Computer Networks

Network Layer - Internet Protocol

Nassim Kobeissy - nassim.kobeissy@gmail.com

The Network Layer

- Role
  - Getting packets from the source all the way to the destination!!!
- Challenges
  - Many hops contrarily to the data link layer
  - Must know about the topology of the network
  - Choose appropriate paths
  - Offers addressing over heterogeneous networks!

Service Provided to Transport Layer

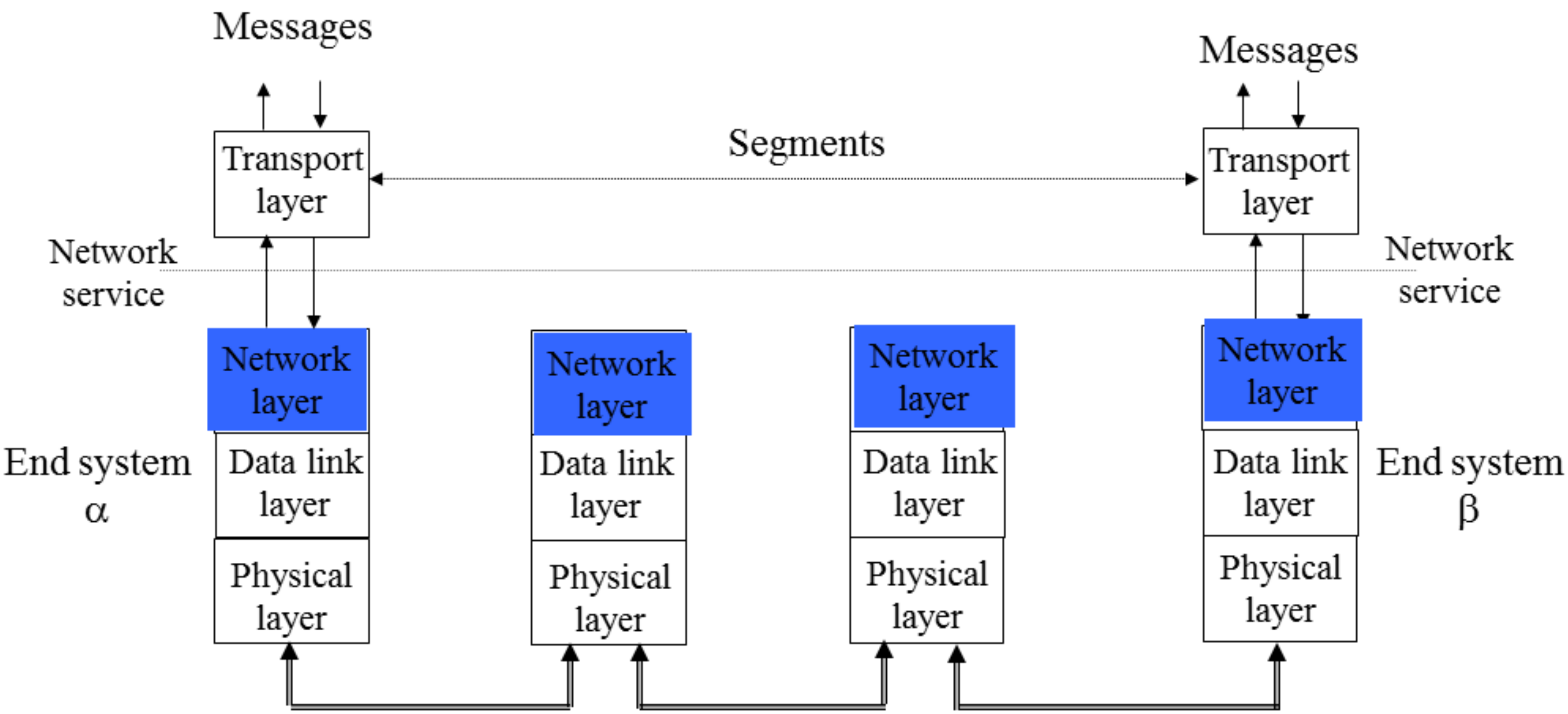
Goals:

1. The services provided by the network layer should be independent of the subnet topology.
2. The Transport Layer should be shielded from the number, type and topology of the subnets present.
3. The network addresses available to the Transport Layer should use a uniform numbering plan (even across LANs and WANs).

Services:

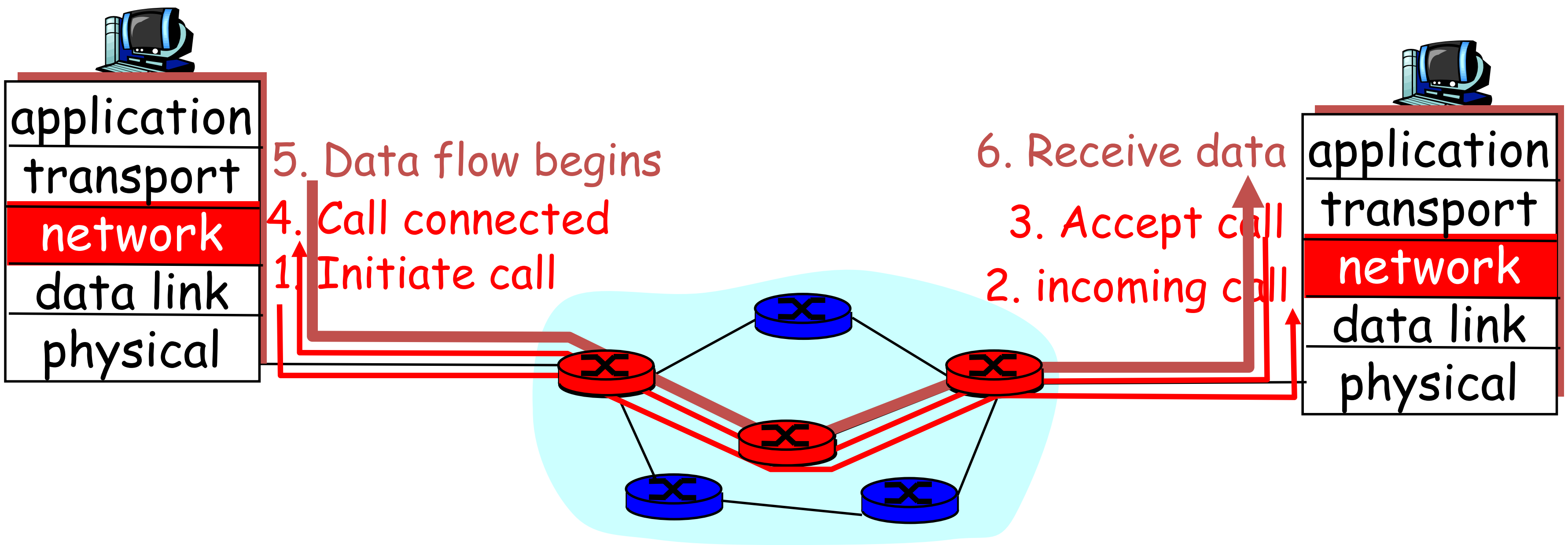
- Reliable Connection-Oriented -> virtual-circuit subnet
  - Connection establishment before communication
- Connectionless -> diagram subnet
  - Just forward packets and forget about them
  - Internet community backing it

Service Provided to Transport Layer



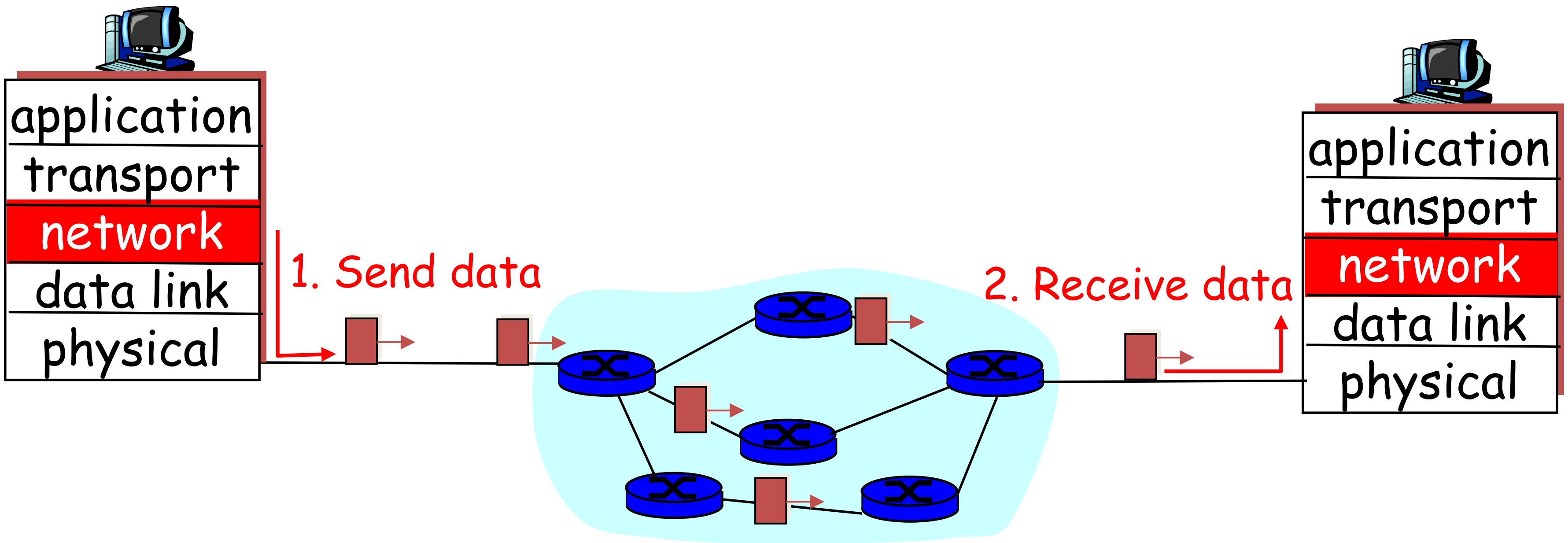
Virtual circuits: signaling protocols

- used in ATM, frame-relay, X.25
- not used in today's Internet



Datagram networks

- no call setup at network layer
- routers: no state about end-to-end connections
  - no network-level concept of “connection”
- packets typically routed using destination host ID
  - packets between same source-dest pair may take different paths
- the Internet model



The Internet Protocol (IP)
IPV4 Datagrams Format

IPV4 Datagrams Format

Version = 4

IHL: **I**nternet **H**eader **L**ength

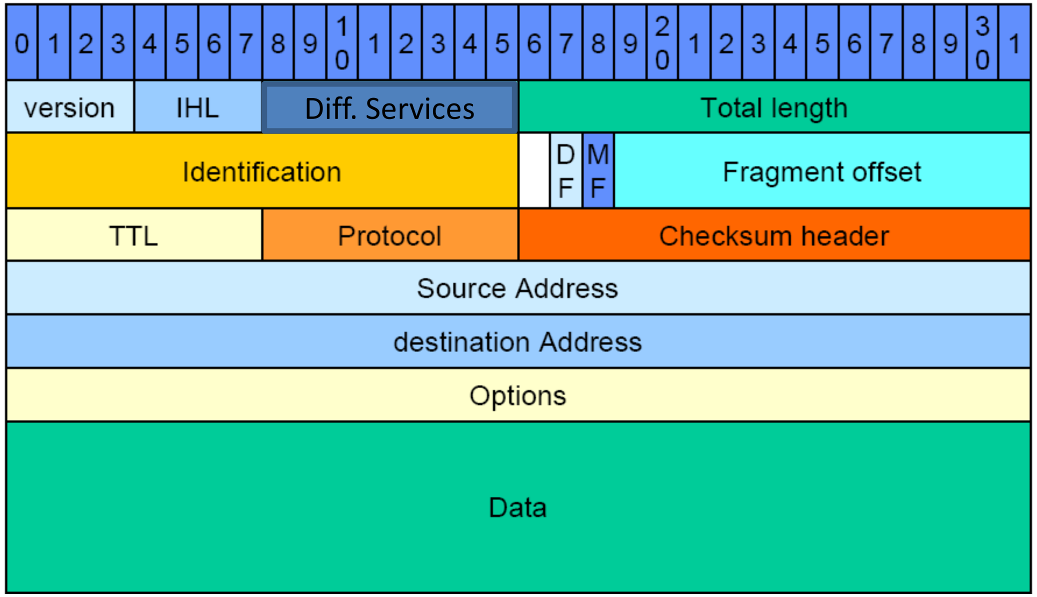
- IPv4 datagram can contain a variable number of Options
- Most IP datagrams do not contain options, so the typical IP datagram has a 20-byte header.

Differentiated Services

- Or **TOS** (Type of Service)
- low delay, high throughput, reliability

Total Length

- Includes Header and Data
- In number of bytes
- The maximum theoretical length is 65535 bytes.
- Datagrams are rarely larger than 1,500 bytes



IPV4 Datagrams Format

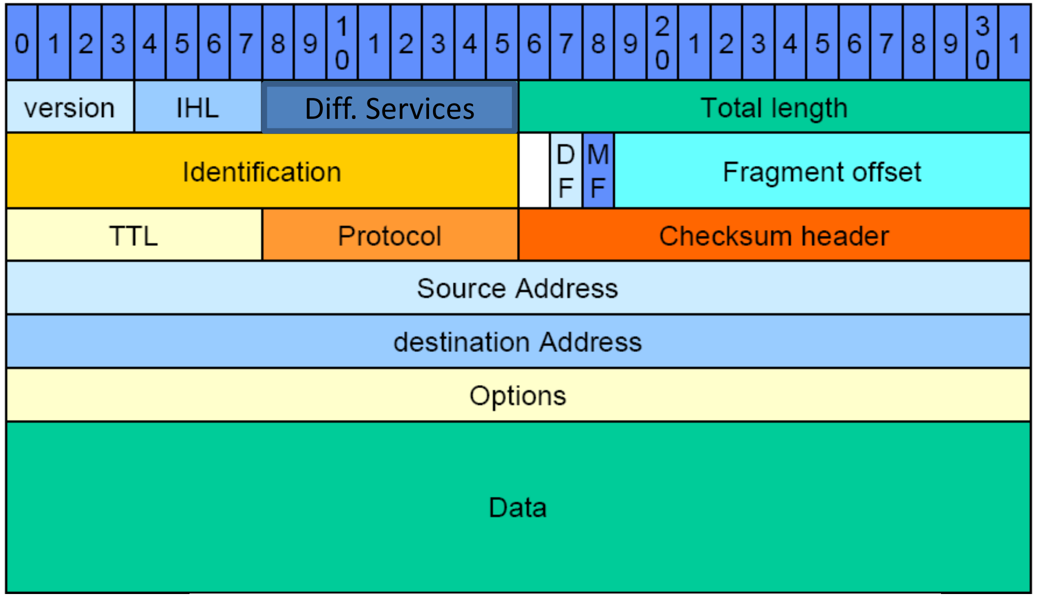
Identification
- All fragments of the same packet have the same ID

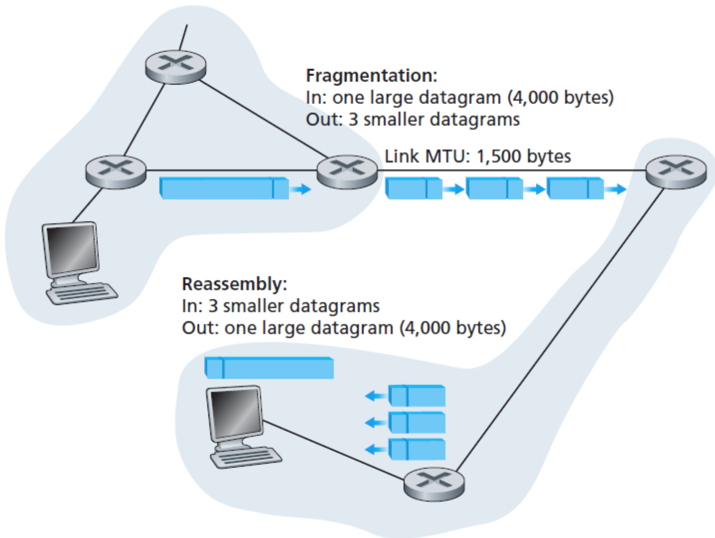
Unused bit
- There had been a proposal to use it to detect malicious traffic
- Not adopted ☹

DF: Don’t Fragment
- Used especially to discover the path **MTU** (Maximum Transmission Unit)
- Either the packet arrives in peace or an error message is returned to the sender

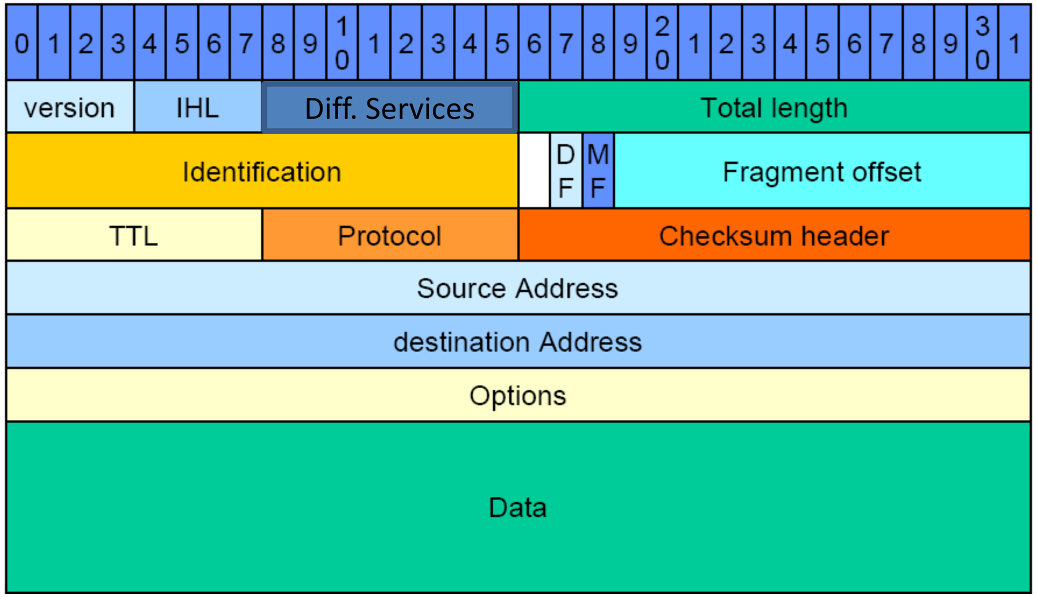
MF: More Fragments
- It is set to 1 in All fragments except the last one

Fragment offset
- Indicates the location of the fragment in the current packet
- All fragments except the last one are multiple of 8 bytes
- Maximum number of fragments per datagram is 8192





IPV4 Datagrams Format



TTL: Time to Live

- Is decremented each time a router forwards the packet
- When it is Zero, it is discarded and a message is returned to the sender
- It prevents a packet to circulate forever in a network especially in cases of corrupted routing tables

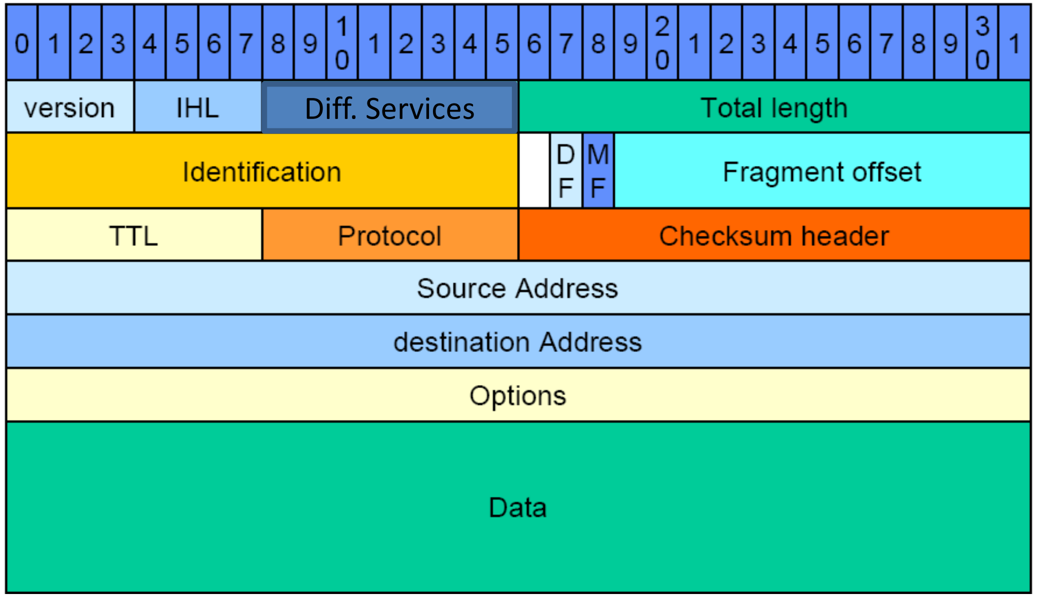
Protocol

- Transport protocol (TCP = 6 /UDP = 17)
- It tells the network layer in which transport process to give the packet to.
- **RFC 1700** for protocol number
- Or: www.iana.org

Checksum header

- Error detection
- Calculated over header only
- 16 bits recalculated by router because of TTL change

IPV4 Datagrams Format



Options:

Security (130: 1 000 0010):

- specifies how secret the datagram is
- Rarely used

Strict source routing:

- Gives the complete path to be followed

Loose source routing:

- Gives a list of router not to be missed
- Ex: From London to Sidney, go west to NYC and Honolulu

Record Route (7: 0000 0111):

- Makes each router appends its IP address
- Was good for tracking anomalies
- Now; very limited since only 9 IPs can be recorded

Timestamp:

- Makes each router appends its IP address + 4B timestamp
- More limited today !!

IPV4 Datagrams Format

Source address

– IP address of the source machine

Destination address

– IP address of the **ultimate** destination machine

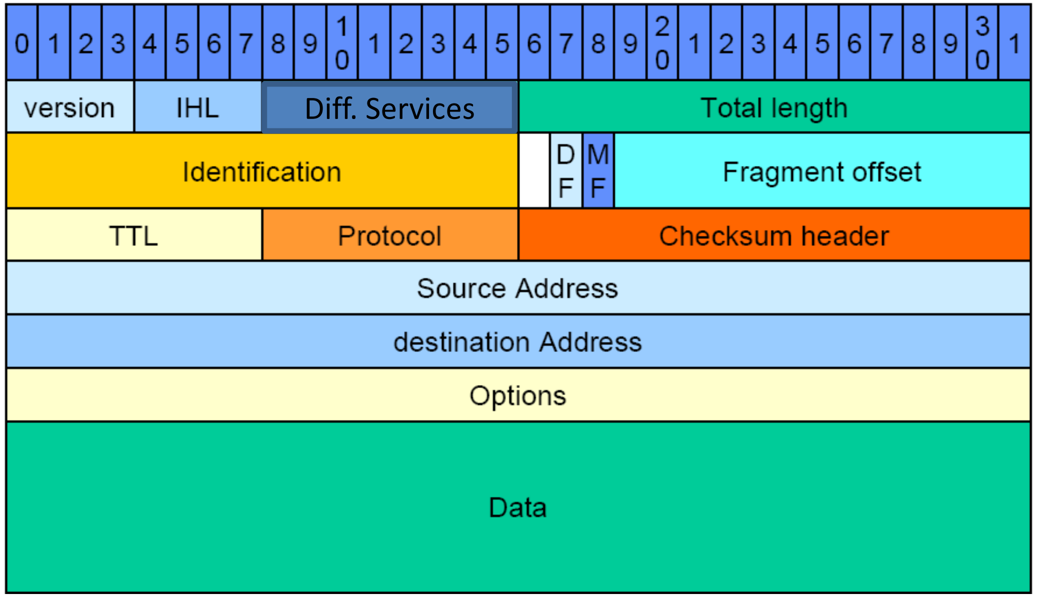
– Often the source host determines the destination address via a DNS lookup (Domain Name System)

Data (payload)

– The **raison d'être** for the datagram

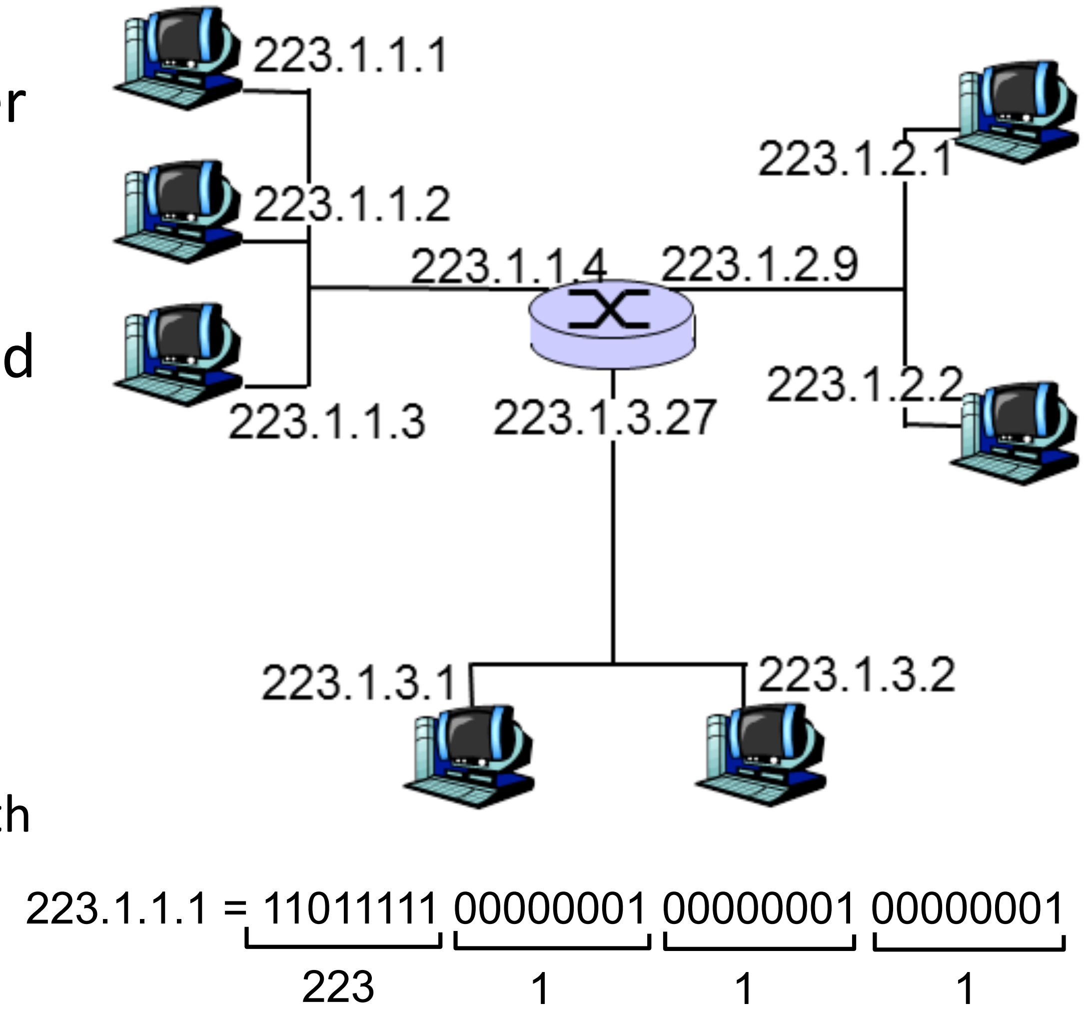
– It mainly contains the transport-layer segment (TCP or UDP)

– Or other types of data, such as ICMP messages



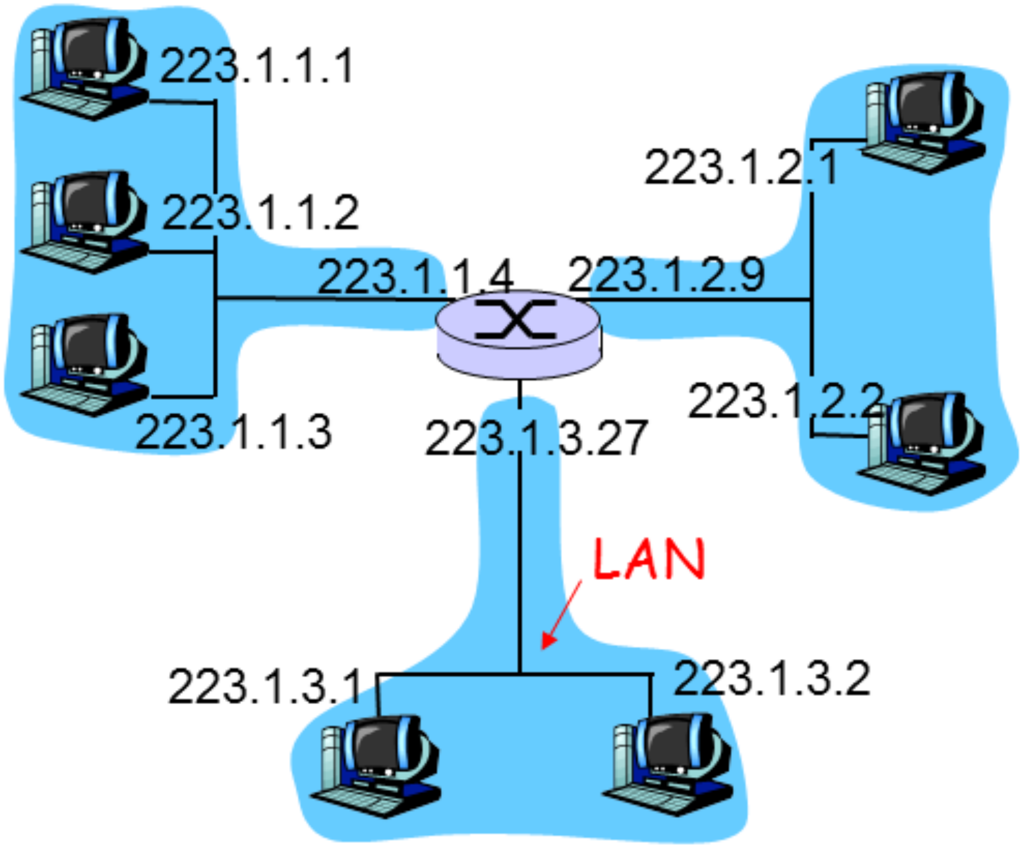
The Internet Protocol (IP)

- IP address: 32-bit identifier for host, router *interface*
- *interface:* connection between host/router and physical link
  - router's typically have multiple interfaces
  - host may have multiple interfaces
  - IP address associated with each interface



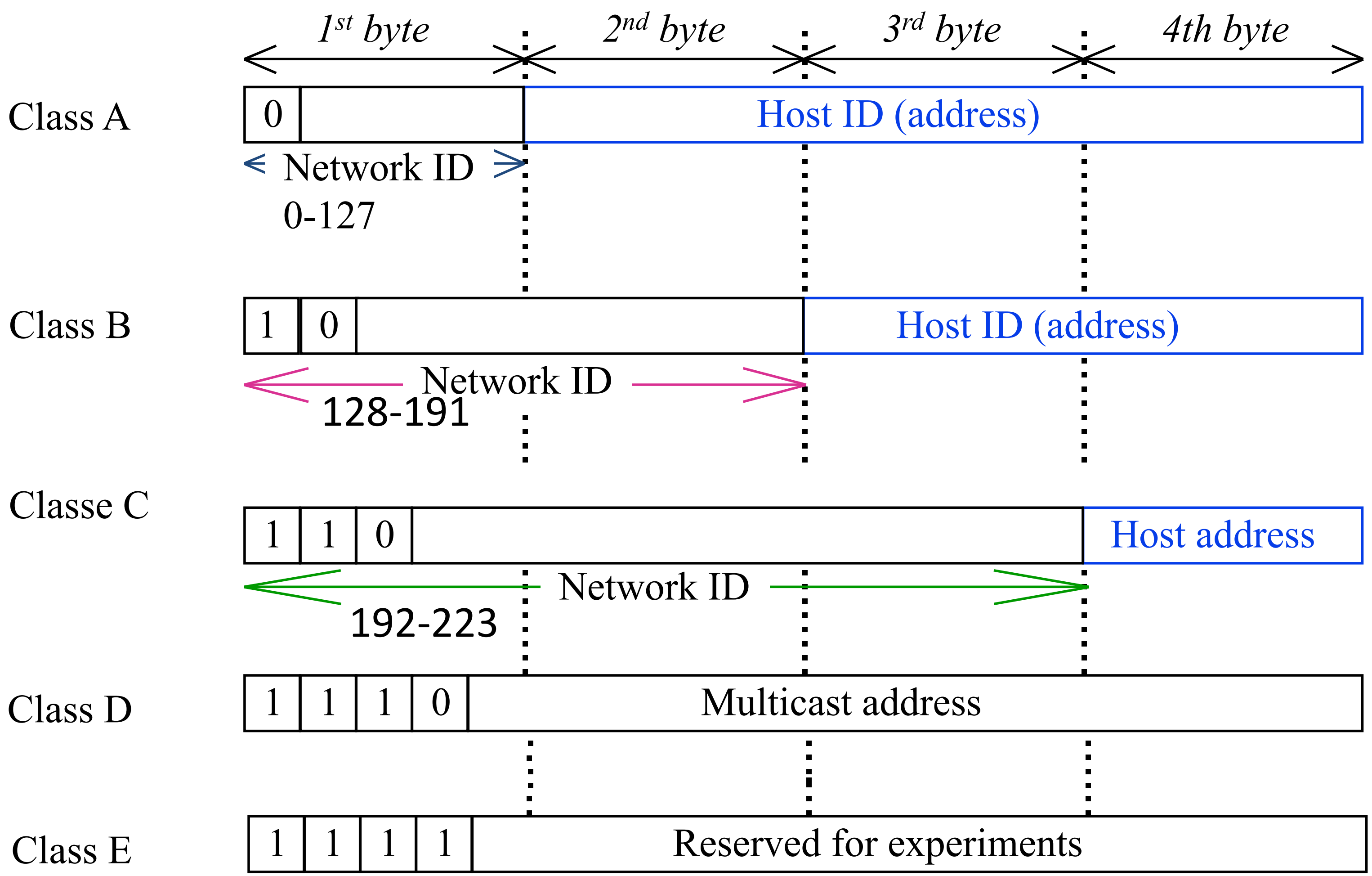
The Internet Protocol (IP)

- IP address:
  - network part (high order bits)
  - host part (low order bits)
- *What's a network ?* (from IP address perspective)
  - device interfaces with same network part of IP address
  - can physically reach each other without intervening router



network consisting of 3 IP networks (for IP addresses starting with 223, first 24 bits are network address)

IP V4 Address classful addressing



Example : 181.3.12.35

Reserved and Private networks

- **0.0.0.0 & 127.0.0.0 are reserved**
- No two machines that connect to a public network can have the same IP address because public IP addresses are global and standardized.
- However, private networks that are not connected to the Internet may use any host addresses, as long as each host within the private network is unique.
- RFC 1918 sets aside three blocks of IP addresses for private, internal use.
- Connecting a network using private addresses to the Internet requires translation of the private addresses to public addresses

Reserved and Private networks

- RFC 1918 had reserved the following IP addresses for private use:

Class	Addresses
A	10.0.0.0 – 10.255.255.255
B	172.16.0.0 – 172.31.255.255
C	192.168.0.0 – 192.168.255.255

IP addressing: CIDR

- Classful addressing:
  - inefficient use of address space, address space exhaustion
  - e.g., class B net allocated enough addresses for 65K hosts, even if only 2K hosts in that network
- CIDR: Classless InterDomain Routing
  - network portion of address of arbitrary length
  - address format: a.b.c.d/x, where x is # bits in network portion of address

network part

host part

11001000 00010111 00010000 00000000

200.23.16.0/23

CIDR : Classless Inter-Domain Routing

- We need to manage subnets and supernets as well
- Supernets are needed when route aggregation is used especially to **reduce routing tables**
- With subnet mask: 182.128.0.0/255.255.0.0
- With CIDR notation: 182.128.0.0/16
  - 16 bits (of the subnet mask) are set to 1 from the left:
- Subnet example:
  - 182.33.200.35/255.255.240.0
  - Or 182.33.200.35/20
- Supernet example
  - 182.128.0.0/9 Subnet mask=255.128.0.0
    - network addresses from 182.128.0.1 to 182.255.255.254

CIDR : Classless Inter-Domain Routing

- Class C addresses are divided into 4 categories, each corresponding to a continent

194.0.0.0 - 195.255.255.255	Europe
198.0.0.0 - 199.255.255.255	North America
200.0.0.0 - 201.255.255.255	South America
202.0.0.0 - 203.255.255.255	Asia (Pacific)

- Each address 194.x.y.z is routed towards European router

Subnetting

- Subnetting means dividing a given network into two or more sub-networks (or simply **subnets**)
- This means the network part in a layer 3 IP address is extended.

The network ID increased by adding more bits from the host space [2]

Subnets

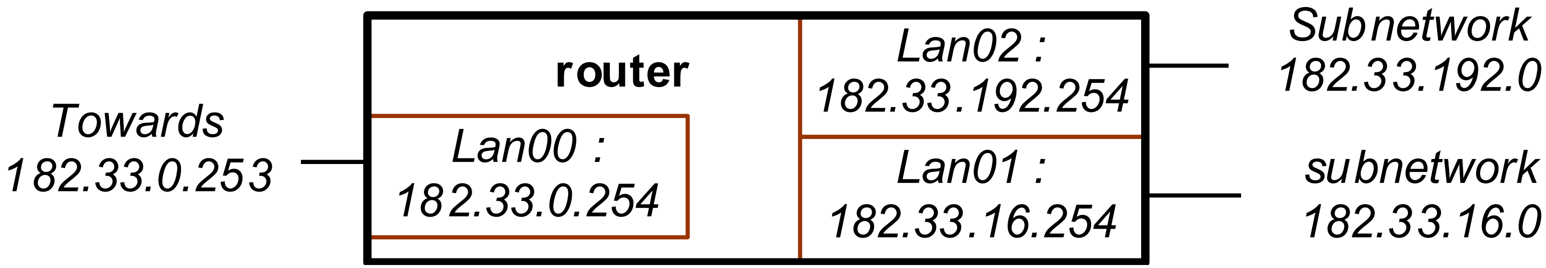
- Problem :
  - Suppose that a university is given a network address of class B
  - The university has 16 big departments and it wants to attribute an address number to each of them
- Solution : *Subnetting*
  - Needs information to tell about the sub-networks addresses

  → **use subnet mask (put 1 the network bits and zero for host part)**
- <u>Example</u> : Network 183.13.0.0
  - 4 bits shall be added from the host part to the network part to be able to represent 16 new sub network.
  - Subnet mask: 11111111 . 11111111 . 1111 0000 . 0000 0000
255 . 255 . 240 . 0
  - The address 183 . 13 . 200 . 35
Belongs to 183 . 13 . XXX . 0

Subnets

- The initial network address is
  - 10110111 . 00001101 . 0000 0000 . 0000 0000 (183. 13 . 0 . 0)
- The 16 obtained sub-networks are:
  - 10110111 . 00001101 . 0001 0000 . 0000 0000 (183. 13 . 16 . 0)
  - 10110111 . 00001101 . 0010 0000 . 0000 0000 (183. 13 . 32 . 0)
  - 10110111 . 00001101 . 0011 0000 . 0000 0000 (183. 13 . 48 . 0)
  - 10110111 . 00001101 . 0100 0000 . 0000 0000 (183. 13 . 64 . 0)
  - 10110111 . 00001101 . 0101 0000 . 0000 0000 (183. 13 . 80 . 0)
  - 10110111 . 00001101 . 0110 0000 . 0000 0000 (183. 13 . 96 . 0)
  - 10110111 . 00001101 . 0111 0000 . 0000 0000 (183. 13 . 112. 0)
  - 10110111 . 00001101 . 1000 0000 . 0000 0000 (183. 13 . 128. 0)
  - 10110111 . 00001101 . 1001 0000 . 0000 0000 (183. 13 . 144. 0)
  - 10110111 . 00001101 . 1010 0000 . 0000 0000 (183. 13 . 160. 0)
  - 10110111 . 00001101 . 1011 0000 . 0000 0000 (183. 13 . 176. 0)
  - 10110111 . 00001101 . 1100 0000 . 0000 0000 (183. 13 . 192. 0)
  - 10110111 . 00001101 . 1101 0000 . 0000 0000 (183. 13 . 208. 0)
  - 10110111 . 00001101 . 1110 0000 . 0000 0000 (183. 13 . 224. 0)
  - 10110111 . 00001101 . 1111 0000 . 0000 0000 (183. 13 . 240. 0)

Routing tables with Subnetting



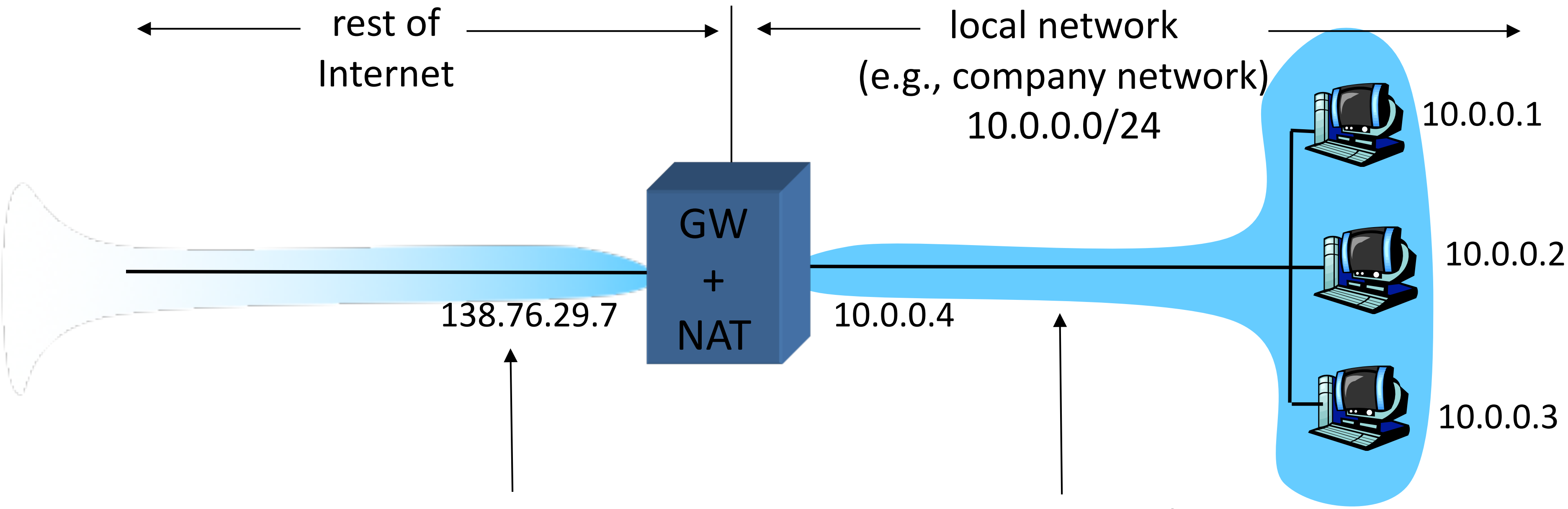
Destination	Subnet mask	Interface
127.0.0.0	255.0.0.0	lo
182.33.0.0	255.255.240.0	Lan00
182.33.16.0	255.255.240.0	Lan01
182.33.192.0	255.255.240.0	Lan02
default	0.0.0.0	Lan00

Network Address Translation

Scalability Problem

- Internet growing very fast
  - People over the world are more than 2^{32}.
    - How many device per user?
    - Tens of billions of devices >> 2^{32}
    - Each device needs an address for communication
    - How do you address each of them?
    - IP par device?
  - IP addressing provides up to 2^{32} different address
    - Are **not enough**

NAT: Network Address Translation



All datagrams *leaving* local network have same single source NAT IP address: 138.76.29.7, different source port numbers

Datagrams with source or destination in this network have 10.0.0.x/24 address for source, destination (as usual)

Internet Control Message Protocol ICMP

- Report errors and unexpected behavior and to test the Internet
- Networks troubleshooting

Message Type	Description
Destination Unreachable	Router can not locate destination. A packet with DF = 1 can not be delivered
Time Exceeded	TTL=0 Used in **Traceroute** tool to discover paths between sources and destinations
Parameter Problem	Invalid header field
Echo and Echo Reply	Check if a machine is alive Used by the **Ping** utility
Timestamp Request/Reply	Same as echo with timestamp (departure and arrival times recorded) Used to measure network performances

Dynamic Host Configuration Protocol DHCP

- Almost every network has a DHCP server to configure IP addresses of different host automatically.
- RFC 2131 and 2132
- Any host has a MAC address but not an IP.
- To obtain a IP, a machine must know the DHCP server on its connected network
  - For this purpose, it broadcasts a **DHCP DISCOVER (Ethernet Address)**
- The DHCP server allocates a free IP address and sends a **DHCP OFFER**
- If the DHCP server is not on the host's network, the router shall be configured to relay discovers and offers
- **Leasing**
  - is used to assign IPs to fixed durations to prevent losing IP addresses (that are no longer used)
  - The host must ask for a DHCP renewal before the lease expires.

DHCP Handshake

- DHCPDISCOVER **with**
  - **address 0.0.0.0 as source address**
  - 255.255.255.255 as destination address.
- DHCPOFFER
  - **sent by any DHCP server**
- DHCPREQUEST
  - **The client accepts the first proposal**
- DHCPACK
  - **sent by the DHCP server**

Address Resolution Protocol ARP

- Network Interface Cards (NICs) do not understand IPs.
- An Ethernet NIC sends and receives based on MAC addresses (48 bits) and does not know anything about IP.
- ARP helps hosts to find the MAC addresses of destinations
- Broadcasts are used on the Ethernet asking for the MAC address on an IP-known host

ARP Handshake:

Host A needs to learn the MAC address of host B(The IP of B is: x.y.z.t)

✓ A broadcasts an **ARP Request**

What is the hardware address (MAC address) of B (x.y.z.t)?"

✓ Host B replies with **ARP Reply**

This is my hardware address (MAC address)

ARP

- Proxy ARP:
  - It is a host or a router that responds to **ARP Requests** that arrive from one of its connected networks for other hosts on different connected networks.
  - **Request by A:** What is the MAC address of B (x.y.z.t)
  - The Proxy knows that B is on another network
  - **Reply by the proxy:** The MAC address of B(x.y.z.t) is X where X is the proxy’s MAC address.

IP Version 6

IPv6 Goals

- Support billions of hosts
- Reduce routing table sizes
- Simplify protocol to allow routers to process packets faster
- Better security (authentication and privacy)
- More attention to type of service
- Enhance multicasting
- Support mobility (roaming without changing its address)
- Allow future protocol evolution
- Permit the old and new protocols to exist together

IPv6 History

- IN 1990 IETF started working on IPv6.
- One proposal was to use CLNP (OSI network protocol). It uses 160 bits addresses
- IN 1993, two main proposals were merged resulting in SIPP (Simple Internet Protocol Plus)…. Now called IPv6
- In general IPv6 is not compatible with IPv4 but with other internet protocols like TCP, UDP, BGP, ICMP, OSPF, DHCP (Modified)…
- Relevant references: RFC 2460 and RFC 2466
- IPv6 is an Internet standard from 1998

IPv6 Header Format

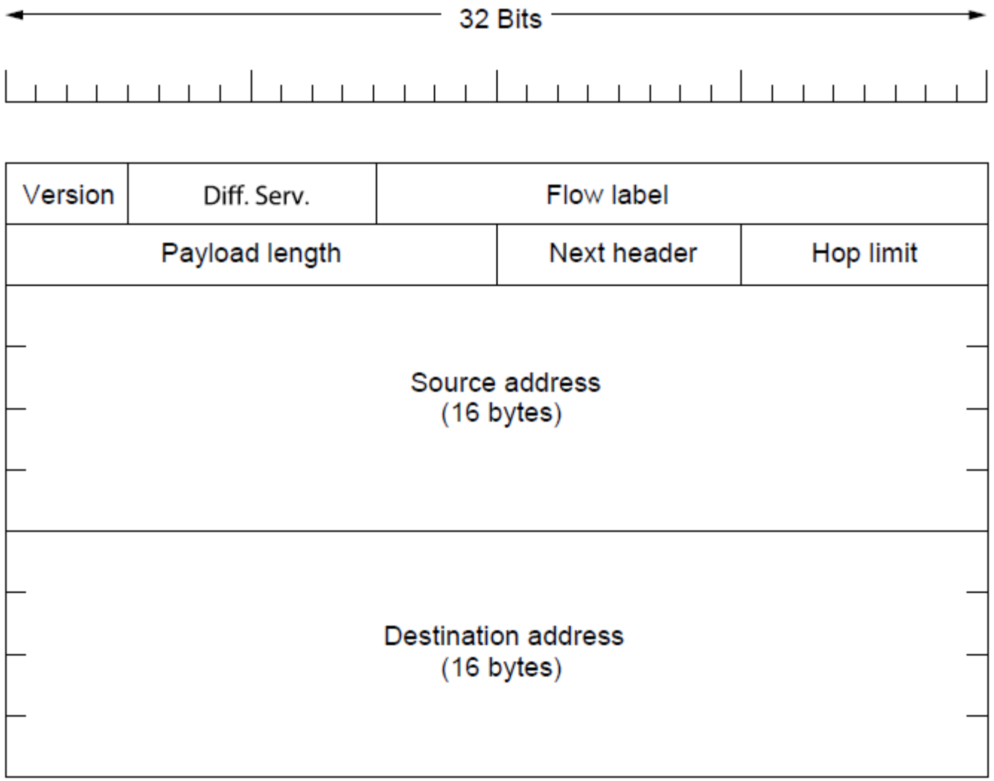
Version = 6

Differentiated Services

- Also called traffic class
- Distinguish the class of service for packets with different real time requirements
- Similar to IPv4
- 2 bits for Early Congestion Notification information

Flow Label

- Used for quality of service QoS
- Mark flows of data (groups of packets with same requirements)
- When non zero, routers look up for special treatment
- A kind of packet switching



IPv6 Header Format

Payload Length

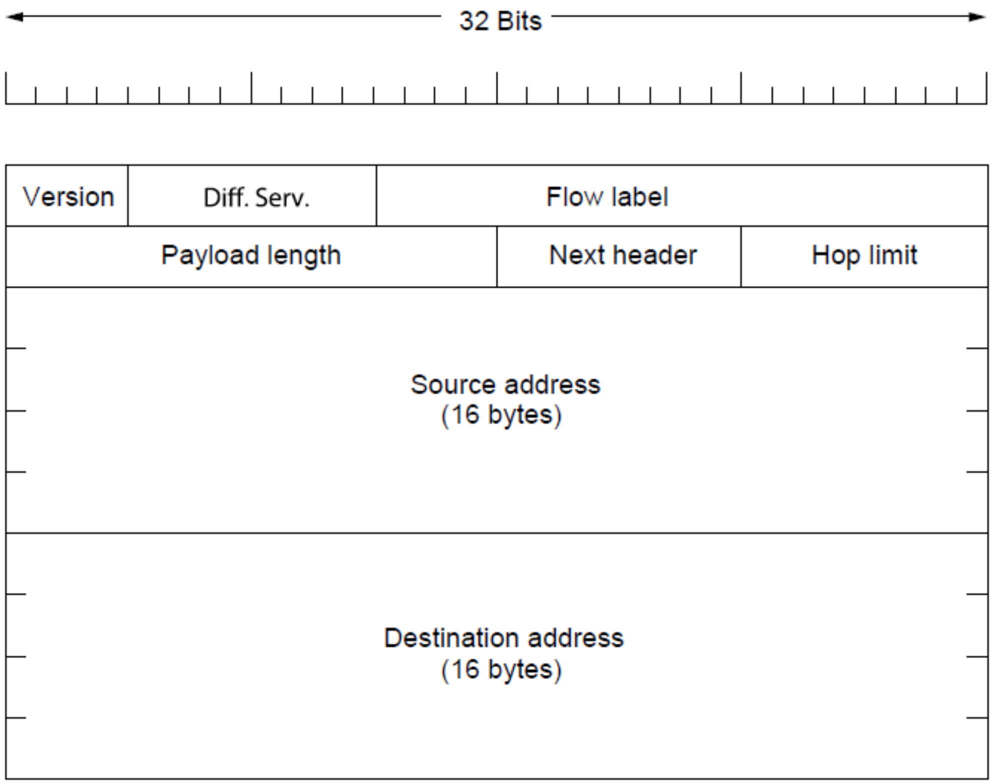
- In bytes
- Does not account for the 40 Byte fixed header length

Next header

- Identifies next header
- Tells which of the extension headers is used
- Otherwise, it identifies the transport protocol
- Provides simplification w.r.t IPv4

Hop Limit

- Same as TTL of IPv4
- Used to avoid that packets live forever



IPv6 Addresses

- $2^{128} \approx 3 * 10^{28}$
- Each meter square on the Earth (including oceans) can have:
  $3 * 10^{13}$ IP
- 8 parts separated by **:**
- Each part represents 16 bits in hexadecimal
  2001 : 0db8 : 0000 : 85a3 : 0000 : 0000 : ac1f : 8001
- Zeros can be regrouped
  2001 : 0db8 : 0: 85a3 : 0 :0 : ac1f : 8001
- Zero groups can be removed
  2001 : 0db8 : 0: 85a3 :: ac1f : 8001
- The double colon may only be used once in an address

IPv6 Addresses

Network addresses

- **CIDR** notation is used to distinguish network addresses

- **2001 : 0db8 : 85a3 ::/48** represents the addresses from
  2001 : 0db8 : 85a3 : 0 :0 : 0 : 0 : 0 to **2001 : 0db8 : 85a3 : ffff : ffff : ffff : ffff : ffff**

- The prefix 2000::/3 represents the addresses from
  2000 : 0: 0: 0 :0 : 0 : 0 : 0 to **3fff : ffff: fffff: ffff : ffff : ffff : ffff : ffff**

- The prefix fD00::/8 represents the addresses from
  fD00 : 0: 0: 0 :0 : 0 : 0 : 0 to **fdff : ffff: ffff: ffff : ffff : ffff : ffff : ffff**

IPv4 over IPv6 representation

- ::ffff:192.0.2.128 represents the IPv4 address 192.0.2.128.
- A deprecated format for IPv4-compatible IPv6 addresses was ::192.0.2.128

IPv6 Versus IPv4

- **IHL** was removed
  - In fact it was merged with the transport layer protocol identifier
- All the **fragmentation fields** removed
  - IPv6 assumes no fragmentation at routers
  - Hosts are supposed to discover the MTU and do fragmentation if necessary
  - Routers that can not transfer a packet returns a report back to sender
  - The minimum-size packet at routers was raised from 576 to 1280 bytes
- **Checksum** field was removed:
  - Calculating reduces performance
  - IPv6 supposes that the networks are more reliable !!
- **Conclusion**
  - ☺ Fast and flexible protocol ☺
  - ☺ Plenty of addresses ☺
  - ☹ Some headers are occasionally needed ☹

    → **use optional extension headers**

IPv6: The Future

- Finally IPv4 and IPv6 are not compatible
- In 2010, 1% of internet users uses IPv6
- In 2011, some companies like Google started using IPv6 in their networks
- In 2013, the use is estimated at **☹ 2% ☹**

- The migration will never be brutal
- IPv6 islands will emerge every where. Communicate using IPv4 tunnels…. Until one day?????

References

[1] Andrew S. Tanenbaum, David J. Wetherall, *Computer Network*

[2] Jim Kurose, Keith Ross, Computer Networking: A Top Down Approach

[3] Wikipedia