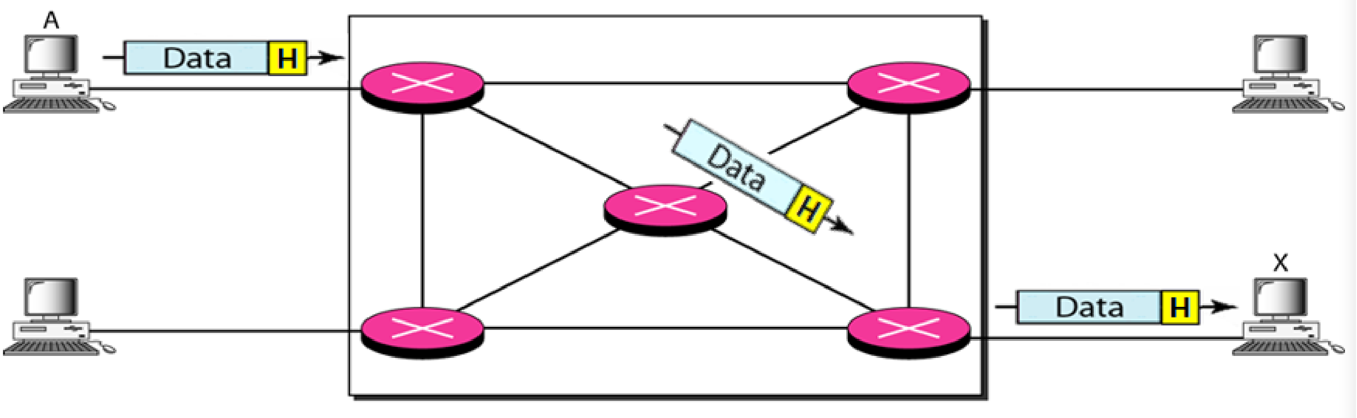
C4 Internet Protocol

1. Network Taxonomy

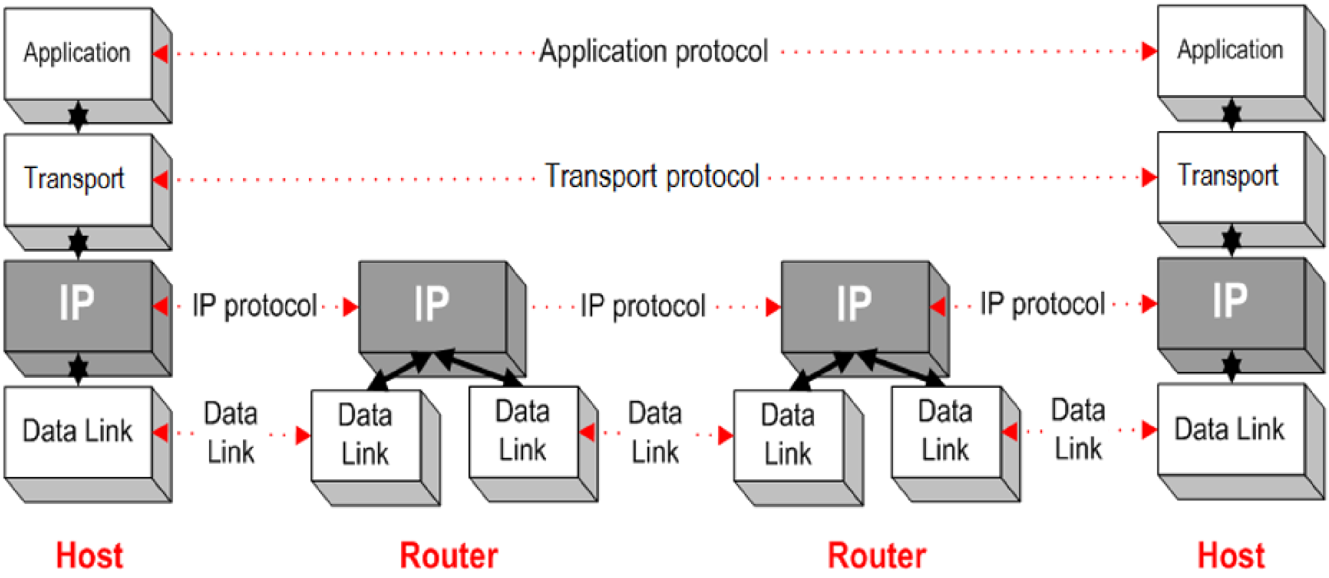
- In packet-switched datagram networks, transmission begins without establishing a communication path
    - Addressing is needed in each packet



- Immediate transmission can be done from A to X but A needs to append a header containing X's address so that intermediate packet-switched nodes know how to send it to X
- Packet-switched nodes are also shared with others for transmission
- Each data can also be divided into suitable sized packets according to MTU (Maximum Transfer Unit) of individual network (this process is known as fragmentation)
    - Each packet can take a different path depending on load of packet-switches which may result in out of sequence / lost / corrupted packets
    - Handled by Transport layer (TCP)
- MTU is the size of the packet including IP and TCP header (20 bytes each)
    - MSS is maximum size of the segment, so if MTU is 1500bytes, MSS is 1460bytes

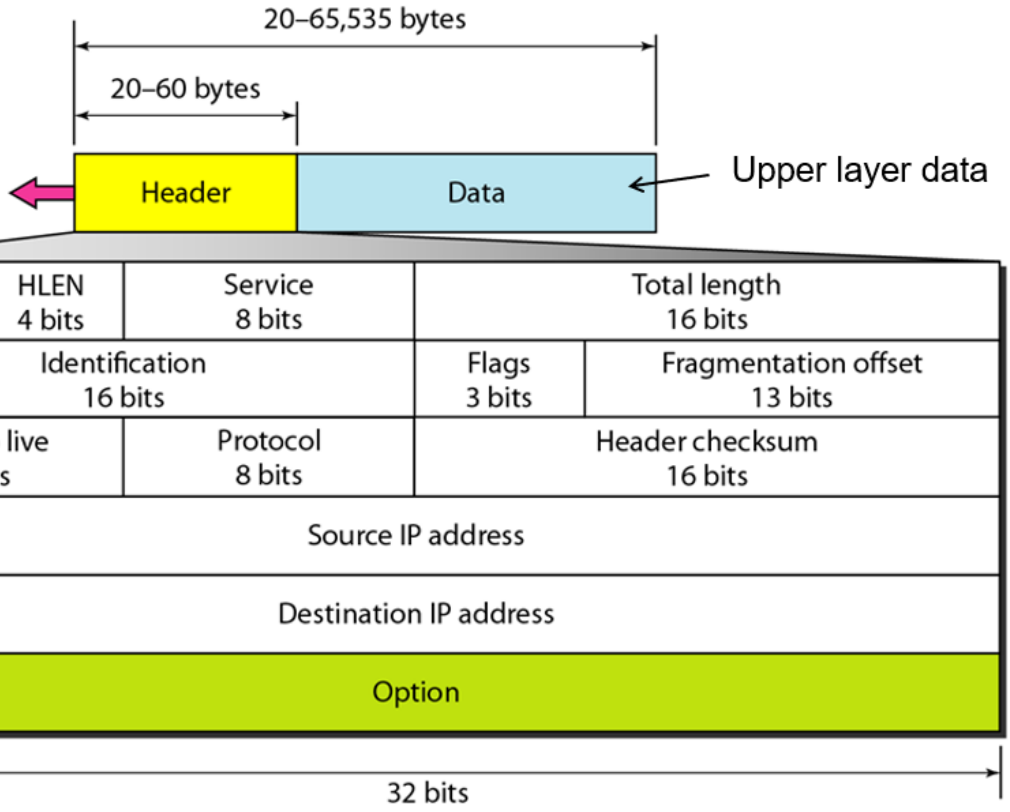
2. IPv4

Internet Protocol



- A packet will travel from host (S) to host (R) through intermediate routers via 'hopping'
- Protocol Functions include **addressing** and **fragmentation**
- Provides
    - **Connectionless** - each packet is handled independently
    - **Unreliable** - no error control
    - **Best Effort** - no throughput / delay / Quality of Service guaranteed

2.1 IPv4 Header



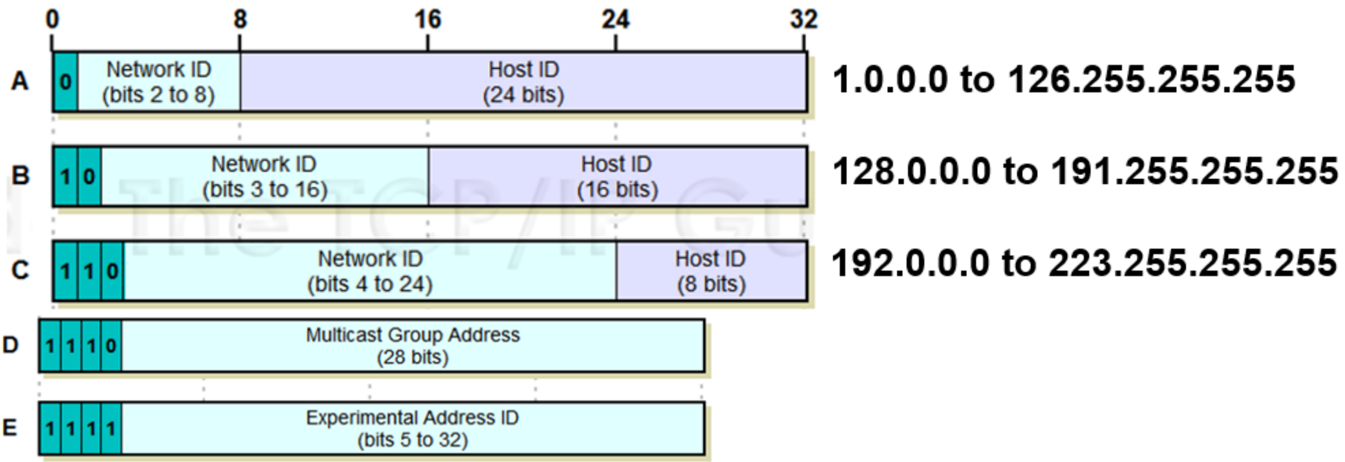
- IP header is 20 bytes
- VER: version number of IP

- HLEN: header length (to be multiplied by 4)
- Total Length: datagram length including header (max 65,535)
- Protocol: indicates protocol IP is carrying
  - 01_{16} ICMP
  - 06_{16} TCP
  - 11_{16} UDP
- Header Checksum: verify header is free from error
- Source / Dest. IP addresses: trivial

Notation

`[8bits] [8bits] [8bits] [8bits]`

- Classes are defined by the leading (first 4) bits
  - A: leading bit is 0
  - B: leading two bits is 10
  - C: leading three bits is 110
  - D: leading 4 bits is 1110
  - E: leading 4 bits is 1111



Special Use Cases

- Network and / or Host id all 0s
  - Can only be used as source address
  - `0.0.0.0` → host is on this network
  - `115.69.0.0` → network / subnet ID
- Network and / or Host id all 1's
  - Can only be used as destination address
  - `255.255.255.255` → limited broadcast within this network

- `155.69.255.255` → directed broadcast on `155.69.x.x`

- Loopback Address (`127.x.y.z`)
  - Internal loopback to same host, useful for self testing

2.2 IPv4 Address

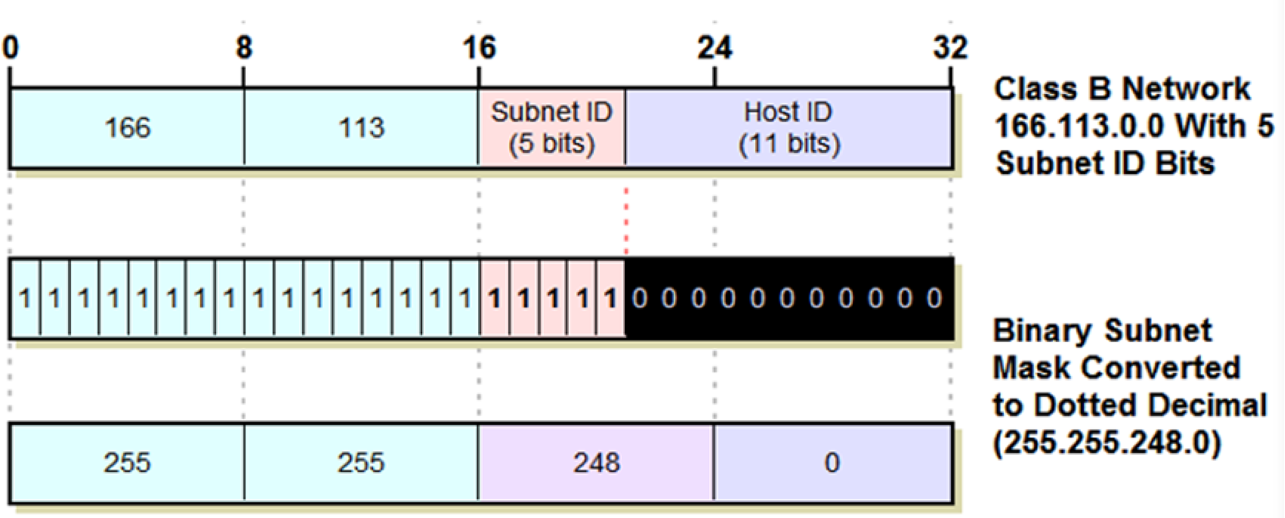
- IP address assigned to host / router's interface
  - Interface - connection between host / router and physical link
- A network consists of device interfaces with same network id of IP address which can physically reach each other without immediate router

2.3 Subnetting

- Partitioning of a single large network into smaller networks
- IP address structure has network prefix extended, length of host bits sacrificed
- Subnets are only visible within the organisation
  - Free to decide number of bits for subnet and host numbers

Subnet Masks

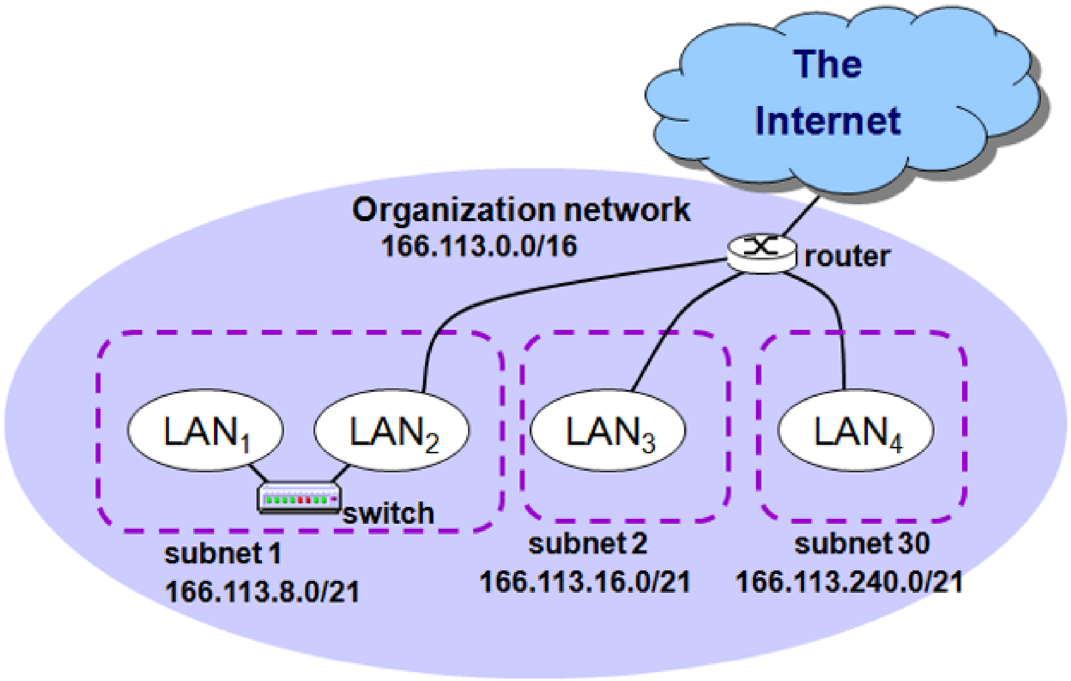
- Subnet mask `w.x.y.z` used to indicate length of extended network prefix; bits corresponding to extended network prefix set to 1's and 0's otherwise



Subnet Address Calculation

- 'slash' notation: `a.b.c.d/x`, x indicates number of bits for **extended network prefix**
- **Subnet address is different from IP address**
  - e.g. Class B address `166.113.0.0/21`, subnet address will be `255.255.248.0` (`11111111.11111111.11111000.00000000`)
- Maximum number bits available for hosts = 32 - x
- Let x = 21, maximum number of hosts in each subnet $= 2^{32-21} - 2 = 2046$

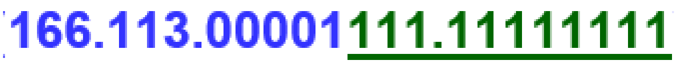
  - host id = `000 00000000` = network/subnet id number
  - host id = `111 11111111` = broadcast address



- For organisation network, 16 bits used for network prefix, remaining 16 bits set to 0 0 (hex) to represent organisation network id
- **is x fixed for different subnets under the same organisation network???**
  - Yes, since we are not dealing with CIDR (mentioned below) subnet in this case

Subnet Broadcasting

- Sets host address to all 1's, using example above, `166.113.15.255/21` means a broadcast to the subnet `166.133.8.0/21`

  166.113.00001111.11111111

  - Every bit after the network prefix is set to 1
  - **This is because 15 is used, if 255 is used, broadcast is sent to all hosts in organisational network (all subnets),** `166.133.11111111.11111111`
  - **if 7 is used, broadcast to subnet** `166.133.0.0/21`**???**
    - In reality this will not happen since subnets are partitioned properly

Note

- For subnets, everything after the n^{th} bit (if /n) needs to be 0 for broadcast, so if network id is required to be 172.121.13x.x/22, everything after the 22nd bit needs to be 0 for it to be the network id, but because 10000000 is 128 and not 13x, we need to increment the 22nd bit to 1 so that we get 132 (10000100)

2.4 IPv4 Address Exhaustion

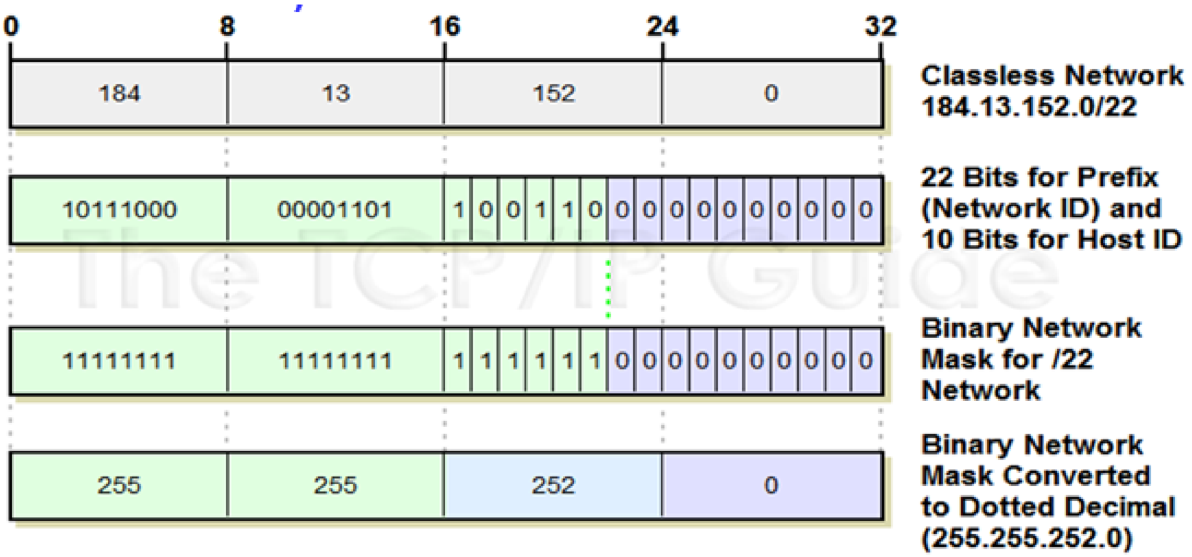
- 32-bit IP address not enough for today's global network and future growth

Solutions

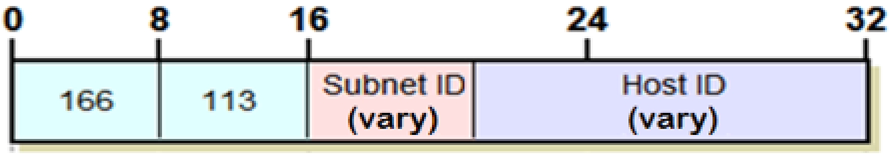
- Classless InterDomain Routing (short-term)
  - Reduces wastage in address allocation
  - Organisations given adequate but not excessive address space
- Network Address Translation (NAT) using IP addresses (eases problem, doesn't solve it)
  - Single machine (router) with IP address representing many computers / hosts behind it; IP addresses require translation
- IPv6, 128-bit space (long-term)
  - Large enough to install several billion computers on each square meter

2.5 Classless Inter-Domain Routing

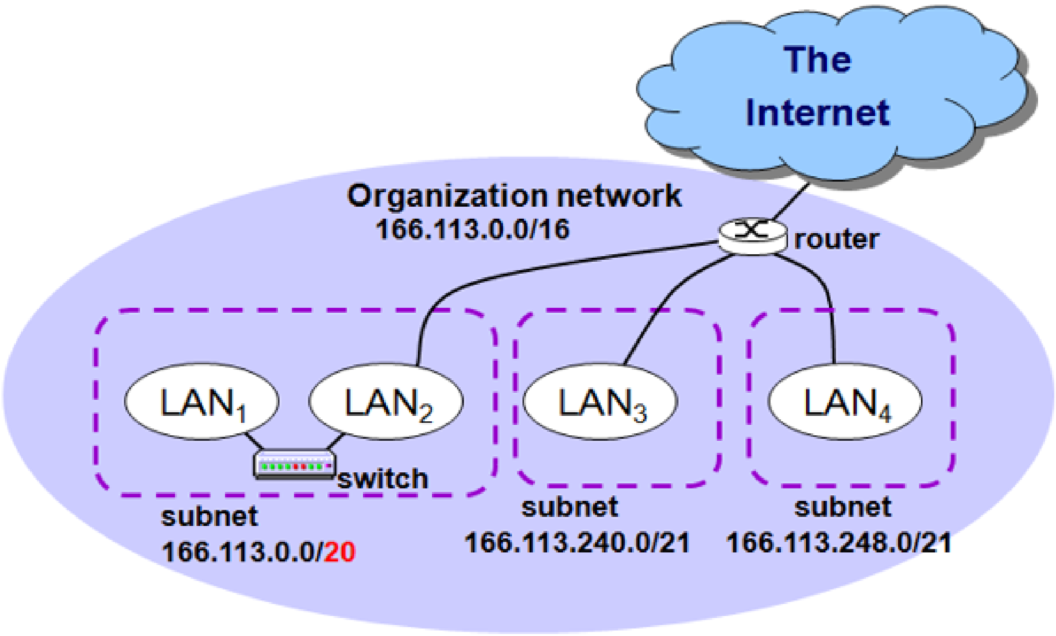
- No more classful (A,B,C,D,E) addressing
  - Length of network prefix can now be any length, have to add network mask (similar to subnet mask)


0
8
16
24
32
184
13
152
0
Classless Network 184.13.152.0/22
10111000
00001101
1 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
22 Bits for Prefix (Network ID) and 10 Bits for Host ID
11111111
11111111
1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
Binary Network Mask for /22 Network
255
255
252
0
Binary Network Mask Converted to Dotted Decimal (255.255.252.0)

- 'slash' notation again, `w.x.y.z/n`, n is number of bits allocated to network prefix
- Can be extended to subnetting, enabling variable length subnet masks



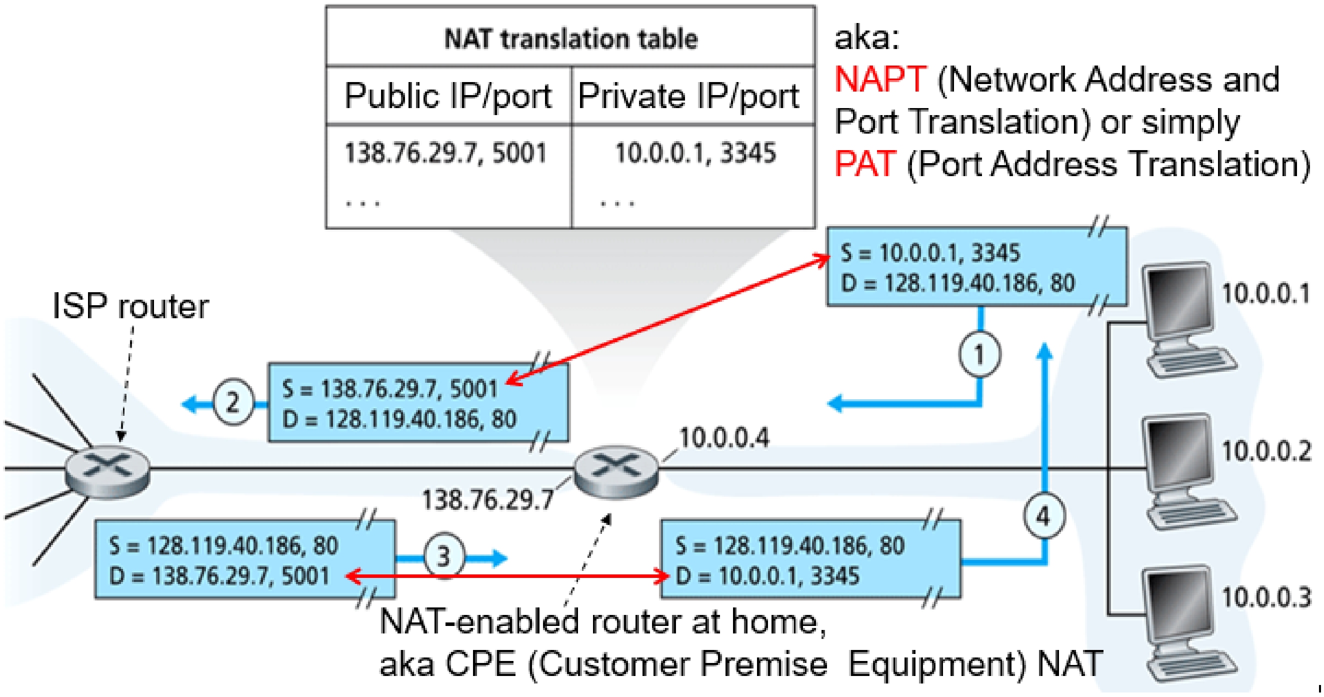
- no more subnet broadcast, subnet numbers including all 0's and 1's can be used



- We can see that the first subnet has a different length subnet mask compared to the rest

2.6 Network Address Translation

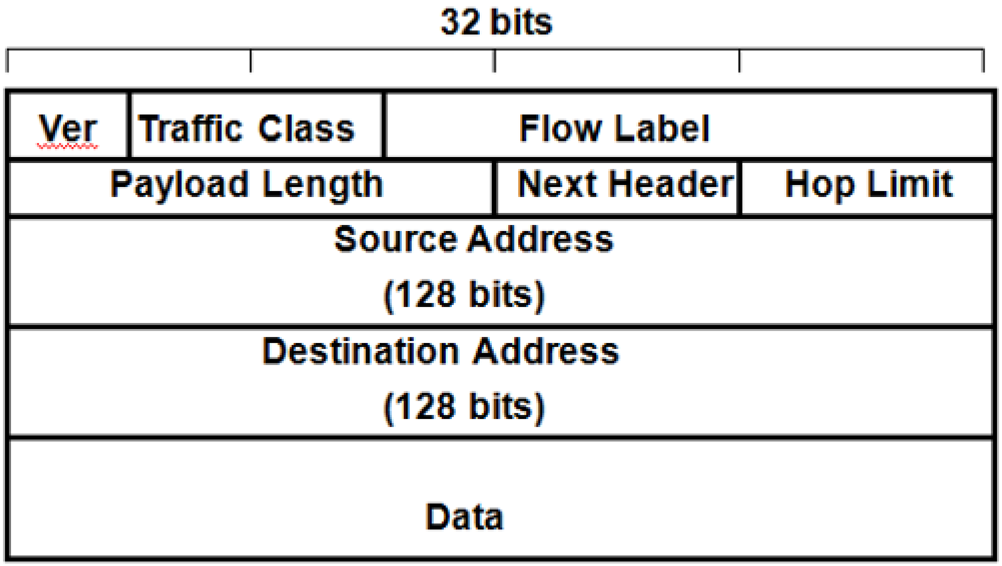
- Standard private addresses for private internet
  - `10.0.0.0/8` (`10.0.0.0` - `10.255.255.255`)
  - `172.16.0.0/12` (`172.16.0.0` - `172.31.255.255`)
    - `172.00010000.0.0` - `172.00011111.255.255`
  - `192.168.0.0/16` (`192.168.0.0` - `192.168.255.255`)
- Private addresses are not forwarded into the internet
  - Different private networks can re-use same private IP addresses
- A router (NAT enabled) is used to identify a network; only one IP address is required from ISP to support the whole private network to connect to the internet



3. IPv6

- Expanded address space (128 bits)
- Colon hexadecimal notation
    - Abbreviated notation
        - Within each 16-bit value, `0000` can be written as `0`, consecutive groups of 0s can be replaced by `::`, can only have `::` appearing once in a IPv6 address

3.1 IPv6 header

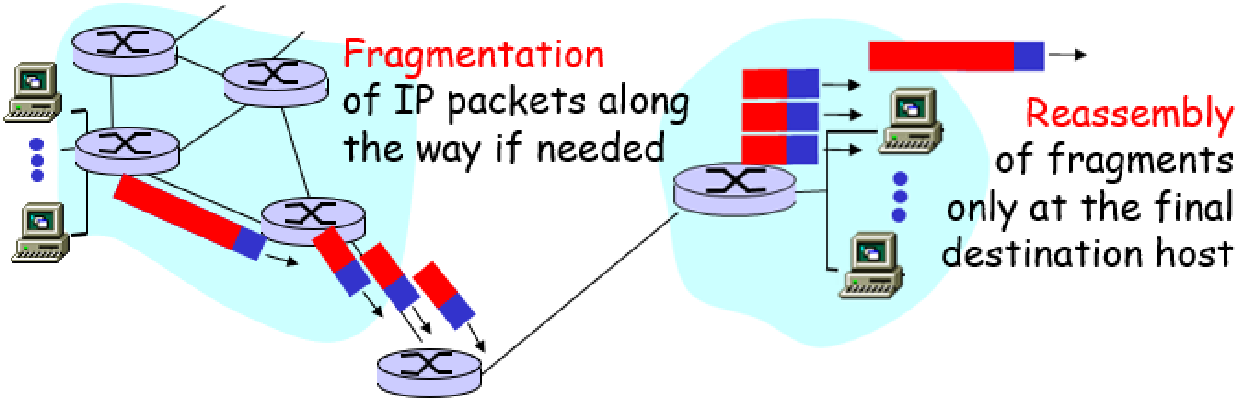


Ver	Traffic Class	Flow Label
Payload Length	Next Header	Hop Limit
Source Address (128 bits)		
Destination Address (128 bits)		
Data		

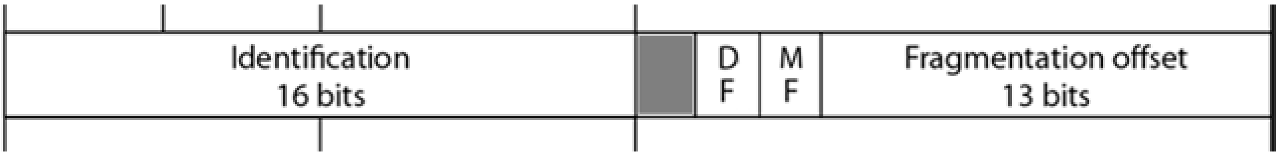
- Also 20 bytes
- IPv4 and IPv6 are not compatible, only `ver` field are the same to distinguish them

4. Fragmentation and Reassembly

- Different networks have different MTU, size of IP packet in each network must be $\leq$ MTU of that network; done by fragmenting packets into smaller packets and then reassembling fragmented packets at final destination host

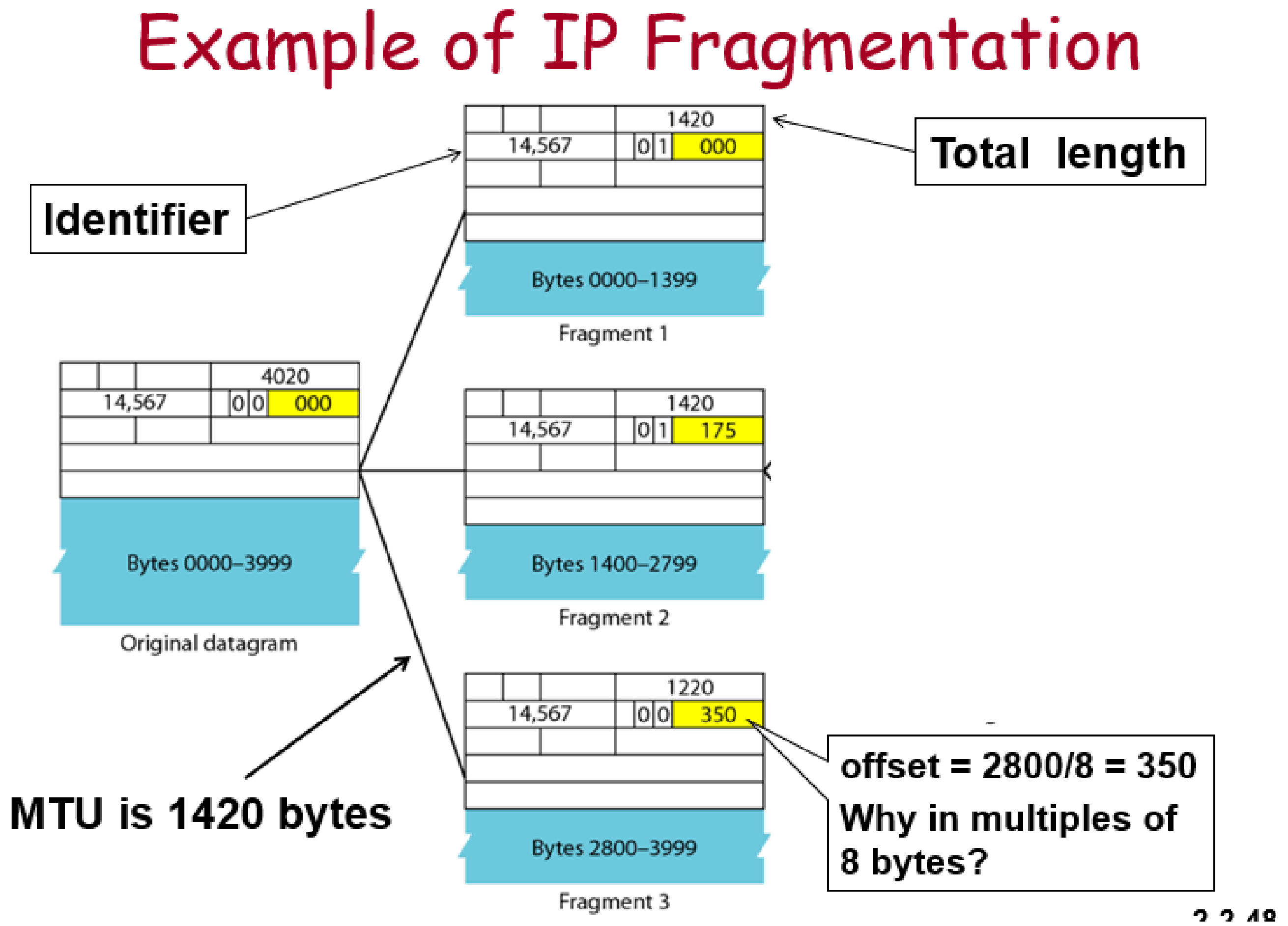


4.1 IPv4 Header Fragmentation Fields



Identification 16 bits		DF	MF	Fragmentation offset 13 bits

- Identification - for reassembly purpose; all fragments of a datagram contain same identification value
- DF (don't fragment) - datagram not fragmented if flag is set
- MF (more fragment) - set as 1 for all fragments except last
- Fragmentation offset - tells the position of fragment in the datagram, all fragments except last one must be a multiple of 8 bytes



Example of IP Fragmentation

Identifier

Total length

MTU is 1420 bytes

offset = 2800/8 = 350
Why in multiples of 8 bytes?

  - 175 is the position of the first segment inside the fragment, given by 1400/8, similar for 350
  - 1400 bytes for first fragment, 1400 bytes for second fragment, 1200 bytes for third fragment
- **IP headers are always 20 bytes and are included in the size of MTU; ICMP headers 8 bytes**

5. IP Routing

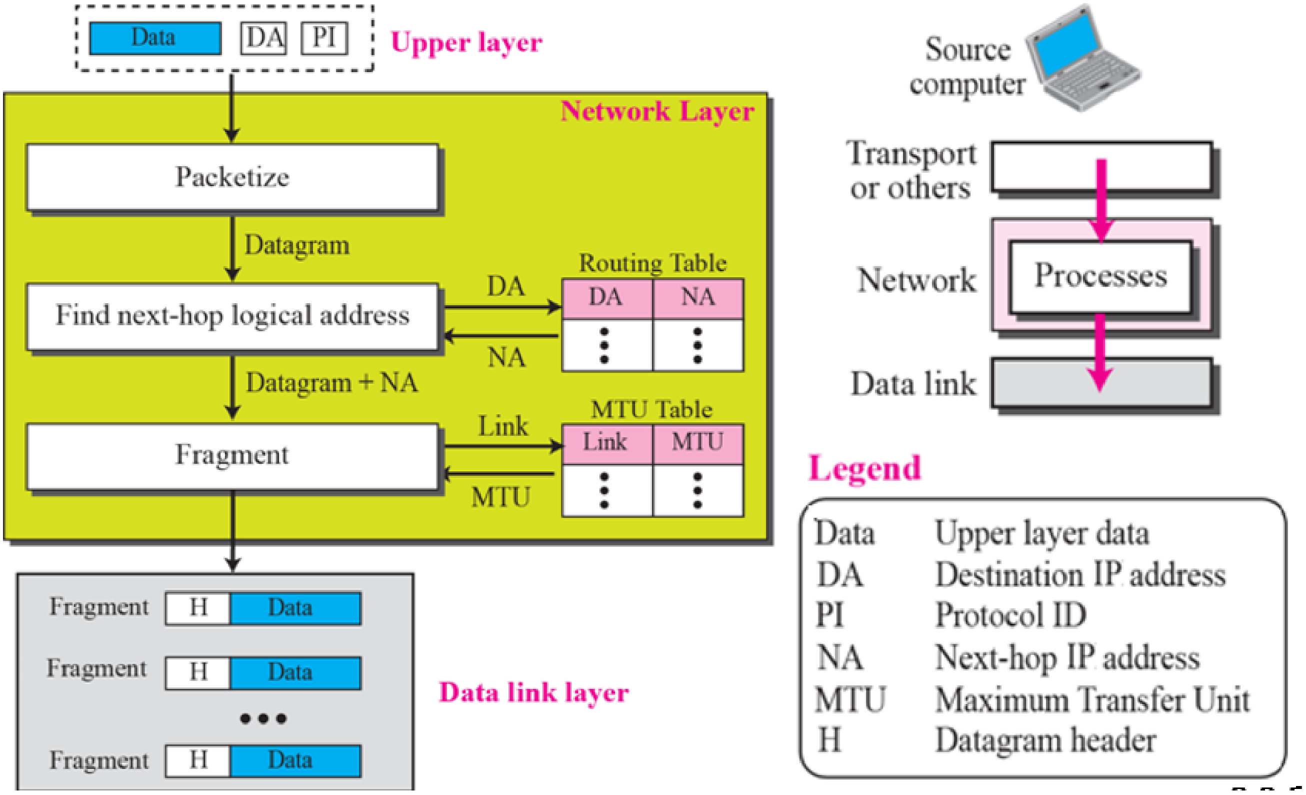
5.1 Routing Process

- Hosts usually do not know how to send packets to destinations outside its own network, uses a router (default gateway) to assist with forwarding
- Each router has a routing table consisting of 3 columns:
  - Network address (destination)
  - Cost (number of hops)
  - Next hop (who to pass to next)
- Router will look up destination address and proceed down the chain
  - Destination IP is directly connected, send on the specified interface
  - Destination IP appear as host specific route, route data as specified
  - Destination Network address is in routing table, send to specified IP address through specified interface

  - Send through default route (usually points to ISP router)

Details

- At source host, IP encapsulates upper layer data into packet, then determine route and MTU, fragment if needed



- At each router, datagram is first checked if valid (discard if not). Then IP determines next route and MTU, further fragmenting packets if necessary.
- At destination, IP reassembles packets if fragmented before returning data to upper layer

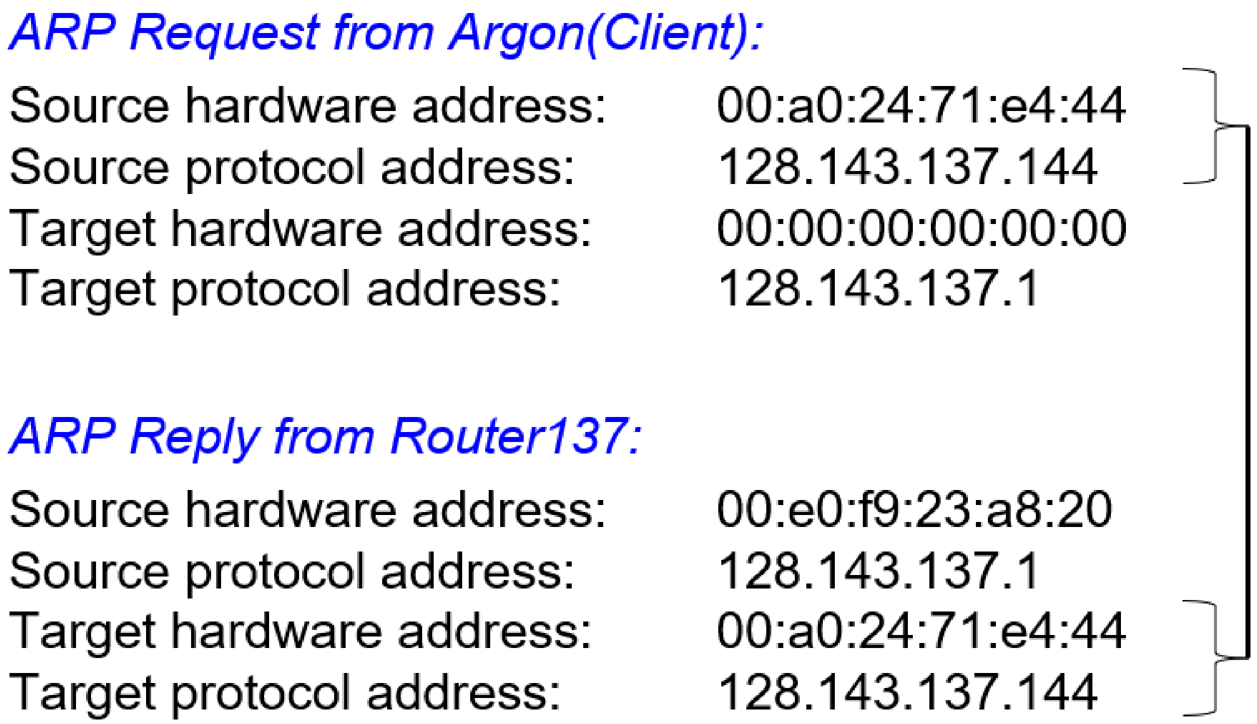
5.2 IP over Ethernet

- Ethernet requires MAC address, but only IP address is known most of the time

Address resolution protocol (ARP)

- An ARP request is when host broadcasts to all stations on the network asking for MAC address of an IP (e.g. router A).

- Router A would respond with an ARP reply which contains its own MAC address



ARP Request from Argon(Client):

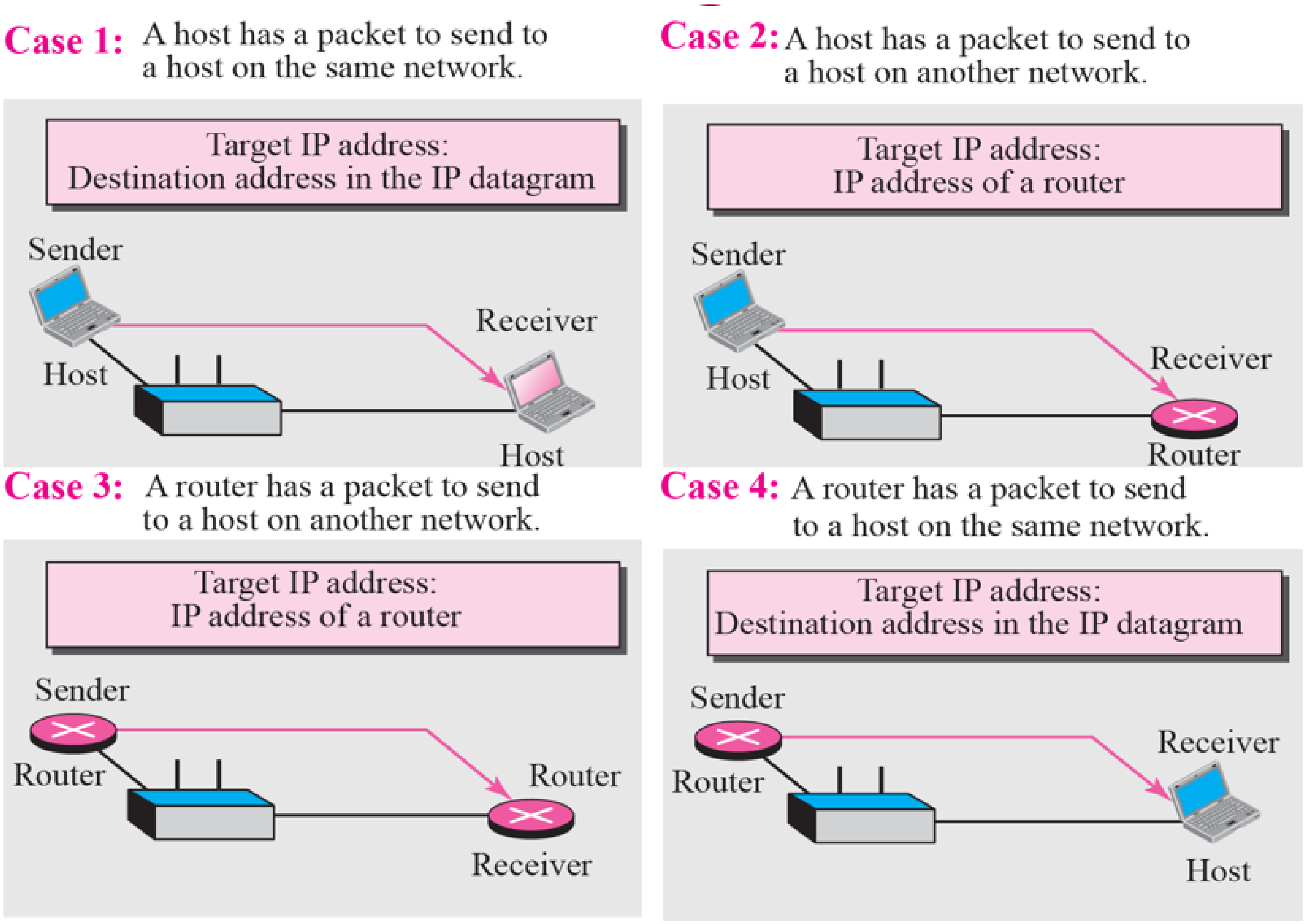
Source hardware address:	00:a0:24:71:e4:44
Source protocol address:	128.143.137.144
Target hardware address:	00:00:00:00:00:00
Target protocol address:	128.143.137.1

ARP Reply from Router137:

Source hardware address:	00:e0:f9:23:a8:20
Source protocol address:	128.143.137.1
Target hardware address:	00:a0:24:71:e4:44
Target protocol address:	128.143.137.144

  - Target = Router A, host does not know Router A's MAC address so it sends `00:00:00:00:00:00` which Router A will replace with a valid address

- 4 cases


Case 1: A host has a packet to send to a host on the same network.
Target IP address: Destination address in the IP datagram
Sender
Host
Receiver
Host
Case 2: A host has a packet to send to a host on another network.
Target IP address: IP address of a router
Sender
Host
Receiver
Router
Case 3: A router has a packet to send to a host on another network.
Target IP address: IP address of a router
Sender
Router
Router
Receiver
Case 4: A router has a packet to send to a host on the same network.
Target IP address: Destination address in the IP datagram
Sender
Router
Receiver
Host

- Sending ARP request / reply for each IP datagram is inefficient → host will maintain an ARP cache of current entries which are expired after 2-20mins according to configuration

6. IP Related Protocols

6.1 ICMP

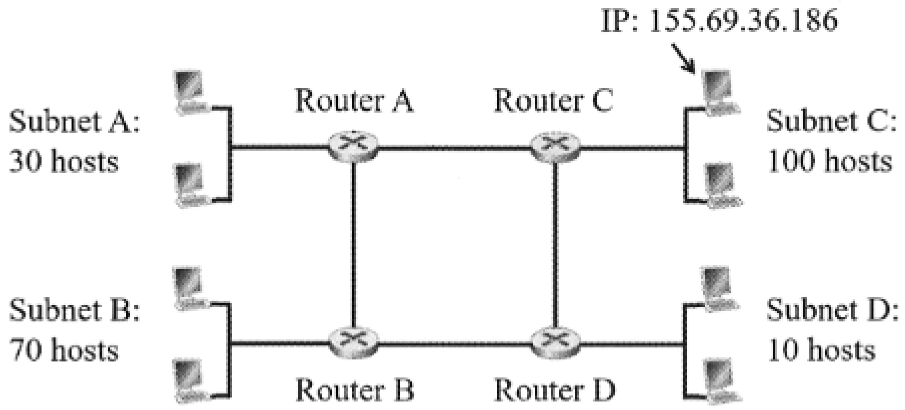
- Used by network devices to send error messages and operational information indicating that a requested service is not available / host (or router) cannot be reached
- ICMP packet is sent over IP packet, which is sent over data link layer protocol (Layer-3 / Ethernet)
- **Echo Request**
    - Sent by source host to query a router / destination host which will respond with ICMP Echo reply message
- **Ping**
    - Debugging tool for testing reachability of a host / router
- **Error message**
    - Sent by router / destination host to inform source host that their datagram has been received in error and discarded
- **Port unreachable**
    - Sent by Server when a client requests for a service from server at a specific port but there is no process waiting at port 80
- **Time to Live**
    - Field in IP header
    - Counter used to limit the lifespan of IP datagram, checked at every router, decremented at every hop
    - When value reaches 0, router discards IP datagram

6.2 Tracert

- Debugging tool used for tracing a path from source to destination host, operated by sending sequence of ICMP echo request over IP with Time to Live set to 1,2,... until destination is reached
- **Source address field of the IP that carries the ICMP message is the address of the interface that sends it**

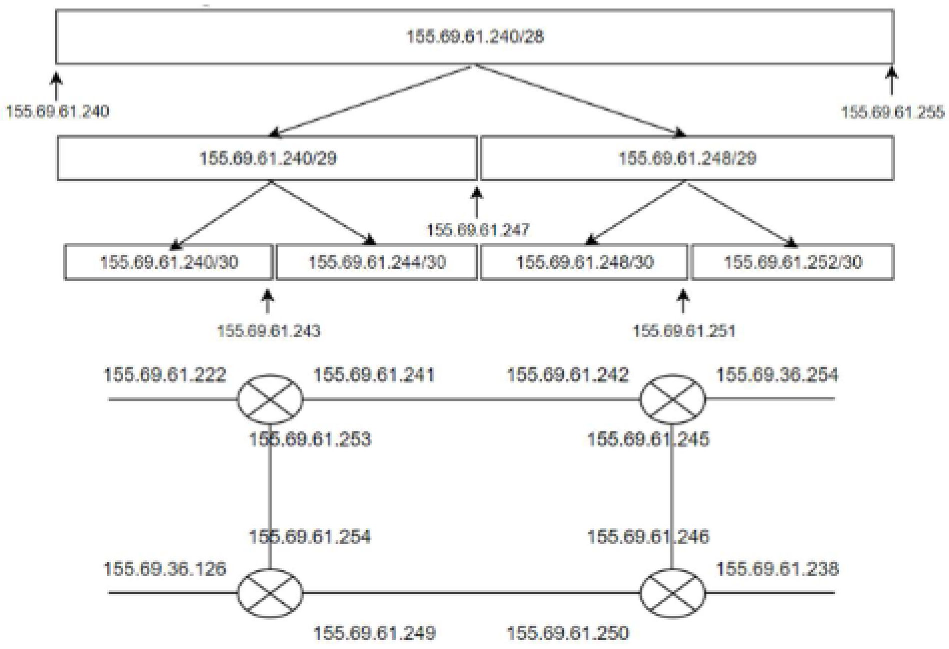
7 Questions

1. Subnet allocation

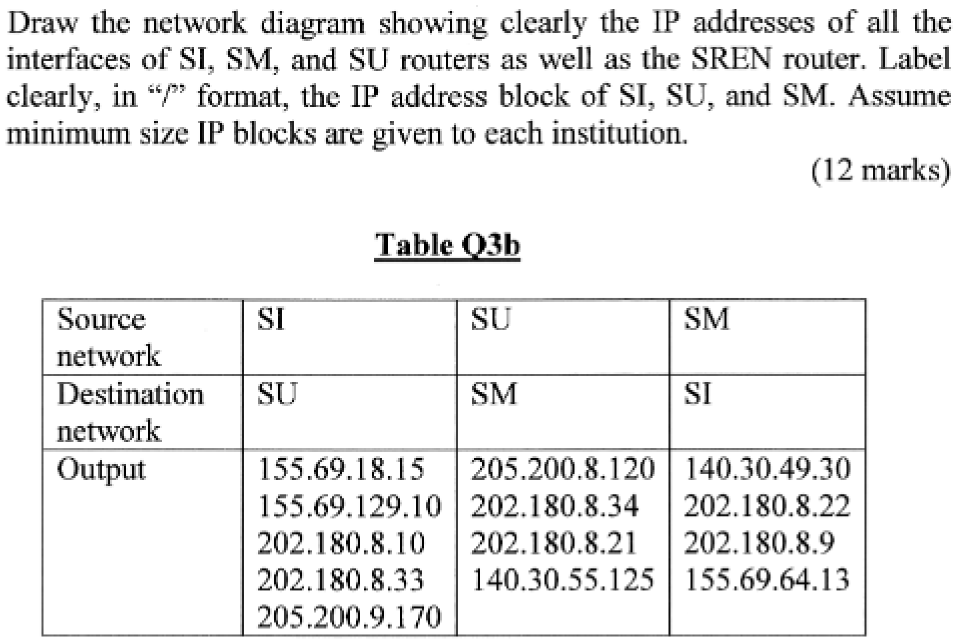


Given two blocks of IP addresses, allocate suitable blocks to each subnet and assign suitable IP for each router. Blocks: 155.69.36.0/24 and 155.69.61.192/26.

- Usually need to satisfy constraints like the host IP in subnet C
- First find number of bits and corresponding mask required for each subnet
    - A: 5 / 27, B: 7 / 25, C: 7 / 25, D: 4 / 28
- Satisfy conditions, C needs last 7 bits to accommodate host, just nice satisfied by 155.69.36.128/25, ip range:155.69.36.128 - 155.69.36.255
    - Remaining blocks are 155.69.36.0 - 155.69.36.127 and 155.69.61.192/26
- Start from the biggest subnet (B) only first block can satisfy so B is assigned subnet 155.69.36.0/25, ip range: 155.69.36.0 - 155.69.36.127
- Subnet A gets allocated 155.69.61.192/27, range: 155.69.61.192 – 155.69.61.223
- Subnet D gets allocated 155.69.61.224/28, range: – 155.69.61.224 - 155.69.61.239
- **Masks assigned to subnets are determined by the size of the subnet, not the initial mask of the block / anything else**
- To assign ip addresses to router, take the remaining IP block and partition it into n blocks of 4 ip addresses, first and last are reserved, second and third are used to connect two routers together.



2. Draw network diagram



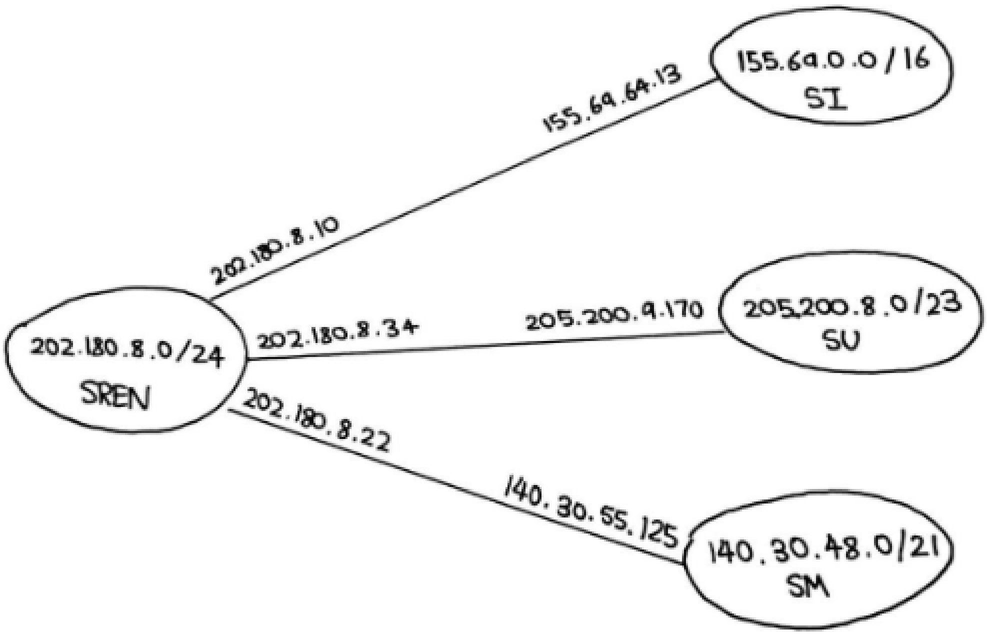
Draw the network diagram showing clearly the IP addresses of all the interfaces of SI, SM, and SU routers as well as the SREN router. Label clearly, in "/" format, the IP address block of SI, SU, and SM. Assume minimum size IP blocks are given to each institution.

(12 marks)

Table Q3b

Source network	SI	SU	SM
Destination network	SU	SM	SI
Output	155.69.18.15 155.69.129.10 202.180.8.10 202.180.8.33 205.200.9.170	205.200.8.120 202.180.8.34 202.180.8.21 140.30.55.125	140.30.49.30 202.180.8.22 202.180.8.9 155.69.64.13

- Find block allocated to each network, by comparing IP addresses + masks
- Label gateways as the IP addresses used for entry into the network



- 155.69.64.13 is used to enter SI block, 202.180.8.10 is used to enter SREN block etc.