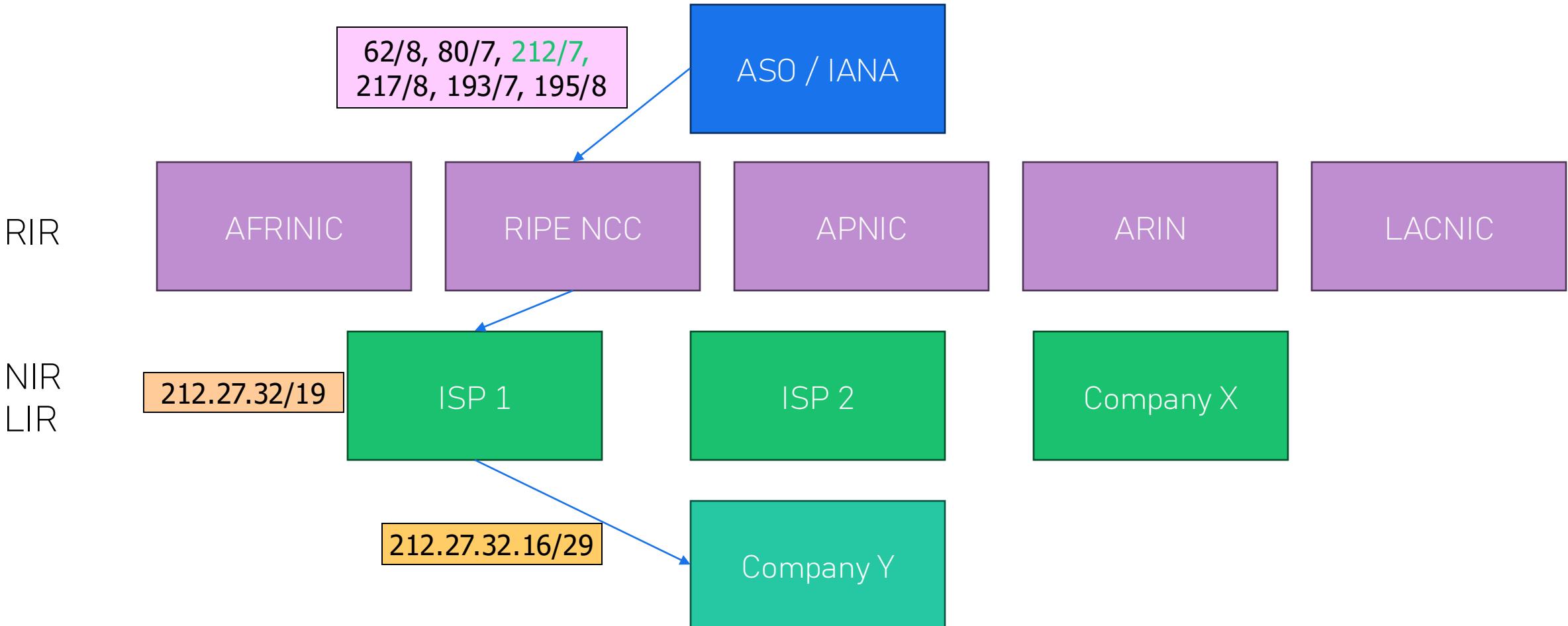


# Network layer: IP Architecture, Addressing

# Internet Routing Architecture: Address Assignment



# Internet layers - Encapsulation

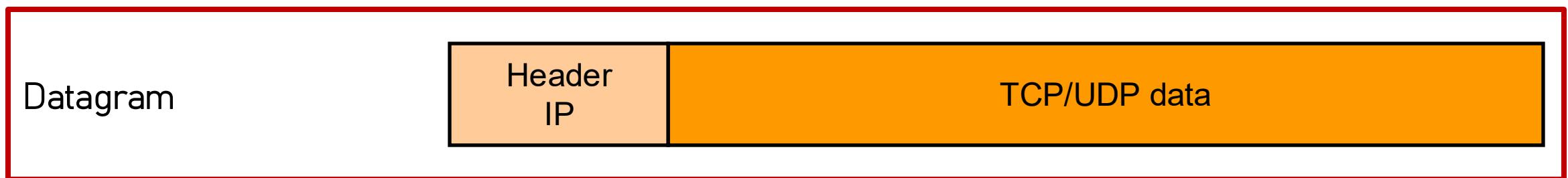
Message



Segment



Datagram



Frame



# IPv4 addresses

- IPv4 hosts are uniquely identified by a 4 bytes address
- IPv4 addresses are hierarchical, they are divided into two parts:
  - Network part: identify the network to which the host is connected
  - Host part: identify the host inside the network
  - In the Internet, only the network part is used by the routers for routing and forwarding
- IP addresses notation:
  - Each byte is represented by its decimal value and the decimal expression of the bytes are separated by dots.
  - Ex: 192.200.0.0
- The addresses are grouped into 5 classes

# IPv4 Addresses: Classes

Class A (/8)

0	Network (7b)	Host (24 bits)	1.xxx.xxx.xxx to 127.xxx.xxx.xxx
---	--------------	----------------	----------------------------------

Class B (/16)

10	Network (14 bits)	Host (16 bits)	128.0.xxx.xxx to 191.255.xxx.xxx
----	-------------------	----------------	----------------------------------

Class C (/24)

110	Network (21 bits)	Host (8 bits)	192.0.0.xxx to 223.255.255.xxx
-----	-------------------	---------------	--------------------------------

Class D (/4)

1110	Multicast address (28 bits)	224.xxx.xxx.xxx to 239.255.255.255
------	-----------------------------	------------------------------------

Class E (/4)

1111	Experimental address (28 bits)	240.xxx.xxx.xxx to 255.255.255.255
------	--------------------------------	------------------------------------

# IP packet forwarding

Incoming Packet: 137.194.40.37

What to do:

- IP table look-up
- Longest match prefix

Route	Next Hop	Port	Calculation
137.194.10.0/24	R5	Tk0	
137.194.20.0/24	local	Tk0	
137.194.30.0/24	local	Eth0	
137.194.40.0/24	R4	fddi0	
137.194.50.0/24	local	fddi0	
137.194.0.0/16	G2	fddi0	
200.0.0.0/10	G2	fddi0	
Default route	G1	fddi0	

# IP packet forwarding

Incoming Packet: 137.194.40.37

What to do:

- IP table look-up
- Longest match prefix

Route	Next Hop	Port	Calculation
137.194.10.0/24	R5	Tk0	$(137.194.40.37 \& 255.255.255.0) = 137.194.40.0 \neq 137.194.10.0$
137.194.20.0/24	local	Tk0	$(137.194.40.37 \& 255.255.255.0) = 137.194.40.0 \neq 137.194.20.0$
137.194.30.0/24	local	Eth0	$(137.194.40.37 \& 255.255.255.0) = 137.194.40.0 \neq 137.194.30.0$
137.194.40.0/24	R4	fddi0	$(137.194.40.37 \& 255.255.255.0) = 137.194.40.0 = 137.194.40.0 \checkmark$
137.194.50.0/24	local	fddi0	$(137.194.40.37 \& 255.255.255.0) = 137.194.40.0 \neq 137.194.50.0$
137.194.0.0/16	G2	fddi0	$(137.194.40.37 \& 255.255.0.0) = 137.194.0.0 = 137.194.0.0 \checkmark$
200.0.0.0/10	G2	fddi0	No match (different range)
Default route	G1	fddi0	Always matches (lowest priority) $\checkmark$

# IPv4 addressing limitations

- IPv4 addressing limitations:
  - IP Address exhaustion: No more class B addresses
  - Explosion of BGP Routing tables
- Proposed solutions:
  - Long term: IPv6
  - Short term:
    - CIDR (Class-Less Inter Domain Routing)
      - Introduced in BGP-4
      - Associated Strict Address Assignment Policy
    - Subnetting: born with the introduction of the routing hierarchy
    - Private Addressing / NAT
      - Possible reuse of some addresses ranges

# Subnetting

Class C (/24)



- Number of possible Host in a Class C address:  $2^8 = 256 \text{ Addresses}$
- Possible approaches:
  - Fixed Length Subnet Masking (FLSM)
  - Variable Length Subnet Masking (VLSM)
- For each subnet:
  - 1 address to identify the subnet
  - 1 broadcast address per subnet

# Subnetting: Fixed Length Subnet Masking(FLSM)

Steps:

1. Minimum number of required bits to partition:  
 $\log_2 \text{max\_Host}$
2. Calculate new subnet mask and addresses
3. Assign usable host ranges

Solution:

1.  $[\log_2(60 + 1 + 1)] = 6$
2. New submask =  $32 - 6 \Rightarrow /26$
3. Subnet addresses:
  - CIDR Notation Subnet A: 192.168.1.0/26 (1-62)
  - CIDR Notation Subnet B: 192.168.1.64/26 (65-126)
  - CIDR Notation Subnet C: 192.168.1.128/26 (129-190)
  - CIDR Notation Subnet D: 192.168.1.192/26 (193-254)

Example:

- Class C network: 192.168.1.0/24
- Subnets:
  - A. 60 hosts
  - B. 30 Hosts
  - C. 10 Hosts

# Subnetting: Variable Length Subnet Masking(VLSM)

## Steps:

1. Determine the smallest subnet mask for each requirement
2. Allocate address blocks one by one
3. Update the available pool after every assignment

NB: Always use the smallest partition available

- Class C network: 192.168.1.0/24
- Subnets:
  - A. 60 hosts
  - B. 30 Hosts
  - C. 10 Hosts

## Example:

- Subnet A (60 hosts):
  - Needed:  $60 + 2 = 62$  addresses
  - Host bits (h):  $\lceil \log_2(62) \rceil = 6$
  - Subnet mask:  $32 - 6 \Rightarrow /26$
  - CIDR Notation: 192.168.1.0/26
- Subnet B (30 Hosts)
  - Needed:  $30 + 2 = 32$  addresses
  - Host bits (h):  $\lceil \log_2(32) \rceil = 5$
  - Subnet mask:  $32 - 5 \Rightarrow /27$
  - CIDR Notation: 192.168.1.64/27
- Subnet C (10 hosts):
  - Needed:  $10 + 2 = 12$  addresses
  - Host bits (h):  $\lceil \log_2(12) \rceil = 4$
  - Subnet mask:  $32 - 4 \Rightarrow /28$
  - CIDR Notation: 192.168.1.96/28

What do you notice between FLSM and VLSM?

# Private Addressing

- A host connected to the Internet needs a “public” IP address (unique routable address) to be reached.
- A host located in a private network can have a “private” address (non routable, non necessarily unique) if it does not access the public Internet
- To be able to differentiate between internal and public hosts, private IP address blocks networking have been defined:
  - CIDR Notation: 10.0.0.0/8 (10.0.0.0 to 10.255.255.255)
  - CIDR Notation: 172.16.0.0/12 (172.16.0.0 to 172.31.255.255)
  - CIDR Notation: 192.168.0.0/16 (192.168.0.0 to 192.168.255.255)

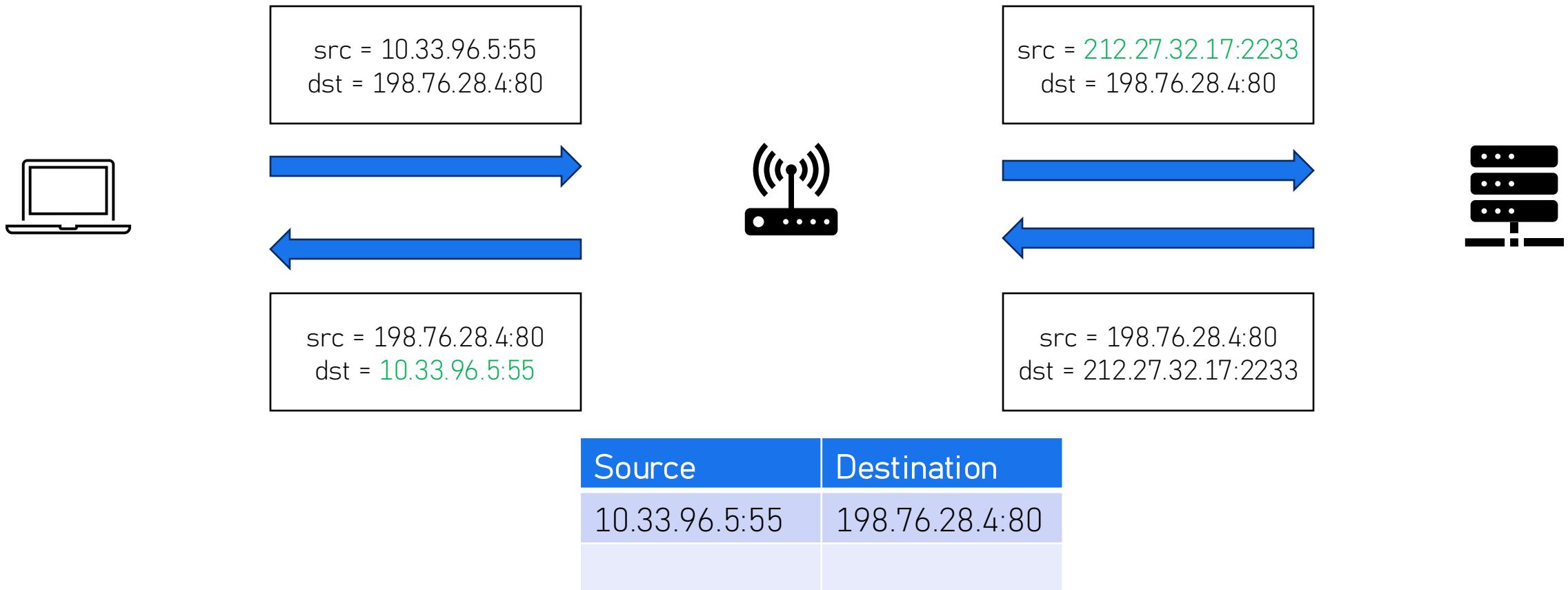
# NAT (Network Address Translation)

NAT protocol is utilized in a bridge router to enable devices within a private network to access the internet

Devices use private IP addresses internally

1. Traffic destined for the Internet is sent to the NAT router
2. The router replaces the source IP and and the entry in its lookup table
3. The router tracks the mapping in a NAT table
4. For each response reaching NAT router
  1. Finds the matching entry in its table
  2. Swaps the destination IP back to the internal (private) IP address
  3. Forwards the packet to the original device

# NAT: Example



# NAT: Approaches

Listed in increased security order:

**Full Cone NAT:** Once an internal IP and port are mapped to an external IP and port, any external host can send packets to the internal host by sending to that external address and port

**Restricted Cone:** External hosts can send packets to the internal host only if the internal host has previously sent packets to that external host IP (regardless of port)

**Symmetric NAT:** Each outbound connection from the same private IP and port to a different public IP and port is mapped. Only the external host that received the initial packet can send back.

# NAT traversal: issues for IP telephony

NAT allows outgoing communications but not incoming ones:

- A device with private address cannot be called whatever the caller has a private or public address
- Port “forwarding technique” is not applicable. Suitable only with one telephone or requires a heavy non scalable configuration of the whole IP telephony network (a different non standard port has to be assigned to each telephone)

NB: Based on the application various solution exist

# IP Addressing Evolution

- CIDR and NAT extend the survival of IPv4
- They do not constitute long term solutions:
  - Performances reasons
  - Increasing number of IP addresses needed
- On the long/medium term, IPv6 will take over

## Nortel, in bankruptcy, sells IPv4 address block for \$7.5 million

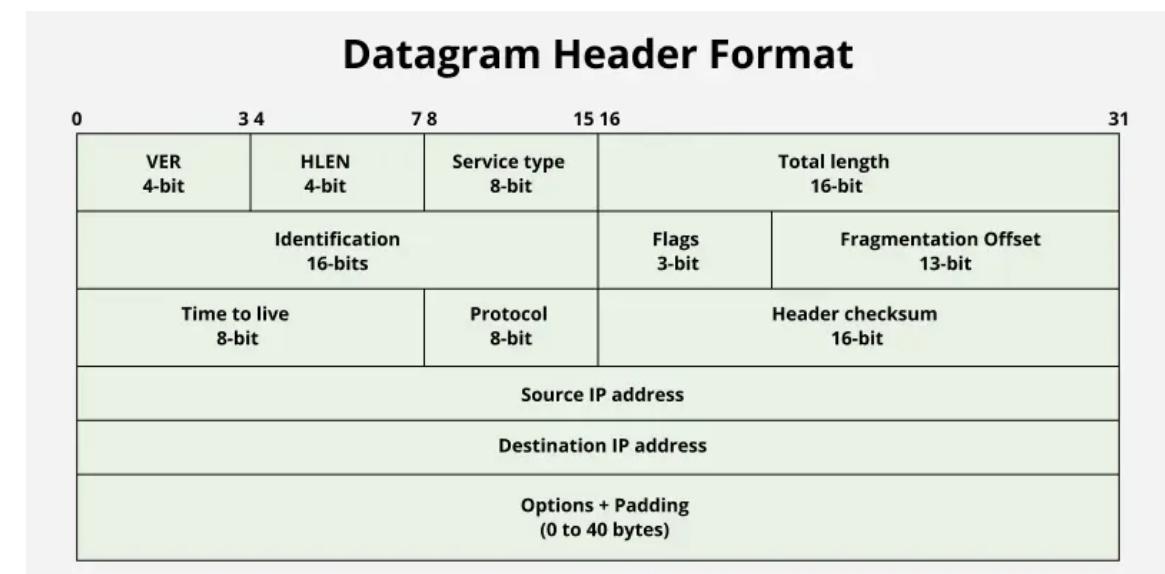
by [Milton Mueller](#) on Wed 23 Mar 2011 10:30 PM EDT | [Permanent Link](#) |



Wake up call for our friends in the Regional Internet Registries. Nortel, the Canadian telecommunications equipment manufacturer that filed for bankruptcy protection in 2009, has succeeded in making its legacy IPv4 address block an asset that can be sold to generate money for its creditors. The March 23 edition of the Dow Jones Daily Bankruptcy Report has reported that Nortel's block of 666,624 IPv4's was sold for \$7.5 million – a price of \$11.25 per IP address. The buyer of the addresses was Microsoft. More information is in its filing in a Delaware bankruptcy court. Now the interesting question becomes, does the price of IPv4s go up or down from here? As the realities of dual stack sink in, I'm betting...up.

# IP datagram

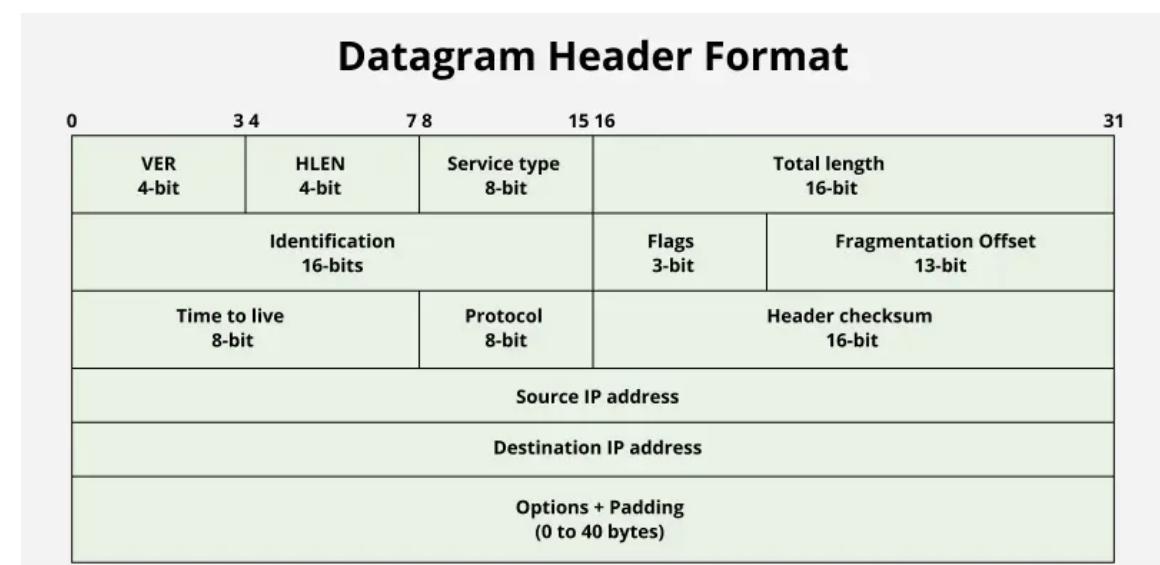
- VER (4 bits): Protocol version used
- HLEN (4 bits):
  - Total length of Header
  - Number of 32-bit words (4 bytes) in the header
  - From 5 to 15 words
- Total Length:
  - Header + Data
  - Expressed in bytes
  - Maximum 65,535 bytes



# IP datagram

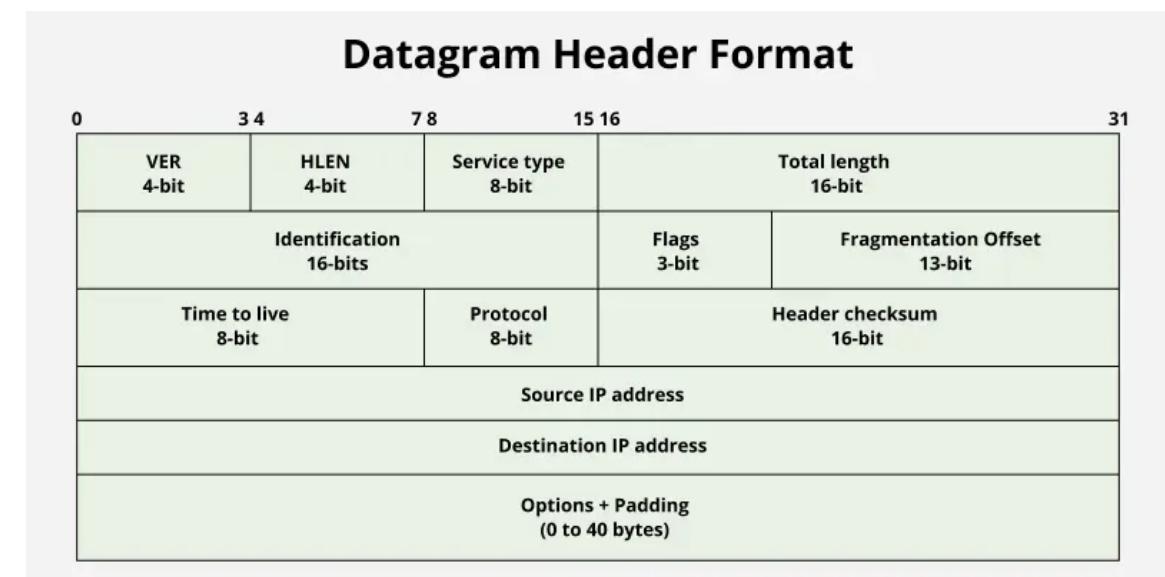
## Type of Service (8 bits):

- Precedence (3 bits): Priority level of the packet.
- Delay (1 bit): Low delay requested.
- Throughput (1 bit): High throughput requested.
- Reliability (1 bit): High reliability requested.
- Cost (1 bit): Minimize monetary cost.
- Reserved (1 bit): Reserved for future use.
- **Differentiated Services (8 bits)[NEW]:**
  - Differentiated Services Code Point (DSCP) (6bits): classify packets into traffic classes
  - Explicit Congestion Notification (ECN) (2 bits)



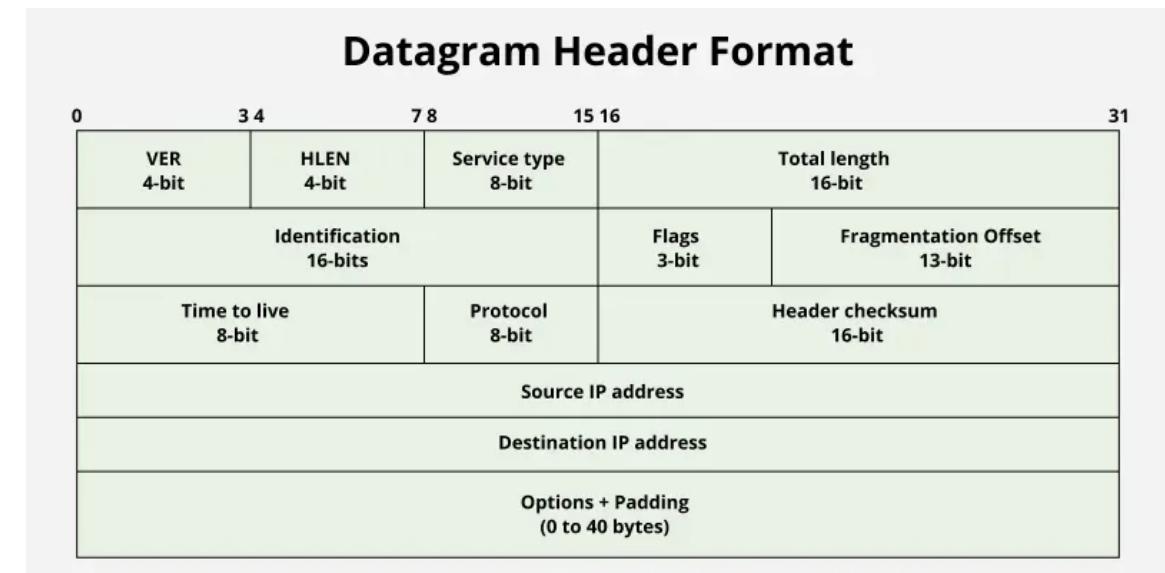
# IP datagram

- Identification (16 bits): identifying fragmented packets
- Flags (3 bits):
  - Reserved and not used
  - Do Not Fragment (DF)
  - More Fragments (MF)
- Fragment Offset ( 13 bits):
  - Offset position of the fragment
  - Measured in bytes
  - Payload/8



# IP datagram

- TTL (Time to Live)(8 bits): Max hops before packet is discarded
- Protocol (8 bits):
  - 6 TCP
  - 17 UDP
  - 1 ICMP
- Header Checksum ( 13 bits):
  - Offset position of the fragment
  - Measured in bytes
- Source/Destination IP Address (32 bits)



# IP datagram: Fragmentation

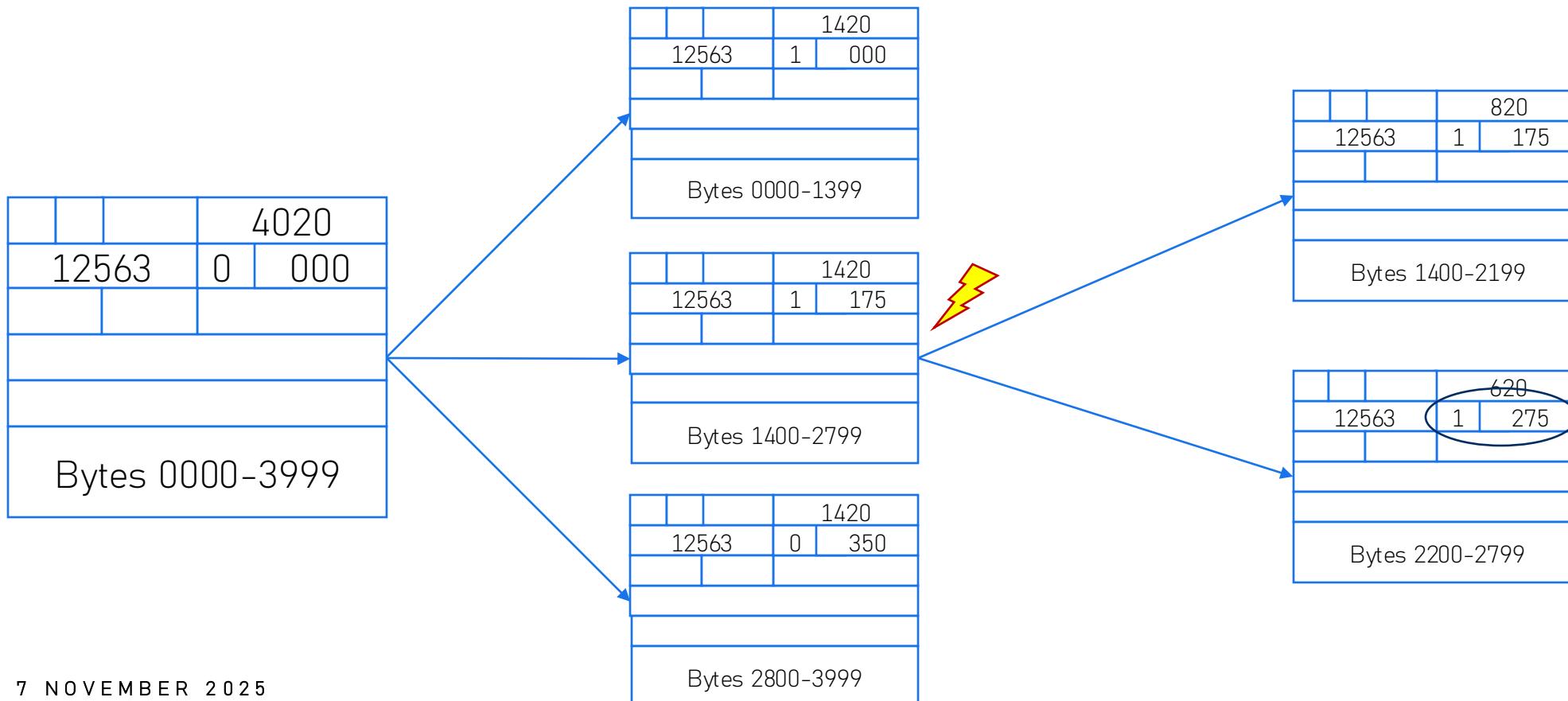
When a large datagram is created by the source computer, it may need to travel through multiple networks to reach the destination. Each of these networks might support different maximum packet sizes. This leads to the need for fragmentation.

## Reasons:

- **Different Network Limits:** Each network has its own MTU.
- **Unknown Path MTU:** Source doesn't always know the smallest MTU along the route.
- **Error Handling:** Smaller fragments reduce retransmission if a part is lost or corrupted.
- **Protocol & Hardware Limits:** Some devices and protocols limit maximum packet size.
- **Transmission Efficiency:** Large packets take longer in the network, delaying others

# IP datagram: Fragmentation

Example: We want to send a picture of 4000 bytes, but the datagram is too large for network, because the MTU is 1400bytes. Datagram needs to be divided into 3 fragments.



# IPv4 vs IPv6: Key Differences

## Address:

- Length: 128 bits
- Format: Hexadecimal e.g., 2001:db8::1
- Space: About  $3.4 \times 10^{38}$  addresses
- Notation: 8 groups, colon-separated hexadecimal, ":" replaces one or more groups of zeros

**Header:** Fixed 40 bytes

**NAT:** Not required

**Broadcast:** No -> Multicast

**Security:** IPSec built-in

# Address Resolution Protocol (ARP)

**GOAL:** Mapping an IP address to a physical MAC address on a private network

How Does ARP Work?

1. Check ARP Cache for the target IP is stored locally
2. Broadcast ARP Request     "Who has IP x.x.x.x? Tell me your MAC."
3. Receivers Process Request and only the device matching the IP replies
4. Unicast ARP Reply sending its MAC address to requester
5. Update ARP Cache and stores it

**NB: Multiple MAC mapped to a single IP address is possible, but it creates conflicts!**

# Internet Control Message Protocol (ICMP)

**GOAL:** Providing error notification and information request

**Network Diagnostics:**

- Ping to test the availability of an IP address
- Traceroute to analyze the path and measure response time

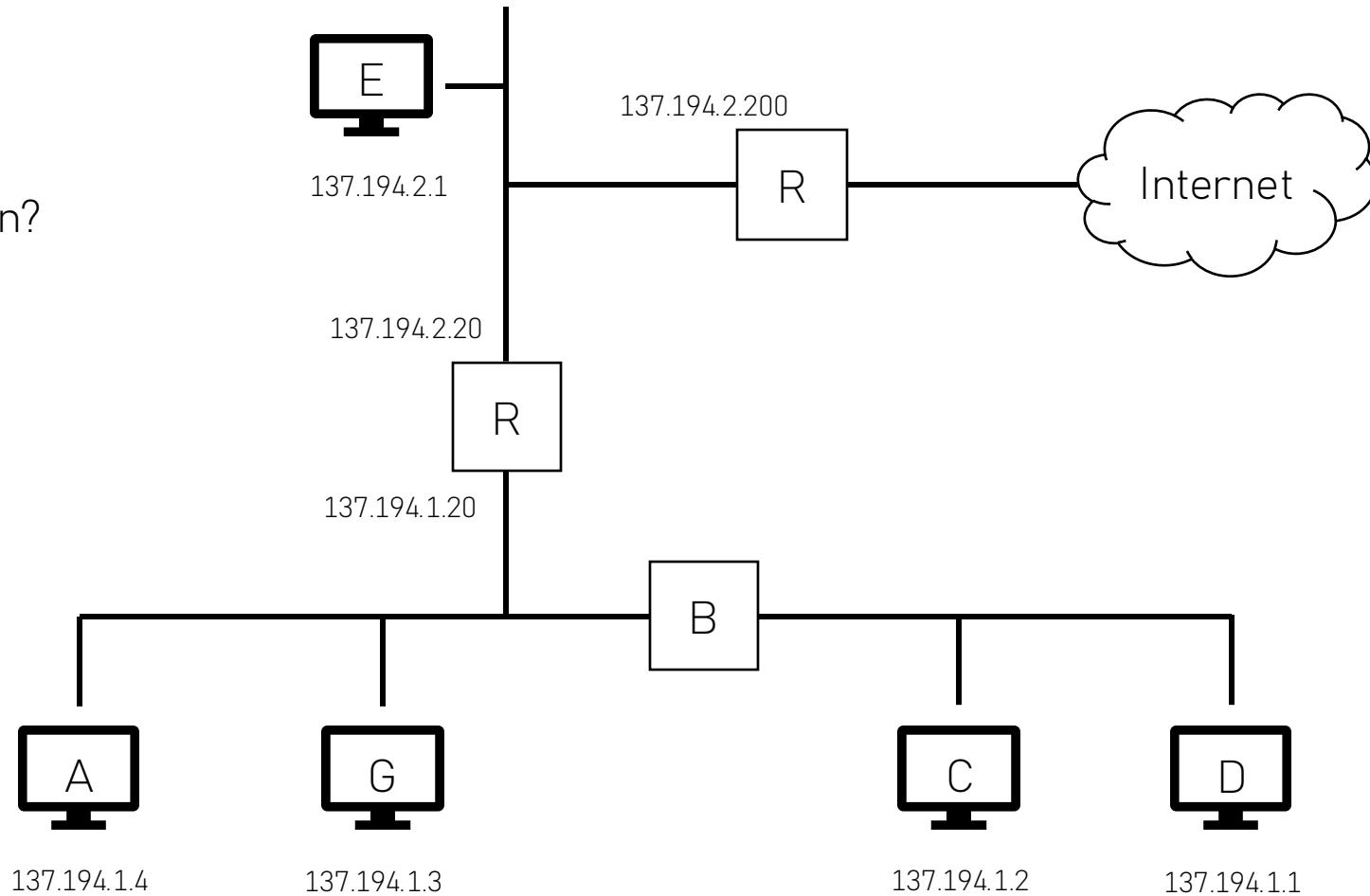
**Error Reporting:**

- Destination Unreachable
- Time Exceeded
- Packet Too Big

# Example: Private network

How is transmission?

- From D to C
- Form A to D
- Form C to G
- From G to E



# Exercise: ARP and ICMP

## Instructions:

1. Inspect Your ARP Cache
2. Identify Your Private IP Address and Netmask
3. Ping all devices
4. Review and Compare the ARP Cache Again
5. Perform a Traceroute to a Local Device
6. Perform a Traceroute to an External Host (roc.cnam.fr and cnam.fr)

NB: Do not include this exercise in the lab report!