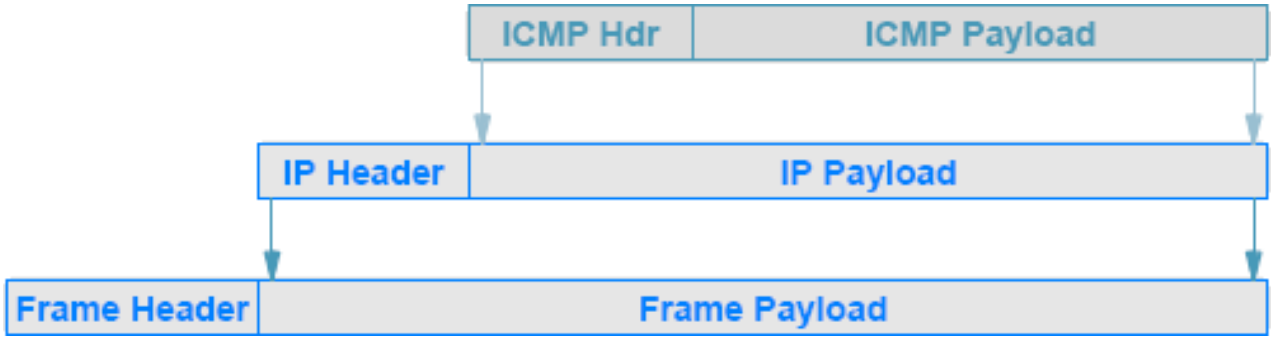
- Besides, ICMP also includes some control messages that help to improve network performance. These are:
- ICMP re-direct
- ECHO Request/Reply

Internet Control Message Protocol (ICMP)

- ICMP uses IP to transport each error message:
  - when a router has an ICMP message to send
    - it creates an IP datagram and encapsulates the ICMP message in it
  - the ICMP message is placed in the payload area of the IP datagram
  - the datagram is then forwarded as usual
    - with the complete datagram being encapsulated in a frame for transmission
- ICMP messages do not have special priority
  - They are forwarded like any other datagram, with one minor exception
- If an ICMP error message causes an error
  - no error message is sent
- The reason should be clear:
  - the designers wanted to avoid the Internet becoming congested carrying error messages about error messages

© 2009 Pearson Education Inc., Upper Saddle River, NJ. All rights reserved.

ICMP Encapsulation



Number	Type	Purpose
0	Echo Reply	Used by the ping program
3	Dest. Unreachable	Datagram could not be delivered
5	Redirect	Host must change a route
8	Echo	Used by the ping program
11	Time Exceeded	TTL expired or fragments timed out
12	Parameter Problem	IP header is incorrect
30	Traceroute	Used by the traceroute program

© 2009 Pearson Education Inc., Upper Saddle River, NJ. All rights reserved.

Private IP Addresses

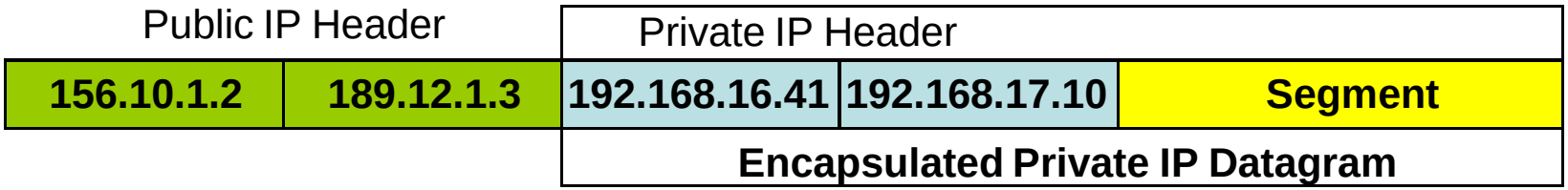
- IANA reserves certain blocks of IP addresses (called private IP address) for use by the private internets. The **private ip address blocks are:**
  10.0.0.0 - 10.255.255.255
  172.16.0.0 - 172.31.255.255
  192.168.0.0 - 192.168.255.255

- The same set of private IP addresses can be used at different organizations (i.e., a private IP address has to be only locally unique); where as a public IP address has to be globally unique.
- Private IP addressing is one of the solutions to reduce the exhaustion of IP address space.
- The private ip addresses are not routable in the public internet (i.e., packets bearing private ip addresses are not forwarded by routers in the Internet).
  - Because if more than one host (located in the networks of two different organizations/domains) has the same IP address, then to which host do the routers forward a packet? – Not possible to resolve.

Virtual Private Network (VPN)

- Organizations setup network sites at different locations, have each of them assigned a private IP address space that is unique among the hosts within the entire organization.
- Hosts at the different sites communicate with each through a VPN setup across the public Internet.
- This is accomplished through IP-in-IP encapsulation. There will be a public gateway (that has a public IP address) setup at each of these sites.
- The private datagram (containing the IP header with the source and destination private IP addresses plus the segment) is encapsulated inside a public IP header containing the public IP addresses of the gateways at the two sites as the source and destination IP addresses.
- Routers across the Internet will transfer the datagram based on the public IP addresses. For security reasons, the inner private IP datagram may be even encrypted by the end-gateways so that the contents of the inner encapsulated datagram are not seen by anyone in the public Internet.

Virtual Private Network (VPN)



IP-in-IP Encapsulation

Network Address Translation (NAT)

- The idea is to have a pool of public IP addresses assigned to one or more gateway routers.
- The internal hosts with private IP addresses go through one of these gateway routers.
- At the gateway router,
  - For outgoing traffic: the private IP address of the internal host is replaced (translated) to the public IP address of the gateway router.
  - For incoming traffic: replace the public gateway IP address with the private IP address of the internal host
  - A translation table between the public/private IP addresses is maintained.
- Drawbacks:
  - Not a scalable solution when the number of connections (either the number of internal hosts and/or the number of applications running on the internal hosts) exceed the number of public IP addresses available for the gateway router(s).

Network Address Port Translation (NAPT)

- The idea is to use the public IP address(es) in combination with the vast range of port numbers (1024 – 65535) to replace the private IP addresses of the internal hosts and the port numbers of the applications running on these hosts, communicating through the Internet.
  - A 4-tuple translation table needs to be maintained at the gateway routers.

- NAPT is more scalable than NAT, because, with this scheme, we can cover several TCP/UDP connections of the internal hosts and their applications with one public IP address and the different port numbers (1024 – 65535).

- Though fundamentally different, given the limitations of NAT, NAPT has been commonly referred to as NAT due to the former's widespread usage of communicating to/from private IP addresses.

6.5 IPv6

IPv6

Address Space Allocation

- 128 bits
- No classes. But the leading bits specify different uses of the IPv6 addresses.

Address Notation

- The standard representation is x:x:x:x:x:x, where each “x” is a hexa-decimal representation of the 16-bit piece of the address.
- Example: 47CD:1234:4422:AC02: 0022:1234:A456:0124

Aggregatable global unicast addresses: The entire functionality of IPv4’s three main classes is contained in the 001 prefix.

Link local use and site local use addresses: The “link local use” address enables a host to construct an address that will work on the network to which it is connected without being concerned about the global uniqueness of the address. Site local use addresses are intended to allow valid addresses to be constructed on a site that is not connected to the global Internet.

IPv4-compatible-IPv6 address: Obtained by zero-extending a 32-bit IPv4 address to 128 bits.

Example: IPv4-compatible IPv6 Address

Compute the IPv4-compatible IPv6 address for 143.132.10.45

Represented as : : 143.132.10.45 (prefixed with 128-32 = 96 0s)

Writing each of the 8-bit decimal value in binary,

: : 10001111:10000100:00001010:00101101

➔ : : 8F84: 0A2D

IPv6 Prefix Values and their Use

Prefix	Use
0000 0000	Reserved
0000 0001	Unassigned
0000 001	Reserved for NSAP Allocation
0000 010	Reserved for IPX Allocation
0000 011	Unassigned
0000 1	Unassigned
0001	Unassigned
001	Unassigned
010	Provider-Based Unicast Address
011	Unassigned
100	Reserved for Geographic-Based Unicast Addresses
101	Unassigned
110	Unassigned
1110	Unassigned
1111 0	Unassigned
1111 10	Unassigned
1111 110	Unassigned
1111 1110 0	Unassigned
1111 1110 10	Link Local Use Addresses
1111 1110 11	Site Local Use Addresses
1111 1111	Multicast Addresses

IPv6

Address Notation

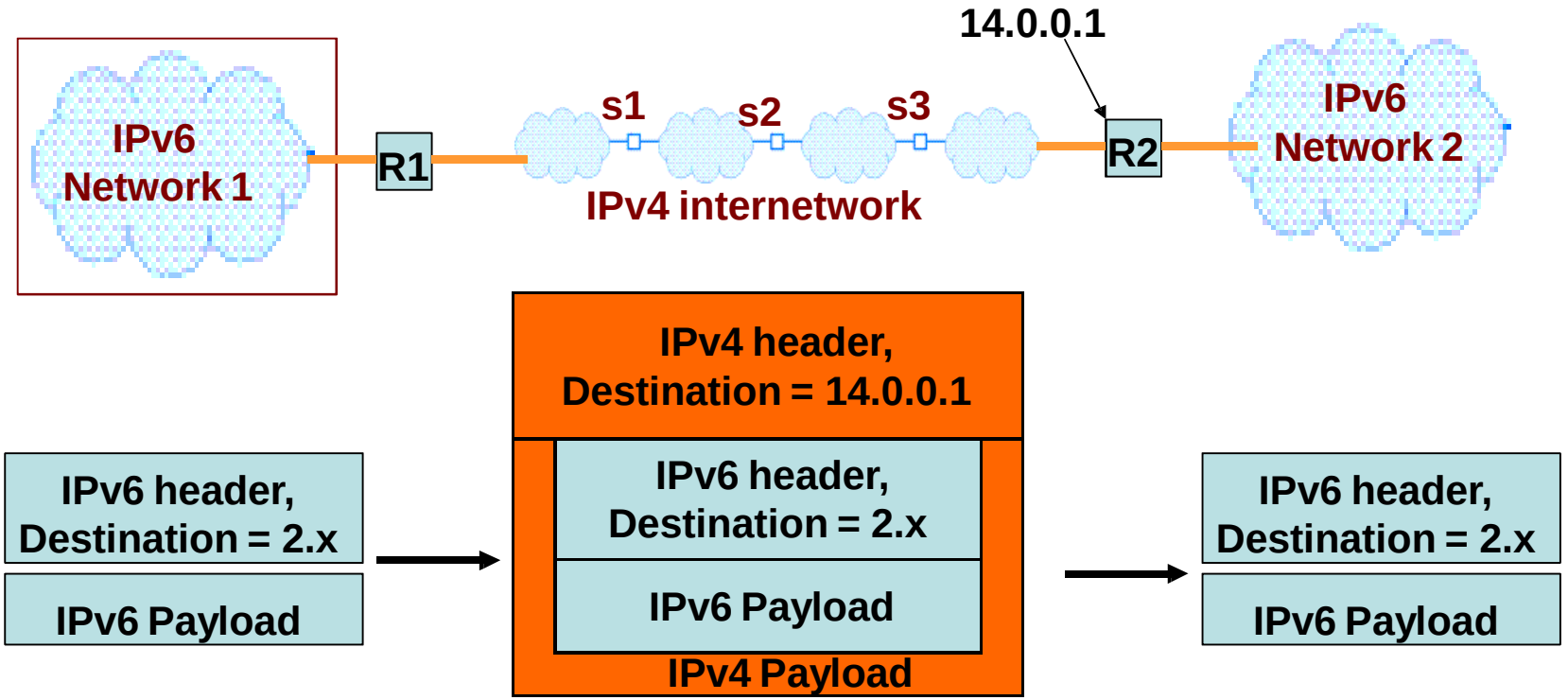
- An address with a large number of contiguous 0s can be written more compactly by omitting all the 0 fields. For example,
  47CD:0000:0000:0000:0000:0000:A456:0124
  can be written as 47CD::A456:0124

Transition from IPv4 to IPv6:

- IPv4-only nodes should be able to communicate with IPv6 nodes and IPv4 nodes.
- Two IPv6 nodes should be able to communicate through an IPV4 network.
- An IPv4 to IPv6 (or vice-versa) goes through a NAT-like gateway at the IPv6 side → The IPv6 address space network is treated like a private network and the IPv6 addresses are replaced with one public IPv4 address.

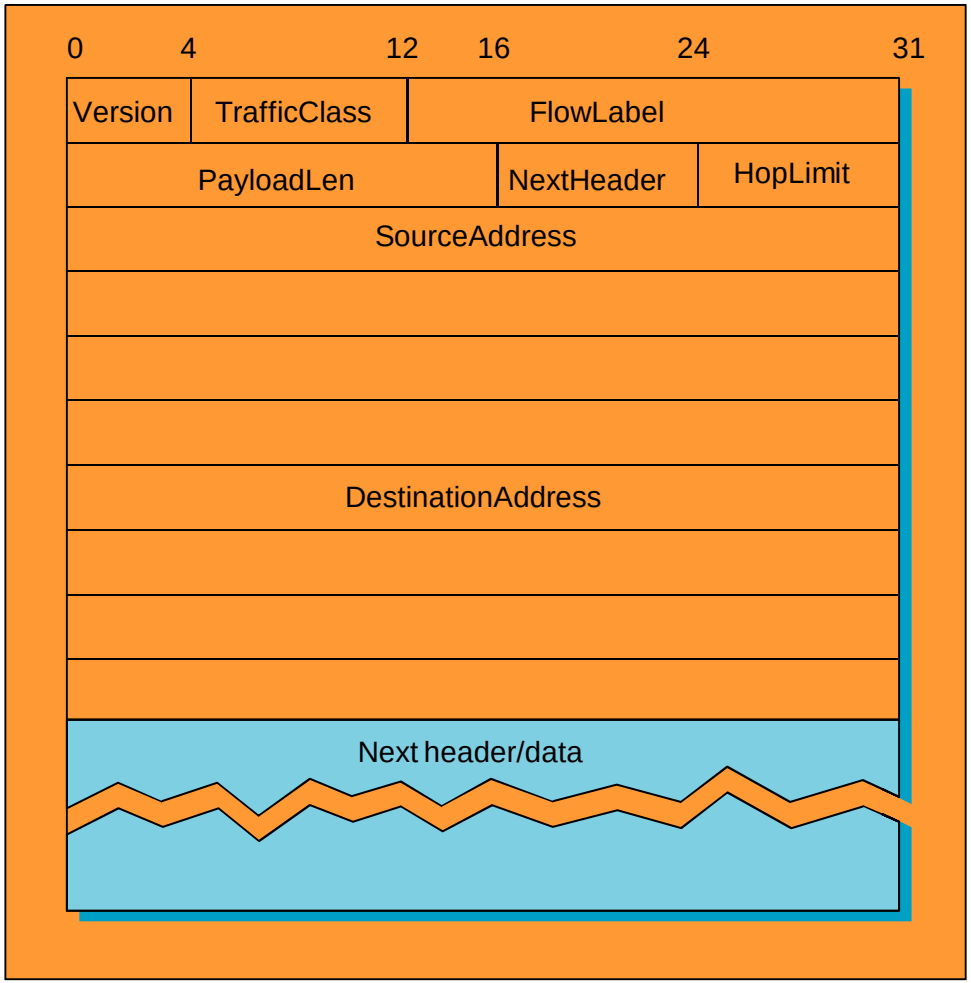
IPv6 in IPv4 Tunneling

- The entire IPv6 datagram is encapsulated inside the IPv4 payload.
- For the IPv6 interface of router R1, the number of hops to the IPv6 network 2 is ONE; while for the IPv4 interface of router R1, the number of hops to the IPv4 interface of router R2 is THREE.



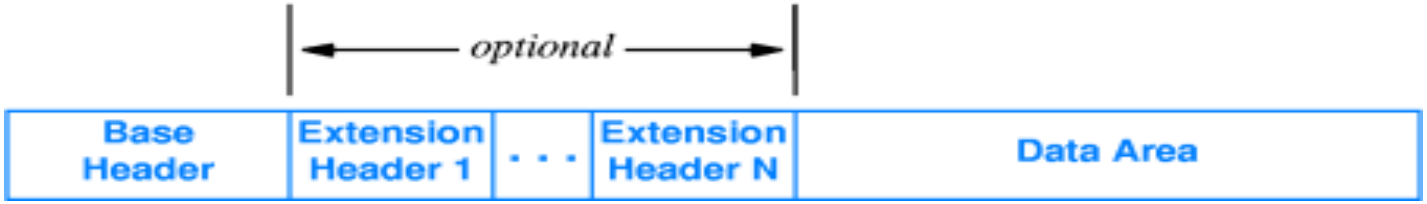
IPv6 Packet Header

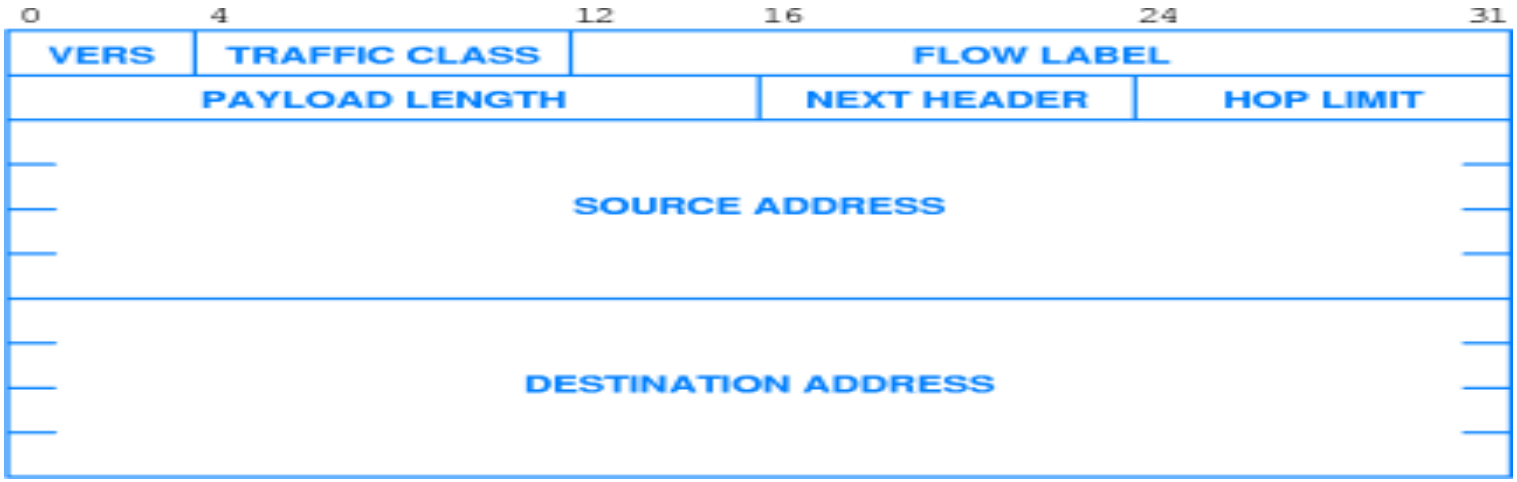
IPv6 Base Header



IPv6 Datagram Format

- Extension headers are optional: at the minimum, a datagram should have the base header followed by data.
- The extension headers are of variable size.





IPv6 Base Header Format

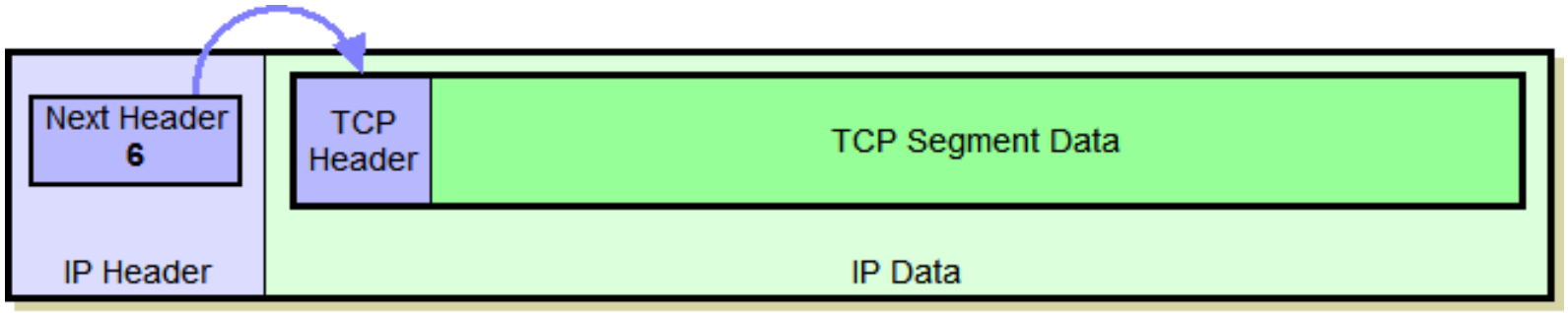
Use of Extension Headers in IPv6

Next Header Value (decimal)	Extension Header Name	Length (bytes)	Description
0	Hop-By-Hop Options	Variable	Defines an arbitrary set of options that are intended to be examined by all devices on the path from the source to the destination
43	Routing	Variable	Defines a method for allowing a source device to specify the route for a datagram
44	Fragment	8	This header is included only with datagrams that contain only a fragment of the original message. This header contains the Fragment Offset, Identification and More Fragment fields
50	Encapsulating Security Payload (ESP)	Variable	Carries encrypted data for secure communications
51	Authentication Header (AH)	Variable	Contains information used to verify the authenticity of the encrypted data
60	Destination Options	Variable	Defines an arbitrary set of options that are intended to be examined only by the destinations of the datagram

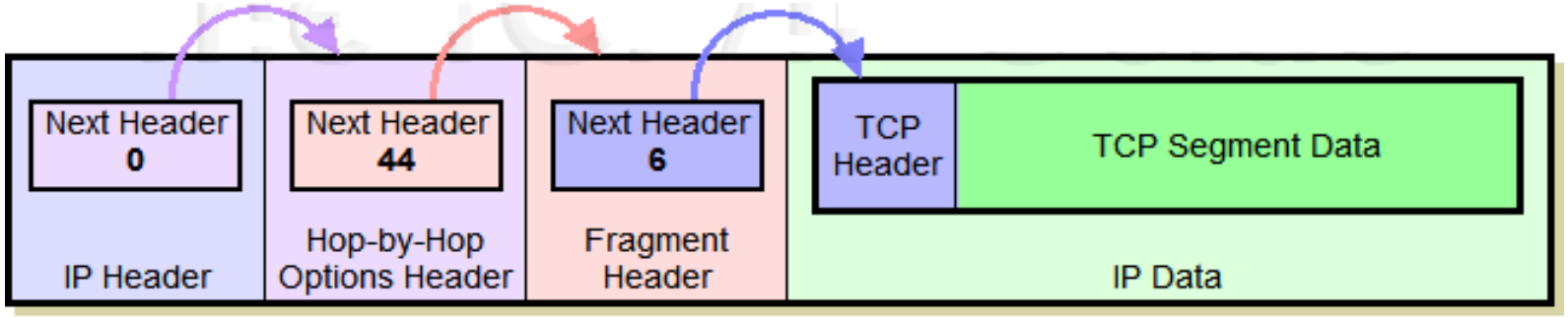
6 TCP Header

17 UDP Header

Use of Extension Headers in IPv6



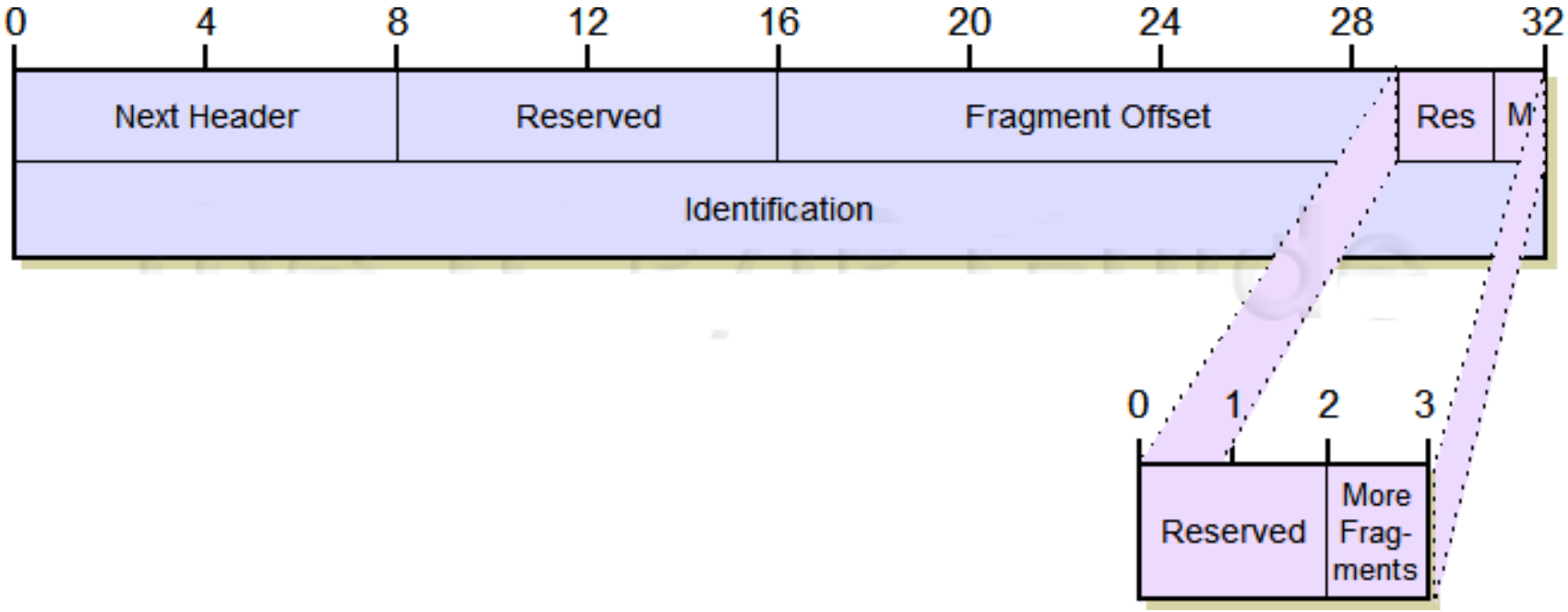
IPv6 Datagram With No Extension Headers Carrying TCP Segment



IPv6 Datagram With Two Extension Headers Carrying TCP Segment

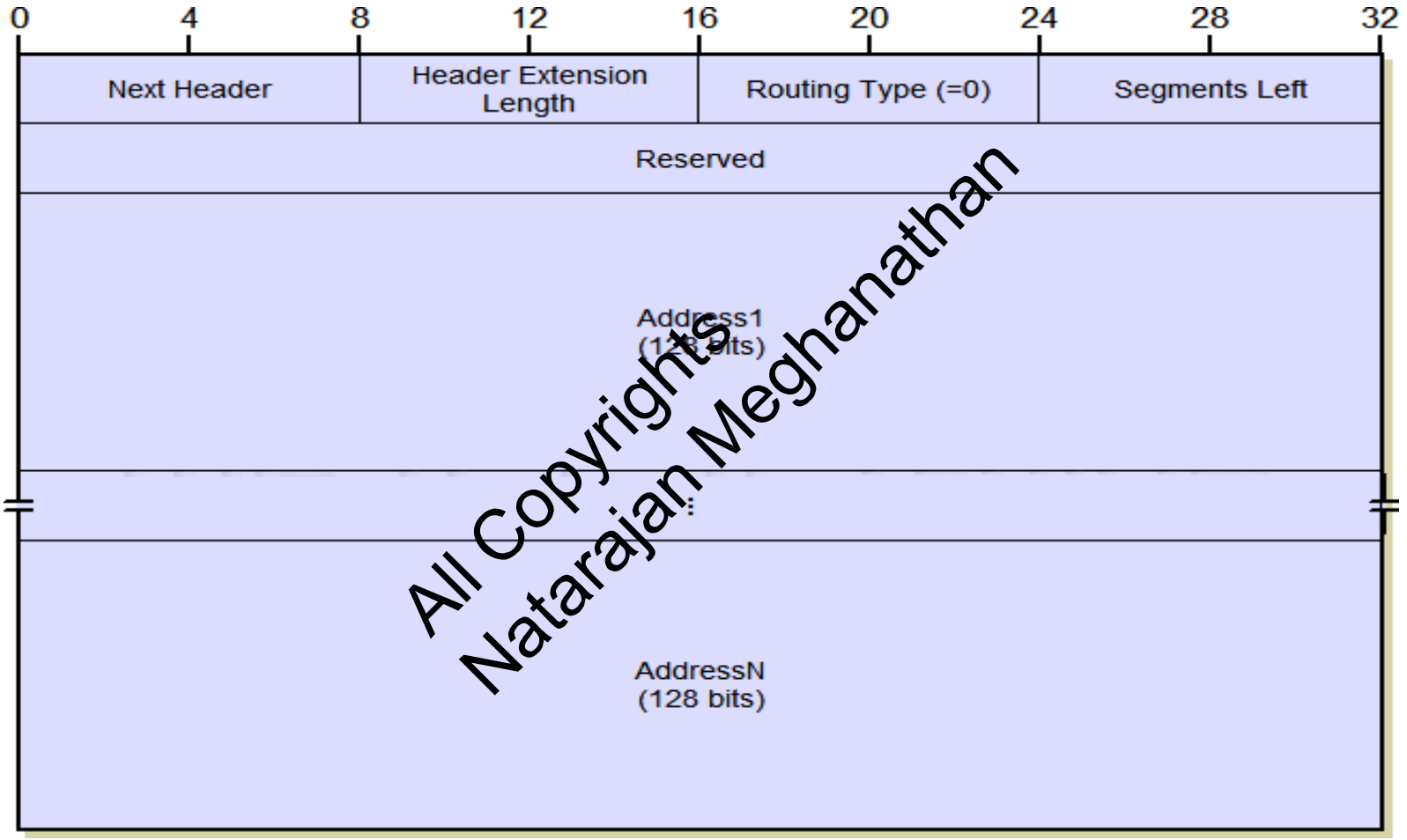
Picture taken from http://www.tcpipguide.com

IPv6 Fragment Extension Header



Picture taken from http://www.tcpipguide.com

IPv6 Routing Extension Header



All Copyrights
Natarajan Meghanathan

Picture taken from http://www.tcpipguide.com