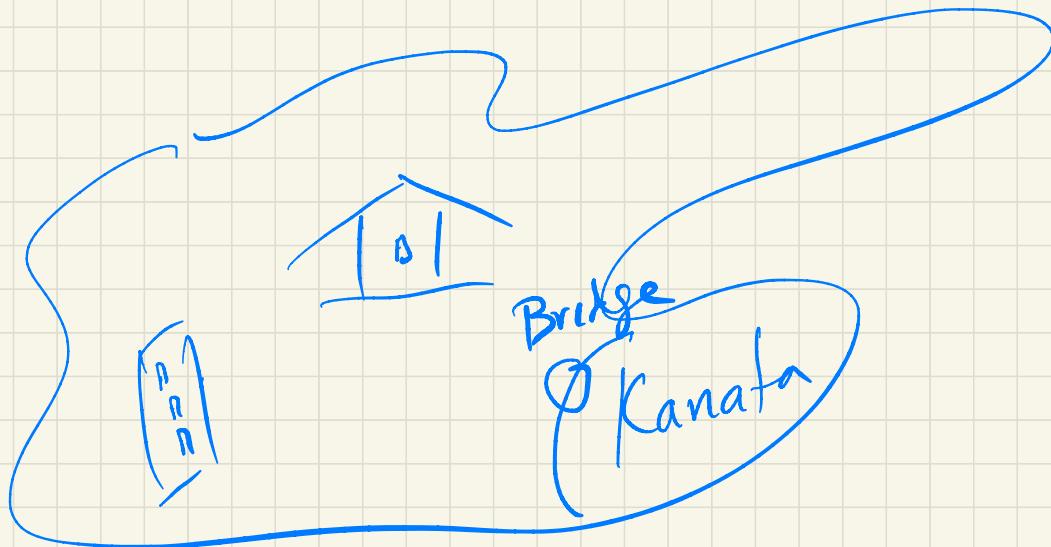


Internetworking (IP)

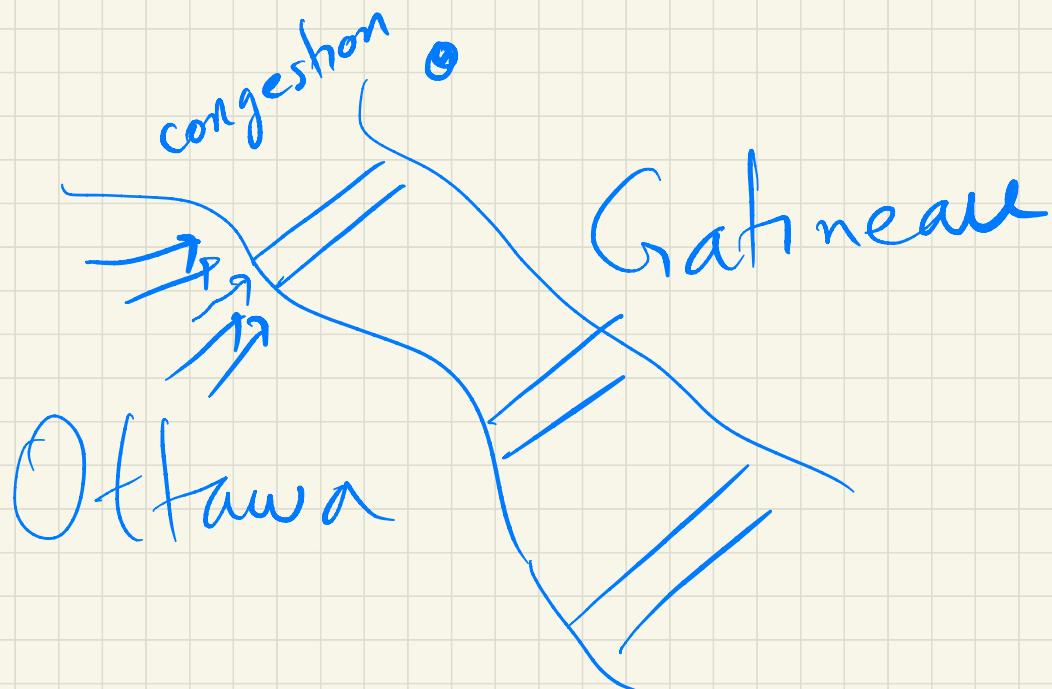
IP/TCP

Big City ; Post Office



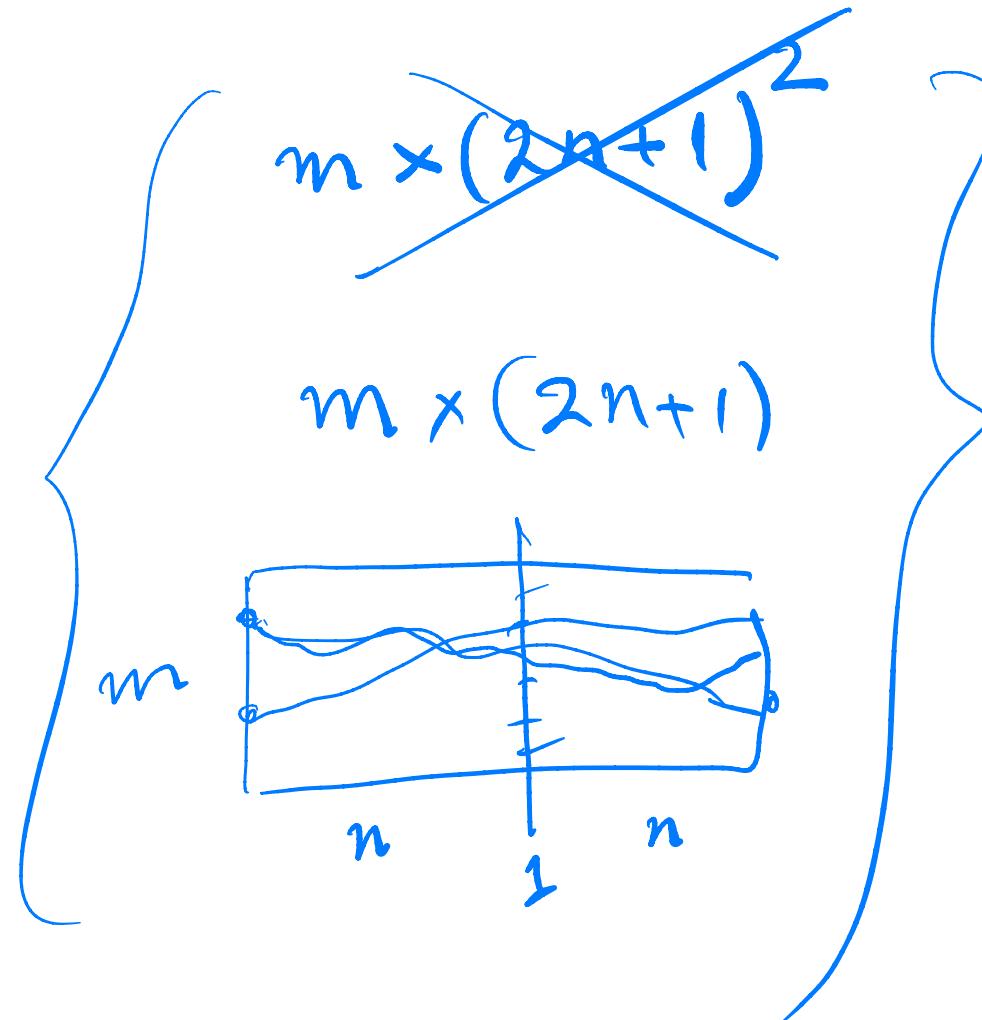
IP: Provides addresses

TCP: Controls traffic



Outline

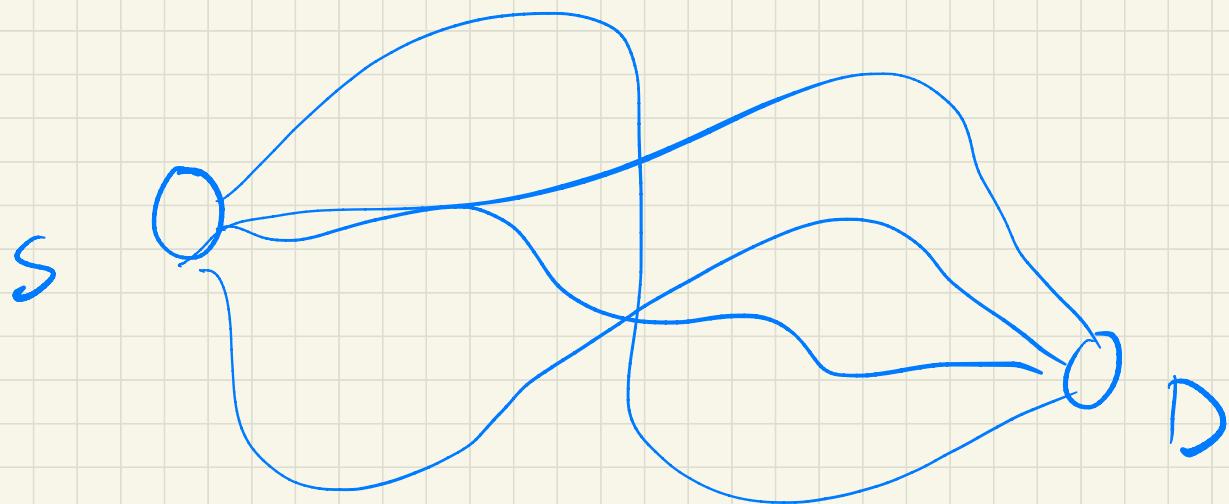
- Internetworks
 - IPv4
- Address Resolution
- IPv6



IP Networks

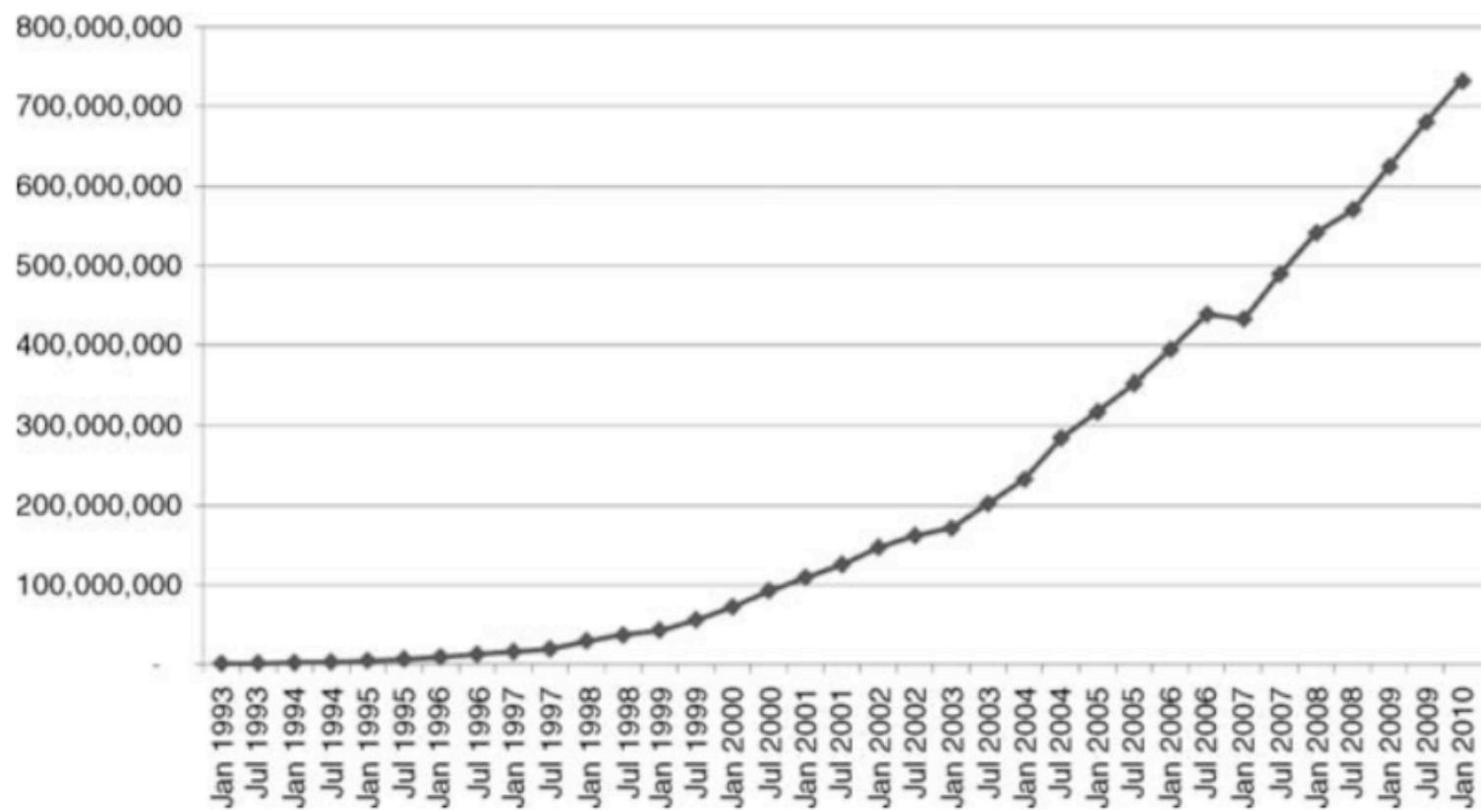
- IP is the most widely deployed network layer protocol worldwide.
- Emerging from a U.S. government sponsored networking project for the U.S. Department of Defense begun in the 1960s, the TCP/IP suite has evolved and scaled to support networks from hundreds of computers to hundreds of millions today.
- The number of devices or hosts on the Internet exceeded 730 million as of early 2010 with average annual additions of over 75 million hosts per year.
- The fact that the Internet has scaled rather seamlessly from a research project to a network of over 730 million computers is a testament to the vision of its developers and robustness of their underlying technology design.

We would use in the past the
"circuit switching" model



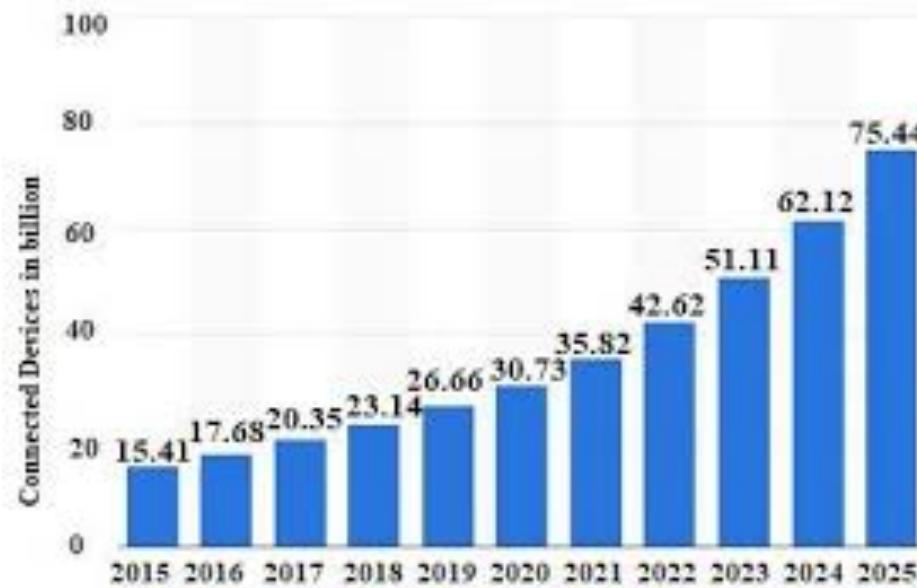
IP Networks

- Growth of Internet hosts during 1993-2010



IoT Devices

- Growth of IoT Devices 2015 to 2025



- The number of devices connected to the internet reached 22 billion worldwide at the end of 2018

IP Networks

- IP was “initially” defined in 1980 in Request for Comments (RFC) 760 and 791, edited by Jon Postel^a
- Postel pointed out in his preface, that RFC 791 is based on six earlier editions of the ARPA^b Internet Protocol, though it is referred to in the RFC as version 4 (IPv4). *v = version*
- RFC 791 states that the Internet Protocol performs two basic functions: addressing and fragmentation.
 - Addressing assures unique addressability of hosts
 - Fragmentation deals with splitting messages into a number of IP packets so that they can be transmitted over networks that have limited packet size constraints, and reassembly of packets at the destination in the proper order.

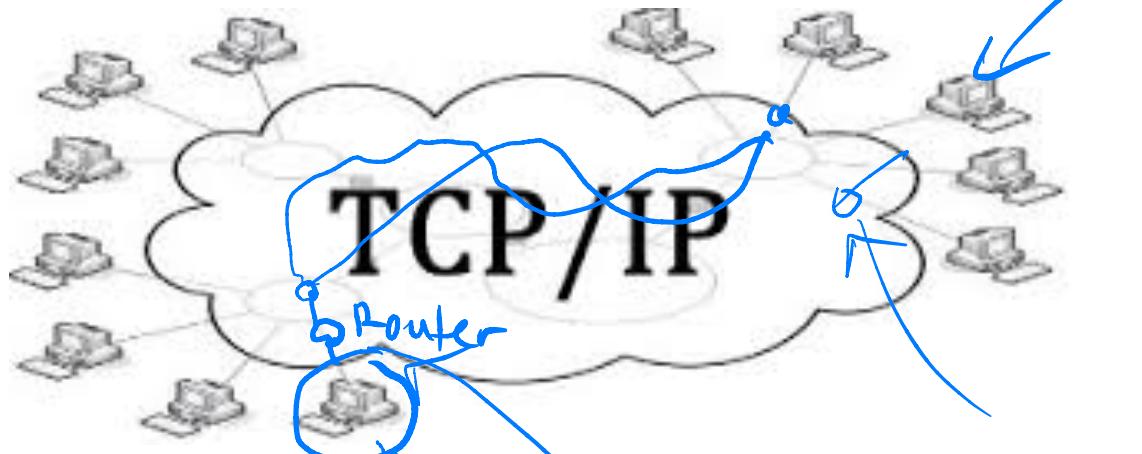
^a1943 to 1998 was also editor of RFC documents; his obituary was published as an RFC by Cerf.

^bAdvanced Research Projects Agency, a U.S. Department of Defense agency

Internetworks

Network

- Hosts are connected via a network cloud.



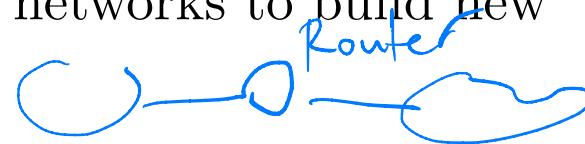
- Its “nerve veins” are based on TCP/IP

IPv4

- Internet Protocol version 4 (IPv4) is the fourth revision in the development of the Internet Protocol (IP) and the first version of the protocol to have been widely deployed.
- Described in IETF (Internet Engineering Task Force) publication RFC 791 (September 1981), replacing an earlier definition (RFC 760, January 1980).
- It is a connectionless protocol for use on packet-switched Link Layer networks
- It operates on a best effort delivery model, in that it does not guarantee delivery, nor does it assure proper sequencing or avoidance of duplicate delivery.
- These aspects, including data integrity, are addressed by an upper layer transport protocol, such as the Transmission Control Protocol (TCP)

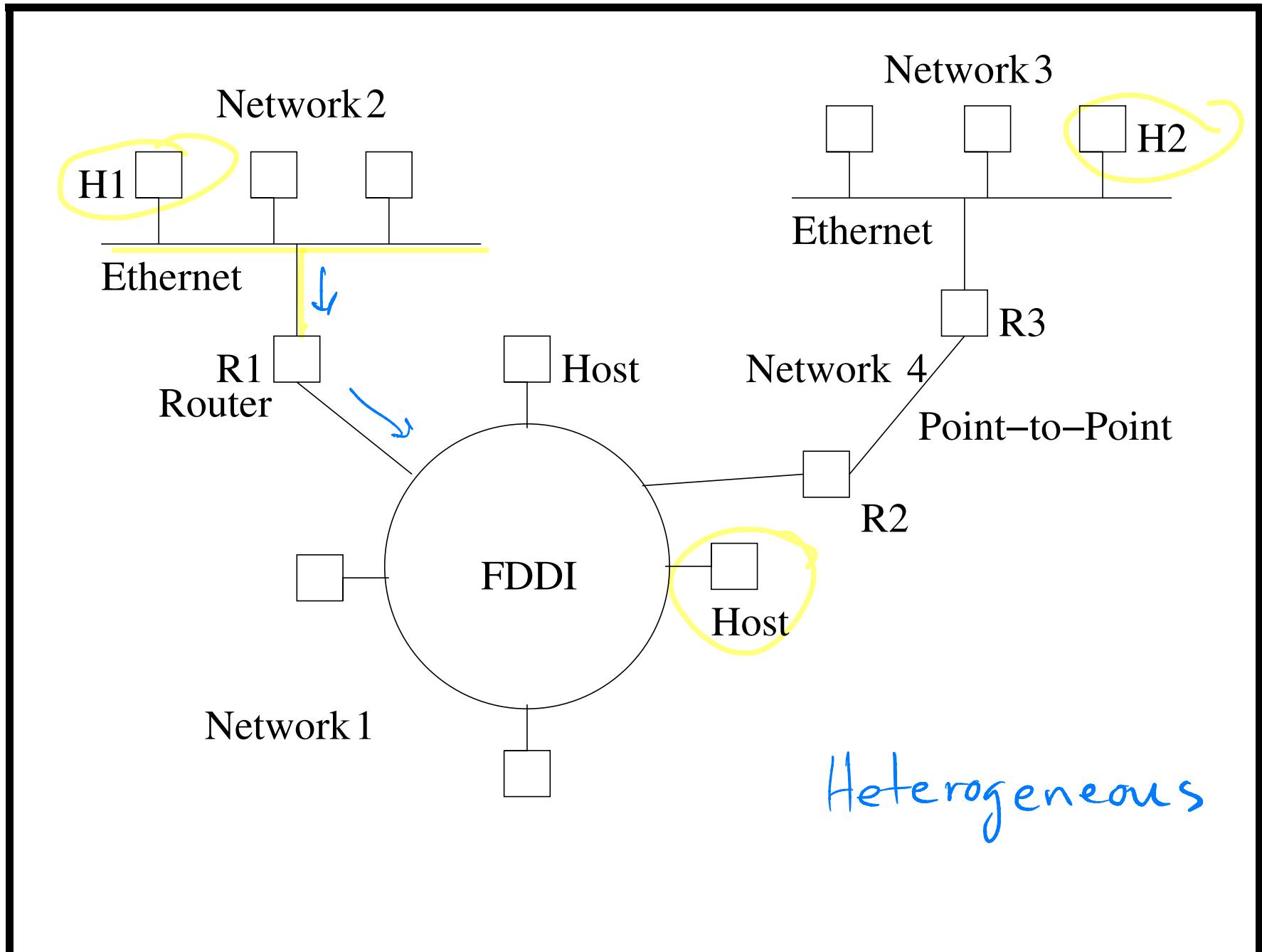
Internetworks

- Important to be able to interconnect networks to build new and larger ones.
- LAN approaches are limited: they do not scale well and they cannot handle heterogeneity.
- An internetwork is an arbitrary collection of networks interconnected to provide host-to-host packet delivery service.
- The networks are interconnected with special nodes called routers, and gateways.
- IP is the main protocol for interconnecting: invented by Kahn and Cerf^a

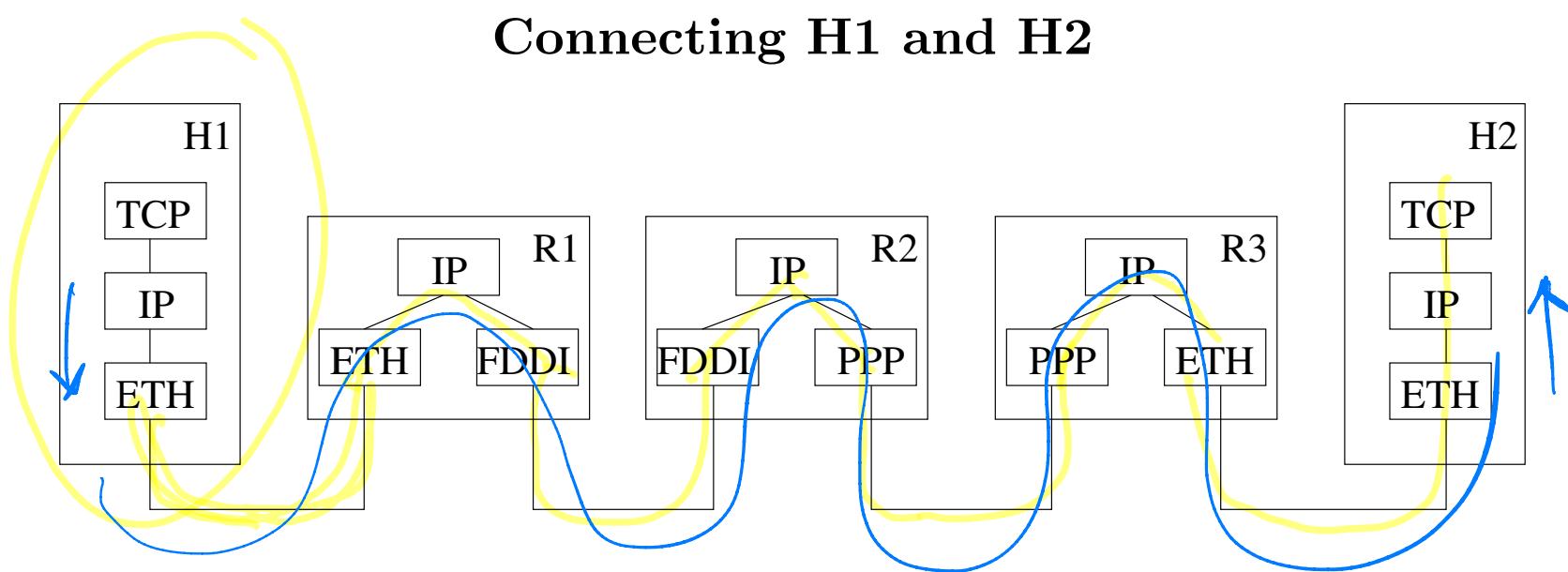


CU Learn

^aDecember 1997: presented with National Medal of Technology by Bill Clinton, “for creating and sustaining development of Internet Protocols and continuing to provide leadership in the emerging industry of internetworking”.



Connecting H1 and H2



1. Leave H1 from Network 1 Ethernet via Router R1
2. From Router R1 to FDDI token ring.
3. From FDDI token ring to R2.
4. From router R2 to router R3 via Point-to-Point Network.
5. From router R3 to Network 3 Ethernet.

Service Model (Ability to Run over Anything!)

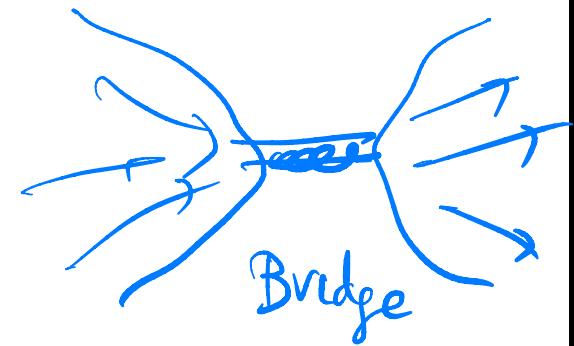
- The philosophy was to give a protocol that will be undemanding enough for just about any existing technology.
- The service model has two parts:
 1. Addressing Scheme.
 2. Datagram delivery connectionless model.
- The addressing scheme provides a way to identify all hosts.
- The datagram delivery scheme is called best effort because it makes no guarantees.
- Datagrams are frames that are sent in a connectionless manner. Of course, they must include sufficient information to enable delivery. **IPv4 Datagrams** consist of a header plus a number of data bytes.

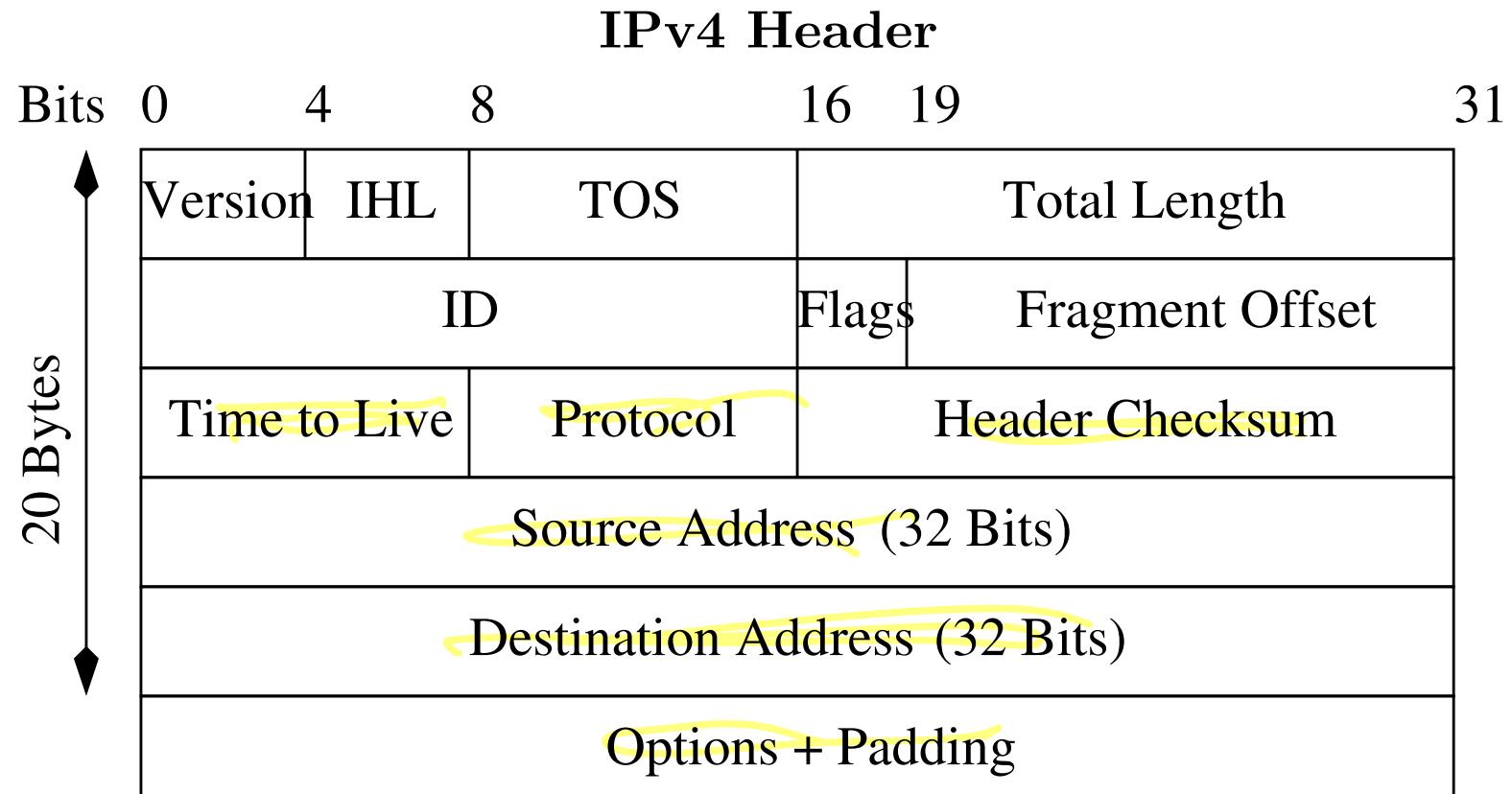
IP

- IP is the vehicle for traffic management.
- IP based internets were designed to support delay insensitive applications.

QoS = Quality of Service

- Today they face the following design requirements.
 - Control congestion
 - Provide low delay
 - Provide high throughput
 - Support QoS
 - Provide fair service
- All these issues fall under the category of traffic management.





IPv4 HEADER

IPv4 Header

- IP with no options is 20 Bytes.
- IHL is header length in 32-bit words.
- TOS (Type Of Service): provides guidance on selecting next hop and relative allocation of router resources.
- TOS subfield: provides route selection, subnetwork service, queuing discipline. These are specified with “certain rules”.
- Precedence Subfield: indicates the degree of urgency from highest level of “Network Control” to lowest level “Routine”. These provide appropriate Queue Service and Congestion Control.
- IPv4 options: Security, Timestamping, Source routing, Route Recording.

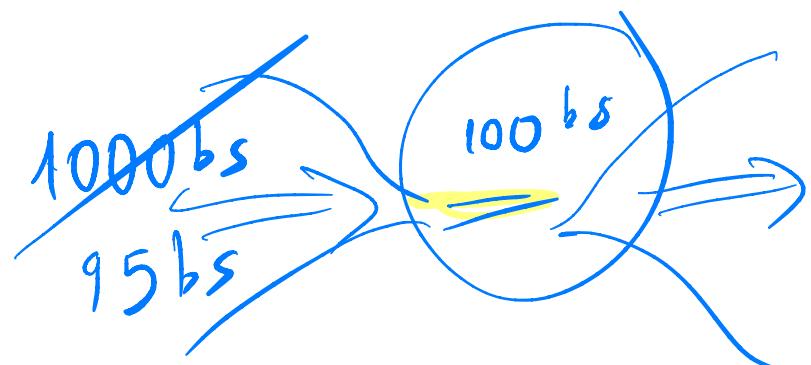
32
2

IPv4 TOS Field

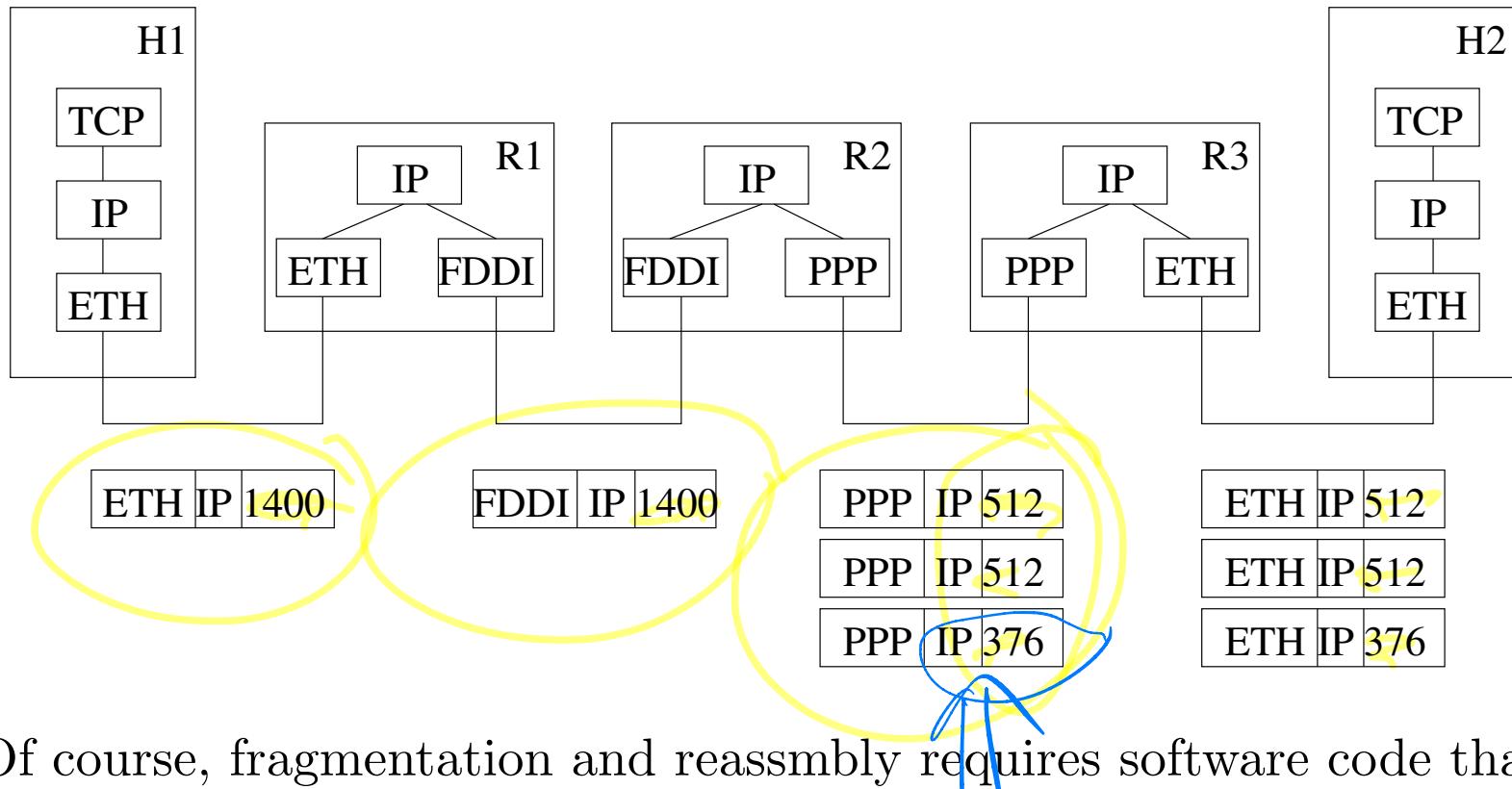
Bits	0	1	2	3	4	5	6	7
	Precedence			TOS			0	

IPv4 TOS-field

Precedence	TOS
111 Network Control	1000 Minimize Delay
110 Internetwork Control	0000 Maximize Throughput
101 Critical	0010 Maximize Reliability
100 Flash Override	0001 Minimize Monetary Cost
011 Flash	0000 Normal Service
010 Immediate	
001 Priority	
000 Routine	



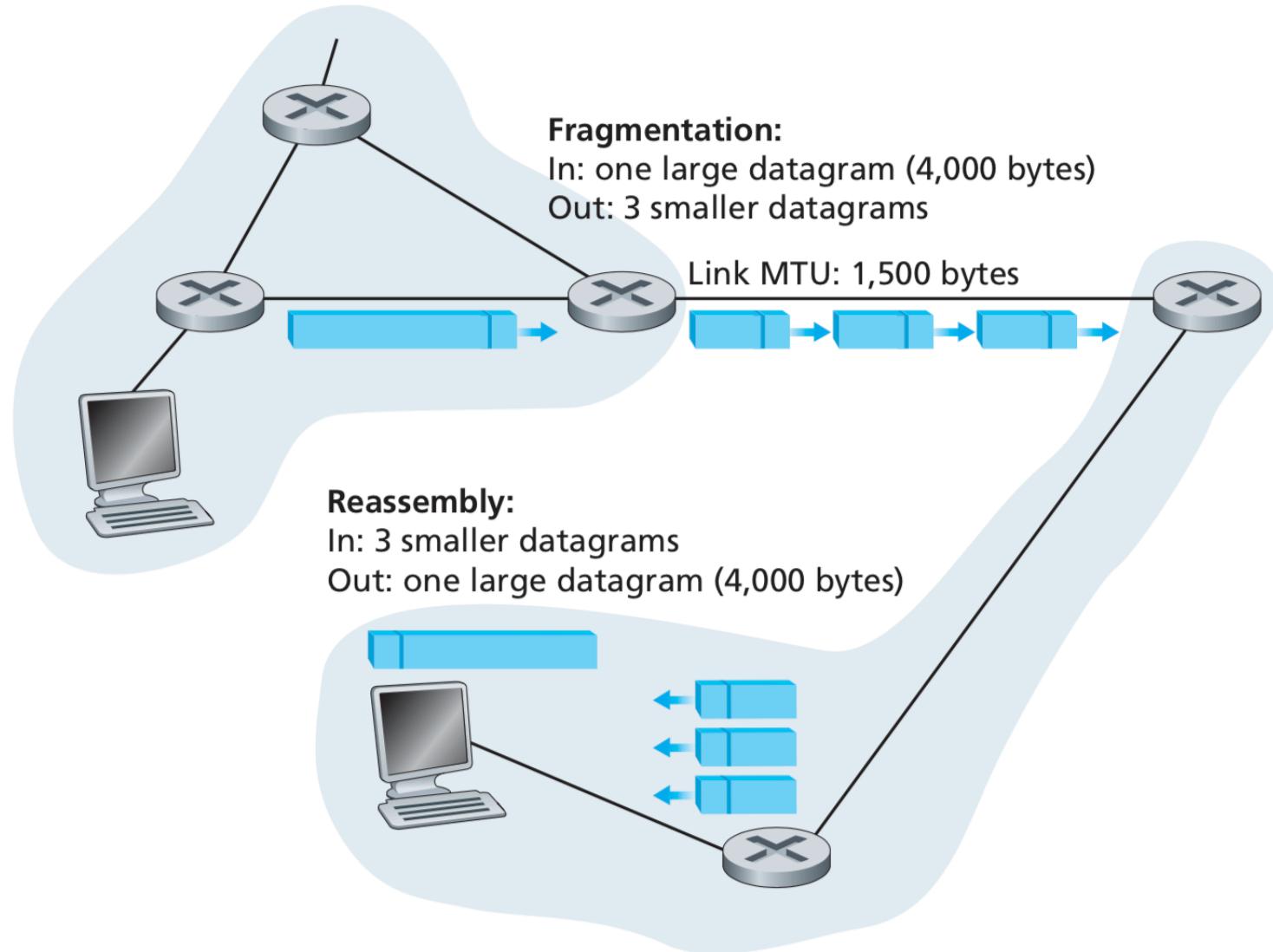
Fragmentation/Reassembly



Of course, fragmentation and reassembly requires software code that performs the required transformations from one packet format to another.

maximum transmission unit

MTU (Max Transmission Unit) Fragmentation/Reassembly



IP Addressing

1,000

$$2^{32} = (2^{10})^3 \cdot 2^2$$

- In addition to physical addresses (contained in NICs) nodes have 32 bit IP addresses.
- It is a two level hierarchy consisting of the net ID and the Host ID: net ID identifies the network the host is connected.
- All hosts connected to the same network have the same net ID.
- IP addresses look like:

Class	Net-ID	Host-ID
-------	--------	---------

- The lengths of Class, Net-ID, and Host-ID are variable, but the total length is 32 bits.



IP Addressing

- There are five classes of addresses: A, B, C, D, E.

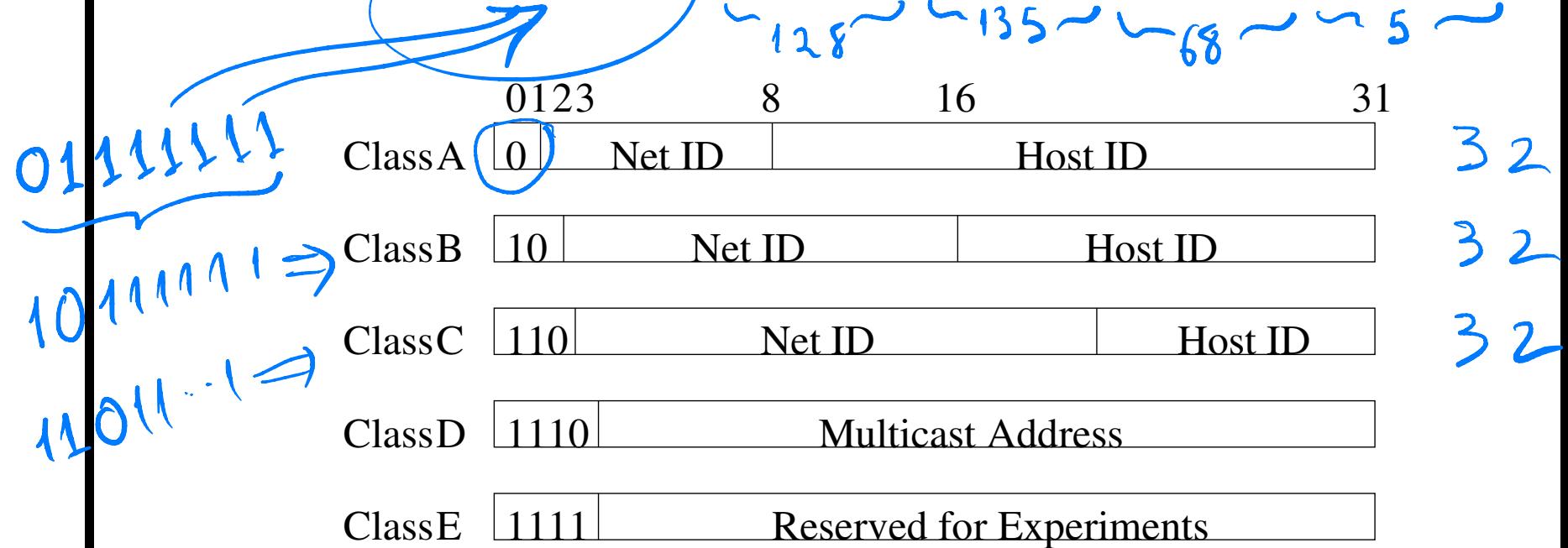
Class	Net ID	Host ID
A	7 bits	24 bits
B	14 bits	16 bits
C	21 bits	8 bits

$$\begin{aligned} & 2^{16} \\ & 1,000 \cdot 2^6 \\ & 64,000 \end{aligned}$$

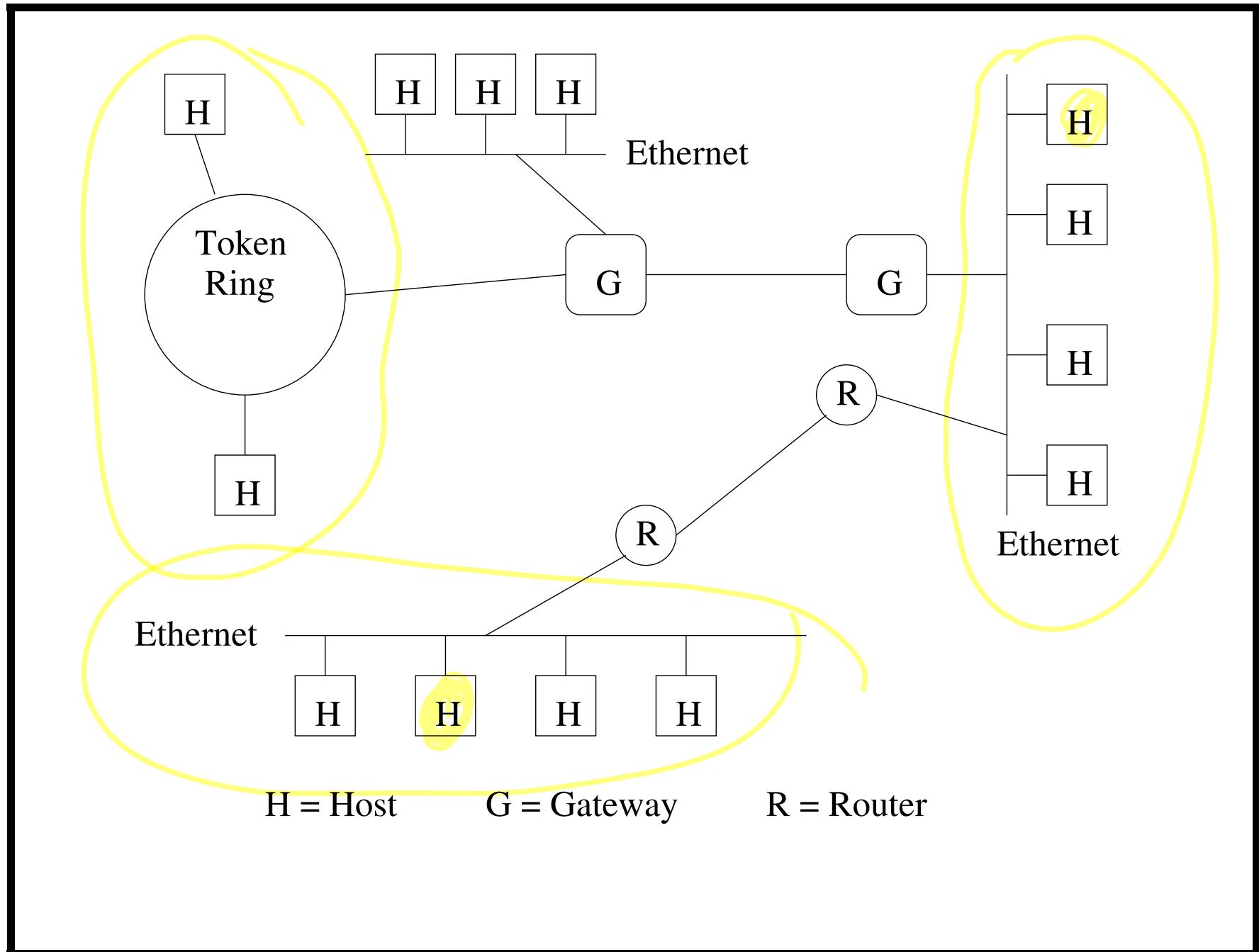
- D is used for multicasting and E for experiments.
- IDs with all 0s or all 1s are used for broadcasting:
 - immediately after booting up a host may not know its ID.
 - So host will transmit packets with all 0s while trying to find out correct ID.

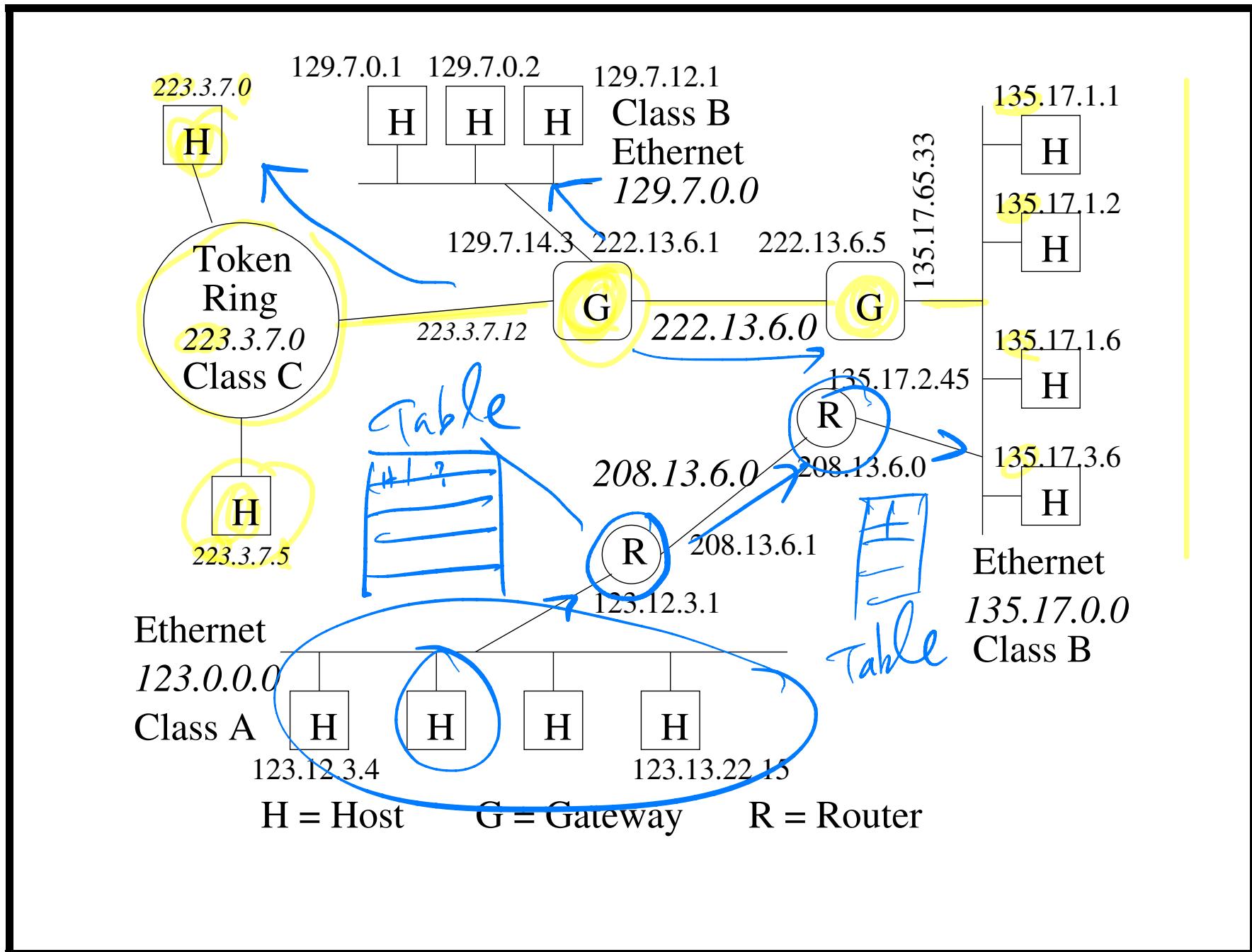
IP Addressing

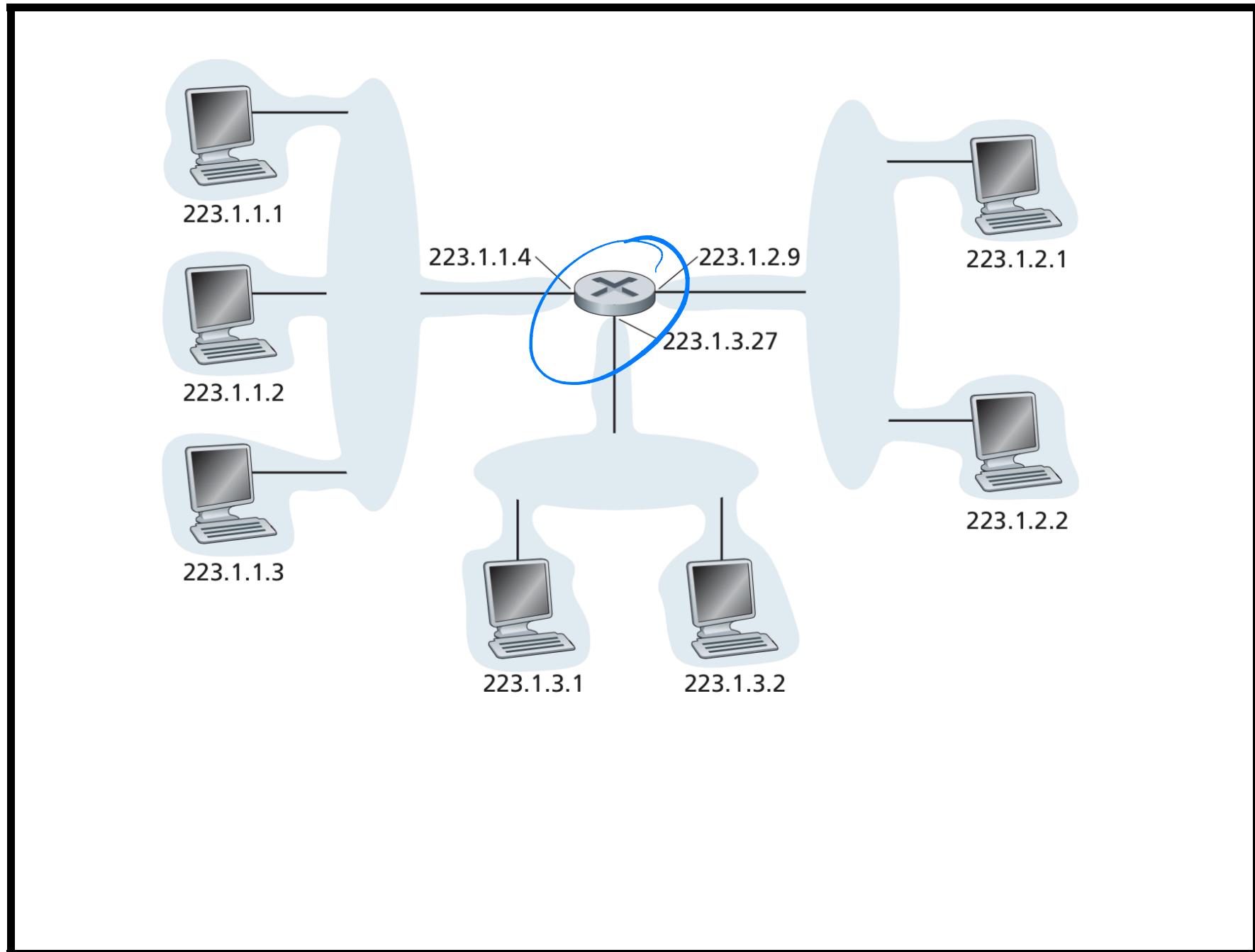
- Addresses broken into four bytes and written as: $X.Y.Z.W$ in decimal: $128.135.68.5 = 10000000.10000111.01000100.00000101$.

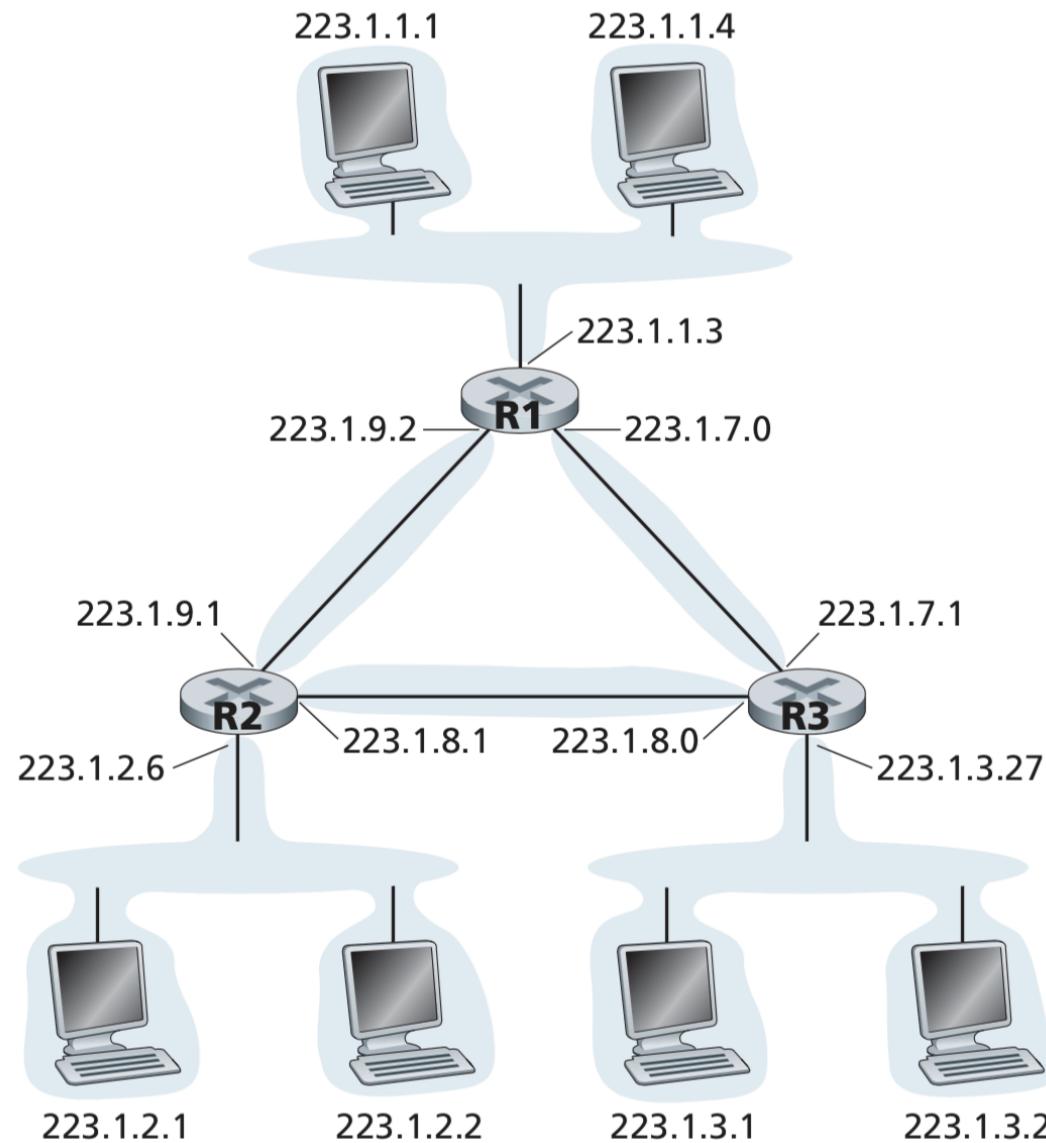


- Class 10, Net ID 00000010000111, Host ID 010001000000101.
- 127.Y.Z.W is a special “loopback” address: packet with this address returns back to host









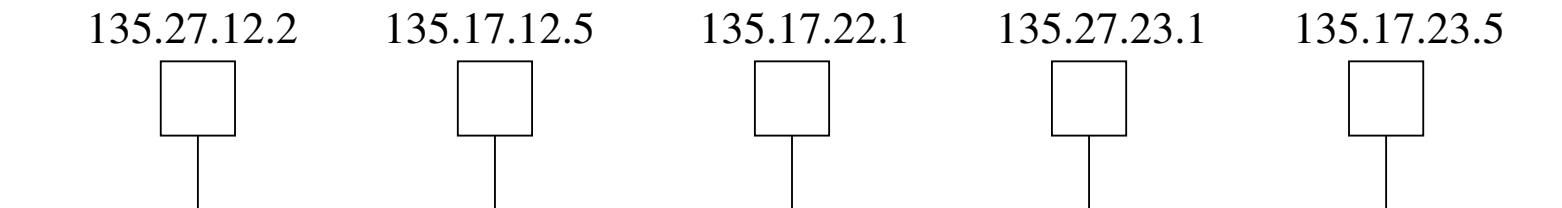
IP Addressing and the Network

- An IP address defines the host's connection to its network.
- A device connected to more than one network (e.g., router) will have more than one internet address.
- In fact, a different address for each network connected to it.
- To reach a host on the internet:
 - first we reach the network using the first portion of the address,
 - second we reach the host itself using the last portion of the address.
- Hence, Classes A, B, C have only two levels!

IP imposes a
hierarchical routing
scheme!

Two-level Hierarchy

A network may have a Class B address.



$135.27.0.0$

Class B

Because it is a two level hierarchy they cannot be grouped into a “less flat” scheme.

Solution: Subnetting!

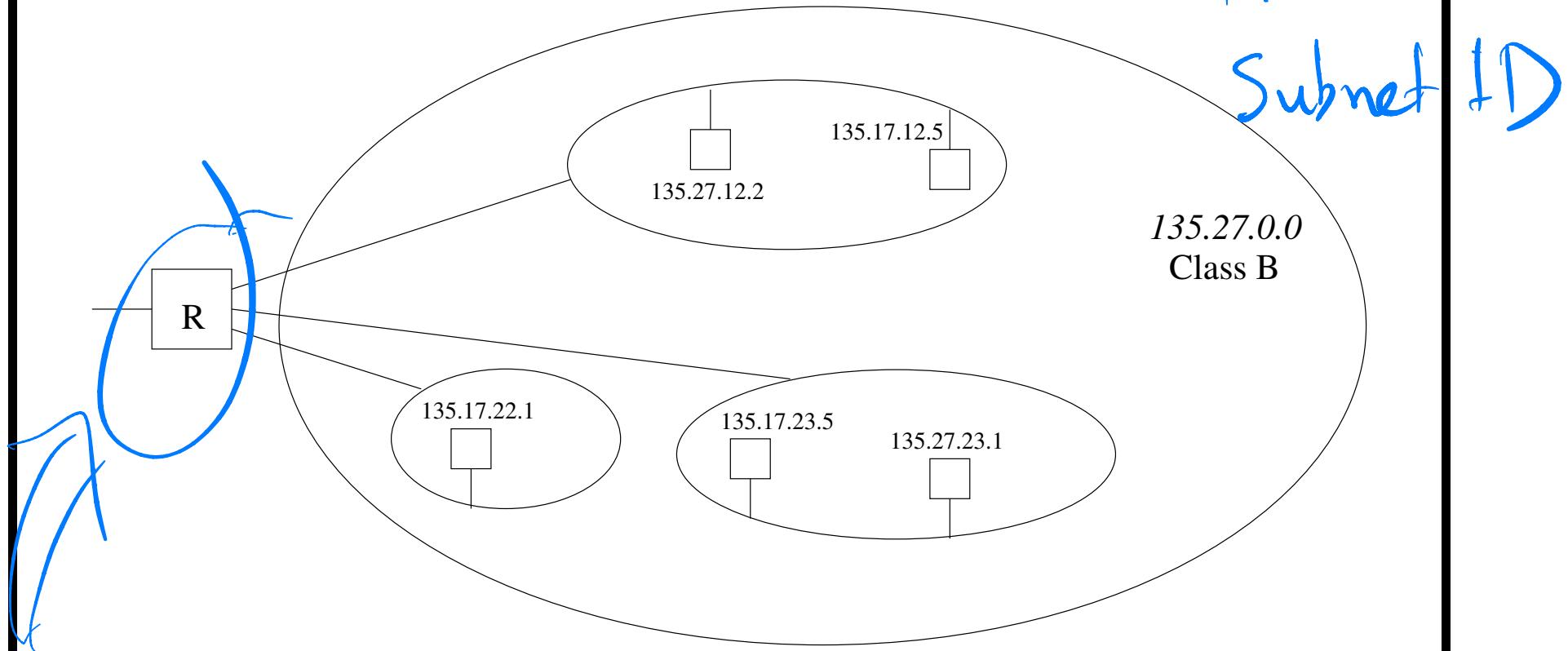
Subnetting

Subnetting

The rest of the internet does not need to be aware of the subnet division! The Router is aware of subnetting!

64,000

Subnet ID

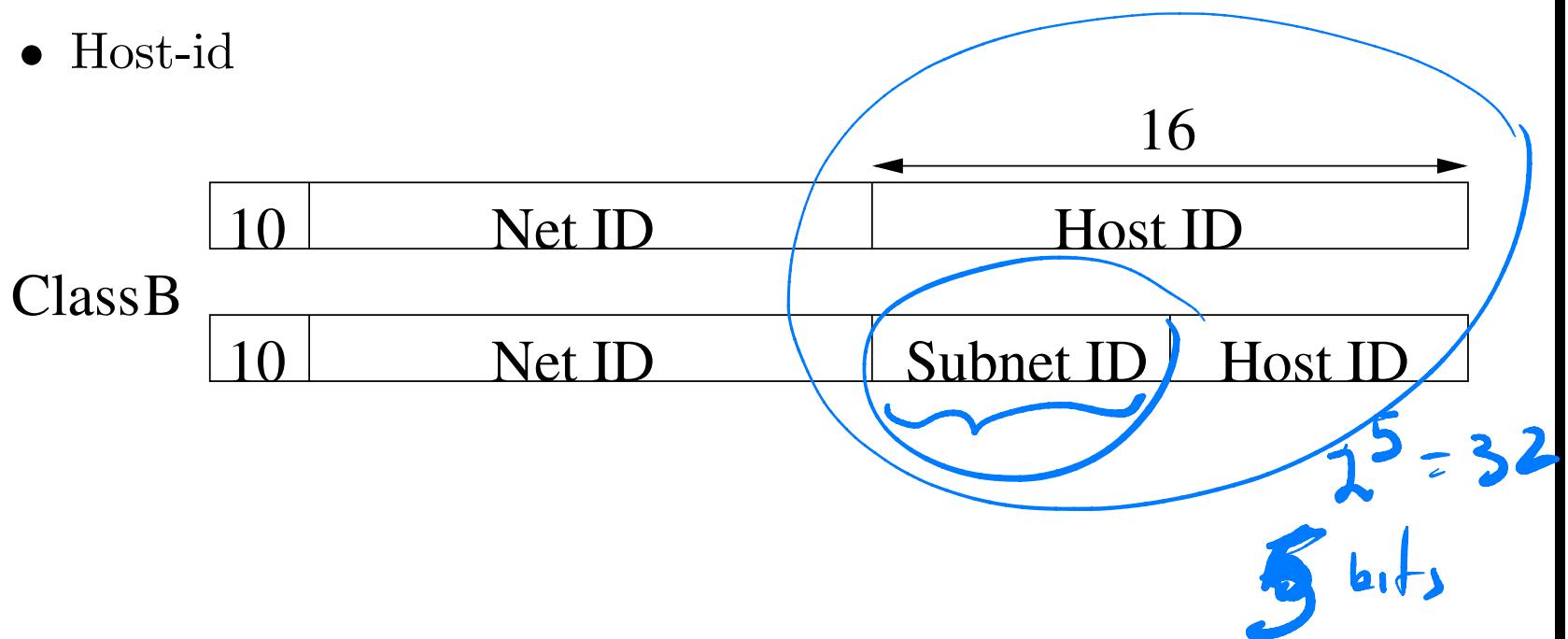


135.27 is the destination net-id, while 22.1 is a host-id.

More Levels

Now we have three levels in the hierarchy:

- Net-id (135.17)
- Subnet-id (12, 22, 23)
- Host-id



More Levels: Example

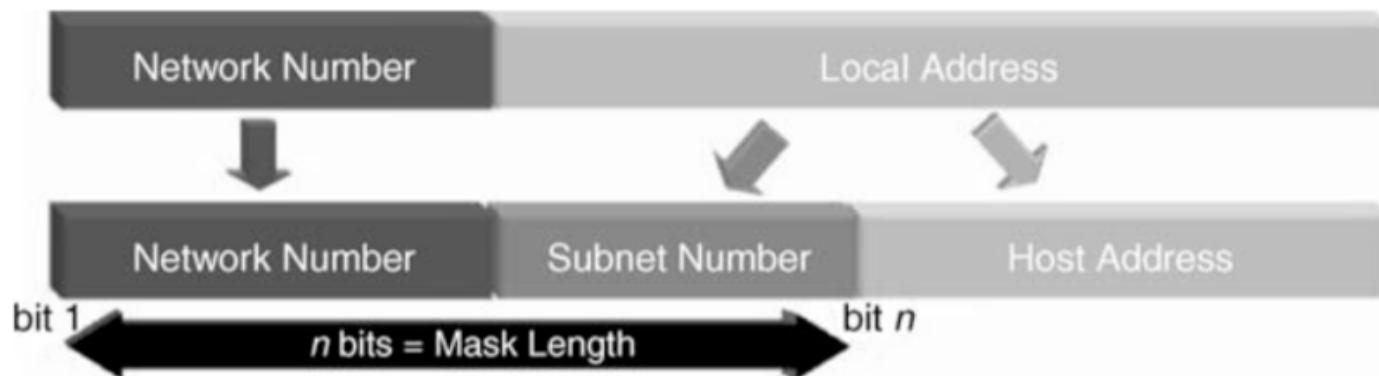
As an example consider an IP address with Net ID 150.100.

If an organization has many LANs each of less than 100 hosts then

- 7 bits suffice to represent hosts within each subnet, and
- remaining 9 bits are left to identify subnets within organization.

IP Subnetting

- A subnetwork or subnet is a logical subdivision of an IP network. The practice of dividing a network into two or more networks is called subnetting.
- Subnetting provides routing boundaries for communications and routing protocol updates.



- Subnetting is facilitated by specifying a network mask along with the network address.

IP Forwarding

IP Datagram Forwarding Algorithm: Source → Destination

To forward IP packets:

- { 1. Participating nodes compare network part of destination address to see if they are connected to same network as destination.
- { 2. If match occurs and destination is in same physical network the packet can be directly delivered.
- { 3. If node is not connected to same physical network all destination datagrams are sent to a router (in general there could be more than one router).
- { 4. The router does this by consulting a forwarding table.
5. There is also a default address used if none of the entries of the forwarding table match the destination address.

In conjunction with TCP

Address Resolution

Address Resolution

- How to re-translate an IP address to an address understood by a local host, e.g. an Ethernet address?
- Each host maintains a table of address pairs:
(IP-address, Physical-address)
Table can be managed either by an administrator or better yet dynamically by the host.
- Translation accomplished by ARP (Address Resolution Protocol).
- ARP enables hosts to build such tables. Moreover, ARP performs updating approximately every 15 minutes.
- ARP performs queries that take advantage of broadcasting capabilities of local networks, like Ethernet.

Mapping IP-Addresses \leftrightarrow Ethernet-Addresses

0	8	16	31		
Hardware Type = 1		ProtocolType = 0x0800			
HLen = 48	PLen = 32	Operation			
SourceHardwareAddress					
SourceHardwareAddress		SourceProtocolAddress			
SourceProtocolAddress		DestHardwareAddress			
DestHardwareAddress					
DestProtocolAddress					

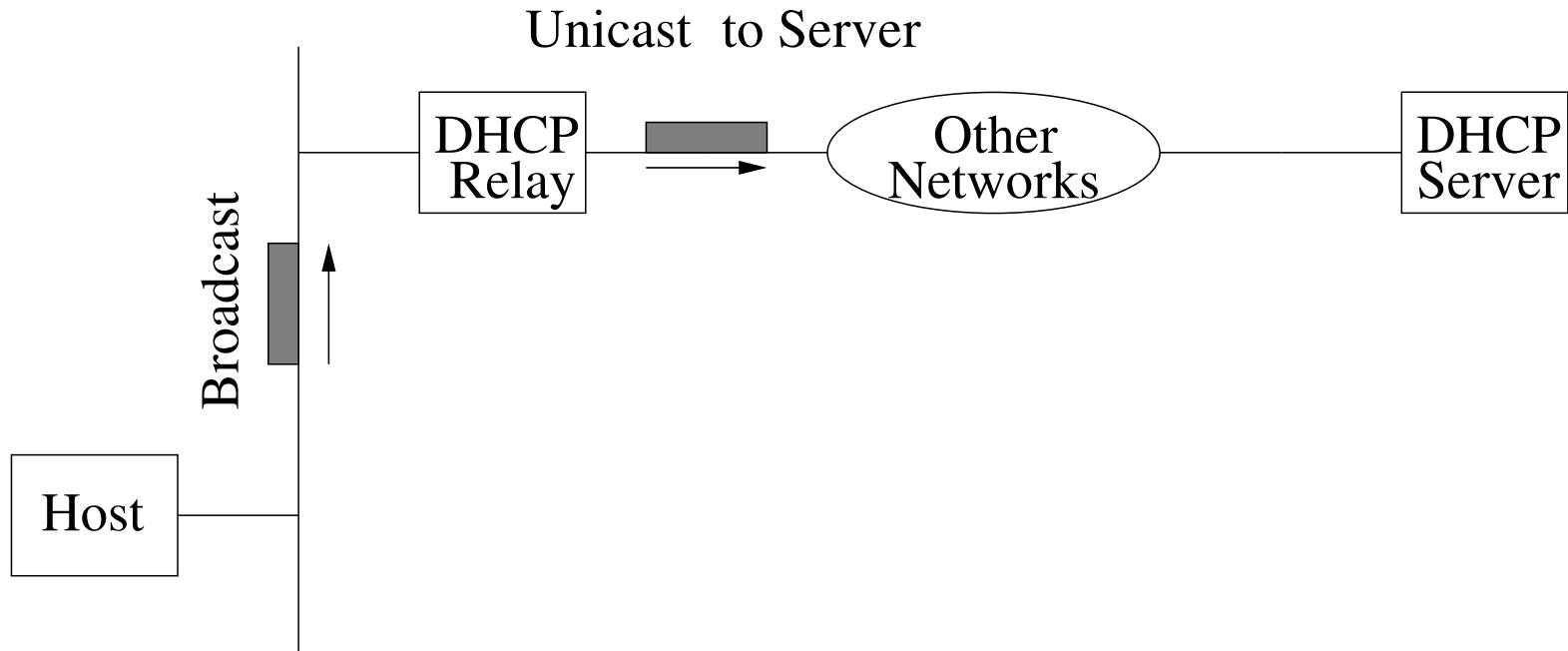
HardwareType is physical network (e.g., Ethernet), ProtocolType is higher level protocol (e.g., IP), HLen, PLen are the Hardware and Protocol address lengths, Operation specifies if this is a request for response.

IPv4: Not Enough Addresses

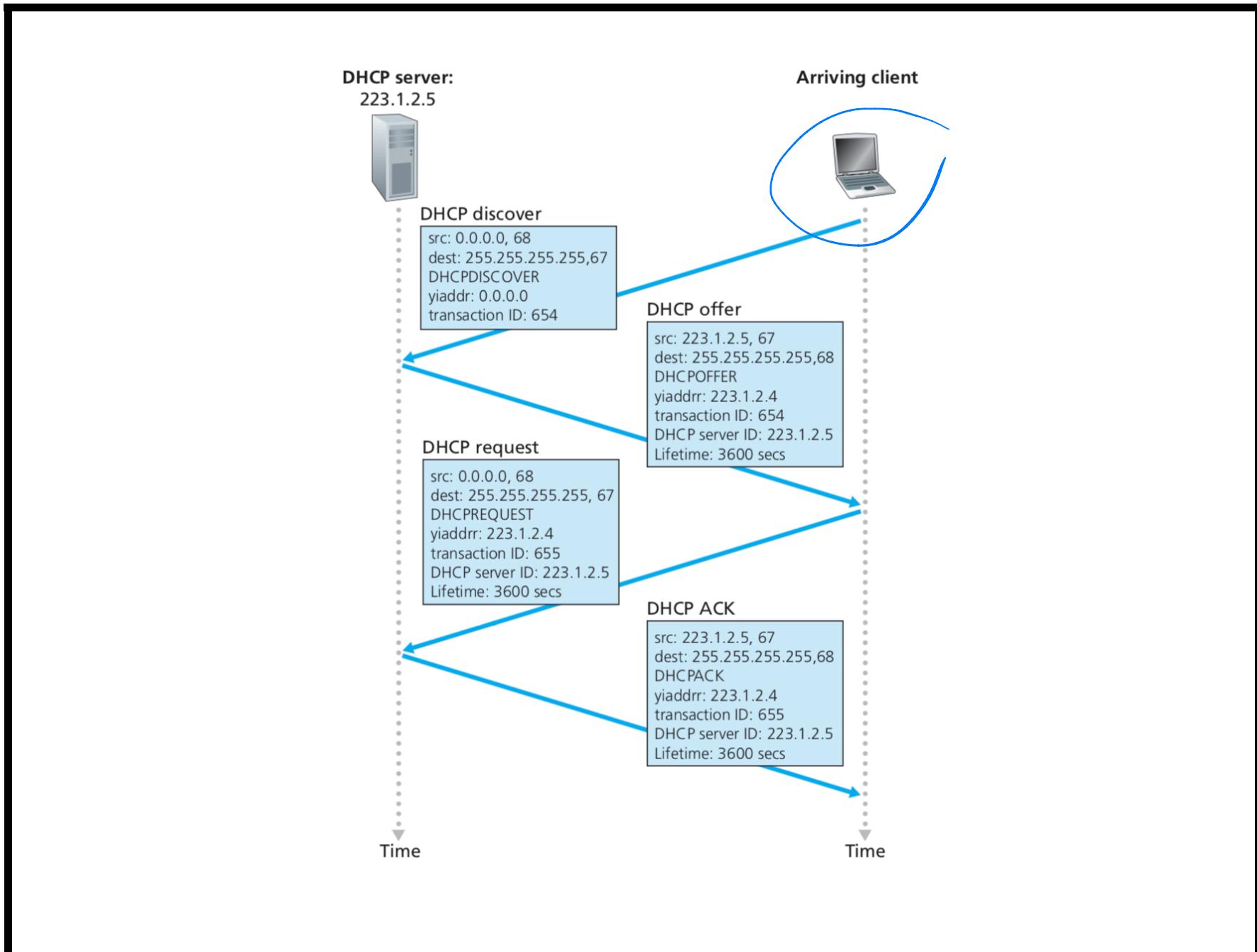
- Unlike Ethernet addresses, IP addresses cannot be configured once and for all into a host at manufacturing.
- Before it can start sending packets the host also needs to know the address of a default router.
- Operating systems enable IP address configuration. Configuring them directly and/or manually can be a lot of work.
- The protocol that automates this process is called DHCP (Dynamic Host Configuration Protocol).
- Every administrative domain (e.g., large company) has at least one DHCP server: DHCP saves the administrator from having to walk around each host.
- Information is stored in a table.

Dynamic Host Configuration Protocol (DHCP)

Newly booted hosts sends a DHCPDiscover message through a DHCP relay.



Used widely by Internet Providers because it maximizes usage of their limited number of IP addresses.



DHCP Packet Format

Operation	HType	HLen	Hops		
Xid					
Secs		Flags			
ciaddr (client IP address)					
yiaddr (your IP address)					
siaddr					
giaddr					
chaddr (client hardware address)					
sname (server name)					
file					
options (defaults, etc)					

Message sent using UDP (runs over IP): it provides a demultiplexing key that says **I am a DHCP packet.**

How DHCP Works

Host broadcasts a DHCP Discover message in its physical network.

Network server(s) respond with a DHCP Offer message.

Host selects one of the offers and broadcasts a DHCP Request message (that includes IP address of server).

Server acknowledges message with a DHCP ACK and assigns IP address for a period of time T with two thresholds T_1, T_2 (usually, $T_1 = T/2$ and $T_2 = 7T/8$).

When T_1 expires host attempts to extend lease by sending DHCP Request to same server. If accepted host also gets new values T', T'_1, T'_2 . If host does not receive DHCP ACK by time T_2 then it broadcasts to any server in the network.

If no ACK is received by time T then host must relinquish old IP address and begin anew.

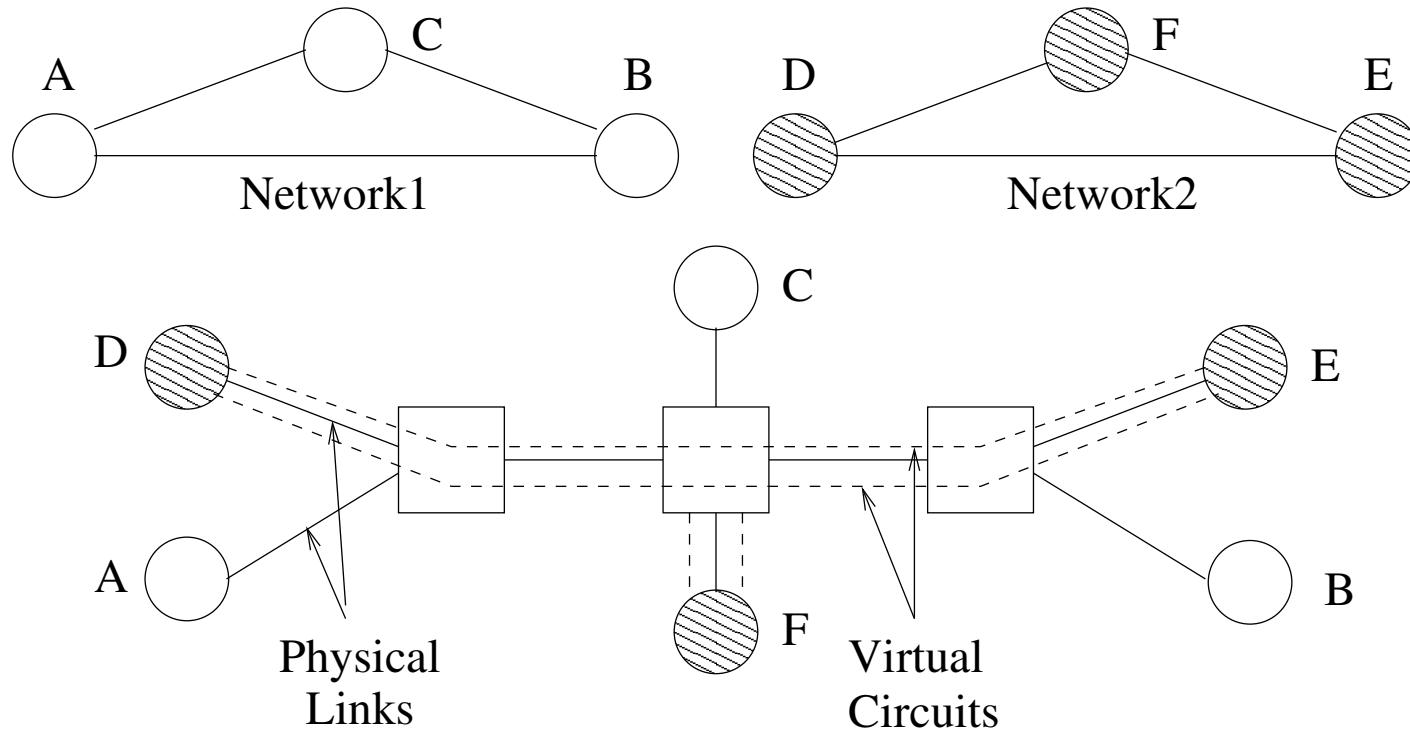
Reporting Errors

- Whenever a router or host is unable to deliver a message, IP reports errors.
- IP has a companion protocol, called Internet Control Message Protocol (ICMP) that defines error messages, including unreachable destination host, failure of datagram arrival, etc
- Check in your UNIX machine the command **ping!**
- ICMP control messages include: ICMP-Redirect, that tells the source host that there is a better route to destination.

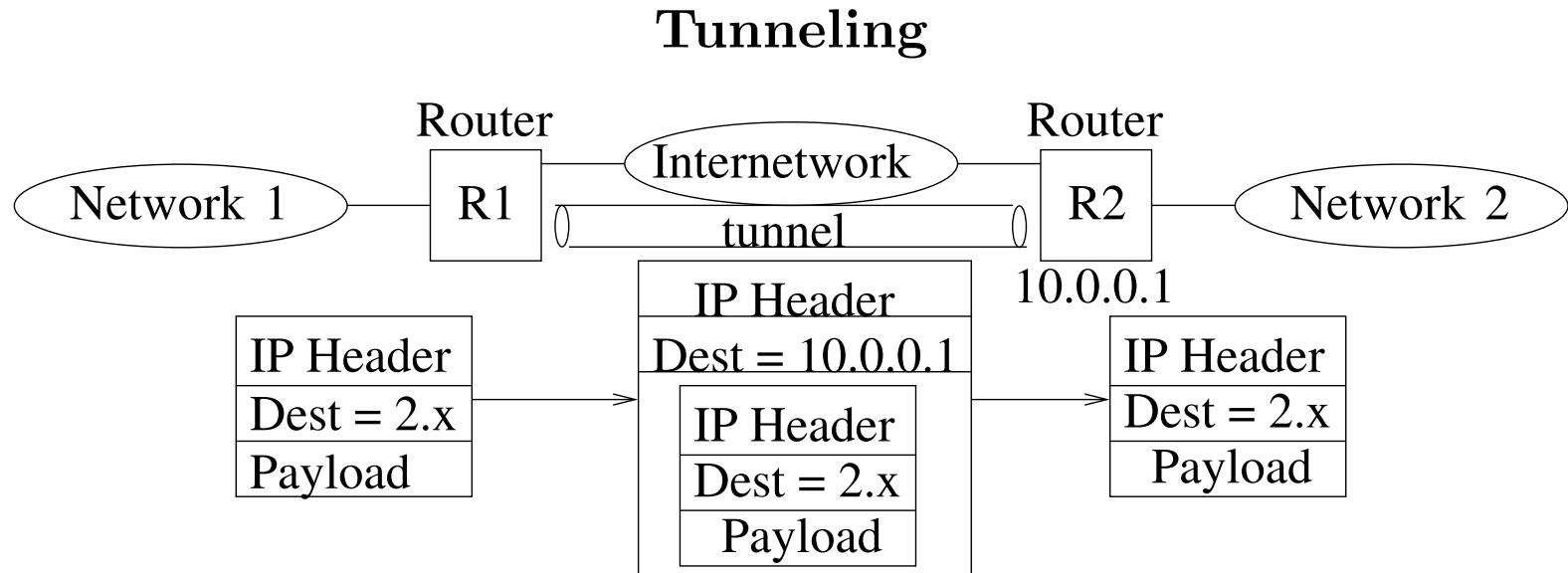
Tunneling

Tunneling

Old Virtual Private Networks



Two separate private networks, Network1 and Network2, can share common switches.



Original packet is encapsulated, receives a new IP header and the address of the destination router. It is then transported through the Internetwork. At the destination router it is decapsulated and forwarded to its “local” destination.

IPv4 to IPv6

IPv6

- Surge in demand for IP address space stimulated the IETF^a to define a new version of the Internet Protocol that would provide more addressing capacity to meet then and anticipated future address requirements.
- Every Regional Internet Registry (RIR) had issued notifications to the Internet community at large that IPv4 space availability is limited and will be exhausted within “a few years.”
- The primary objective for version 6 was essentially to redesign version 4 based on the prior 20 years of experience with IPv4.

^athe engineering and standards body of the Internet

IPv6 Wishlist

- IPv4 not sufficient!
 - Support for real time services.
 - Security support.
 - Autoconfiguration.
 - Enhanced routing functionality (e.g., mobility support).
- Also a transition plan IPv4 → IPv6 was necessary.

IPv4 Address Depletion

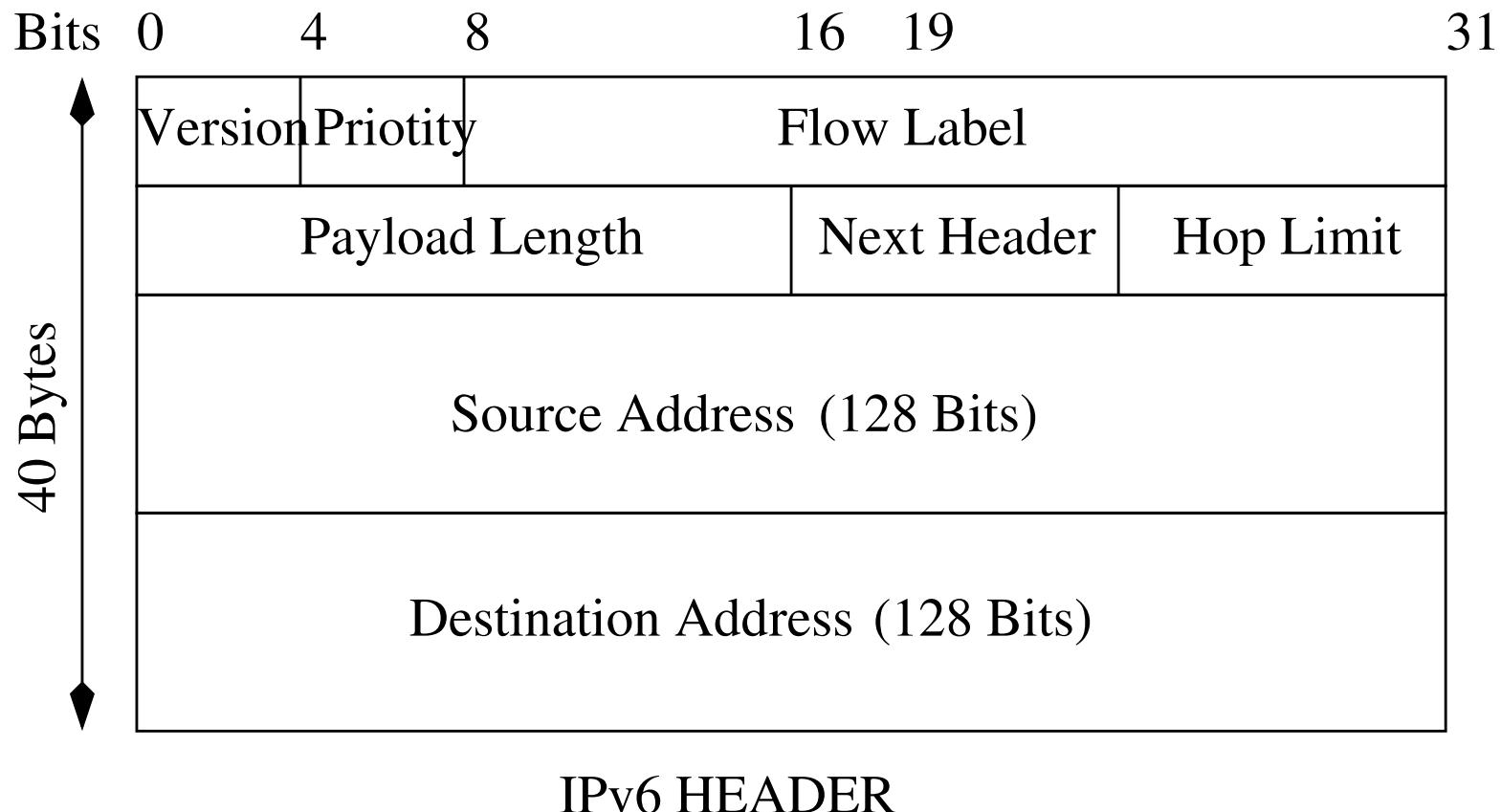
- February 3, 2011: in a ceremony in Miami, the Internet Assigned Numbers Authority (IANA) assigned the last batch of address blocks to the Regional Internet Registries (RIRs): over 80 million combined potential addresses (officially depleting the global pool of completely fresh blocks of addresses).
- APNIC was the first RIR to exhaust its regional pool on 15 April 2011.
- There remains a small amount of address space reserved for the transition to IPv6, which will be allocated in a much more restricted way.

IPv6 Header

- IPv6: designed to accommodate higher speeds. IPv4 uses 32 bit address space. IPv6 uses 128 bit address. IPv4 can address up to 2^{32} (= 4 billion) nodes. IPv6 can address up to $2^{128} = (2^{32})^4$ hosts.
- IPv6 Format: An IPv6 packet has the form: IPv6-header, extension field, . . . , extension header, format level PDU (Protocol Data Unit).
- IPv6 header:
- Priority Field: defines types of traffic.
- Flow labels: e.g. multimedia traffic consists of audio flow, video flow, data flow.

IPv6 Header

Has less fields than IPv4 (for less processing).



IPv6 Addressing (No Classes Being Used)

Prefix	Use	Prefix	Use
0000 0000	Reserved	0001	Unassigned
0000 0001	Unassigned	001	Global Unicast
0000 001	NSAP allocation	1111 1110 10	Link Local Use
0000 010	IPX allocation	1111 1110 11	Site Local Use
0000 011	Unassigned	1111 1111	Multicast
0000 1	Unassigned		

Large address chunks unassigned to allow for future growth.

NSAP used for ISO, **IPX** for Novell. **Link local** and **Site local** enable address construction without concern for global addresses (useful for autoconfigurations), **Multicast** is for multicast addresses, by zero extending with a byte of 0s one assigns IPv4-compatible and IPv4-mapped IPv6 addresses.

IPv6 Packet with Extension Headers

IPv6 treats options as extension headers appearing in certain order.
Thus routers can determine relevant options quickly.

	Size (Bytes)
IPv6 Header	40
Hop-by-Hop Options Header	Variable
Routing Header	Variable
Fragment Header	8
Authentication Header	Variable
Encapsulation Security Header	Variable
Destination Options Header	Variable
TCP Header	20
DATA	Variable

Assigning Addresses

- Three types of addresses: Unicast, Anycast (different interfaces), Multicast (different nodes).
- Hop-by-Hop Options Header: carries optional information that must be examined (like next header, header extension length, options).
- Fragment Header: fragmentation is done only by source nodes not routers. Nodes perform a path discovery algorithm to determine the smaller max transmission unit. With this knowledge source nodes fragments data. Fragment header has several flags and data itself.
- Routing Header: contains a list of one or more intermediate nodes to be visited along the way. Fields include: Next Header, Header Extension Length, Routing Type, etc Destinations
- Options Header: carries optional information.

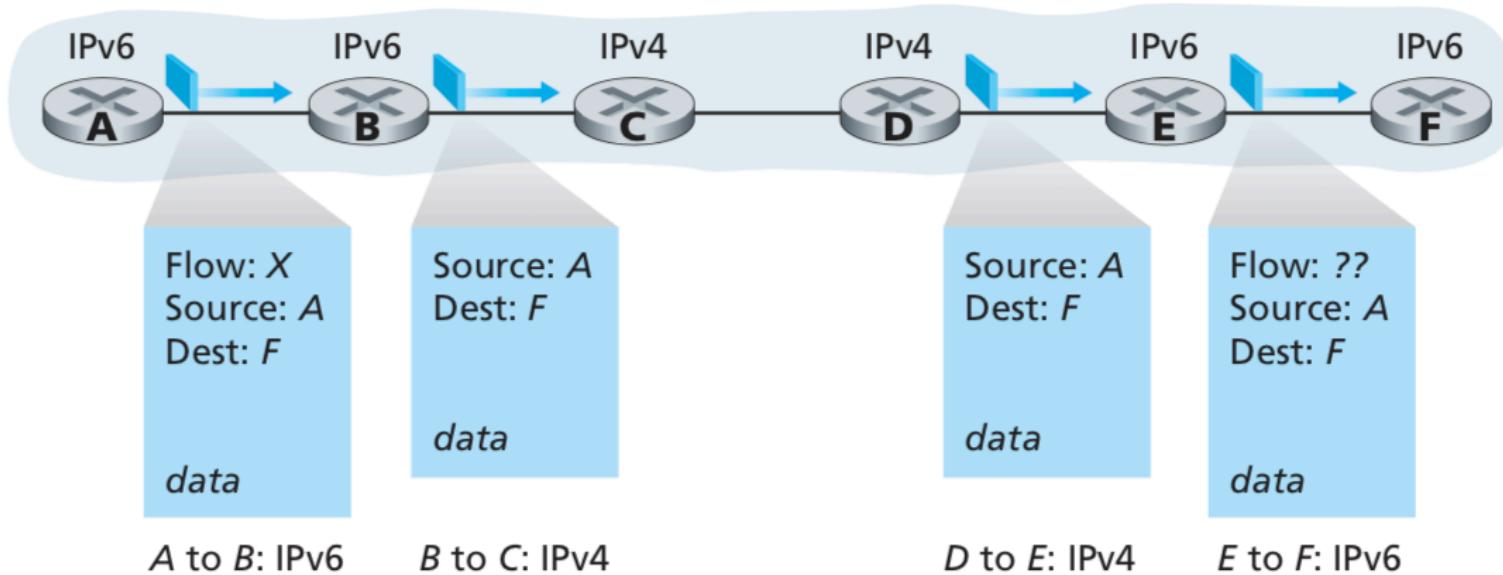
IPv6 Address Notation

- Hexadecimal digits are being used; represented in eight 16-bit blocks. block1:block2:block3:block4:block5:block6:block7:block8
- One set of contiguous 0s can be omitted: block1::block7:block8
- An IPv4-mapped address, like 128.33.87.51 is now written as :: 00FF : 128.33.87.51
- 001 prefix used for global unicast addressing.
- 010 prefix used for IPv6 provider based address. Here, registry IDs are provided as common identifiers, e.g., European, American, etc.

IPv6 Address Notation

- DHCP provides IPv4 autoconfiguration. So does IPv6.. This is done as follows:
 1. obtain correct subnet address prefix (through a router), and
 2. unique interface ID (like Ethernet address).
- IPv6 provides for anycast addresses: selects one of a set of any. Also multicast and security provided.

Transitioning from IPv4 to IPv6

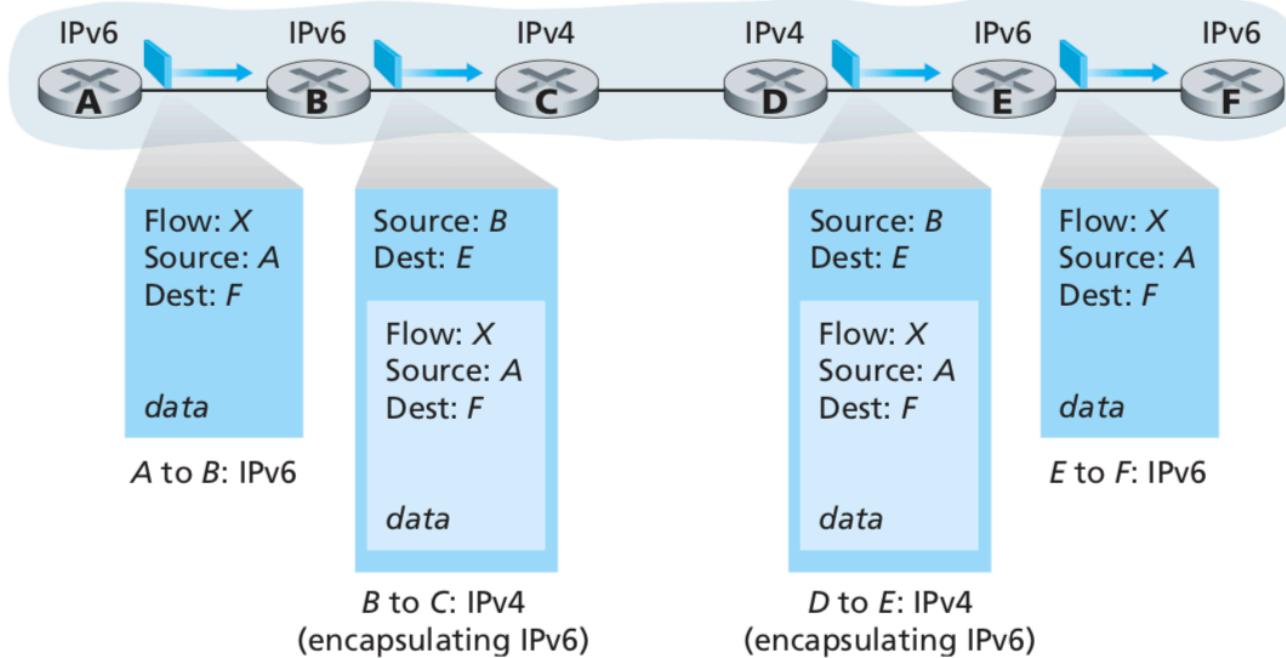


Tunneling in IPv6

Logical view



Physical view



IPv6 Neighbor Discovery

- Enables a node to discover the IPv6 subnet address on which it is connected.
- Enables IPv6 nodes to automatically identify routers on the subnet.
- The discovery process entails each router periodically sending advertisements on each of its configured subnets indicating its IP address, its ability to provide default gateway functionality, its link layer address, the network prefix(es) served on the link including corresponding prefix length and valid address lifetime, as well as other configuration parameters.

IPv6 Deployment (1/4)

- The introduction of Classless Inter-Domain Routing (CIDR) in the Internet routing and IP address allocation methods in 1993 and the extensive use of network address translation (NAT) delayed the inevitable IPv4 address exhaustion, but the final phase of exhaustion started on February 3, 2011.
- Despite a decade long development and implementation history as a Standards Track protocol, general worldwide deployment was still in its infancy: as of October 2011, about 3 % of domain names and 12 % of the networks on the internet have IPv6 protocol support.
- IPv6 has been implemented on all major operating systems in use in commercial, business, and home consumer environments.
- IoT (Internet of Things) is giving a significant boost to IPv6.

IPv6 Deployment (2/4)

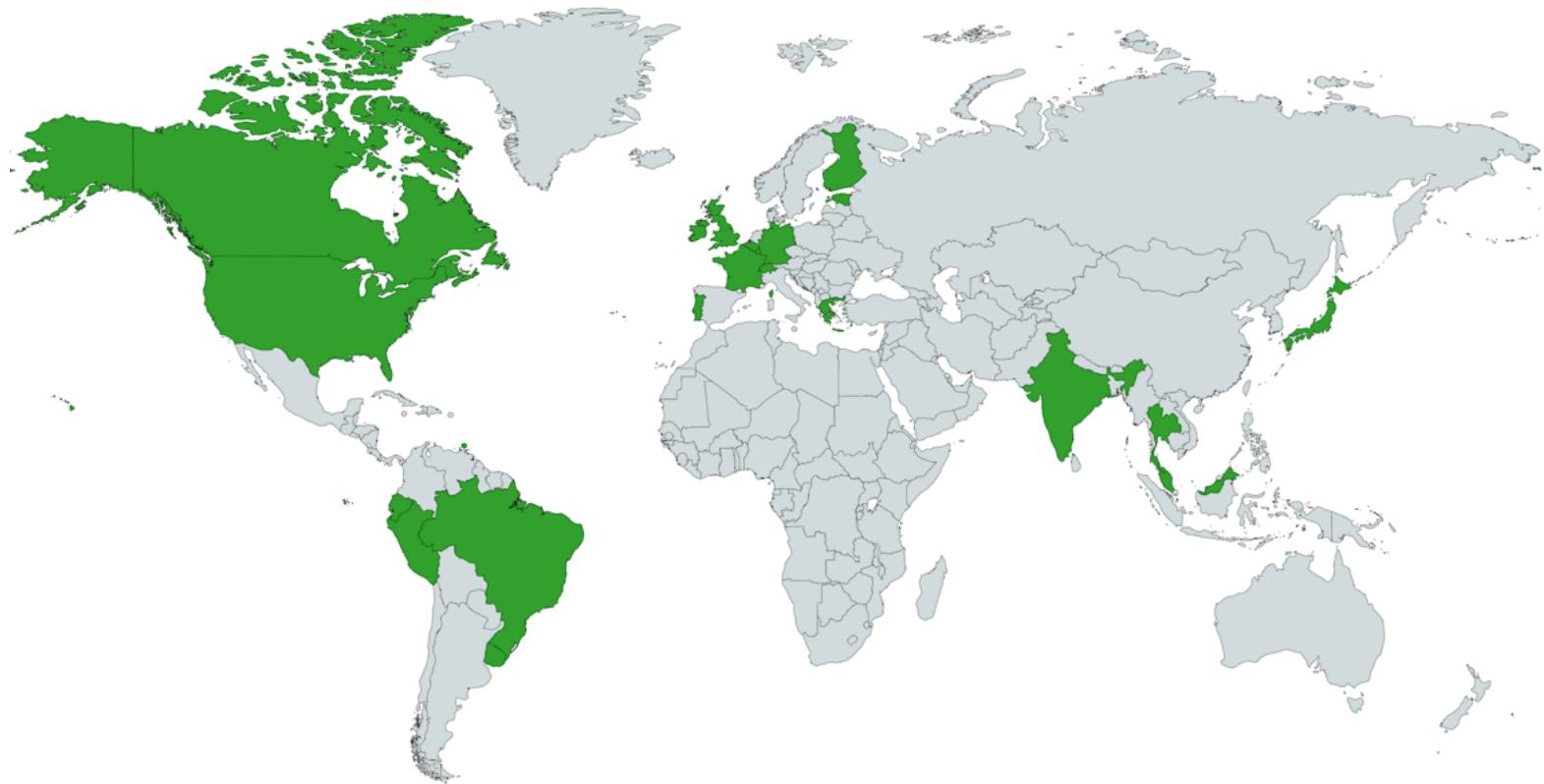
- Since 2008, the domain name system can be used in IPv6 at major web sites like Google, although sometimes with extra configuration.
- IPv6 was first used in a major world event during the 2008 Summer Olympic Games, the largest showcase of IPv6 technology since the inception of IPv6.
- Countries like China or the Federal U.S. Government are also starting to require support for IPv6 on their equipment.
- Finally, modern cellular telephone specifications mandate IPv6 operation and deprecate IPv4 as an optional capability.

IPv6 Deployment (3/4)

- IPv6 deployment continues to increase around the world.
- As of 2018
 - Over 25% of all Internet-connected networks advertise IPv6 connectivity.
 - 49 countries deliver more than 5% of traffic over IPv6, with new countries joining all the time.
 - In 24 countries IPv6 traffic exceeds 15%.

IPv6 Deployment (4/4)

- Countries with Pv6 Deployment greater than 15% as of 2018.^a



^aSource: Internet Society

References

- T. Rooney, Rooney-IP Address Management Principles and Practice, IEEE, 2011.
- RFC 760, DoD standard Internet Protocol, January 1980
- RFC 761, DoD standard Transmission Control Protocol, 1980.

Appendix

Masking

IP Masking

- It is called a subnet masking because it is used to identify network address of an IP address by performing a bitwise AND operation on the netmask.
- A Subnet mask is a 32-bit number that masks an IP address, and divides the IP address into network- and host-address.
- Class A network has a mask length 8, a class B 16, and C 24.
- By essentially extending the length of the network number that routers need to examine in each packet, a larger number of networks can be supported, and address space can be allocated more flexibly.
- For IPv4, a network may also be characterized by its subnet mask or netmask, which is the bitmask that when applied by a bitwise AND operation to any IP address in the network, yields the routing prefix.

IP Addressing and Subnets: Masking

- A subnet mask separates the IP address into the network and host addresses (*<network><host>*).
- Subnetting further divides the host part of an IP address into a subnet and host address (*<network><subnet><host>*) if additional subnetwork is needed.
- Masking extracts the address of the physical network from an IP address.
- If there is no subnet it merely extracts the network address from the IP address.
- If there is subnet division then it extracts the subnet address from the IP address.
- Masking is also used to “hide” addresses.

Global Connectivity and Scalability: Subnets

It is a mistake to assign one network number per physical network.

Address assignment inefficiencies arise:

1. A network with three nodes may be using an entire class C network address (thereby wasting 252 useful addresses).
2. A network with slightly more than 255 hosts is using a class address thereby wasting 64,000 addresses.
3. The more forwarding numbers exist the larger the forwarding tables.

Subnetting

- **Subnetting** takes a single IP network address and allocates it to several physical networks referred to as subnets, e.g., a University Campus, Company, etc.
- Subnets should be physically close to each other.
- Knowing the addresses of a few entry gateways should be enough.

Subnet Masks

1. The mechanism allowing a network number to be shared among multiple networks is called **subnet masking**.
2. This gives rise to the **subnet number**: all hosts on the same network have the same subnet number.
3. Hosts on different physical networks share single network #.
4. Subnet masks introduce another level of hierarchy into IP-addressing.
5. Subnet masks written down like IP addresses, e.g.
255.255.255.3

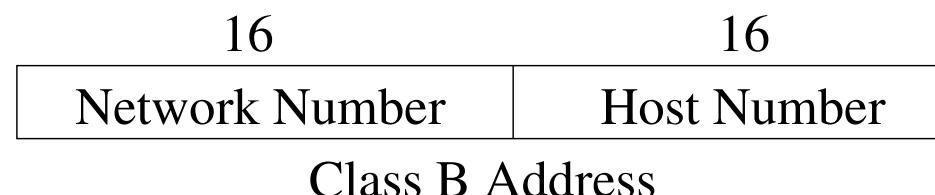
Subnet Masks

To share a single class B address use a subnet mask, say

$255.255.255.0 = 11111111.11111111.11111111.00000000$

I.e., lowest 8 bits are the host numbers.

Address now has three parts: network part, subnet part, and host part.



24 1s	8 0s
-------	------

Subnet Mask: 255.255.255.0



Subnet Masks (Example)

Suppose a router is given a destination address D and a pair (I, M) where I is an IP address and M a mask.

Router verifies the condition $I = D \& M$ (bitwise AND) to test whether or not D belongs to the same subnet.

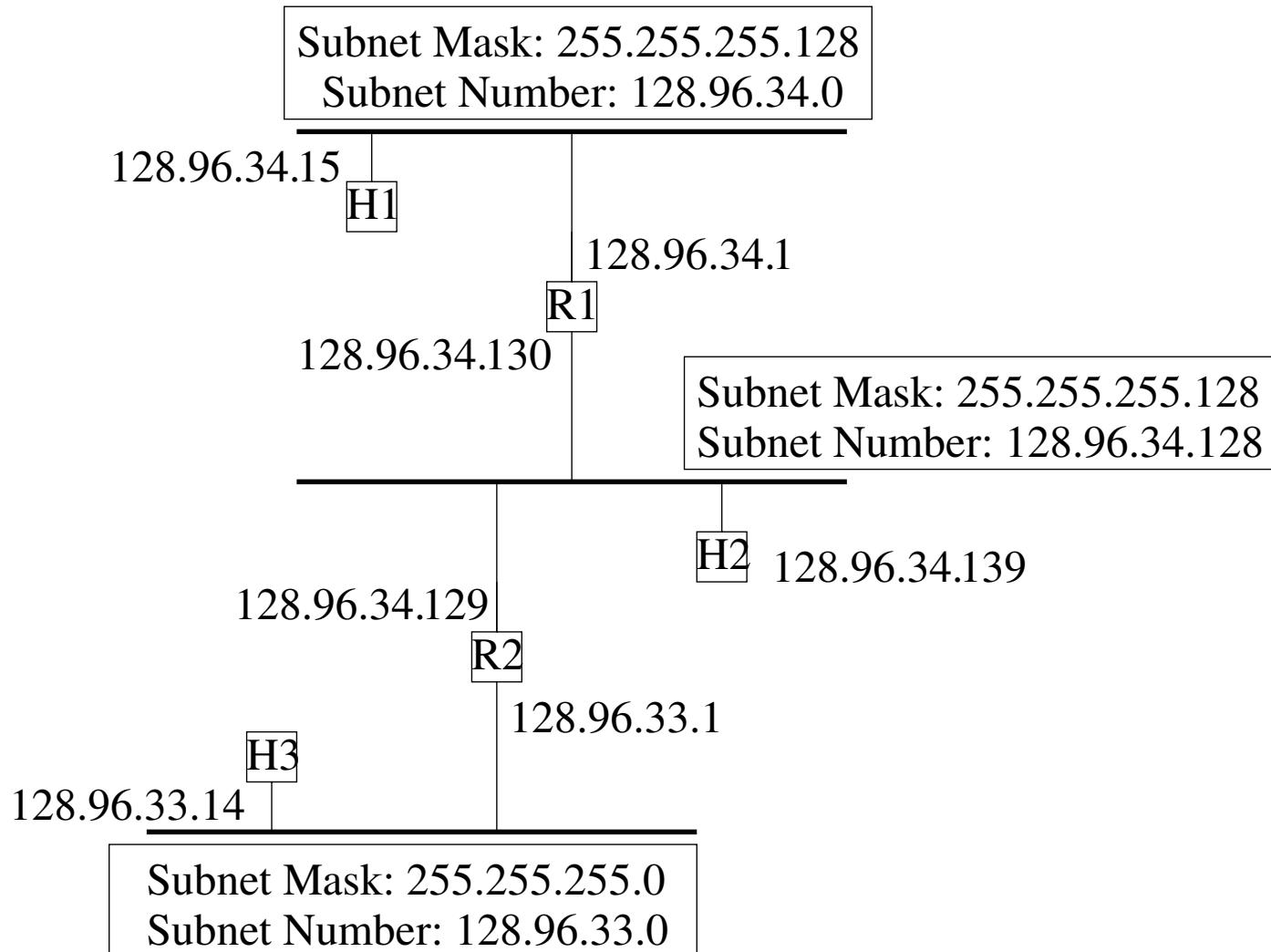
For example, in

$I = 128.10.0.0 = 10000000 \quad 00001010 \quad 00000000 \quad 00000000$

$M = 255.255.0.0 = 11111111 \quad 11111111 \quad 00000000 \quad 00000000$

$D = 128.10.2.3 = 10000000 \quad 00001010 \quad 00000010 \quad 00000011$

the mask M is used as follows: it “matches” the first sixteen bits of the IP address I (ignoring the rest).



Routers R1 and R2 have a different subnet number for each subnet

Subnet Masks and Subnet Number

1. Hosts are configured with an address and the subnet mask.
2. The bitwise AND of these two numbers defines subnet number of the given host as well as all the hosts in the same subnet.

Host	H1	H2
Subnet Number	128.96.34.15	128.96.34.139
Subnet Mask	255.255.255.128	255.255.255.128
BIT-WISE AND	128.96.34.0	128.96.34.128

H1 forwards to H2: H1 calculates AND of H2's subnet address (128.96.34.139) with subnet mask (255.255.255.128). If result is equal to H1's Subnet Number (128.96.34.128) then it is delivered to NextHop for H2 of its forwarding table. If it is not equal to H1's Subnet Number then packet is forwarded to H1's default router.

Masking: Example

If IP address 150.100.12.176 arrives from outside use a binary AND between IP address and mask 255.255.255.125 to determine subnet.

IP address 10010110.01100100.00001100.10110000

mask 11111111.11111111.11111111.10000000

AND 10010110.01100100.00001100.10000000

IP-address	Mask	Network/Subnetwork-address
141.14.3.22	255.255.0.0	141.14.0.0 (without subnetting)
141.14.3.22	255.255.255.0	141.14.3.0 (with subnetting)