ELEN 90061
Communication Networks

Module 4 – Network Layer

Lesson 1: Internet Protocol

Dr. Rajitha Senanayake

Lesson Overview

- Network layer, services, SDN, virtual circuits
- Recent WAN technologies and MPLS
- Internet Protocol (IP), IPv4
- IP Addressing, Subnets, DHCP, multicasting
- Network Address Translation NAT
- IPv6 overview

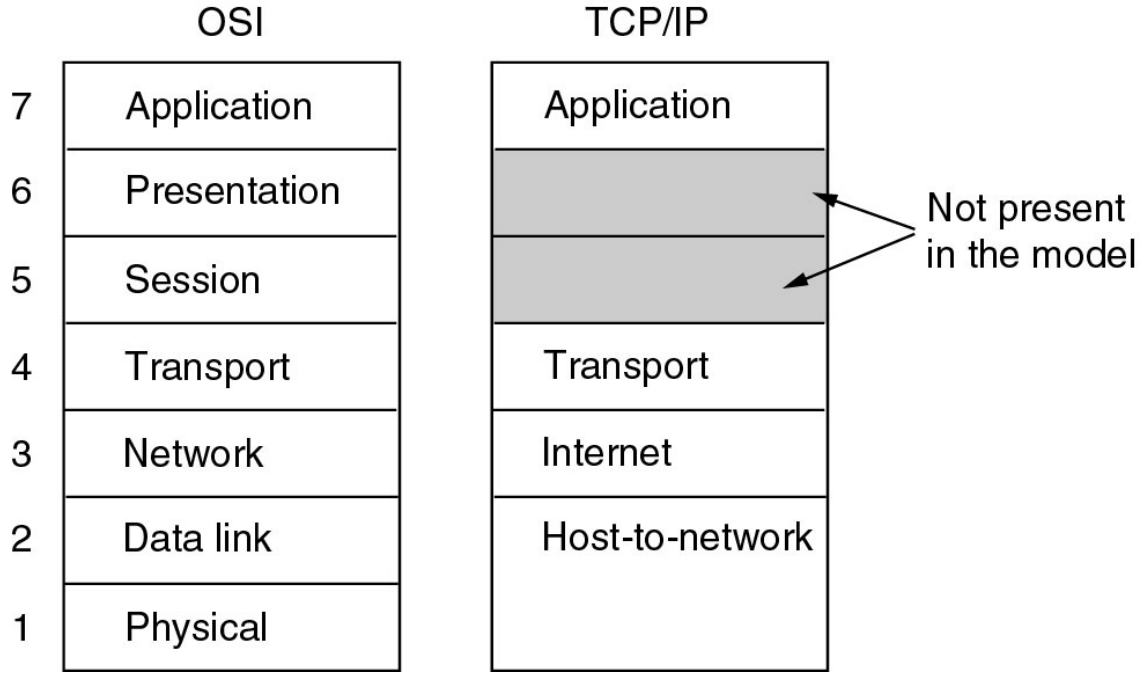
Suggested Reading

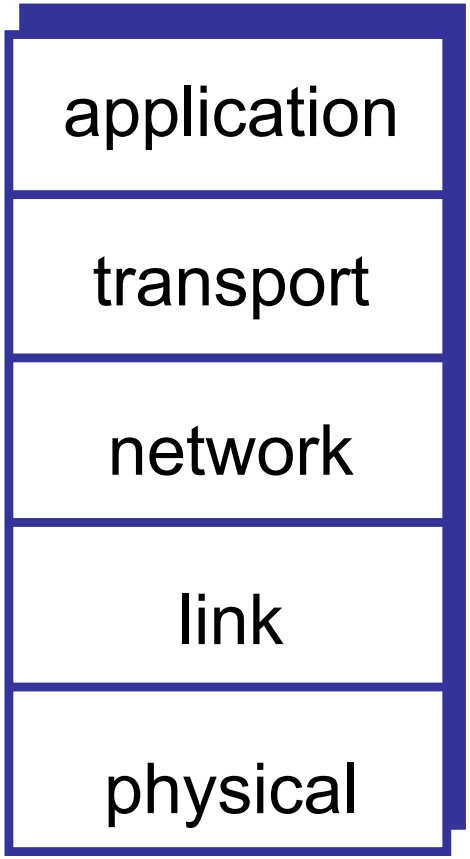
Note that there is overlap between these reading materials. It is a comprehensive list and you can use slides as a guideline for what to focus on.

- Chapter 5 from Tanenbaum
- Chapter 4 from Kurose-Ross

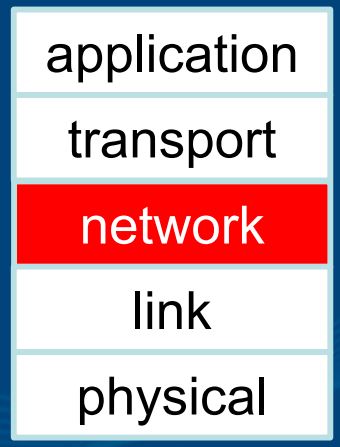
Reminder: Internet protocol stack

- Internet stack “missing”: Presentation, Session from OSI version
- these services, if needed, must be implemented in application



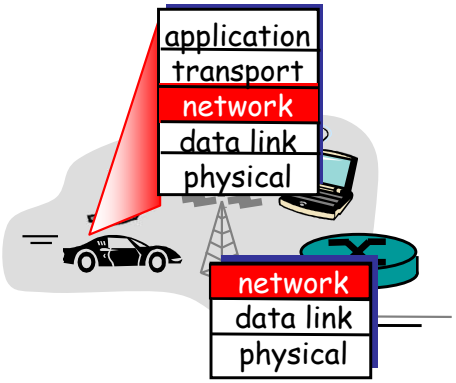


Network Layer: Overview and Services

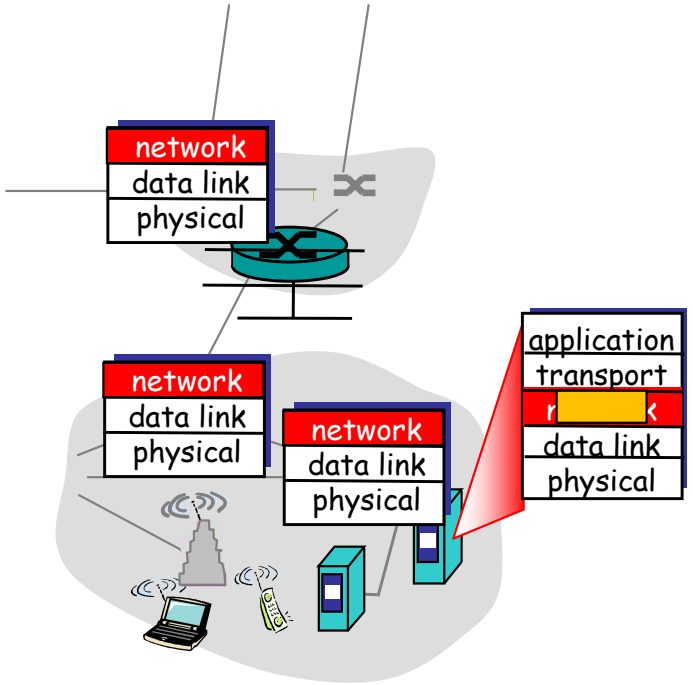


Network layer

- transports data from sending to receiving host
  - on sending side encapsulates data into *datagram (packet)*
  - on receiving side, delivers data to transport layer



- network layer protocols in *every* host, router
  - router examines header fields in all Internet Protocol (IP) *datagrams* passing through it.



Two Key Network-Layer Functions

- *forwarding:* move datagrams from router's input to appropriate router output
- *routing:* determine route taken by datagrams from source to destination
  - ***routing algorithms***

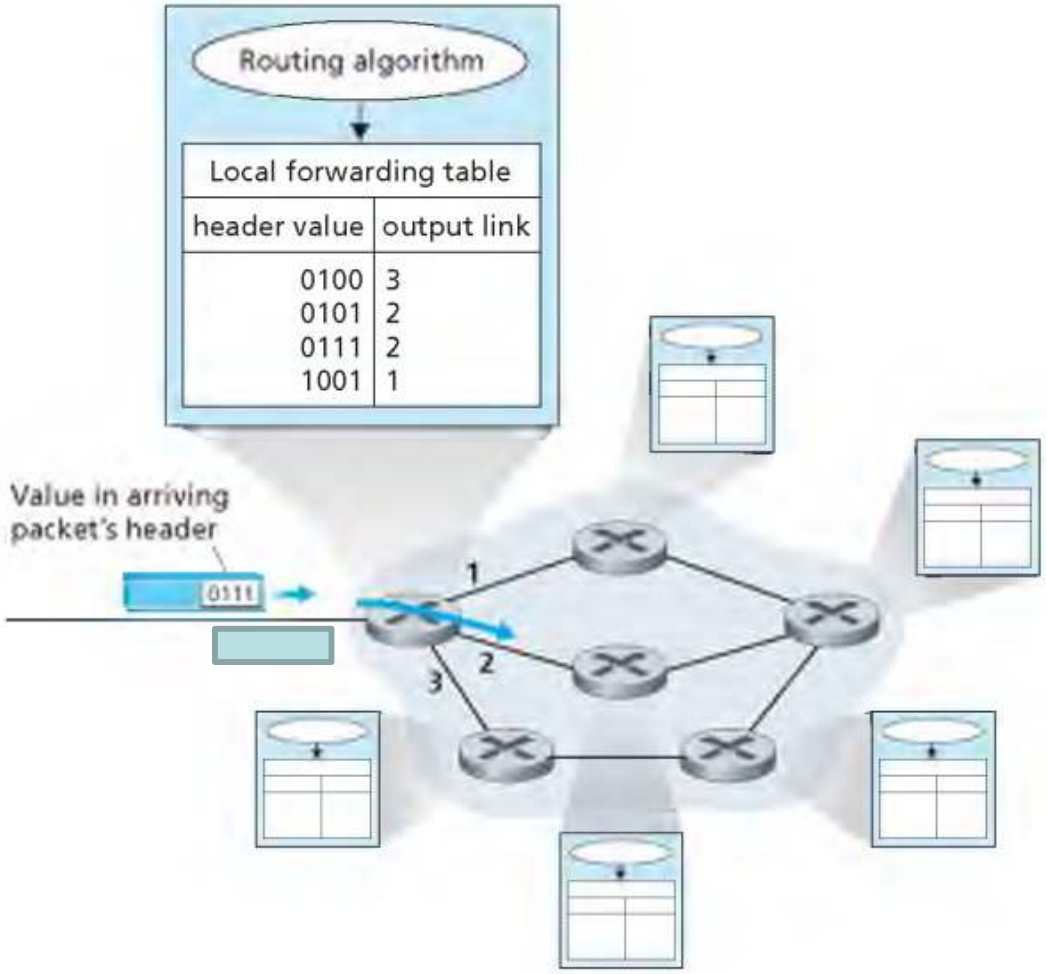
Road trip analogy:

- forwarding: process of driving from one city to another.
- routing: process of planning the whole trip from the source to destination

Question: *when does the analogy break down?*

Routing

- Every router has a forwarding table.
- *A router forwards a packet by using the value in the arriving packet's header and the forwarding table.*
- The values in the forwarding table indicate packets with which header values should be forwarded to which outgoing interface.
- *Routing algorithms* compute forwarding tables

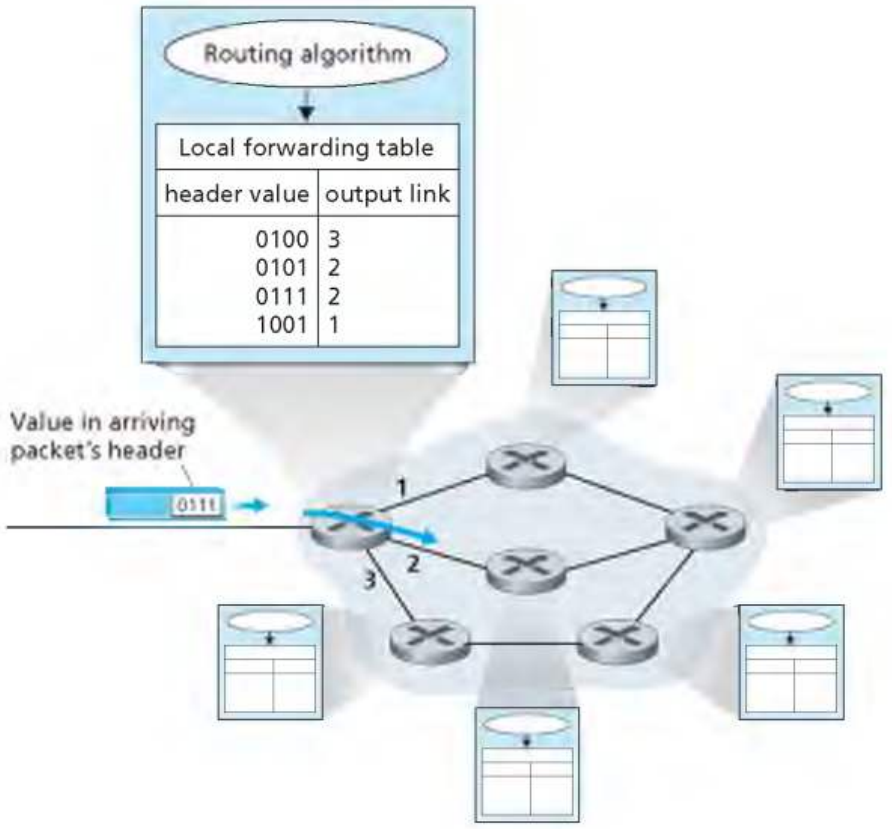


Prefix Matching

Router *matches a prefix of the packet destination address to entries in the forwarding table*.

Prefix Match	Link Interface
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

Destination Address Range	Link Interface
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1



Longest prefix matching

longest prefix matching

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** *********	0
11001000 00010111 00011000 *********	1
11001000 00010111 00011*** *********	2
otherwise	3

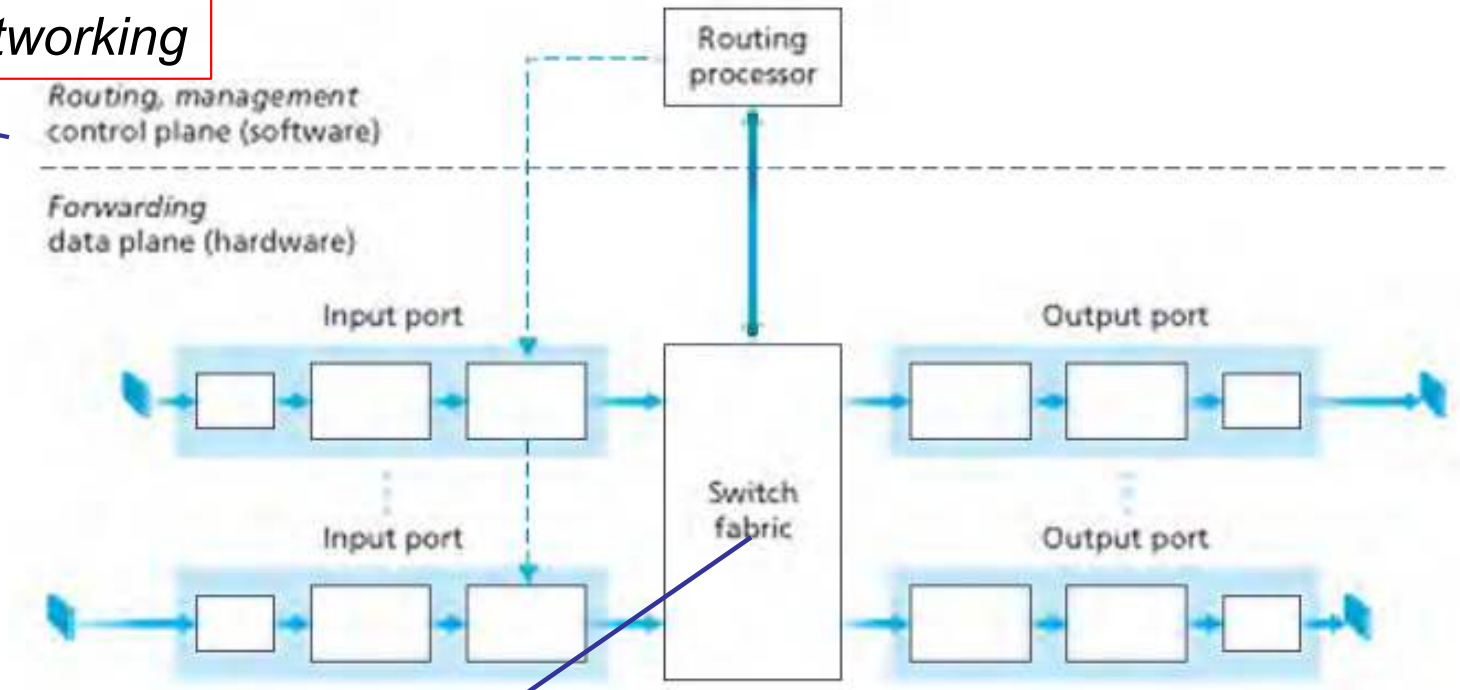
examples:

DA: 11001000 00010111 00010110 10100001 which interface?

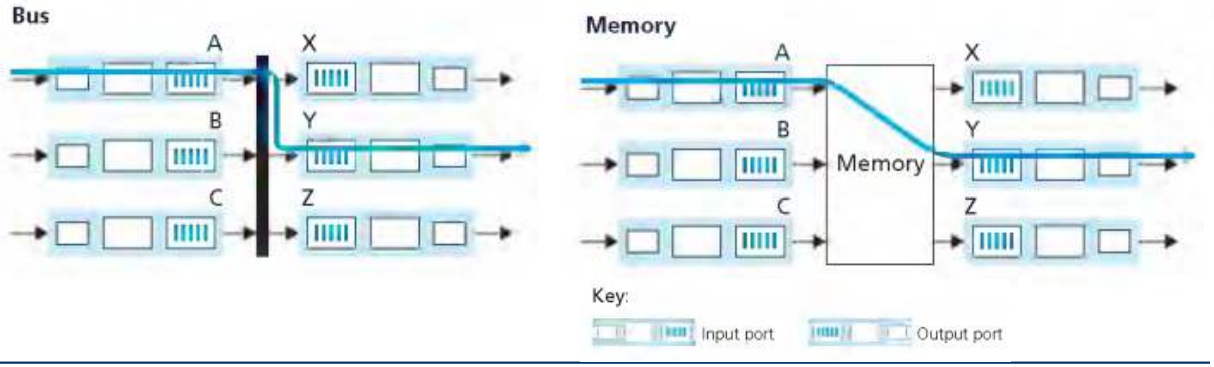
DA: 11001000 00010111 00011000 10101010 which interface?

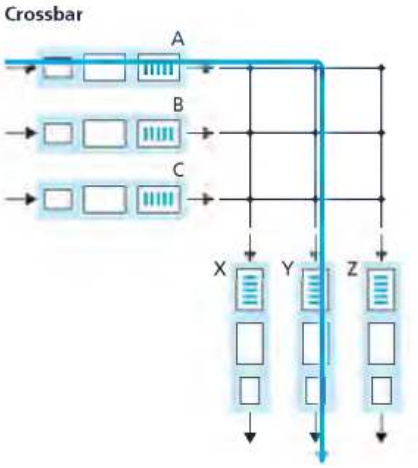
Anatomy of a Router

New development: **SDN**
Software-Defined Networking



Three switching techniques





Network service model

Q: What *service model* to use for the "channel" transporting datagrams from sender to receiver?

example services for individual datagrams:

- guaranteed delivery
- guaranteed delivery with less than, e.g. 40 msec delay (delay QoS guarantee)

example services for a flow of datagrams:

- in-order datagram delivery
- guaranteed minimum bandwidth to flow
- restrictions on jitter (variation in delay)

Network layer service models (ATM perspective)

Network Architecture	Service Model	Guarantees ?				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	**best effort**	none	no	no	no	no (inferred via loss)
ATM	CBR	constant rate	yes	yes	yes	no congestion
ATM	VBR	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

ATM: ***Asynchronous Transfer Mode*** ***(historical)***

CBR: Constant Bit Rate
VBR: Variable BR
ABR: Available BR
UBR: Unspecified BR

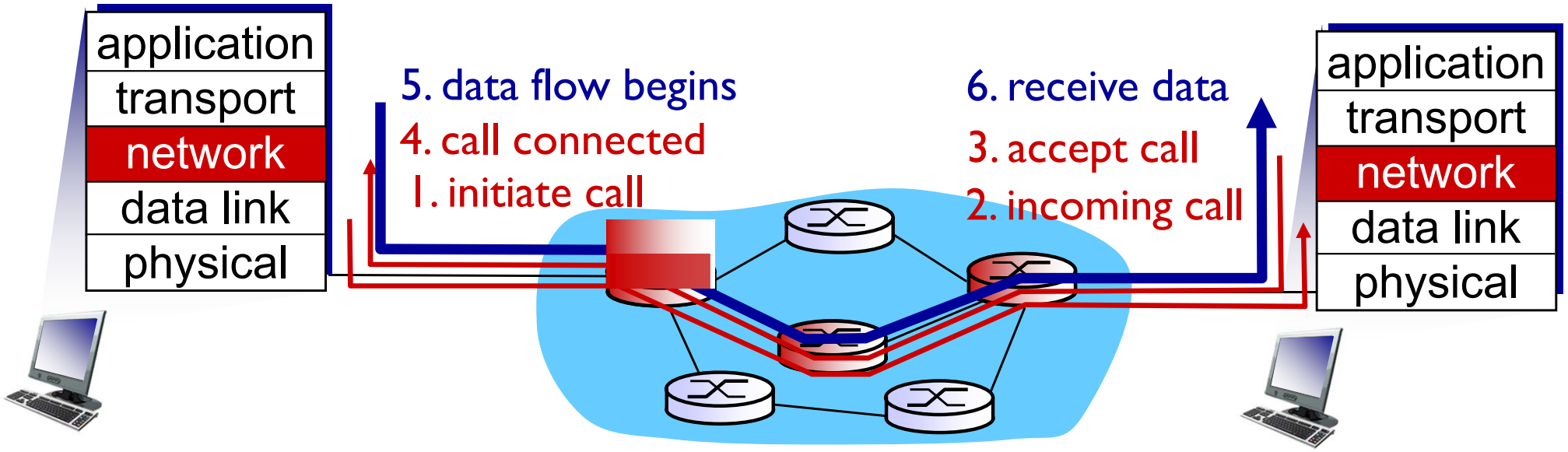
Virtual circuit concept

"source-to-dest path behaves much like telephone circuit"

- call setup, teardown for each call *before* data can flow
- **each packet carries VC identifier** (not destination host address)
- ***every* router on source-destination path maintains "state"** for each passing connection
- **link, router resources (bandwidth, buffers) may be *allocated* to VC** (dedicated resources = predictable service)
- *originated from telephony networks.*

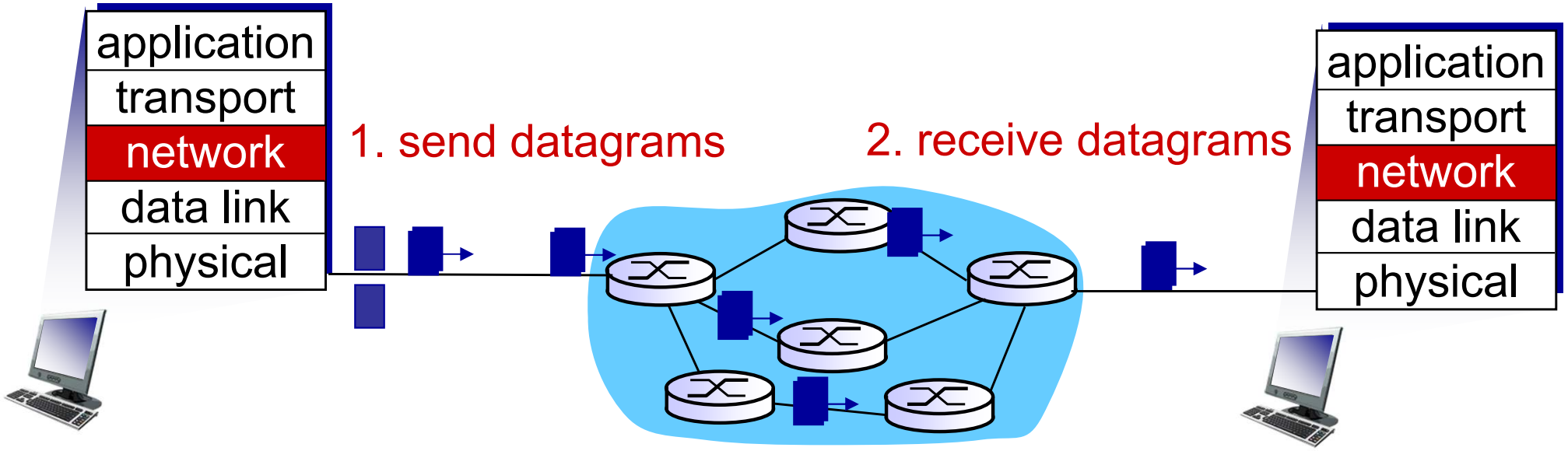
Virtual circuits: signaling protocols

- Used historically in ATM, frame-relay, X.25, which are *not part of today's Internet! But fundamental ideas remain…*
- *MPLS provides a virtual circuit-like solution.*

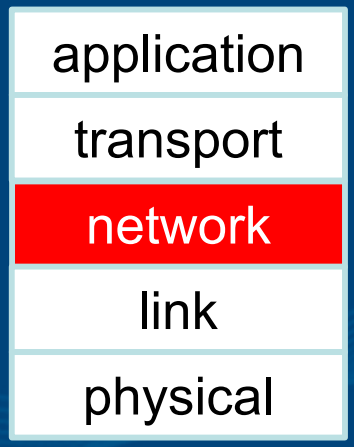


Datagram networks

- **no call setup at network layer**
- routers: **no state about end-to-end connections**
  - **no network-level concept of "connection"**
- datagrams forwarded using destination host address
- Internet!

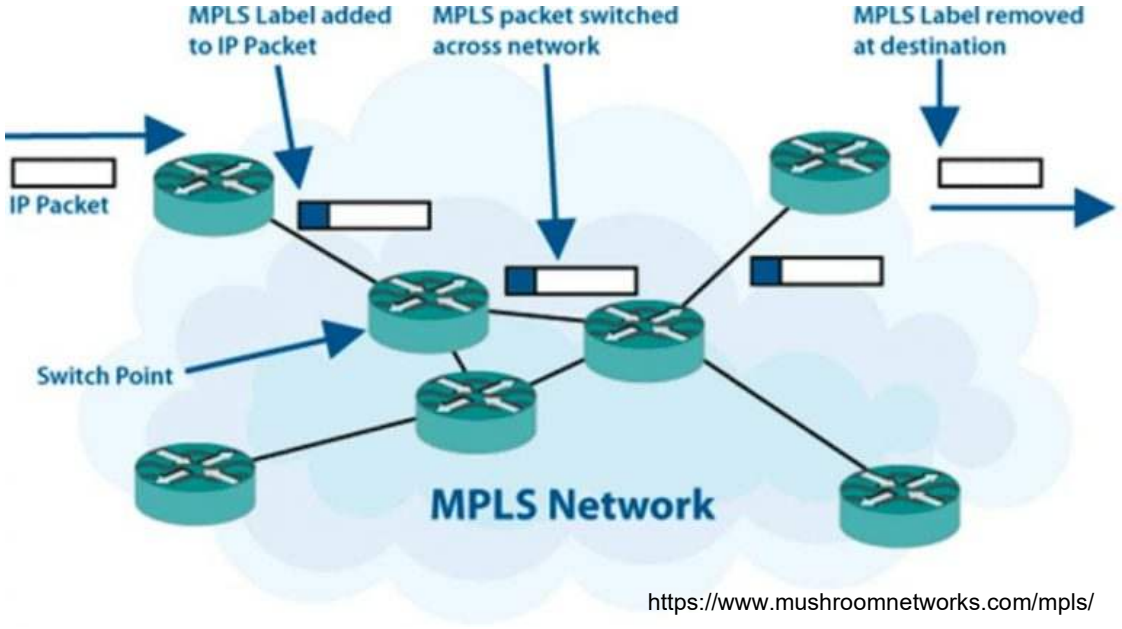


WAN Technologies and MPLS



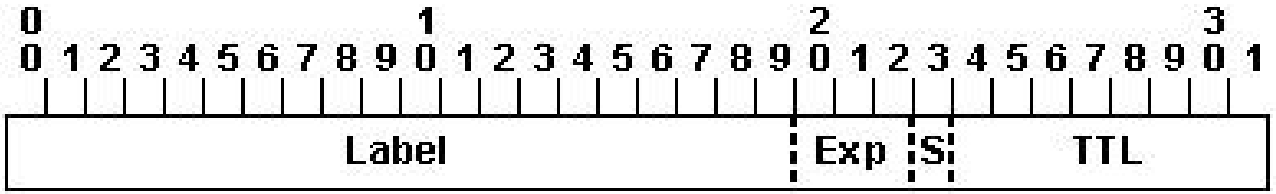
Multi Protocol Label Switching (MPLS)

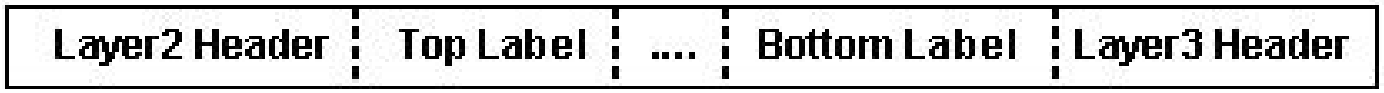
- **Multi Protocol Label Switching, MPLS,** is **a packet-forwarding technology which uses labels in order to make data forwarding decisions** (for any network layer protocol).
- With MPLS, the Layer 3 header analysis is done just once (when the packet enters the MPLS domain). Label inspection drives subsequent packet forwarding.
- MPLS helps with Virtual Private Networking (VPN), Traffic Engineering (TE), Quality of Service (QoS), Any Transport over MPLS (AToM)
- Additionally, it decreases the forwarding overhead on the core routers.



https://www.mushroomnetworks.com/mpls/

A **label** is a four-byte locally-significant identifier used in order to identify a **Forwarding Equivalence Class (FEC),** which is a group of IP packets which are forwarded in the same manner, over the same path, and with the same forwarding treatment.





https://www.cisco.com/c/en/us/support/docs/multiprotocol-label-switching-mpls/mpls/4649-mpls-faq-4649.html

Carrier Ethernet

Classical Ethernet

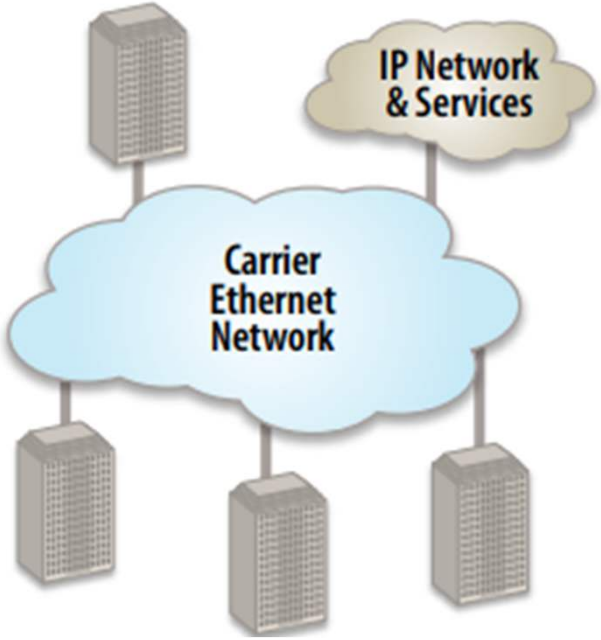
Ethernet, which has been the ubiquitous standard of choice for LAN, is rapidly replacing legacy technologies

Wide Area Network (WAN) providers want to provide their customers with Ethernet services and make use of the volume and cost advantages of Ethernet technologies in their networks.

Carrier Ethernet augments the original set of Ethernet LAN technologies to deliver services. Three fundamental differences between Carrier Ethernet networks from Ethernet LANs:

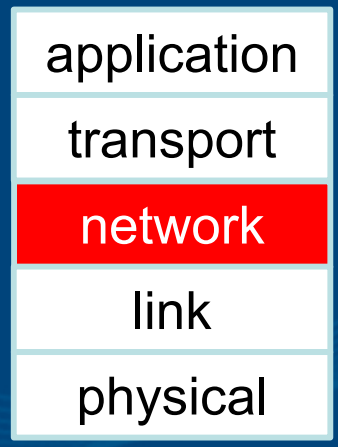
- An entire organization connects to a Carrier Ethernet "port" at a given subscriber location
- The Carrier Ethernet network serves many organizations
- The Carrier Ethernet network is outside the building across a wide area

Carrier Ethernet competes with MPLS in the same broad service space.



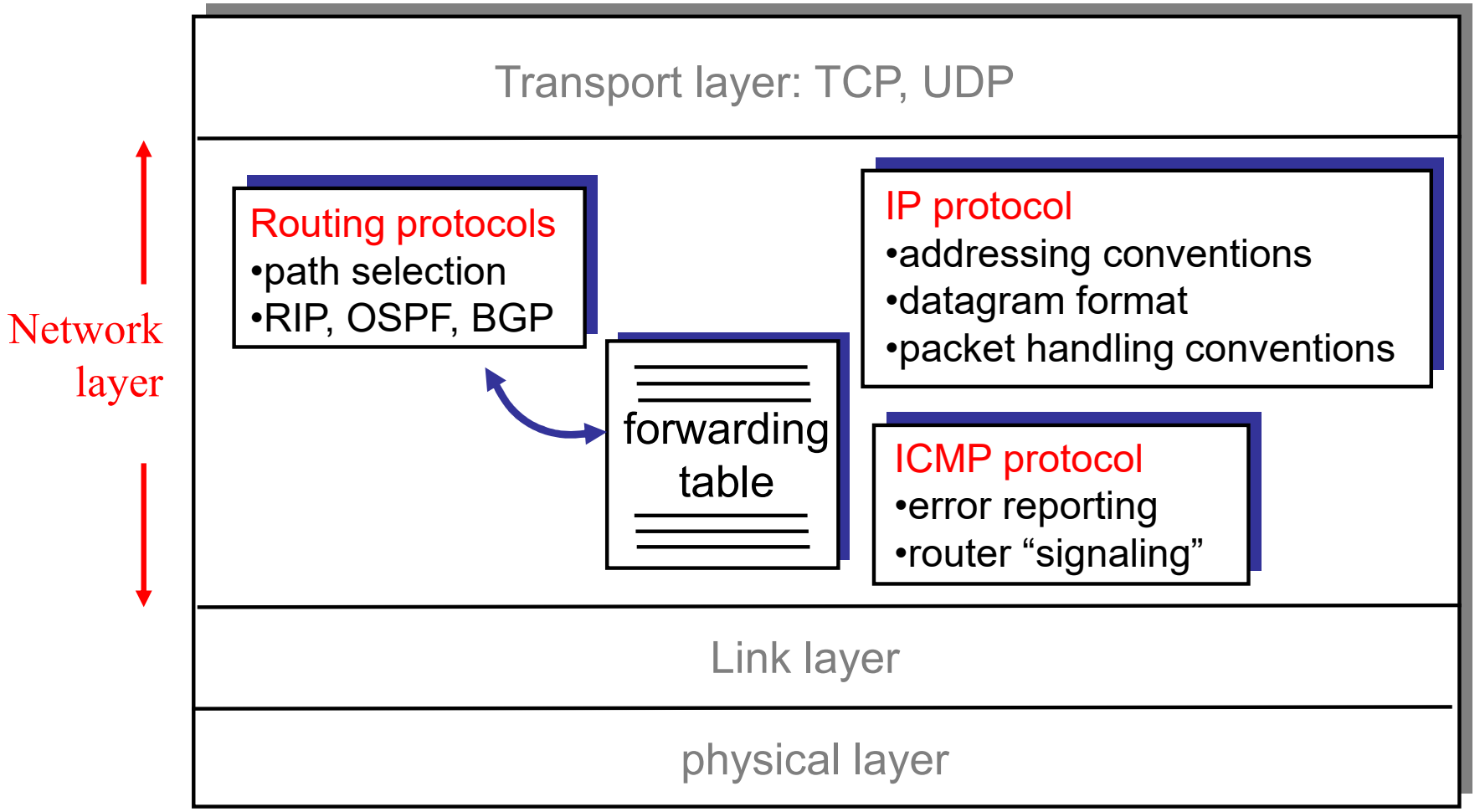
https://www.fujitsu.com/us/Images/CarrierEthernetEssentials.pdf

Internet Protocol (IP)



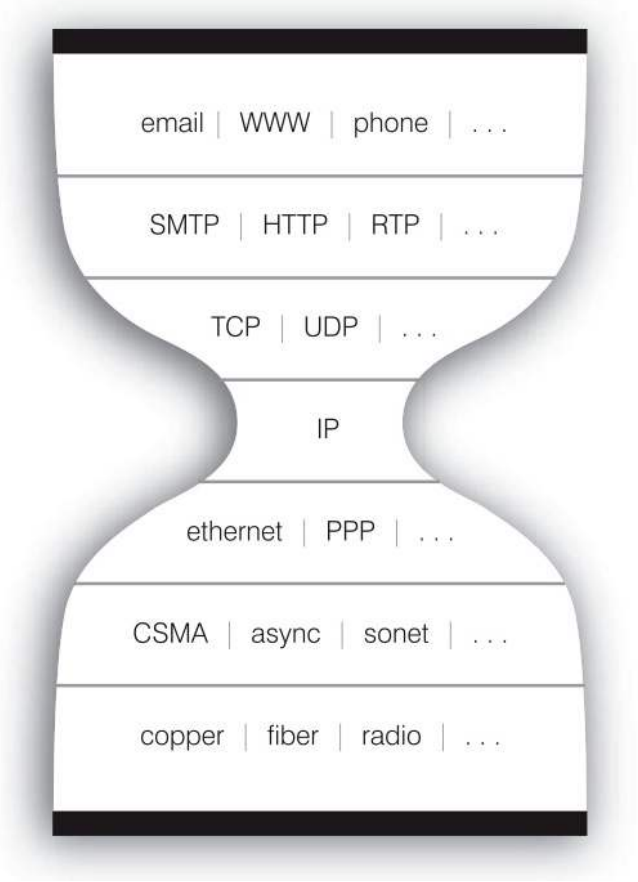
The Internet Network layer

Host, router network layer functions:



Why is IP so important?

Looking back to 20th century history of communications



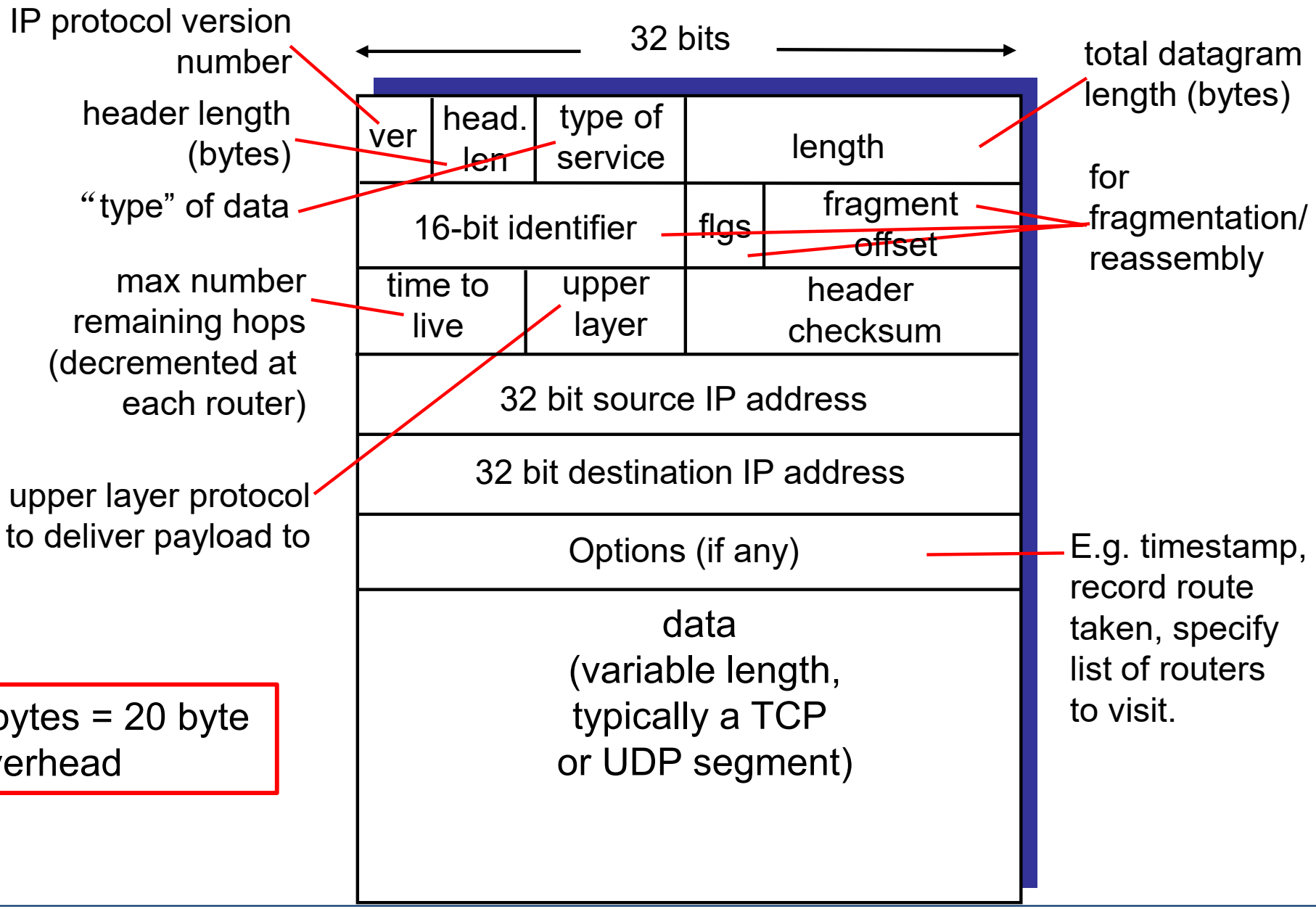
- Telephony (legacy) view vs. Internet approach
- **Netheads vs Bellheads**

(classic reading from 1996: http://archive.wired.com/wired/archive/4.10/atm.html)

- How is it evolving in the 21st century?
  Check the Net neutrality debates!

One *Protocol* to rule them all, One *Protocol* to find them,
One *Protocol* to bring them all and in the darkness bind them

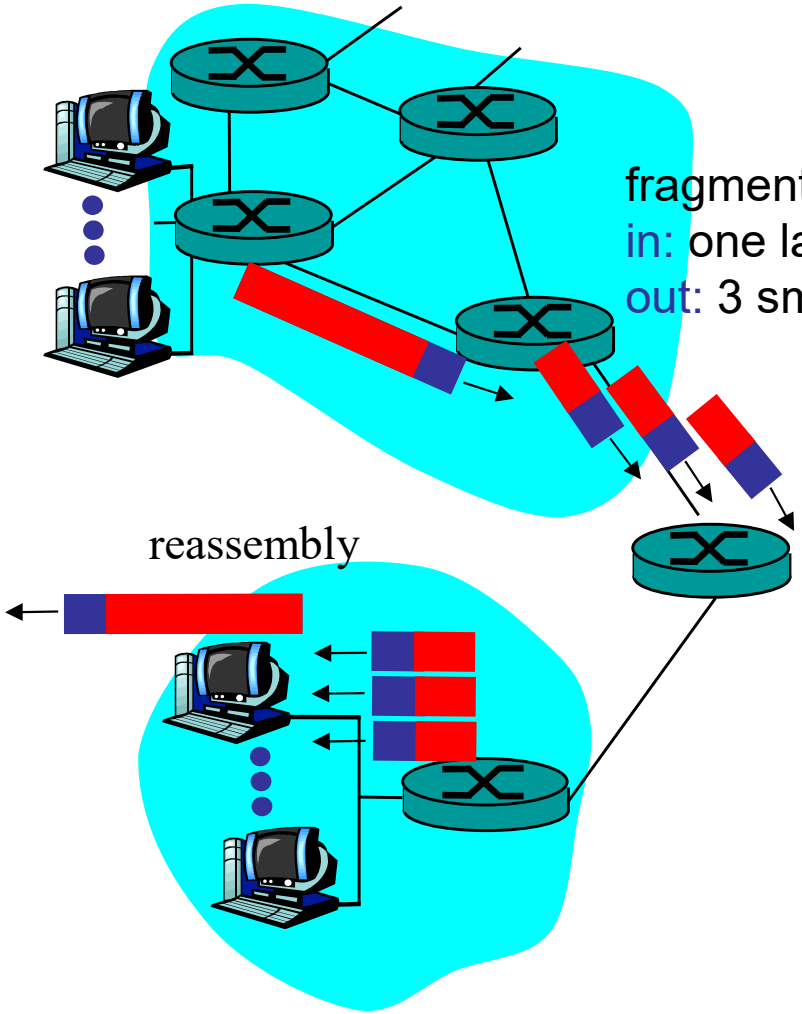
IP (v4) datagram format



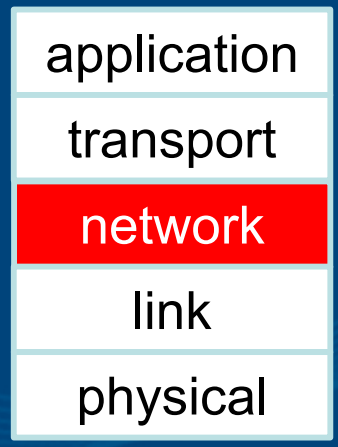
5x4 bytes = 20 byte IP overhead

IPv4 Fragmentation & Reassembly

- network links have **MTU (max. transfer size)** - largest possible link-level frame.
  - different link types, different MTUs
- **large IPv4 datagram divided ("fragmented") within net**
  - one datagram becomes several datagrams to fit into MTU to speed up
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments

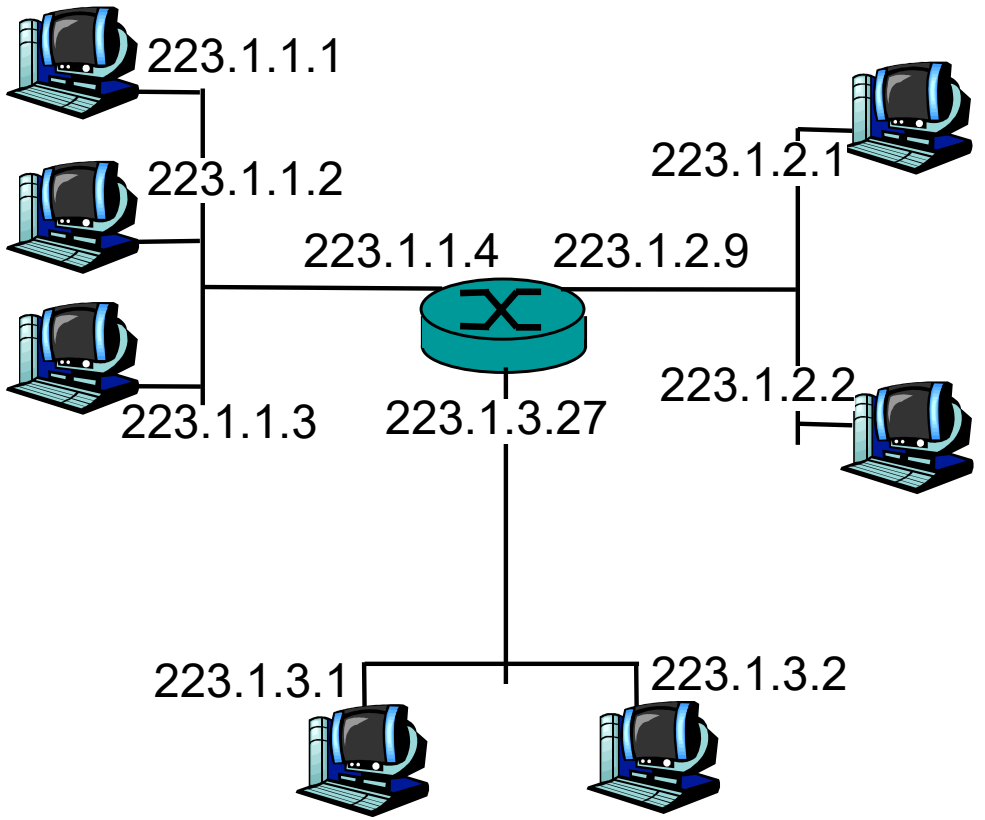


IP Addressing, Subnets, DHCP



IP (v4) Addressing: Introduction

- **IP address**: 32-bit identifier for host, router *interfaces*
- **interface***:* connection between host/router and physical link
  - routers typically have multiple interfaces
  - hosts typically have one or two interfaces
  - **IP addresses are associated with each interface**



Dotted-decimal notation

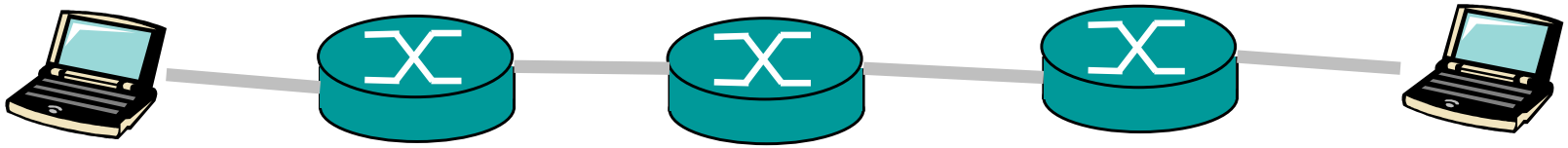
223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Question 1

Suppose there are three routers between a source host and a destination host.

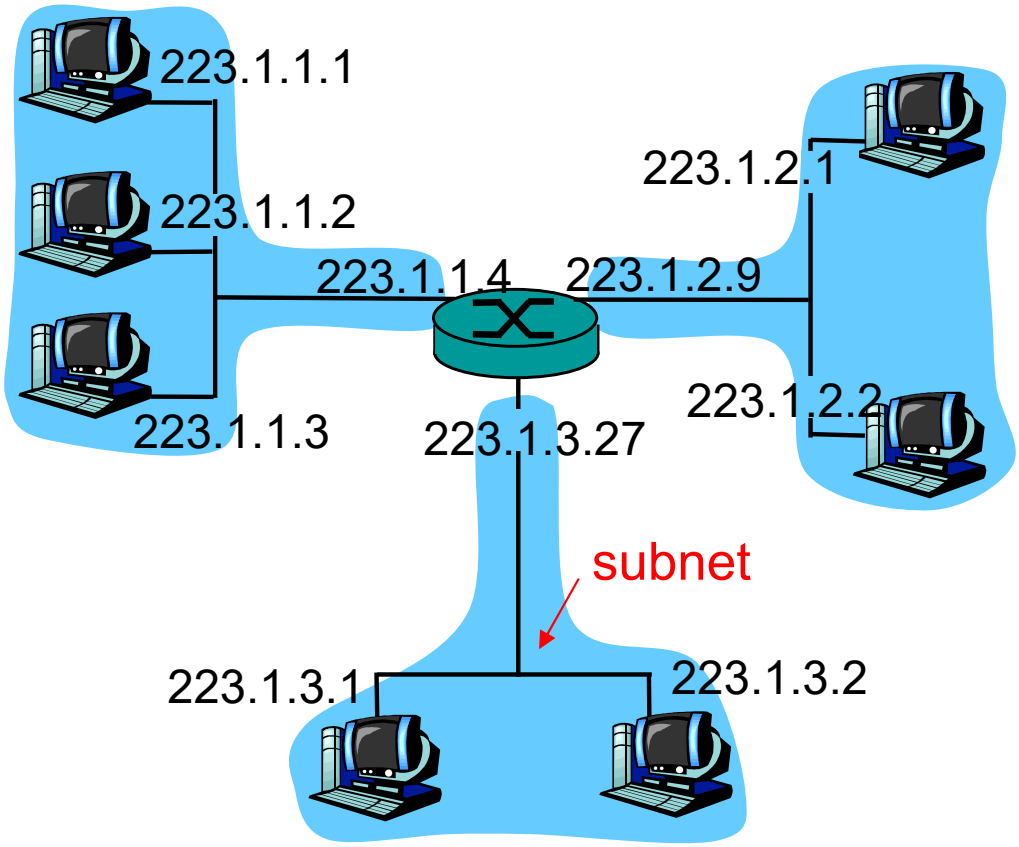
- Ignoring fragmentation, an IP datagram sent from the source host to the destination host will travel over how many interfaces?
- How many forwarding tables will be indexed to move the datagram from the source to the destination?



Subnets

- IP address:
  - subnet part (**high order bits**)
  - host part (**low order bits**)

- *What is a subnet ?*
  - device interfaces with same subnet part of IP address
  - **can physically reach each other without intervening router** (e.g. can be connected by an Ethernet switch).



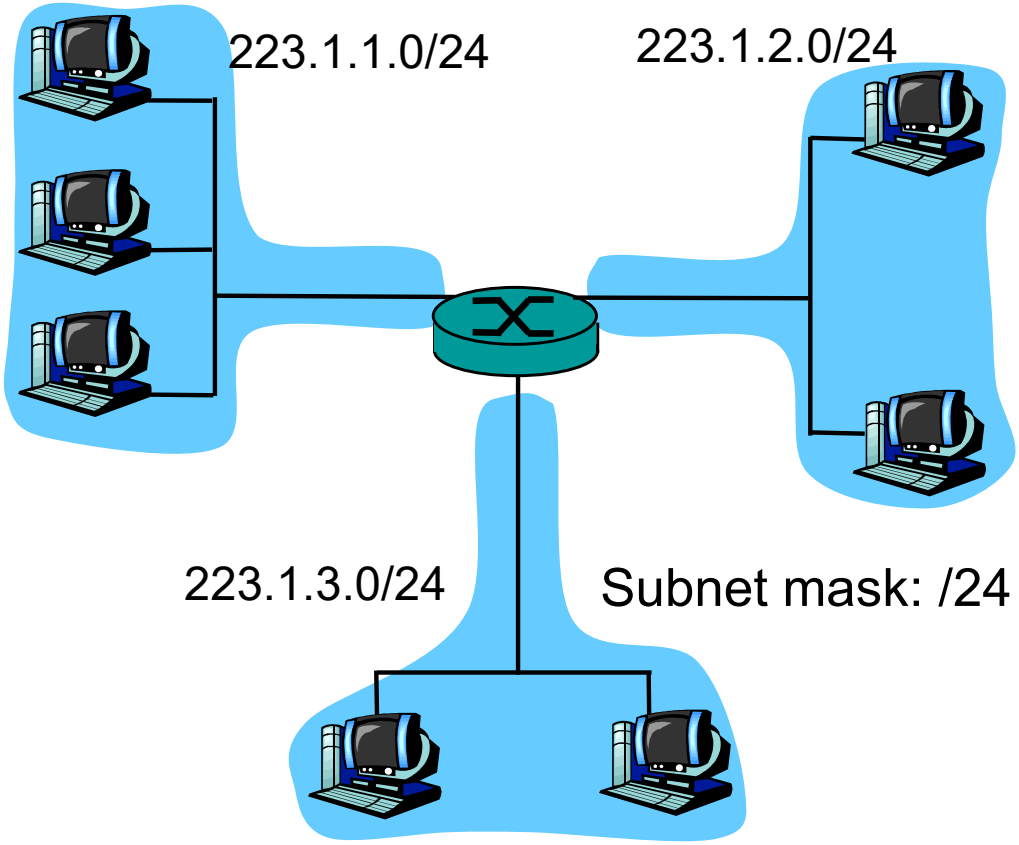
network consisting of 3 subnets

Note: *High order* = *leftmost* = *most significant* **bits** in an address!

Subnets

Recipe:

- To **determine the subnets**, detach each interface from its host or router, creating islands of isolated networks (disconnected graphs).
- Each isolated network is called a subnet.



Leftmost *(most significant) X bits* out of ***32 bit*** defines the subnet address:

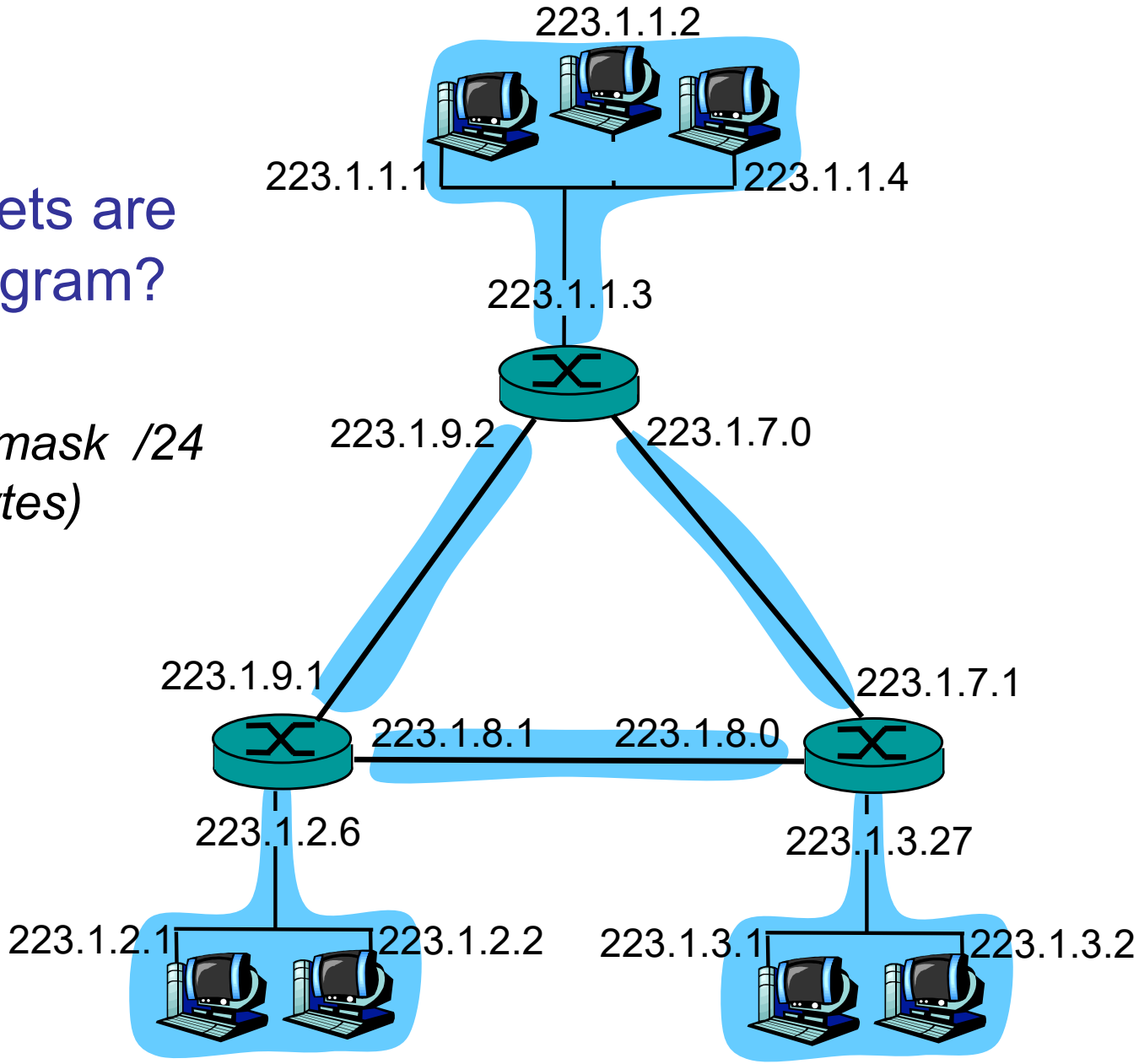
Here: **223.1.1**.xxx ⟺ **11011111.00000001.00000001**.xxxxxxxx

Subnets

Question:

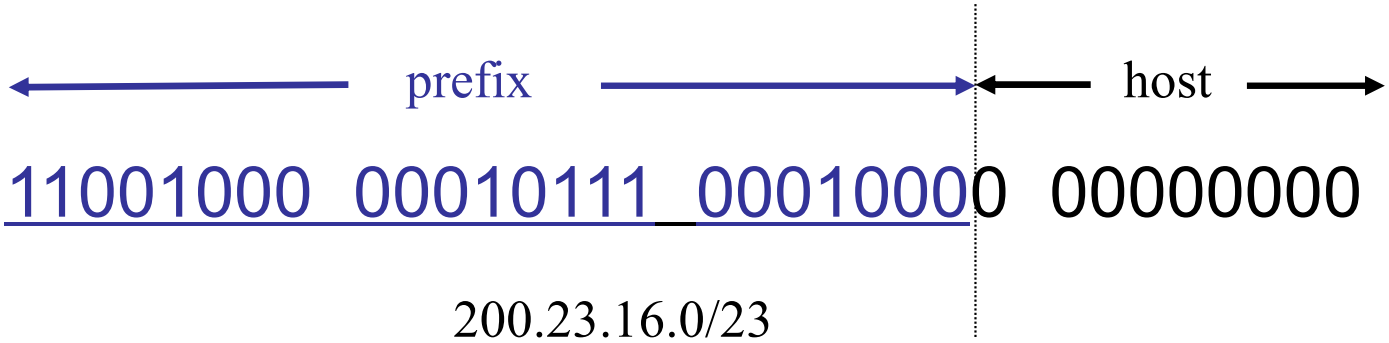
How many subnets are shown in the diagram?

***Hint**: consider subnet mask /24 (i.e. fix first three bytes)*



CIDR: Classless InterDomain Routing (cider)

- flexible subnet addressing: subnet portion of address of arbitrary length
- address format: a.b.c.d/x, where **x is # bits in subnet portion** (leftmost bits) of the address



IP Addressing

- **Classful addressing**: Class A, B, C networks with (8, 16, 24 bit prefixes) – historical, inflexible…
- **Broadcast address**: 255.255.255.255, all datagrams are delivered to all hosts on the same subnet.
- **Subnet (network) mask** designates a subnetwork.

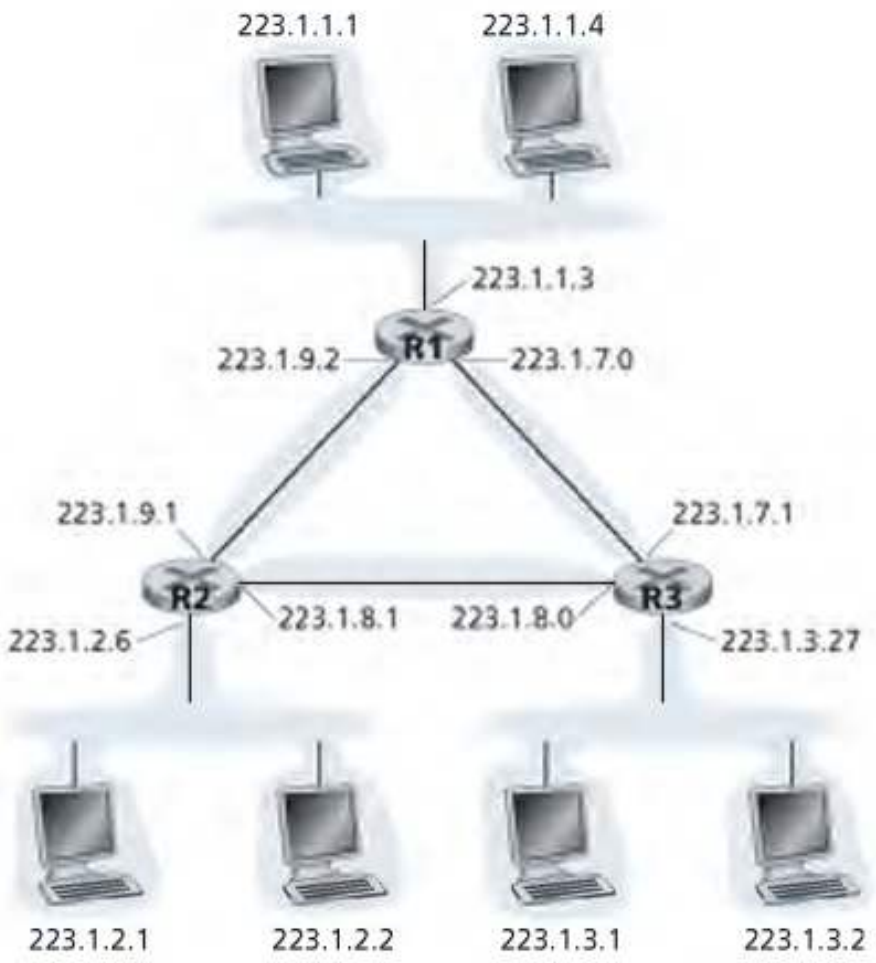
For example:
192.168.0.0 with netmask 255.255.255.0
corresponds to:
192.168.0.0/24 in CIDR notation,
i.e. leftmost 24 bits determine the specific subnetwork which can have up to 256 hosts.

Question 2

Consider the network shown. Denote the three subnets with hosts (starting clockwise at 12:00) as Networks A, B and C. Denote the subnets without hosts as Networks D, E and F.
Assign new network addresses to each of these six subnets, with the following constraints:

- All addresses must be allocated from **214.97.254.0/23**;
- Subnet A should have enough addresses to support 250 interfaces;
- Subnet B should have enough addresses to support 120 interfaces;
- Subnet C should have enough address to support 120 interfaces.
- Of course, subnets D, E and F should each be able to support two interfaces.

For each subnet, the assignment should take the form a.b.c.d/x or a.b.c.d/x – e.f.g.h/y (in CIDR notation)



IP addresses: how to get one?

Q: How does a *host* get its IP address?

- *Manual entry* (rare but possible via config or control panel of operating system)
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from a DHCP server
  - "plug-and-play"

DHCP: Dynamic Host Configuration Protocol

Goal: allow host to *dynamically* obtain its IP address from a network server when it joins network

- Can renew its lease on address in use
- Allows reuse of addresses (only hold address while connection is “on”)
- Support for mobile users who want to join network

DHCP overview:

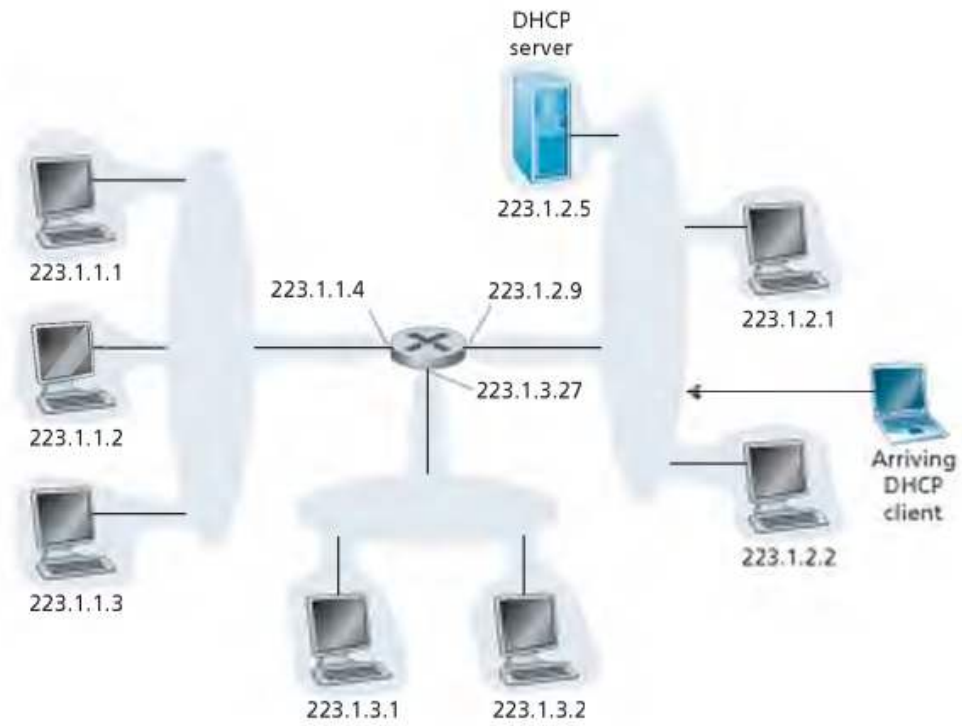
– host broadcasts “DHCP discover” msg

– DHCP server responds with “DHCP offer” msg

– host requests IP address: “DHCP request” msg

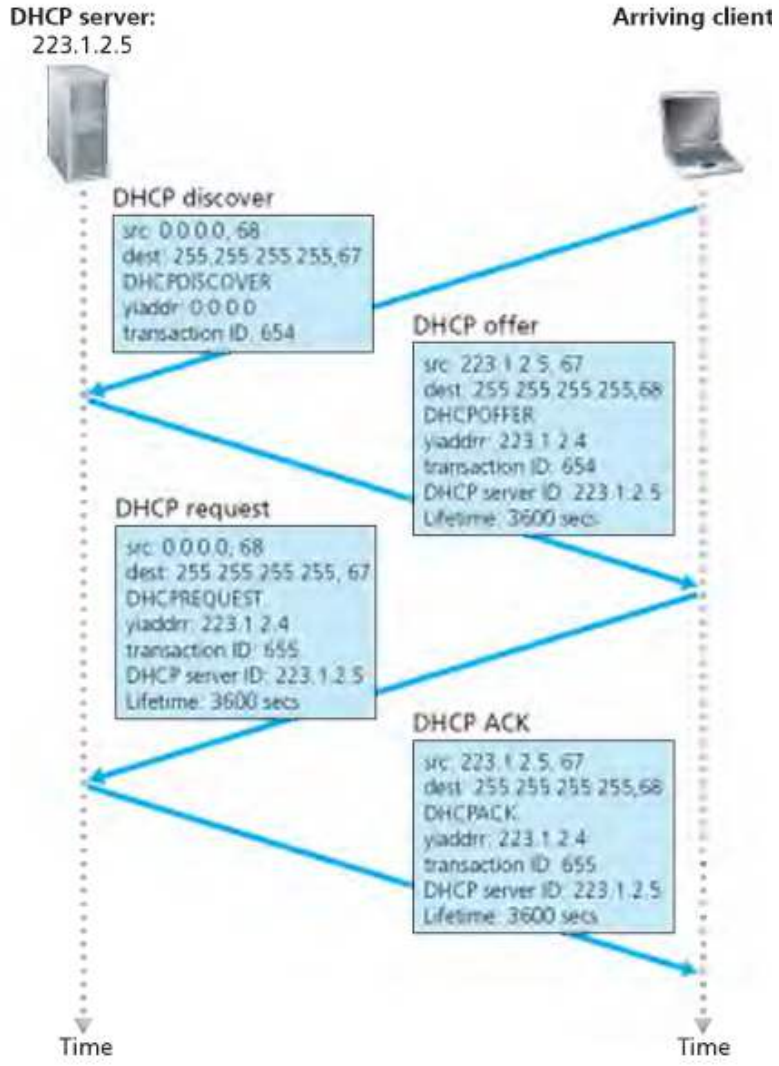
– DHCP server sends address: “DHCP ack” msg

DHCP: Dynamic Host Configuration Protocol



Questions:

1. Why use broadcast in the DHCP protocol?
2. What if the DHCP server is not in the subnet?



yiaddr: your internet addr

IP addresses: how to get one?

Q: How does *network* get subnet part of IP addr?

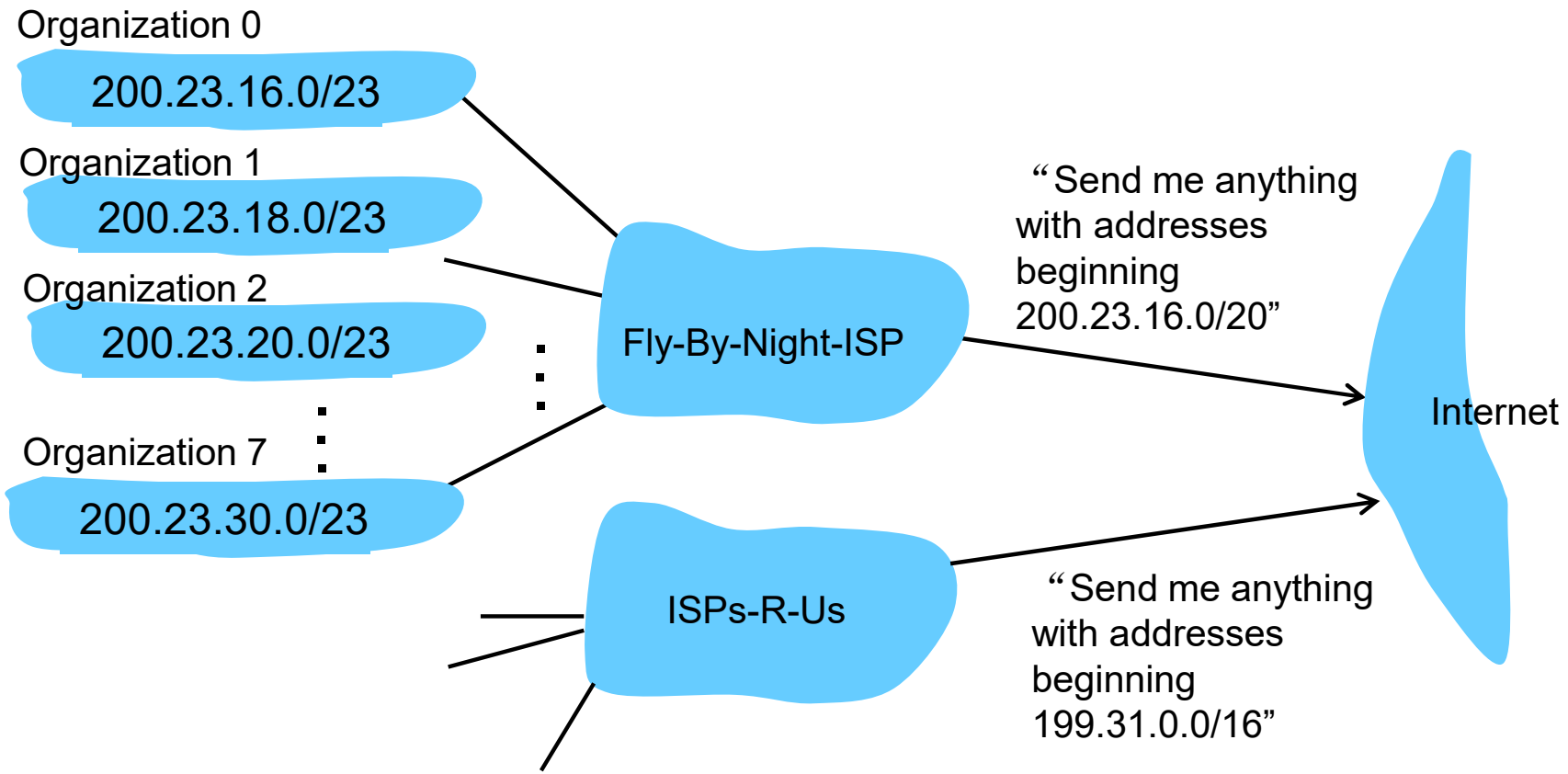
A: gets allocated a portion of its provider ISP's address space

ISP's block	11001000 00010111 00010000	00000000	200.23.16.0/20
Organization 0	11001000 00010111 00010000	00000000	200.23.16.0/23
Organization 1	11001000 00010111 00010010	00000000	200.23.18.0/23
Organization 2	11001000 00010111 00010100	00000000	200.23.20.0/23
...	.....	....	....
Organization 7	11001000 00010111 00011110	00000000	200.23.30.0/23

Example: The University of Melbourne address range https://ipinfo.io/AS10148
Range assigned by **APNIC**, which is the Regional Internet Registry administering IP addresses for the Asia Pacific https://www.apnic.net

Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:



IP address assignment

Q: How does an ISP get block of addresses?

A:

ICANN: Internet Corporation for Assigned Names and Numbers

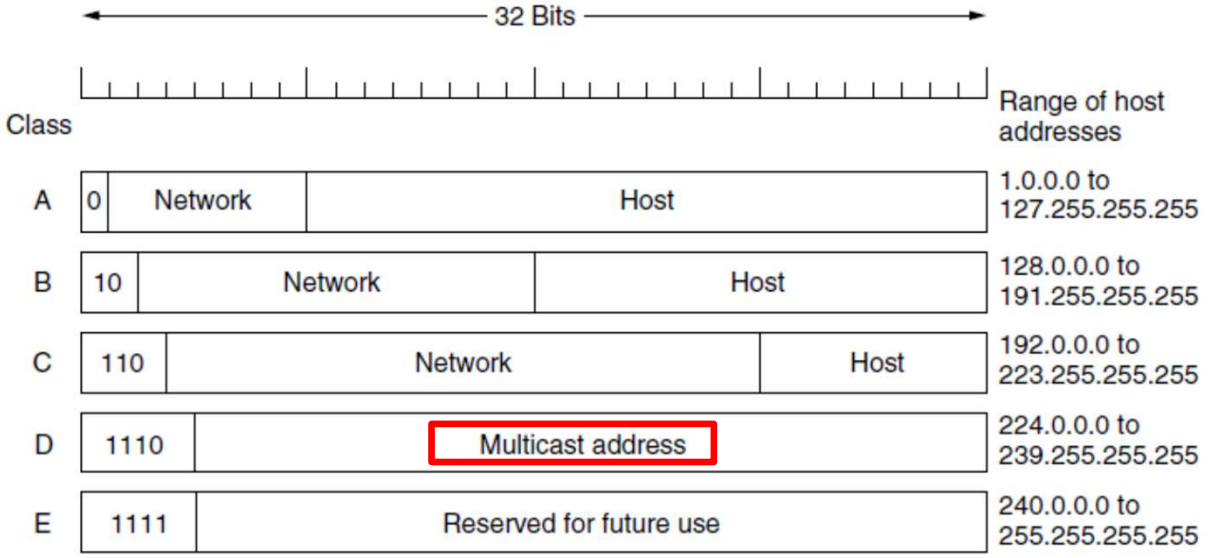
- allocates addresses
- manages DNS
- assigns domain names, resolves disputes

Check recent news about ICANN-related global politics!

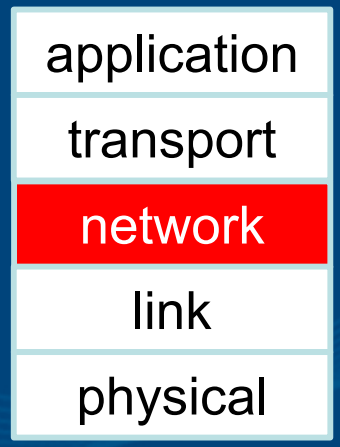
IP Multicast

- Regular IP communication is between one sender and one receiver.
- For some applications, it is useful for a process to be able to send to a large number of receivers simultaneously.
- IP Multicast uses **specific Class D IP Addresses** (see below).
- Each class D address identifies a group of hosts. 28 bits identify groups, so 2^{28} potential groups can be defined.
- When a process sends a packet to a class D address, a best-effort attempt is made to deliver it to all the members of the group addressed, but no guarantees are given.

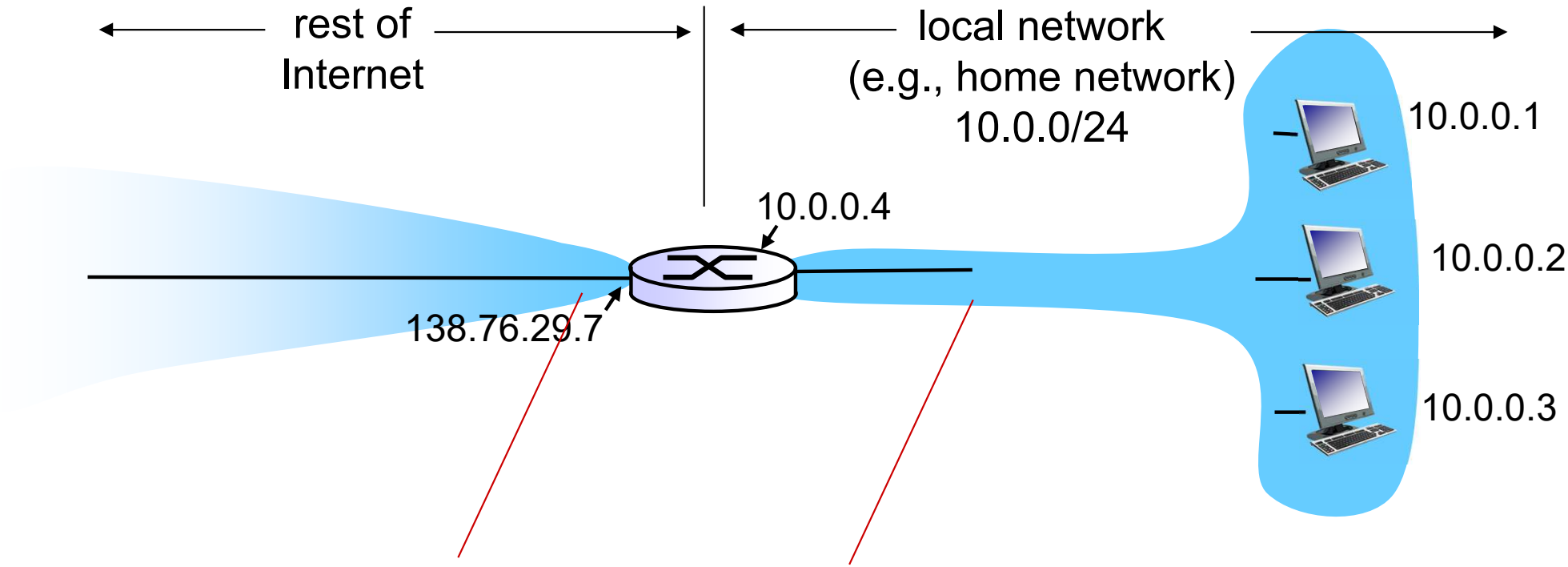
224.0.0.1 All systems on a LAN
224.0.0.2 All routers on a LAN
224.0.0.5 All OSPF routers on a LAN
224.0.0.251 All DNS servers on a LAN



Network Address Translation NAT



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 10.0.0/24 address for source, destination (as usual)

NAT: network address translation

motivation: local network uses just one IP address as far as outside world is concerned:

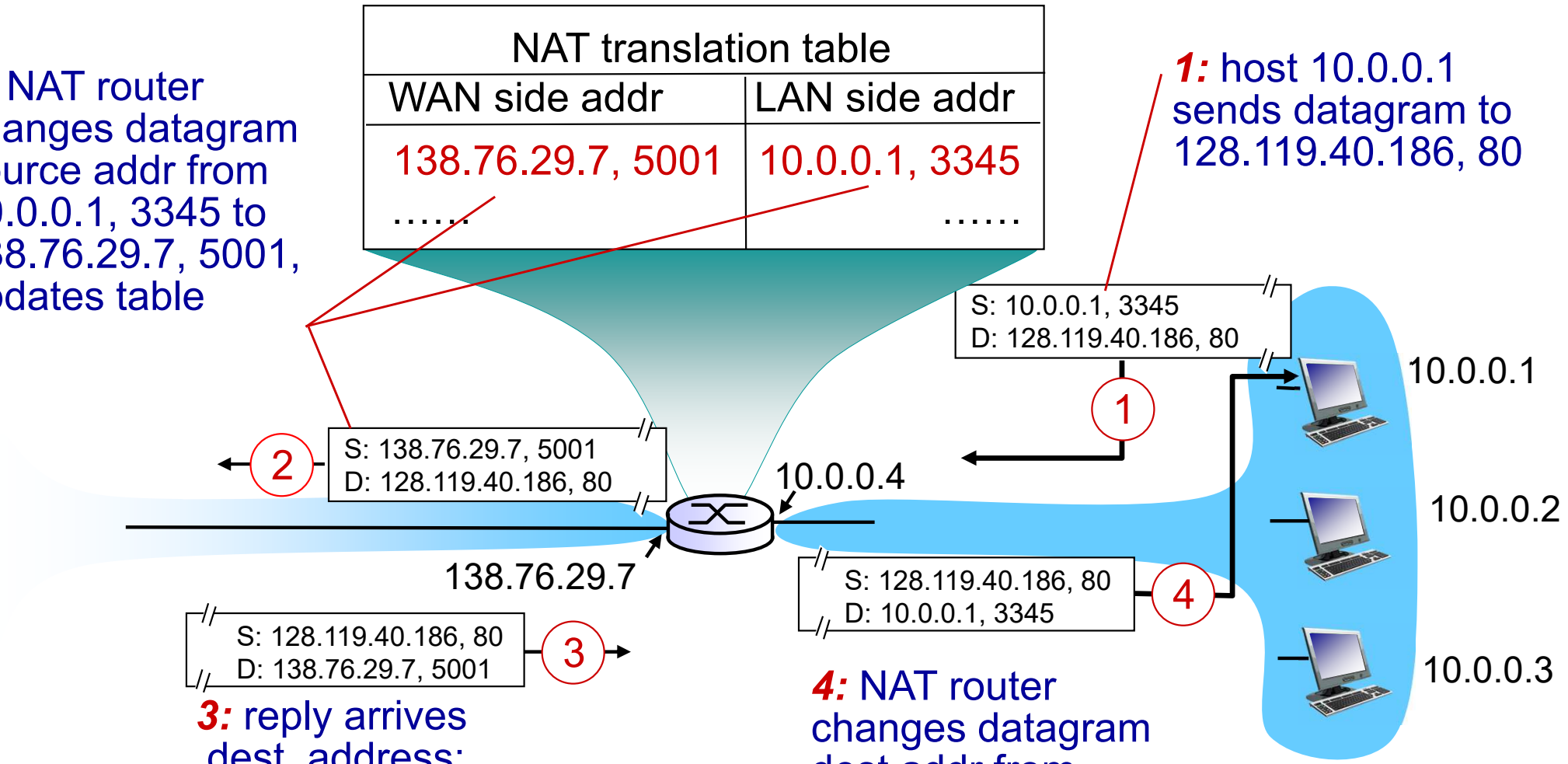
- range of addresses not needed from ISP: just one IP address for all devices
- can change addresses of devices in local network without notifying outside world
- can change ISP without changing addresses of devices in local network
- devices inside local net not explicitly addressable, visible by outside world (**a security plus**)
- NAT is not a firewall and does not secure the subnet!

NAT: network address translation

implementation: NAT router must

- *outgoing datagrams: replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
  - . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr
- *remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair
- *incoming datagrams: replace* (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: network address translation



NAT: Non-unique, local address spaces

- The Internet Assigned Numbers Authority (IANA) reserved the following three blocks of the IPv4 address space for private internets:
  - 10.0.0.0 - 10.255.255.255 (10/8 prefix)
  - 172.16.0.0 - 172.31.255.255 (172.16/12 prefix)
  - 192.168.0.0 - 192.168.255.255 (192.168/16 prefix)

  See the RFC 1918 https://tools.ietf.org/html/rfc1918

- For IPv6, see RFC 4193 https://tools.ietf.org/html/rfc4193
  - FD00::/8 prefix to identify Local IPv6 unicast addresses.
  - See https://en.wikipedia.org/wiki/Unique_local_address
  - The Local IPv6 addresses are created using a pseudo-randomly allocated global ID.

NAT: network address translation

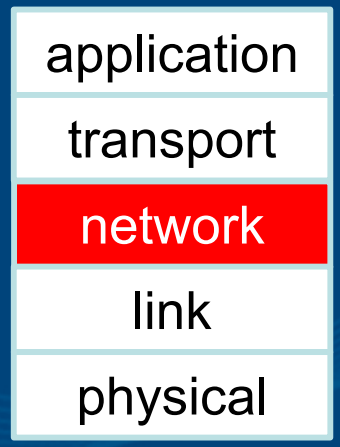
downsides:

- Ports are not meant to be used for addressing processes.
- Routers are supposed to process packets only up to networking layer (**ports are higher**).
- NAT breaks end-to-end communication (e.g. P2P).

reality:

NAT has become an important component of the Internet. Moving to IPv6, necessity of NAT is still being discussed as a hot topic, but in principle there *was supposed to be* no NAT in IPv6!

IPv6 Overview



IPv6: motivation

- Initial motivation*: (from 1990s on)*
  32-bit address space ~~soon to be completely allocated~~ is all used now!
- additional motivation:
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS

IPv5

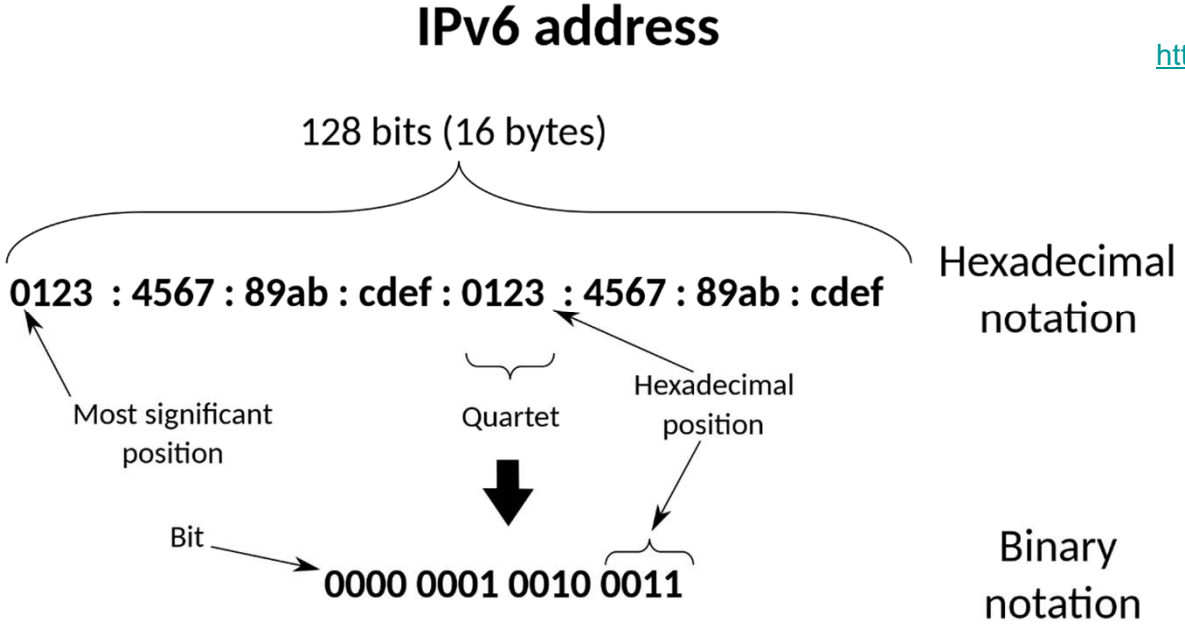
- What happened to IPv5?

- *IPv5* is the experimental Internet Stream Protocol (ST). Therefore, we jump from v4 to v6!
  see http://www.iana.org/assignments/version-numbers/version-numbers.xhtml

IPv6: motivation

IPv6 addressing 128bit address (16 bytes instead of 4)

- ::1/128 loop back address
- fe80::/xx is a link-local prefix
- ::ffff:0:0/96 is an IPv4-mapped Address (128-32=96)
- e.g., **3ffe:1900:4545:3:200:f8ff:fe21:67cf (128 bit)**

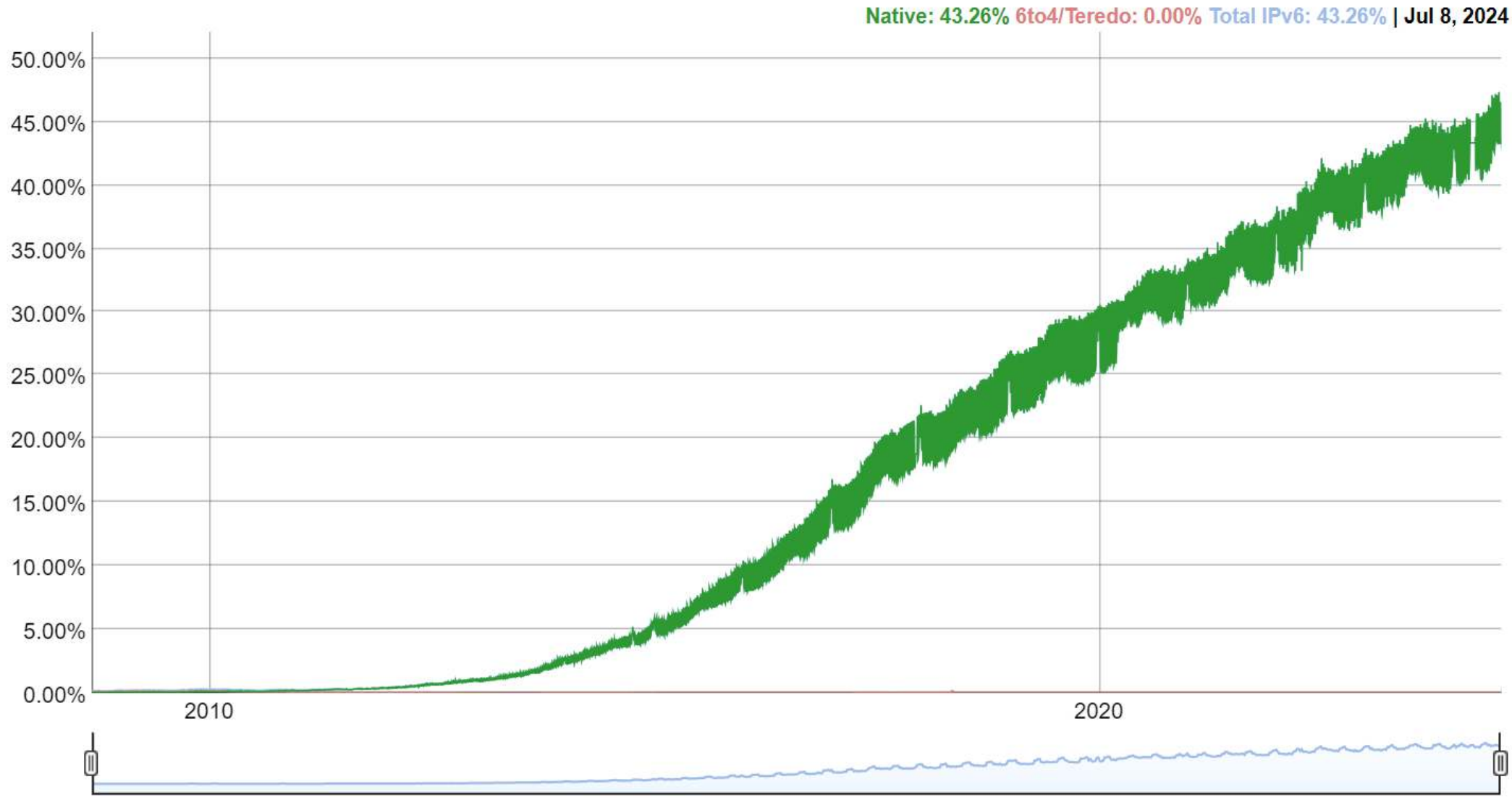


https://en.wikipedia.org/wiki/IPv6

IPv6 Adoption Statistics (Google)

IPv6 Adoption

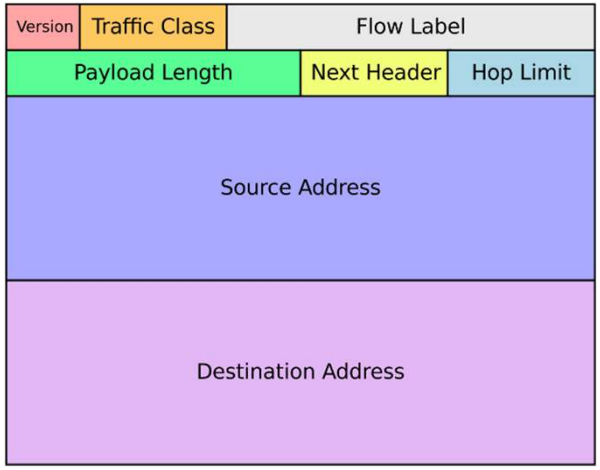
We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.

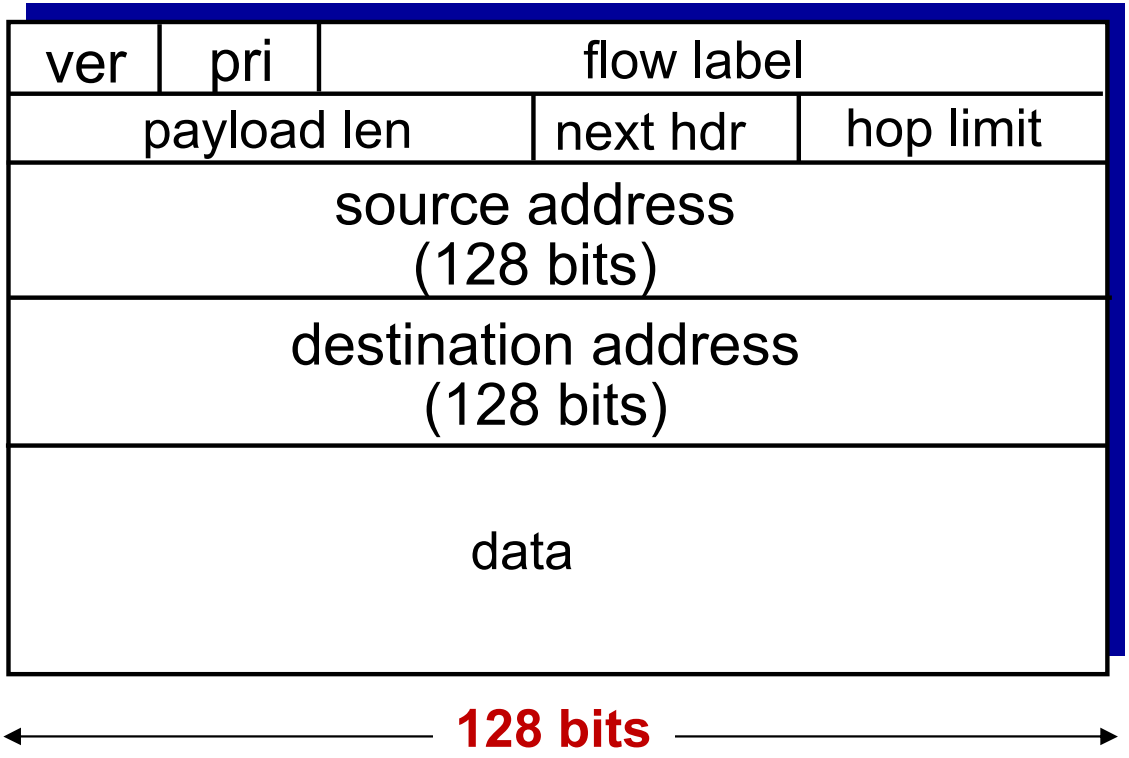


https://www.google.com/intl/en/ipv6/statistics.html

IPv6 datagram format

priority: identify priority among datagrams in flow
flow Label: identify datagrams in same "flow."
(concept of "flow" not well defined).
next header: identify upper layer protocol for data (e.g., TCP or UDP)





IPv6 datagram format:

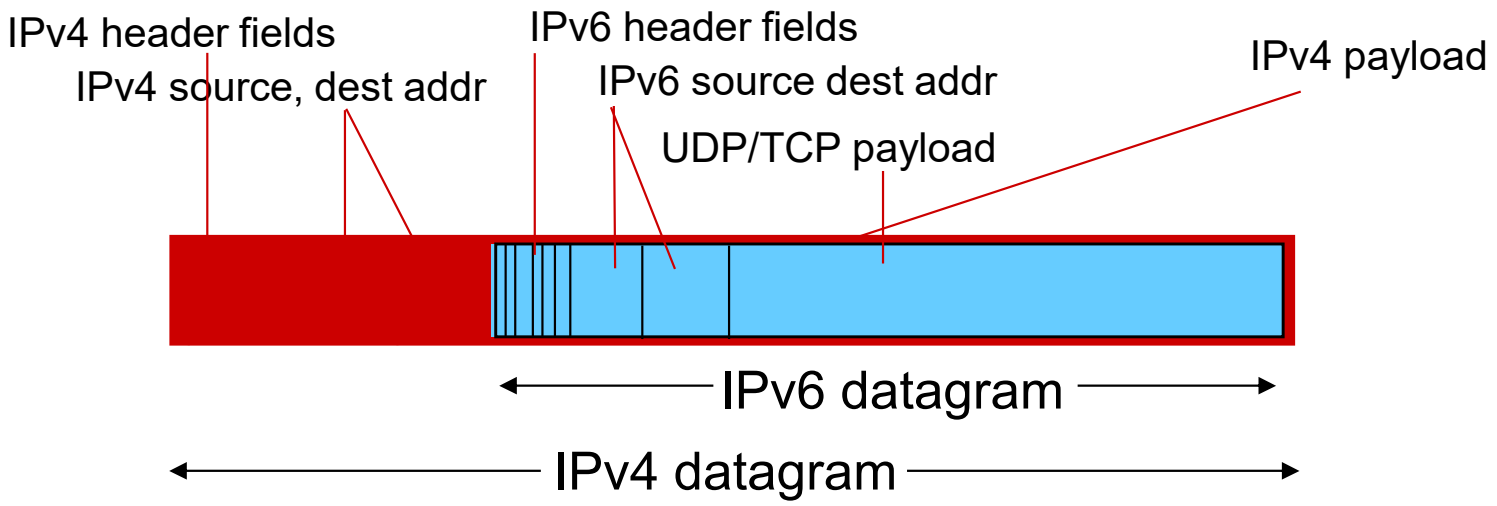
- **fixed-length 40-byte header**
- **no fragmentation built-in**

Other changes from IPv4

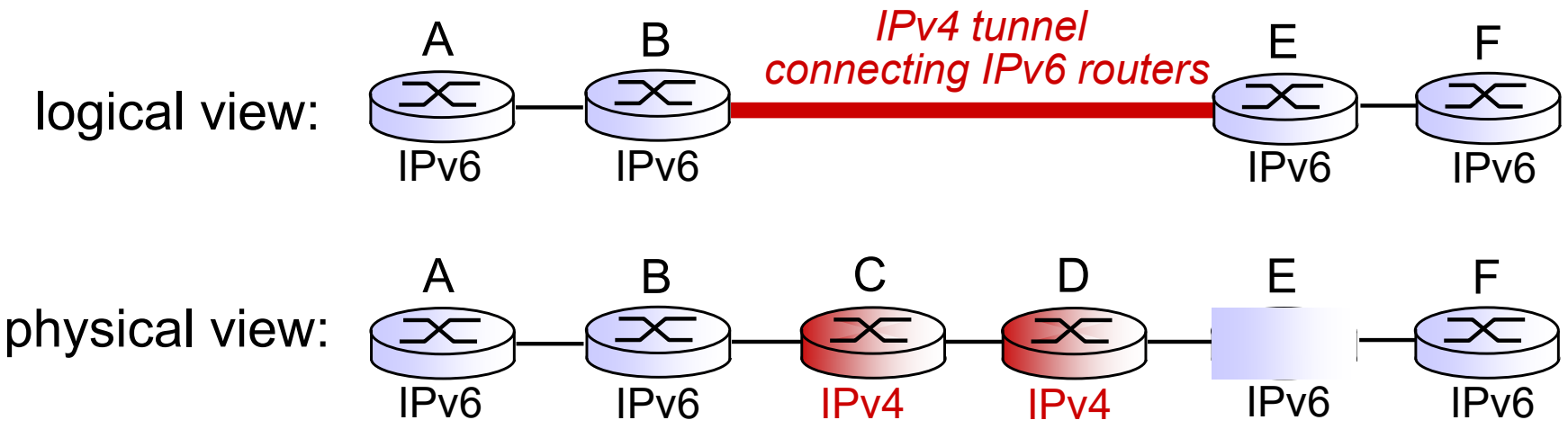
- *checksum*: removed entirely to reduce processing time at each hop
- *options:* allowed, but outside of header, can be indicated by "Next Header" field
- *ICMPv6:* new version of ICMP (*Internet Control Message Protocol* –used for control and diagnostics)
  - additional message types, e.g., "Packet Too Big"
  - multicast group management functions

Transition from IPv4 to IPv6

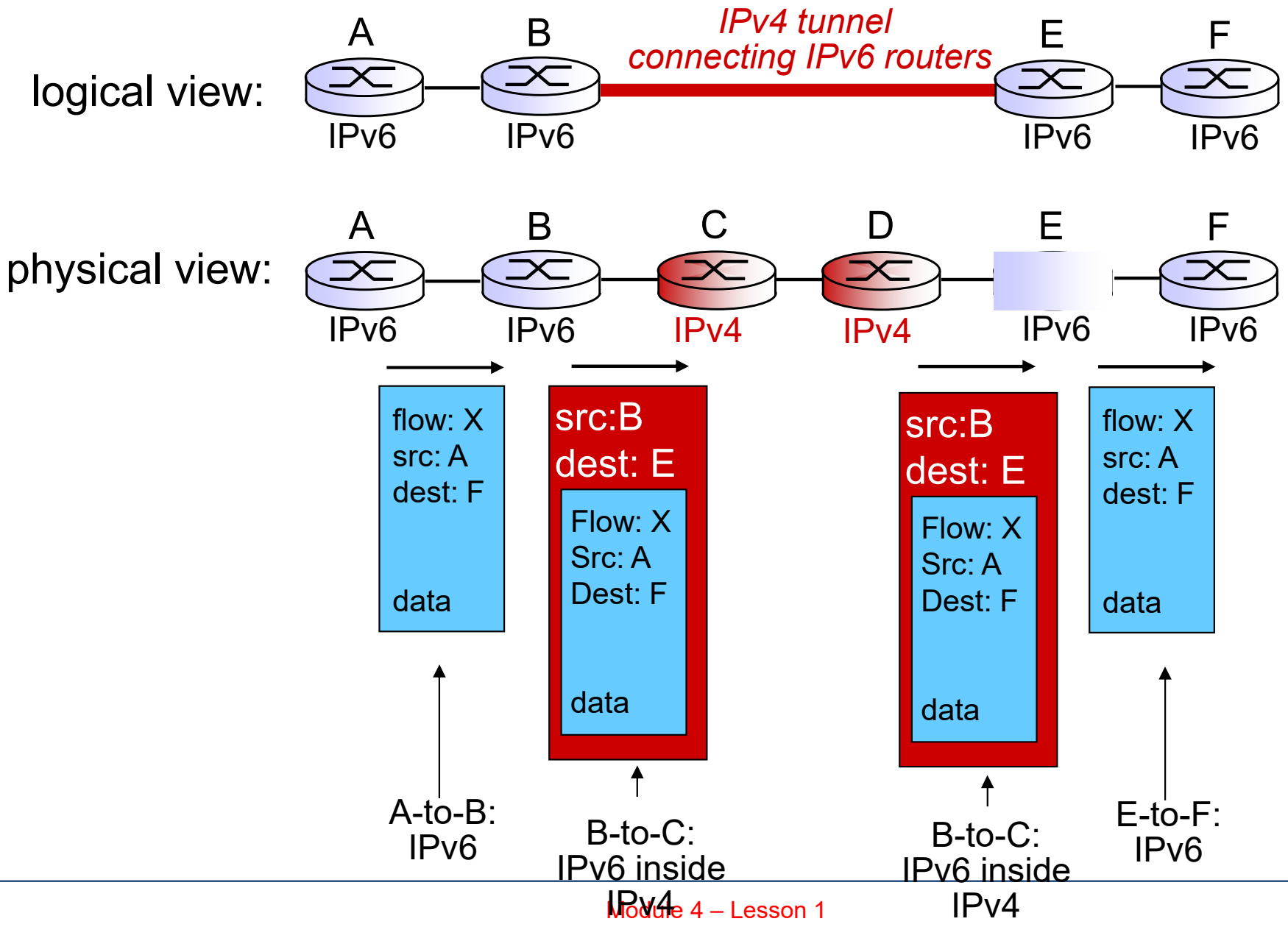
- not all routers can be upgraded simultaneously
  - It is very difficult to change networking protocols.
  - how will network operate with mixed IPv4 and IPv6 routers?
- *tunneling:* IPv6 datagram carried as ***payload*** in IPv4 datagram among IPv4 routers



Tunneling


logical view:
A
B
IPv4 tunnel
connecting IPv6 routers
E
F
IPv6
IPv6
IPv6
IPv6
physical view:
A
B
C
D
E
F
IPv6
IPv6
IPv4
IPv4
IPv6
IPv6

Tunneling


logical view:
A
B
IPv4 tunnel
connecting IPv6 routers
E
F
IPv6
IPv6
IPv6
IPv6
physical view:
A
B
C
D
E
F
IPv6
IPv6
IPv4
IPv4
IPv6
IPv6
flow: X
src: A
dest: F
data
src:B
dest: E
Flow: X
Src: A
Dest: F
data
src:B
dest: E
Flow: X
Src: A
Dest: F
data
flow: X
src: A
dest: F
data
A-to-B:
IPv6
B-to-C:
IPv6 inside
IPv4
B-to-C:
IPv6 inside
IPv4
E-to-F:
IPv6

Question 3

Which two technologies that were covered in previous slides will disappear with full adoption of IPv6?

Learning Objectives

You should now know about

- Network layer services, SDN
- Recent WAN technologies and MPLS
- Internet Protocol (IP), IPv4
- IP Addressing, Subnets, DHCP, multicasting
- Network Address Translation NAT
- IPv6 overview, tunneling