Subject: Internetworking and LAN technologies

- Instructor: Andrey S. Koval
- E-mail: koval#cs.vsu.ru
- Objective: To improve English language skills within the network technologies context during the basic course lessons.
- Notes based on “Computer Networking: A Top Down Approach Featuring the Internet”, 2005, 3d edition, Jim Kurose, Keith Ross, Addison-Wesley.

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

Protocol layers, service models

History

IP addressing and routing

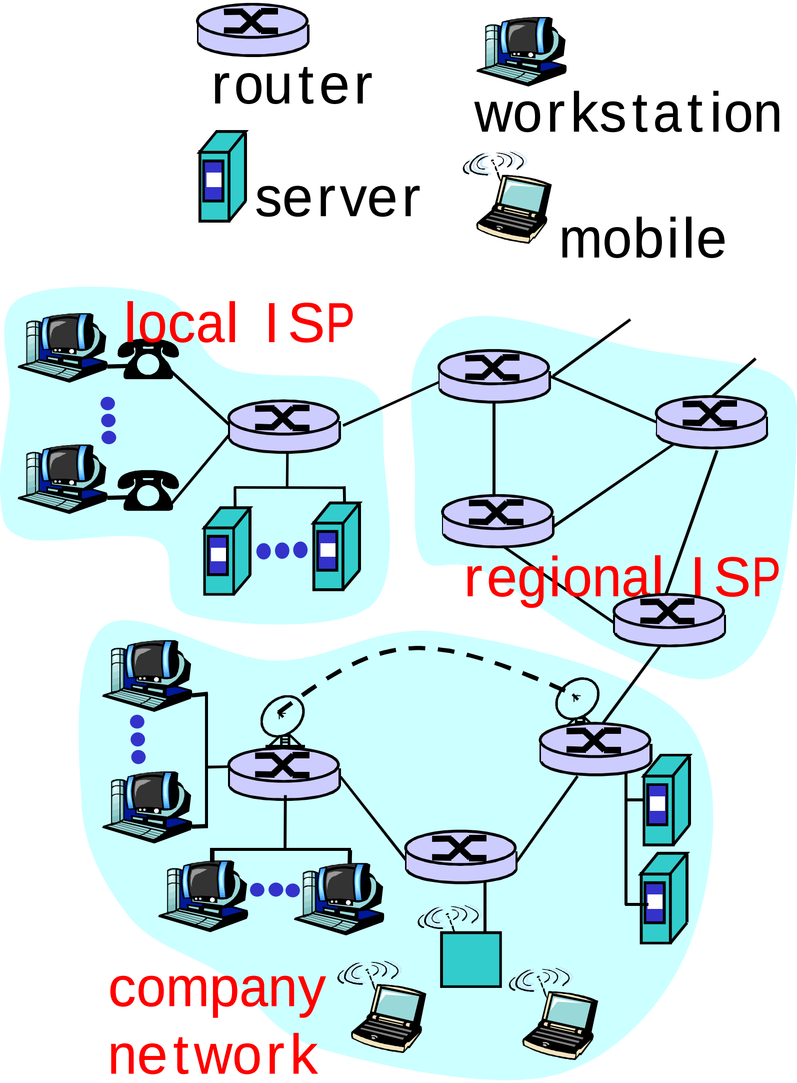
Computer Network?

- "interconnected collection of autonomous computers connected by a *single* technology" [Tanenbaum]
- What is the Internet?
  - "network of networks"
  - "collection of networks interconnected by routers"
  - "a communication medium used by millions"
  - Email, chat, Web "surfing", streaming media

What's the Internet: "nuts and bolts" view

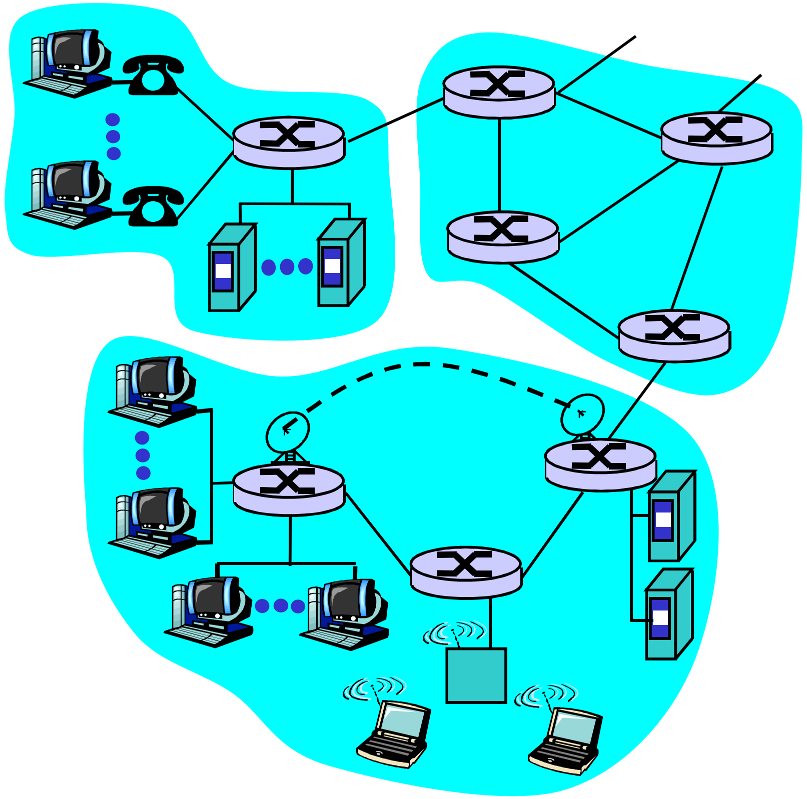
- millions of connected computing devices: *hosts, end-systems*
  - PCs workstations, servers
  - PDAs, phones

  running *network apps*
- *communication links*
  - fiber, copper, radio
  - transmission rate = *bandwidth*
- *routers:* forward packets (chunks of data)



What's the Internet: a service view

- communication *infrastructure* enables distributed applications:
  - Web, email, games, e-commerce, database, file sharing
- communication services provided to apps:
  - connectionless
  - connection-oriented



- cyberspace [W. Gibson]:

  "a consensual hallucination experienced daily by billions of operators, in every nation, ..."

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

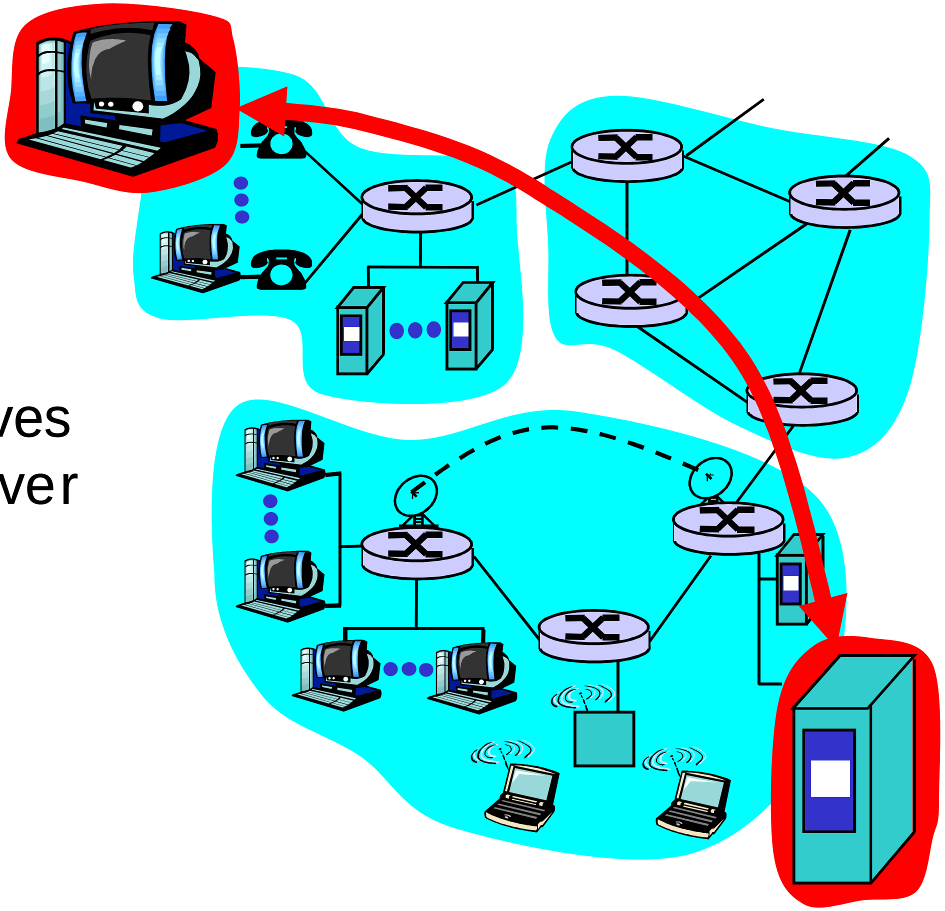
Protocol layers, service models

History

IP addressing and routing

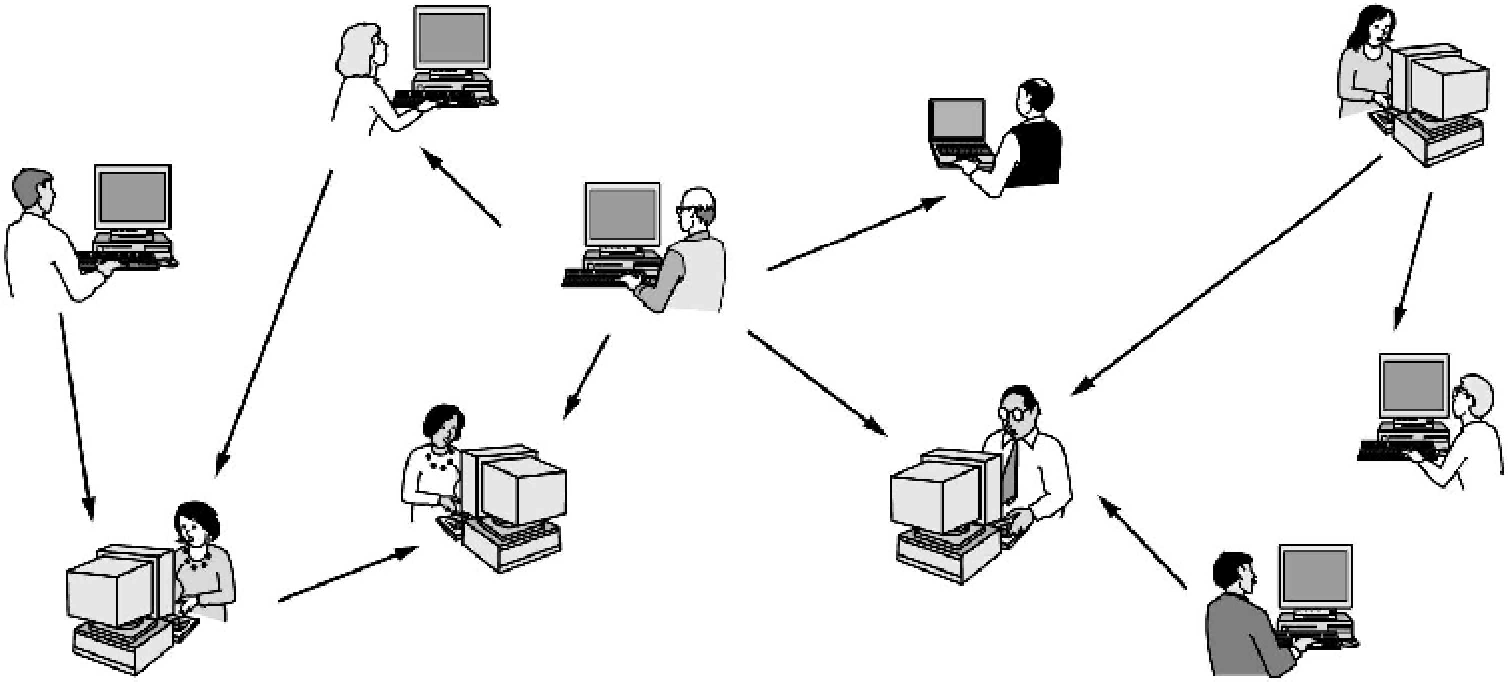
Applications (1)

- end systems (hosts):
  - run application programs
  - e.g. Web, email
  - at "edge of network"
- client/server model
  - client host requests, receives service from always-on server
  - e.g. Web browser/server; email client/server



Applications (2)

- peer-2-peer model:
    - No fixed clients or servers
    - Each host can act as both client & server
- Examples: Napster, Gnutella, KaZaA, Torrent

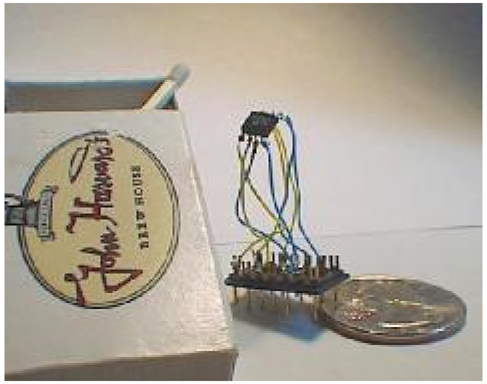


Applications (3)

- WWW
- File Transfer (FTP, SMB, Peer-to-Peer)
- E-mail
- Instant Messaging (Internet chat, text messaging)
- Remote Login and Terminals (SSH, RDP)
- Internet Phone
- Video-on-demand
- Network Games

"Cool" internet appliances

IP picture frame
http://www.ceiva.com/



World's smallest web server
http://www-ccs.cs.umass.edu/~shri/iPic.html

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

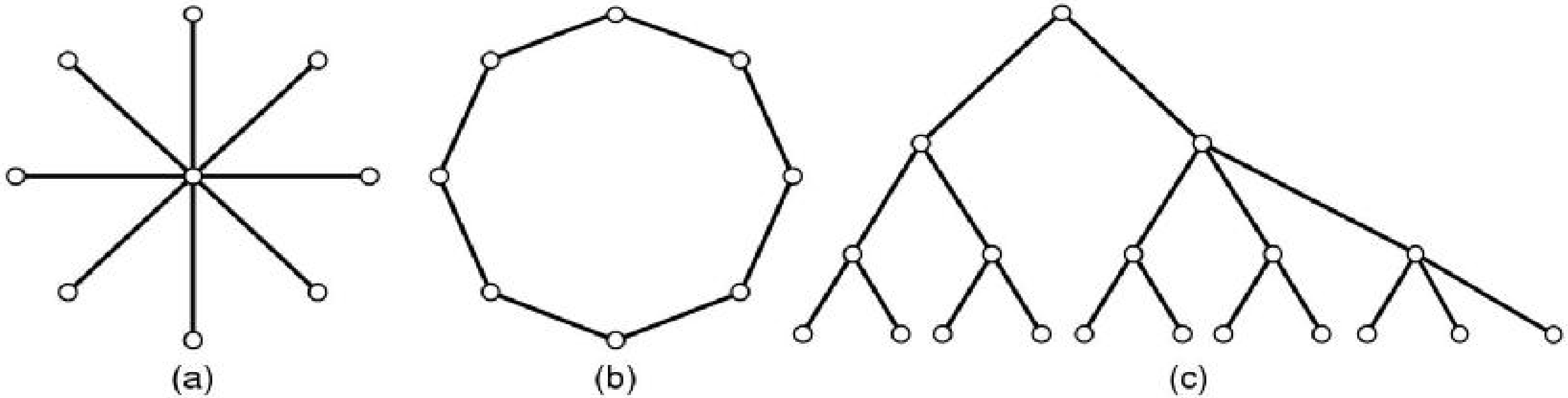
Delay & loss in packet-switched networks

Protocol layers, service models

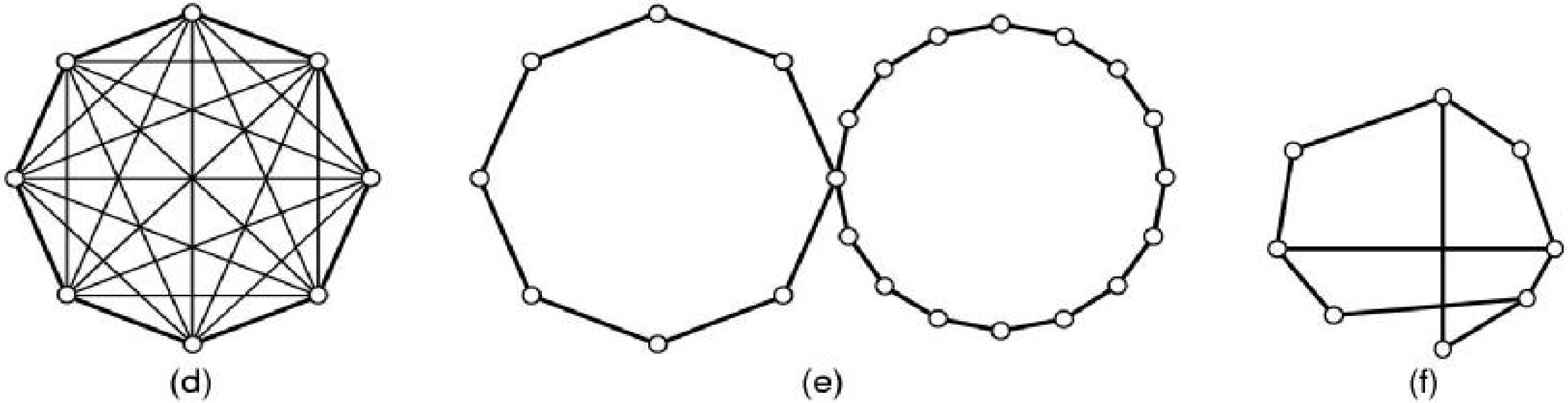
History

IP addressing and routing

Network Taxonomy: Topologies



- Network topology – configuration of nodes interconnection (which is absent at this figure?)



Network Taxonomy: on a network scale

Scale	BER, bit error ratio Nerr/Ntx	Maximum distance between nodes, km	Typical transfer rate, Mbps	Lines are in
PAN*	$10^{-7} – 10^{-9}$	0.1**	1-10	private property
LAN*	10^{-9}	1	10-10000***	private property
MAN*	10^{-6}	50	622-2488	private or municipal property
WAN*	$10^{-3} – 10^{-5}$	> 50	10-2488	property of tel.co.

Notes:

* PAN, LAN, MAN, WAN = Personal, Local, Metropolitan, Wide Area Network

** IEEE 802.15.1, Bluetooth

*** IEEE 802.3ae/an, 10Gigabit Ethernet

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some more Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

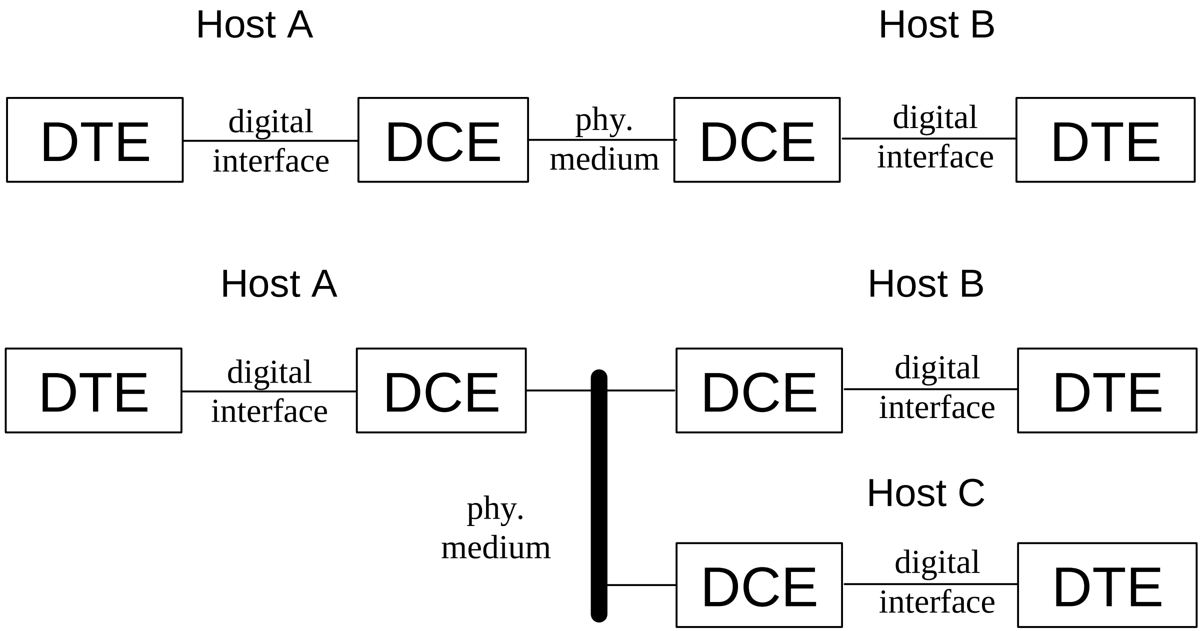
Protocol layers, service models

History

IP addressing and routing

Low-level network structure

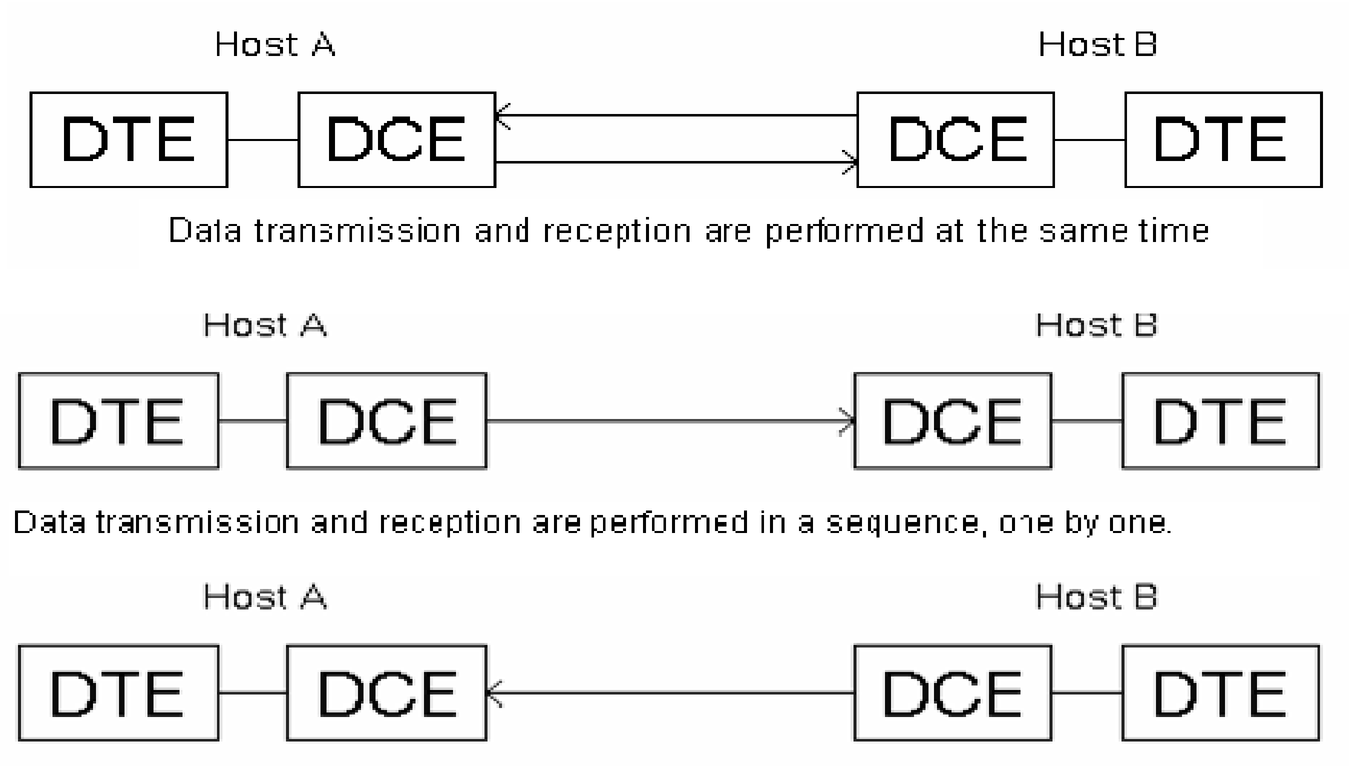
DTE – Data Terminating Equipment
DCE - Data Circuit-terminating Equipment



Point-to-point and multipoint configurations

Data Flows

- Duplex (full-duplex)
- Half-duplex
- Simplex



Data transmission and reception are performed at the same time

Data transmission and reception are performed in a sequence, one by one.

What's a protocol?

human protocols:

- "what's the time?"
- "I have a question"
- introductions

... specific msgs sent

... specific actions taken when msgs received, or other events

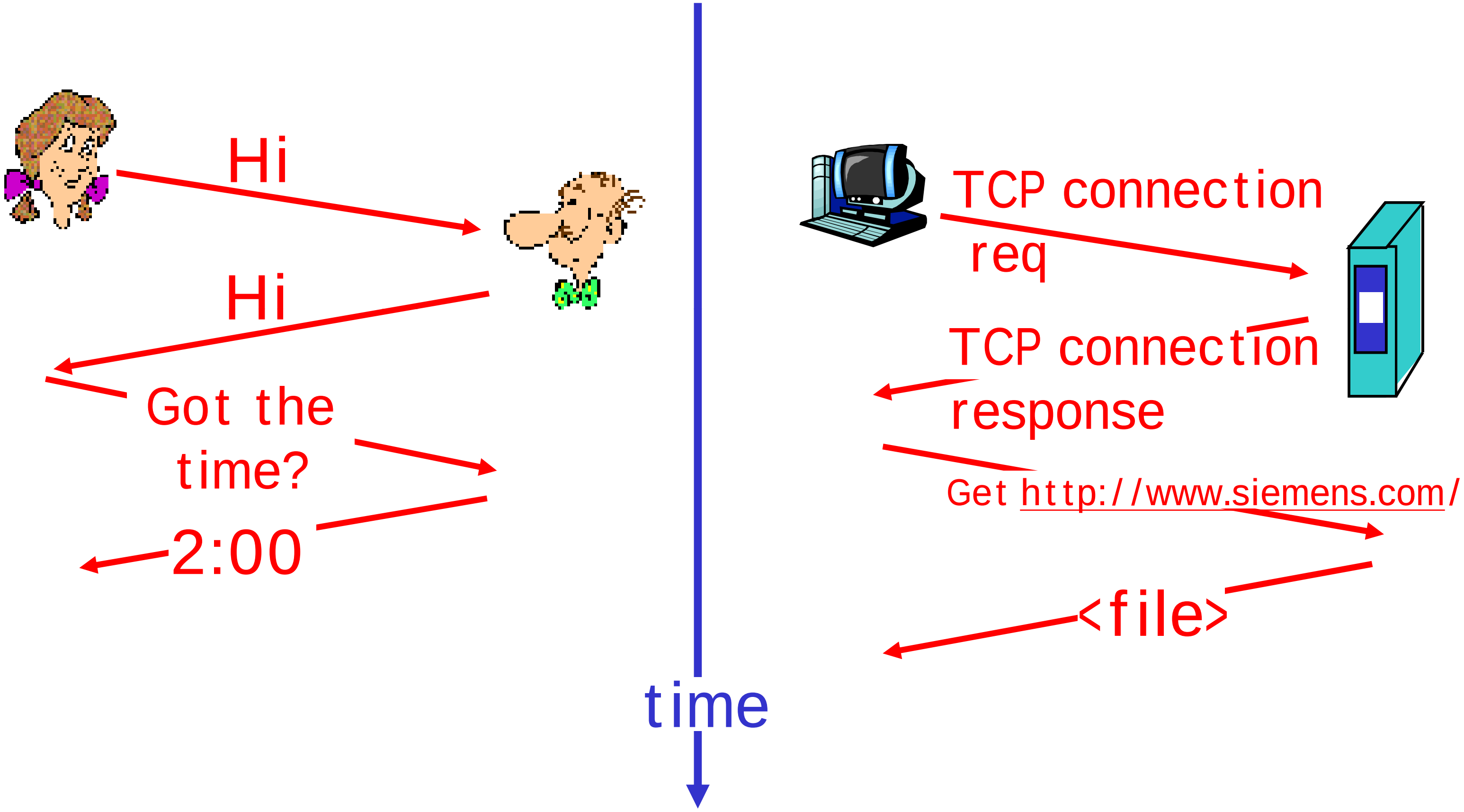
network protocols:

- machines rather than humans
- all communication activity in Internet governed by protocols

protocols define: format, order of msgs sent and received among network entities, and actions taken on msg transmission, receipt

What's a protocol?

a human protocol and a computer network protocol:



Q: Other human protocols?

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

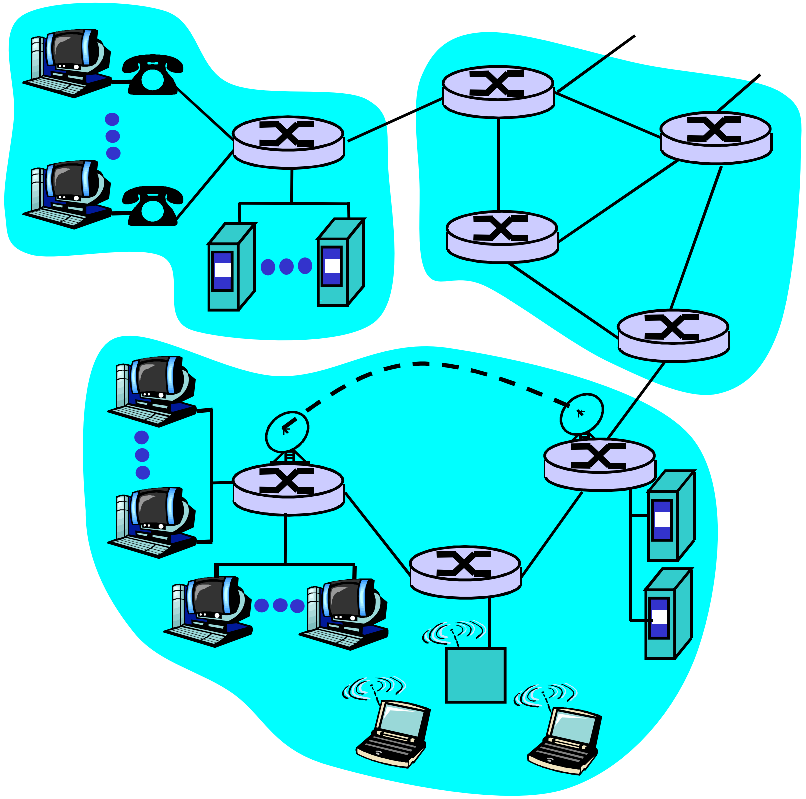
Protocol layers, service models

History

IP addressing and routing

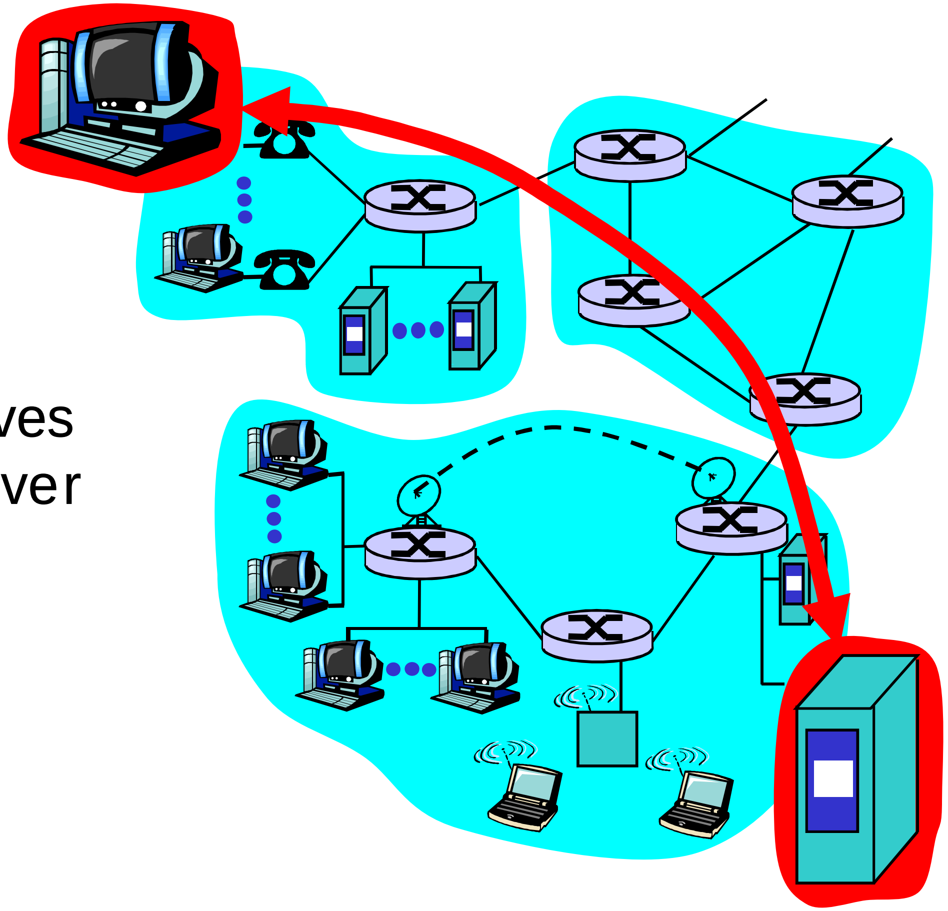
A closer look at network structure:

- network edge: applications and hosts
- network core:
  - routers
  - network of networks
- access networks, physical media: communication links



The network edge:

- q end systems (hosts):
  - m run application programs
  - m e.g. Web, email
  - m at “edge of network”
- q client/server model
  - m client host requests, receives service from always-on server
  - m e.g. Web browser/server; email client/server
- q peer-peer model:
  - m minimal (or no) use of dedicated servers
  - m e.g. Gnutella, KaZaA, Torrent



Network edge: connection-oriented service

Goal: data transfer between end systems

- *handshaking:* setup (prepare for) data transfer ahead of time
  - Hello, hello back human protocol
  - *set up "state"* in two communicating hosts
- TCP - Transmission Control Protocol
  - Internet's connection-oriented service

TCP service [RFC 793]

- *reliable, in-order* byte-stream data transfer
  - loss: acknowledgements and retransmissions
- *flow control:*
  - sender won't overwhelm receiver
- *congestion control:*
  - senders "slow down sending rate" when network congested

Network edge: connectionless service

Goal: data transfer between end systems

- same as before!

- UDP - User Datagram Protocol [RFC 768]: Internet's connectionless service
  - unreliable data transfer
  - no flow control
  - no congestion control

App's using TCP:

- HTTP (Web), FTP (file transfer), Telnet (remote login), SMTP (email)

App's using UDP:

- streaming media, teleconferencing, DNS, Internet telephony
- Q: Why m/media?

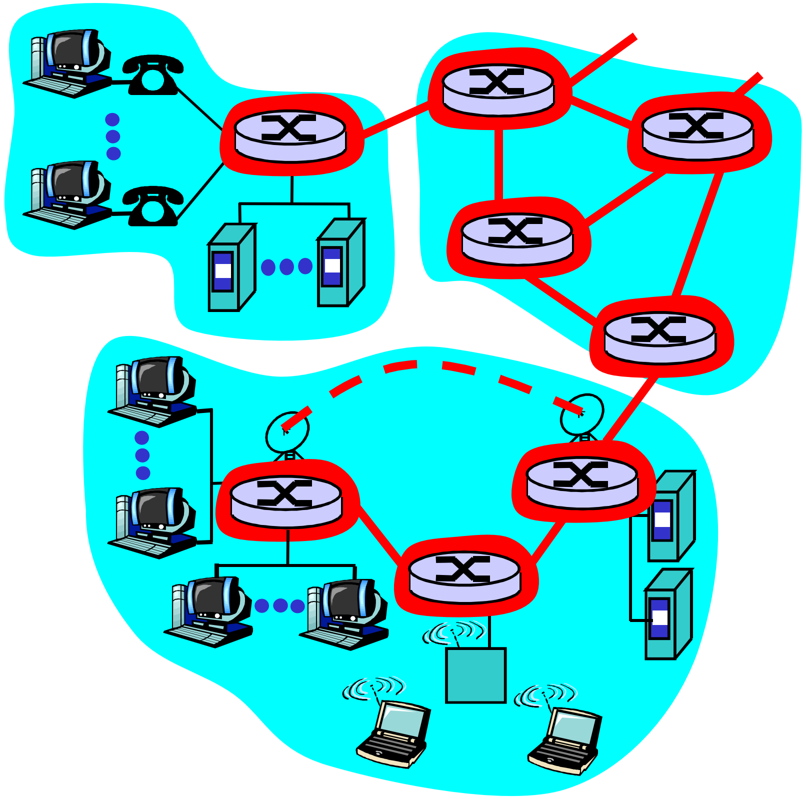
The Network Core

q mesh of interconnected routers

q *the* fundamental question: how is data transferred through net?

  m circuit switching: dedicated circuit per call: telephone net

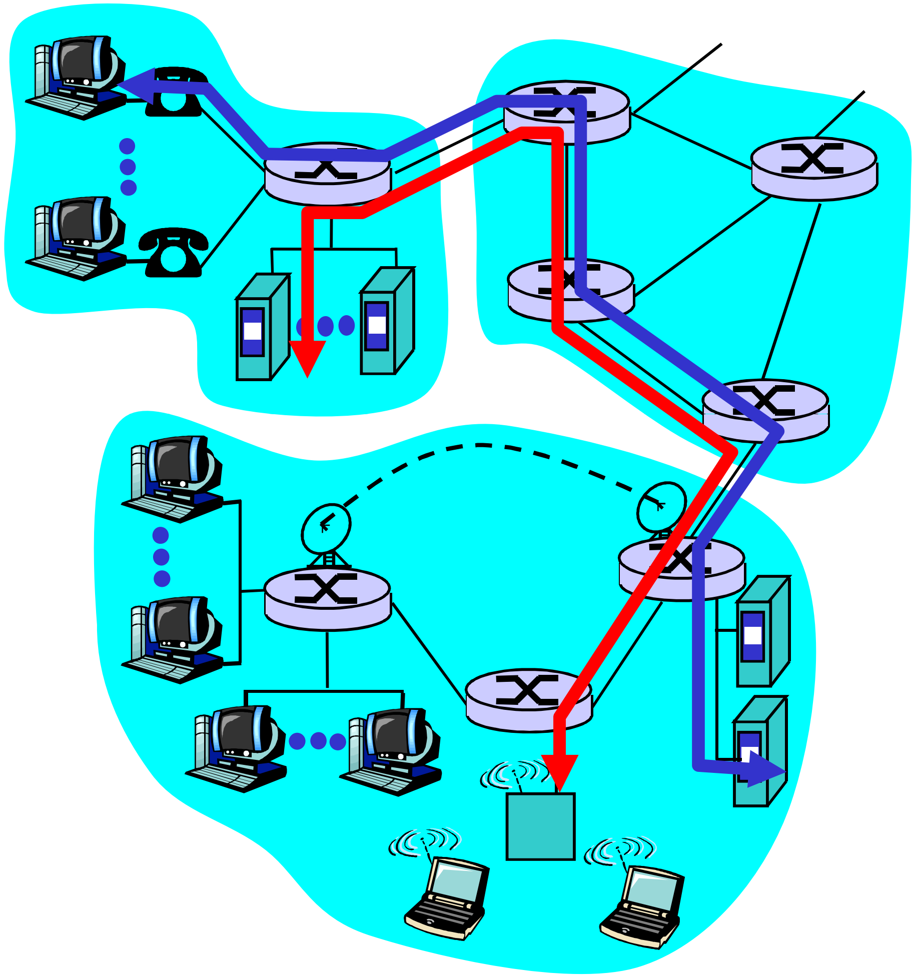
  m packet-switching: data sent thru net in discrete “chunks”



Network Core: Circuit Switching

End-end resources reserved for "call"

- link bandwidth, switch capacity
- dedicated resources: no sharing
- circuit-like (guaranteed) performance
- call setup required



Network Core: Circuit Switching

network resources (e.g., bandwidth) divided into "pieces"

- pieces allocated to calls
- resource piece *idle* if not used by owning call *(no sharing)*

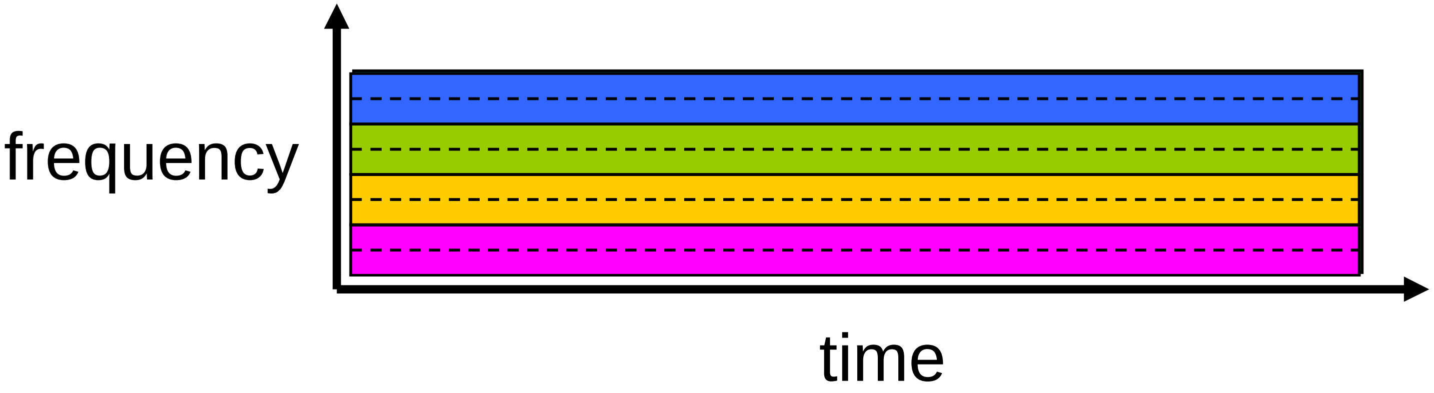
- dividing link bandwidth into "pieces"
  - frequency division
  - time division

Circuit Switching: FDMA and TDMA

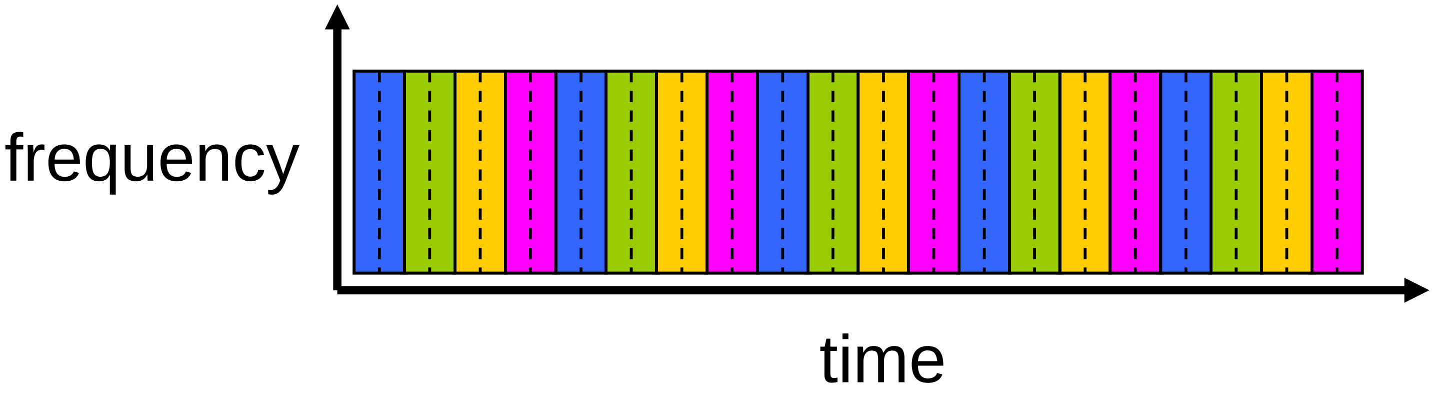
Example:

4 users 🟦🟩🟨🟪

FDMA



TDMA



Network Core: Packet Switching

each end-end data stream divided into *packets*

- user A, B packets *share* network resources
- each packet uses full link bandwidth
- resources used *as needed*

Bandwidth division into "pieces"

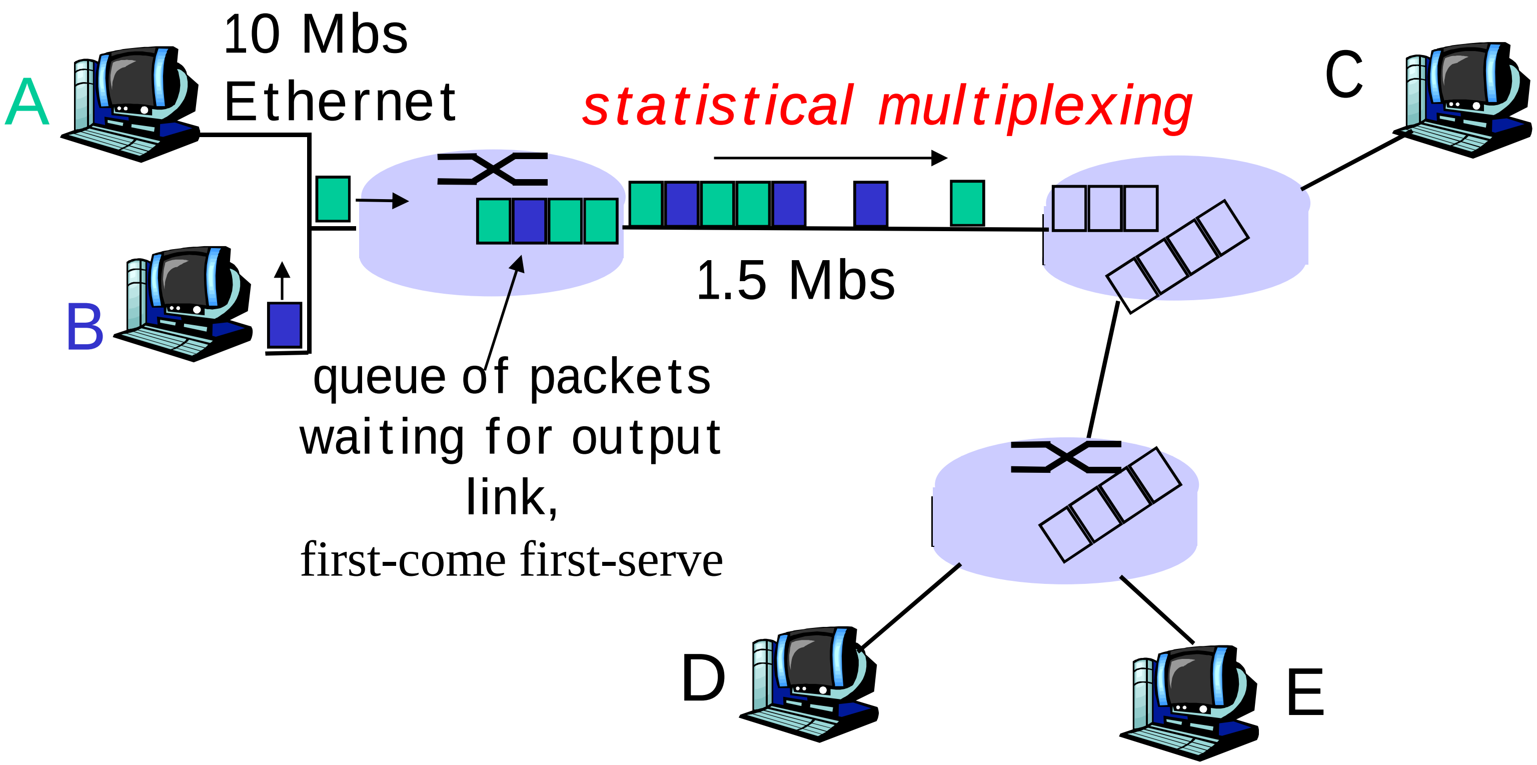
Dedicated allocation

Resource reservation

resource contention:

- aggregate resource demand can exceed amount available
- congestion: packets queue, wait for link use
- store and forward: packets move one hop at a time
  - transmit over link
  - wait turn at next link

Packet Switching: Statistical Multiplexing



Sequence of A & B packets does not have fixed pattern è *statistical multiplexing*.

In TDM each host gets same slot in revolving TDM frame.

Packet switching versus circuit switching

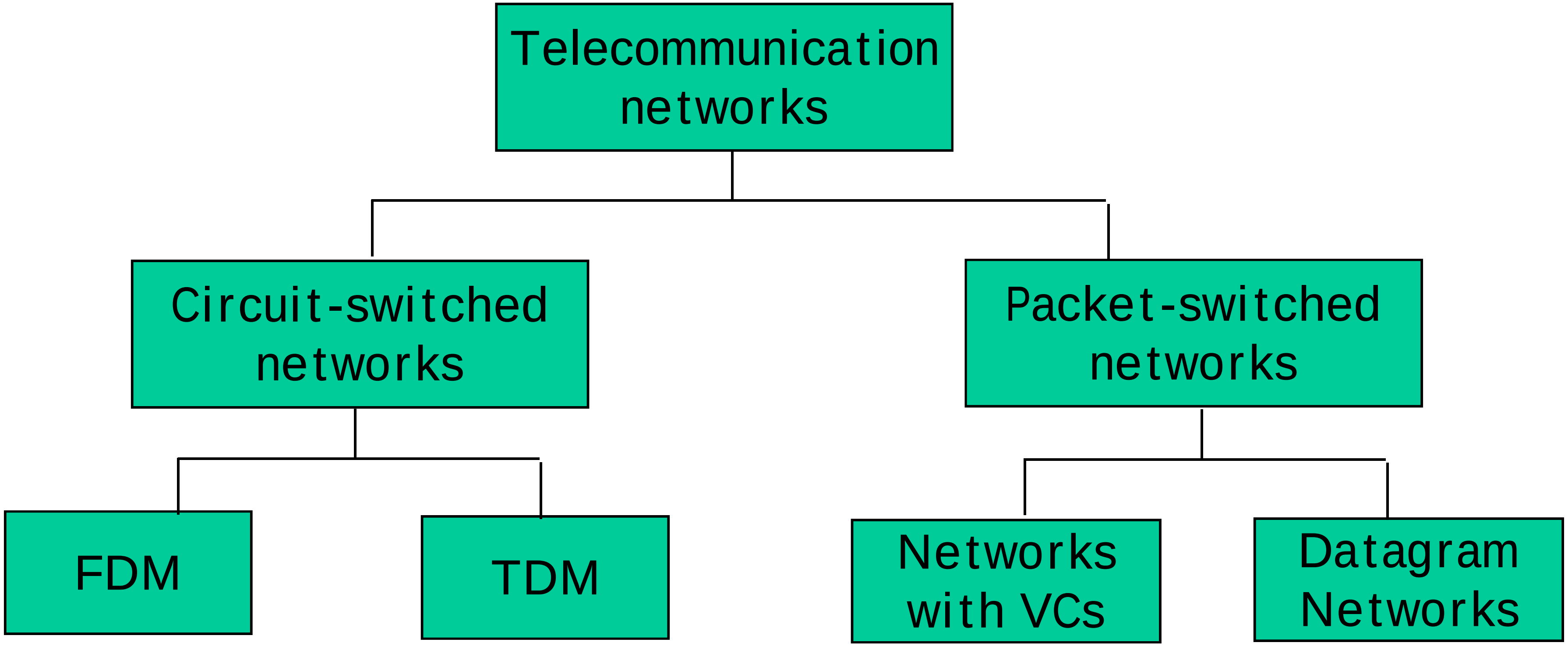
Is packet switching a “slam dunk winner?”

- Great for bursty data
  - resource sharing
  - simpler, no call setup
- Excessive congestion: packet delay and loss
  - protocols needed for reliable data transfer, congestion control
- Q: How to provide circuit-like behavior?
  - bandwidth guarantees needed for audio/video apps

Packet-switched networks: forwarding

- *Goal:* move packets through routers from source to destination
- datagram network:
  - *destination address* in packet determines next hop
  - routes may change during session
  - analogy: driving, asking directions
- virtual circuit network:
  - each packet carries tag (virtual circuit ID), tag determines next hop
  - fixed path determined at *call setup time*, remains fixed thru call
  - *network equipment maintain per-call state*

Yet another Network Taxonomy



- Datagram network is *not* either connection-oriented or connectionless.
- Internet provides both connection-oriented (TCP) and connectionless services (UDP) to apps.

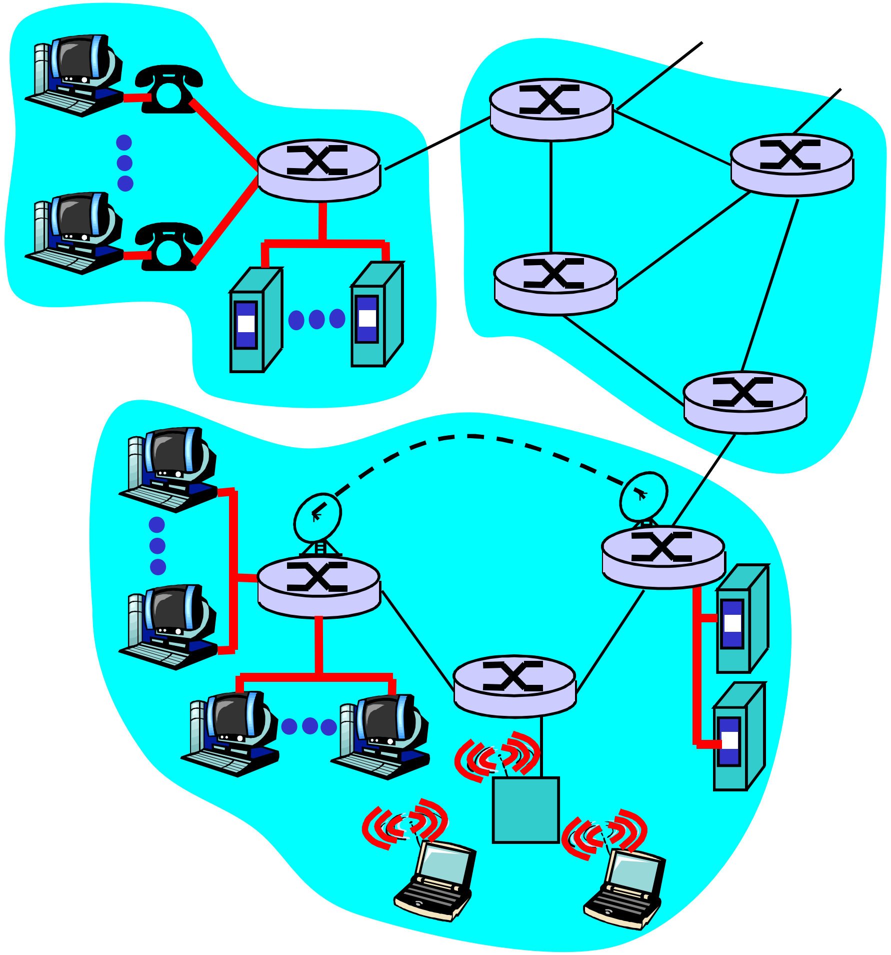
Access networks and physical media

Q: How to connect end systems to edge router?

- residential access nets
- institutional access networks (school, company)
- mobile access networks

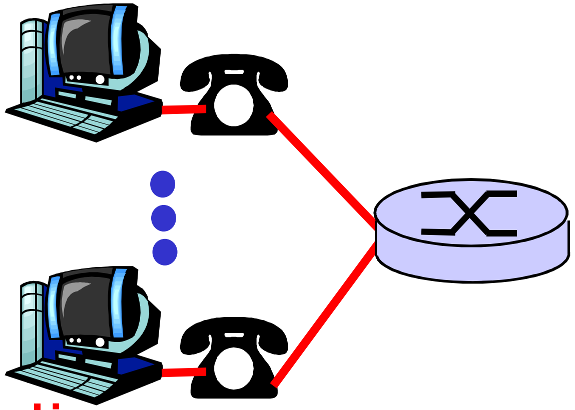
Keep in mind:

- bandwidth (bits per second) of access network?
- shared or dedicated?



Residential access: point to point access

- Dialup via modem
  - up to 56Kbps direct access to router (often less)
  - Can't surf and phone at same time: can't be "always on"



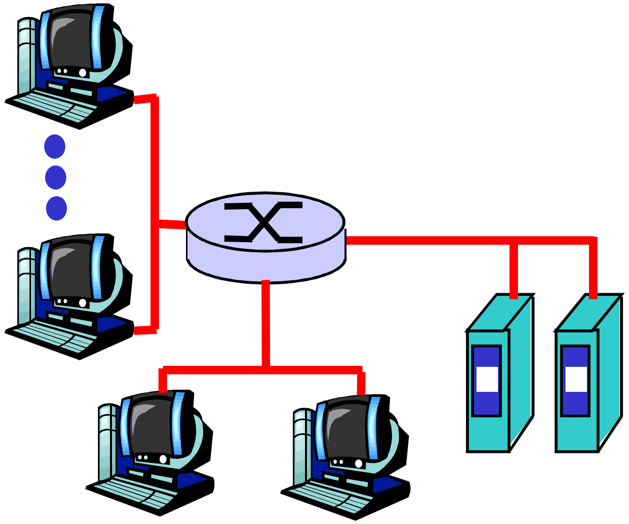
- ADSL: asymmetric digital subscriber line
  - up to 1.5 Mbps upstream (today typically < 1 Mbps)
  - up to 50 Mbps downstream (today typically<24 Mbps)
  - FDM: 50(100) kHz – 1(2) MHz (depends on std) for downstream

    4 kHz – 50(100) kHz for upstream (depends on std)

    0 kHz - 4 kHz for ordinary telephone
- Cable modems (HFC), FTTH

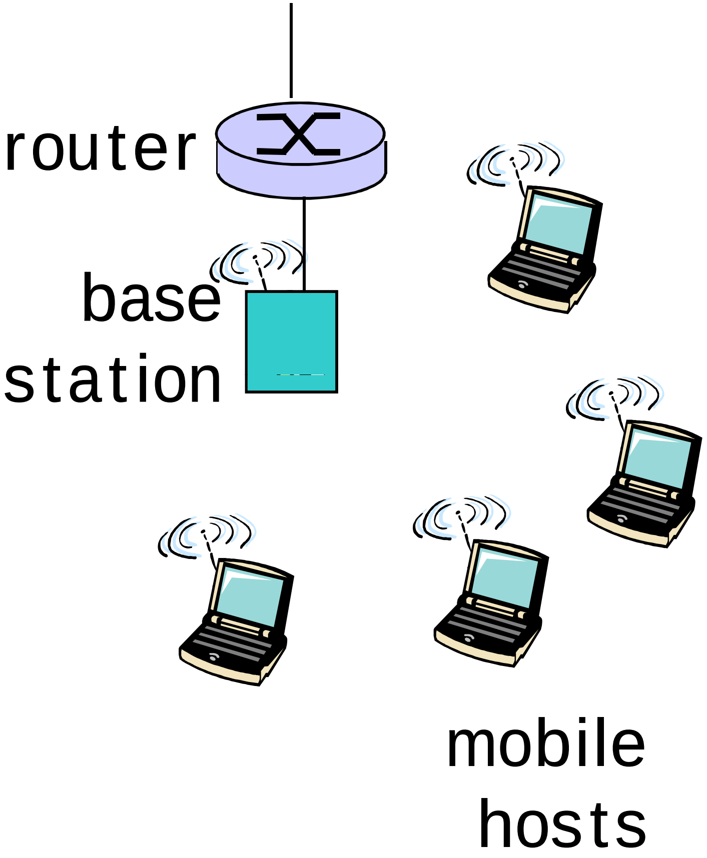
Company access: local area networks

- company/univ local area network (LAN) connects end system to edge router
- Ethernet:
  - shared or dedicated link connects end system and router
  - 10 Mbs, 100Mbps, 1-10 Gigabit Ethernet
- deployment: institutions, home LANs happening now
- LANs: will study further



Wireless access networks

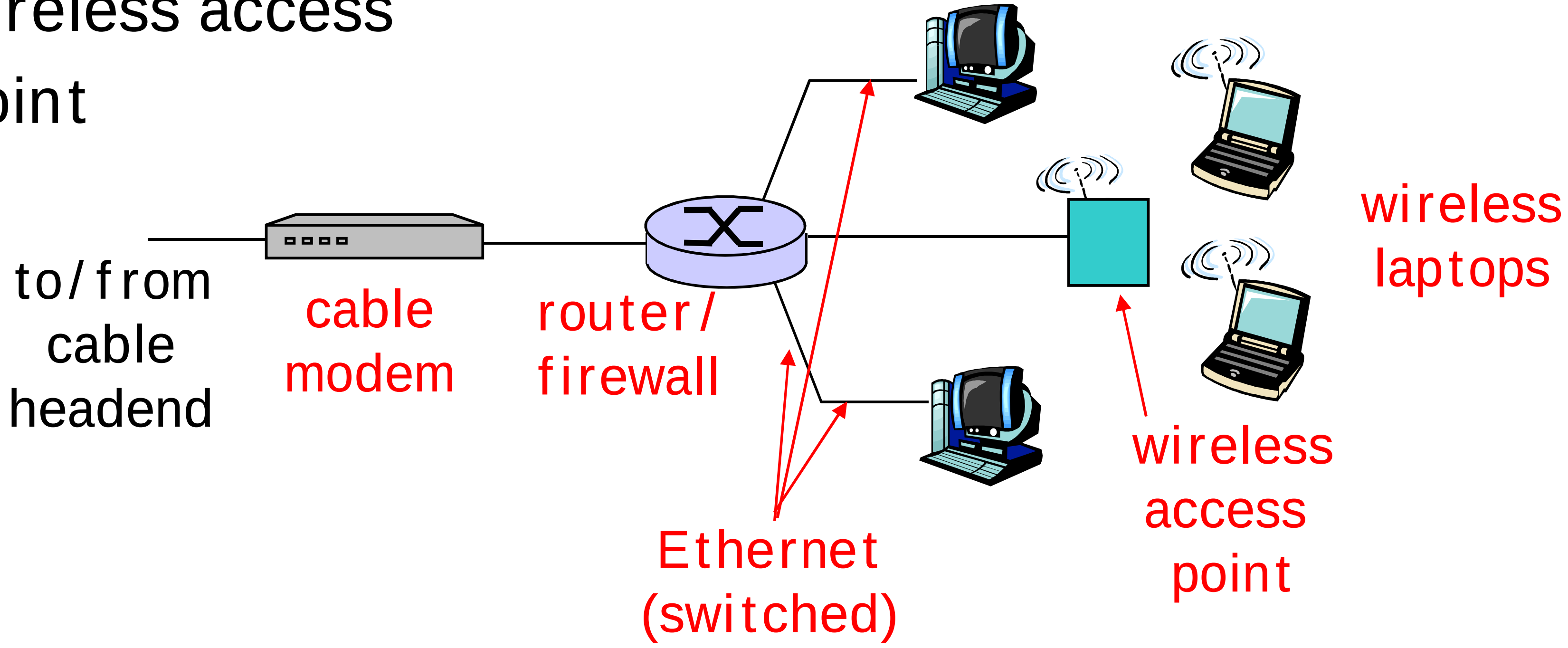
- shared *wireless* access network connects end system to router
  - via base station aka “access point”
- wireless LANs:
  - 802.11b/g/n (WiFi): 11/54/300-600 Mbps
  - Infrastructure and Ad-Hoc modes
- wider-area wireless access
  - provided by telco operator
  - 3G
  - WAP/GPRS



Home networks

Typical home network components:

- ADSL or cable modem
- router/firewall/NAT
- Ethernet
- wireless access point



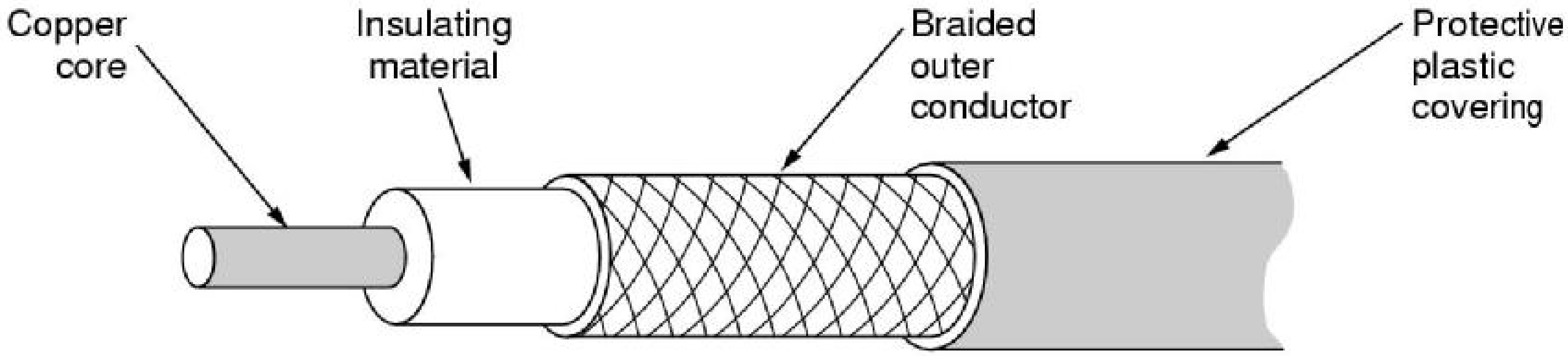
Physical Media

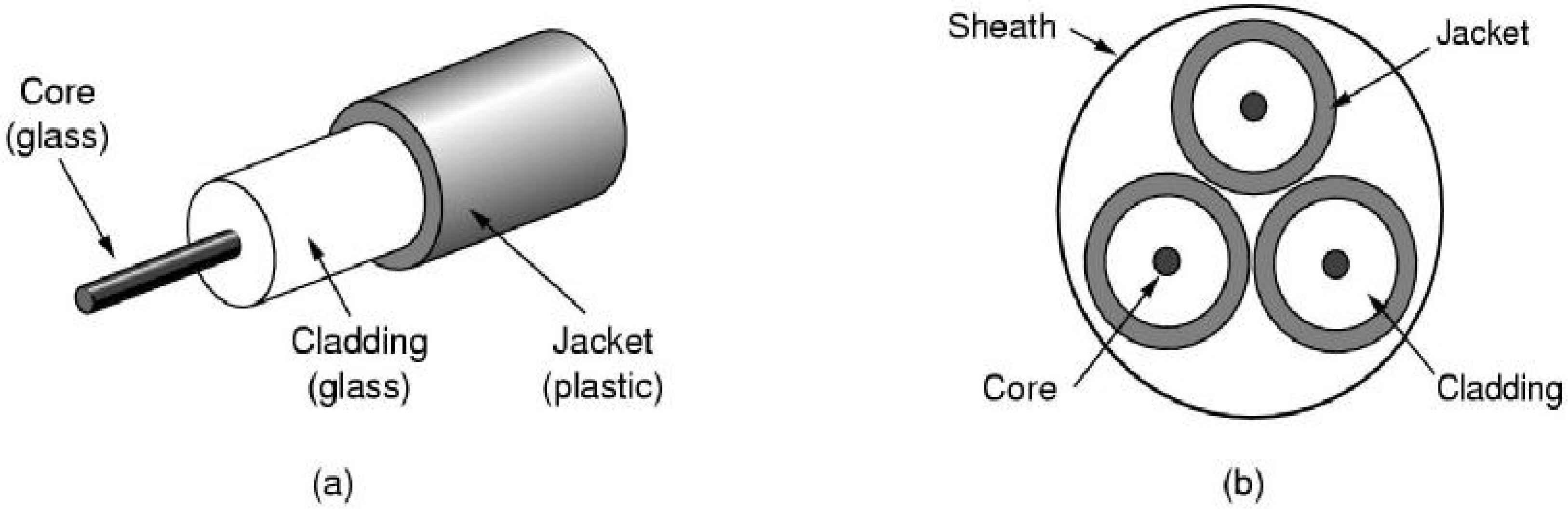
- Bit: propagates between transmitter/rcvr pairs
- physical link: what lies between transmitter & receiver
- guided media:
  - signals propagate in solid media: copper, fiber, coax
- unguided media:
  - signals propagate freely, e.g., radio

Twisted Pair (TP)

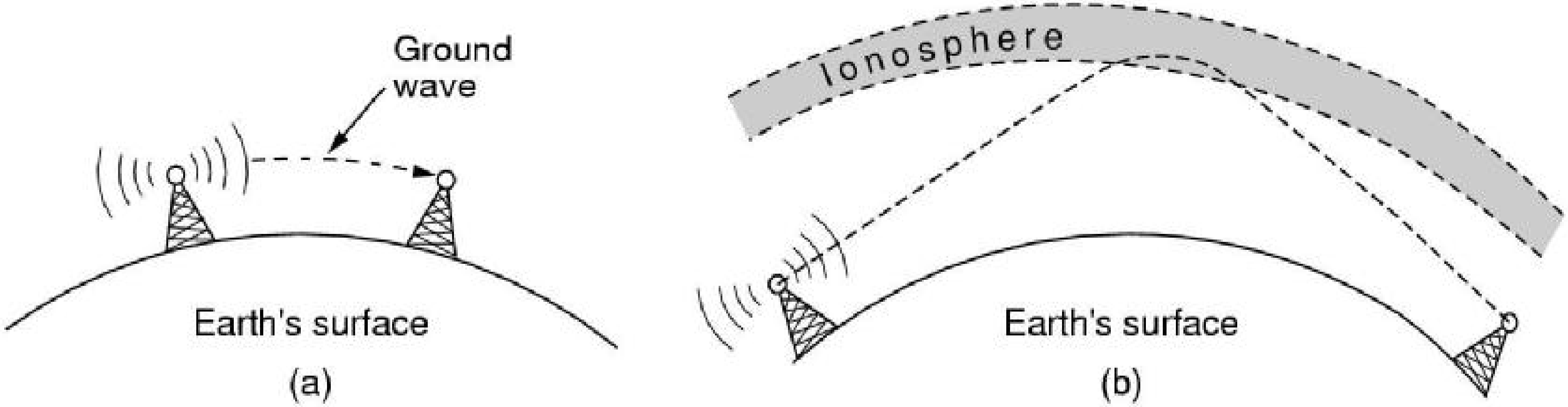
- two insulated copper wires
  - Category 3: traditional phone wires, 10 Mbps Ethernet
  - Category 5, 5e, 6, 6a, 7: 100 .. 10000Mbps

Guided media: coax-cable, TP, FO

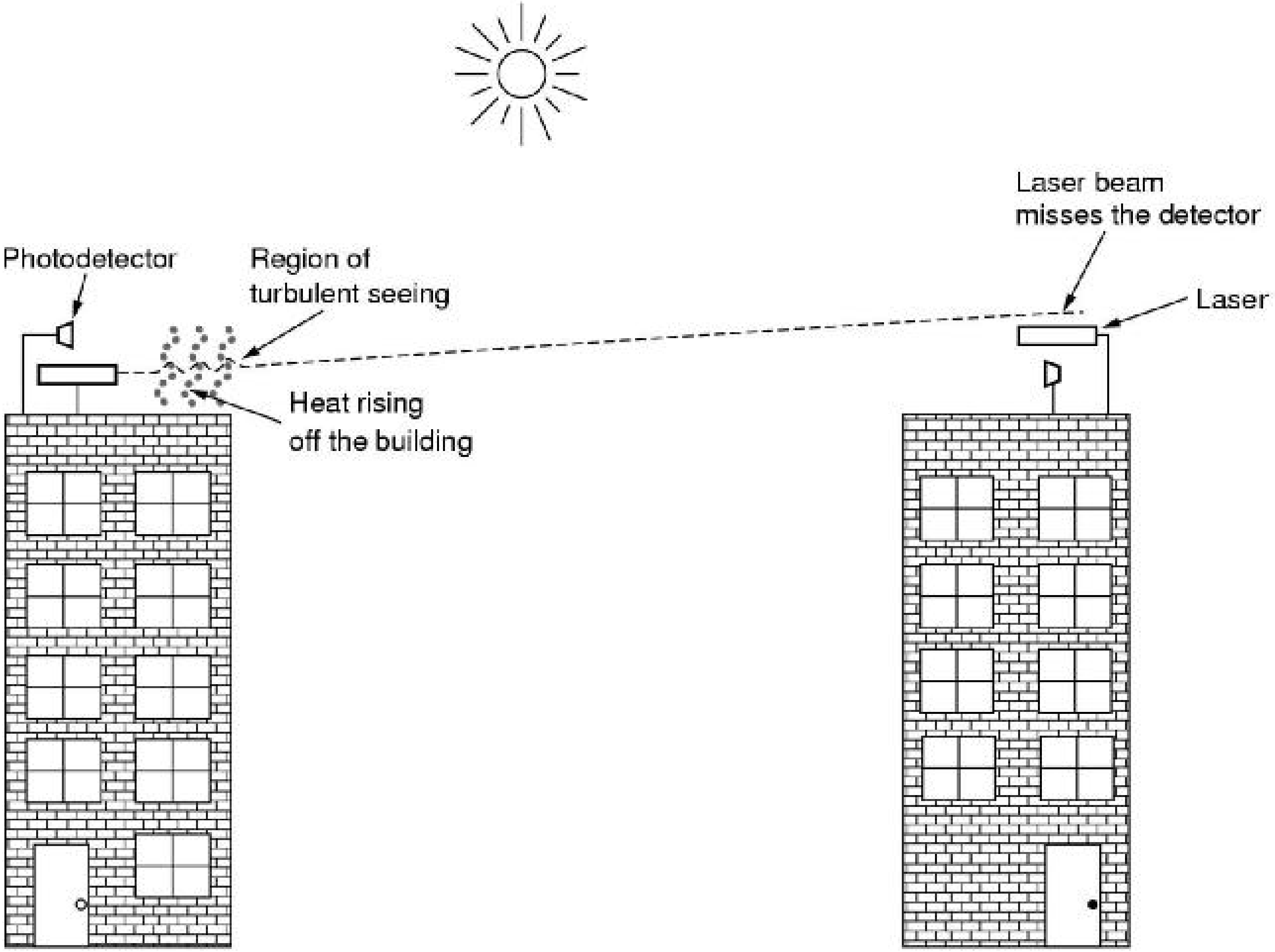




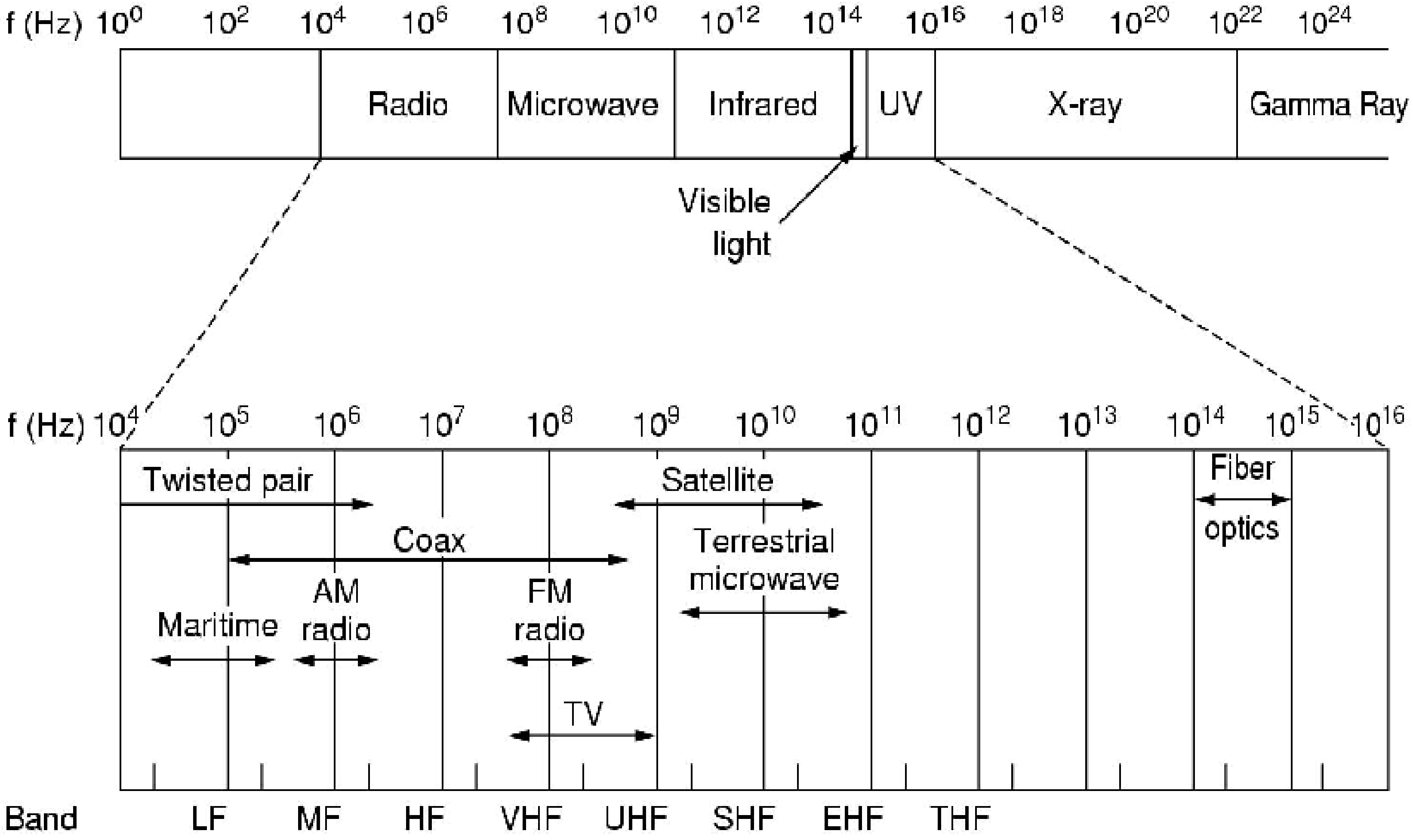
Unguided media



IR lasers

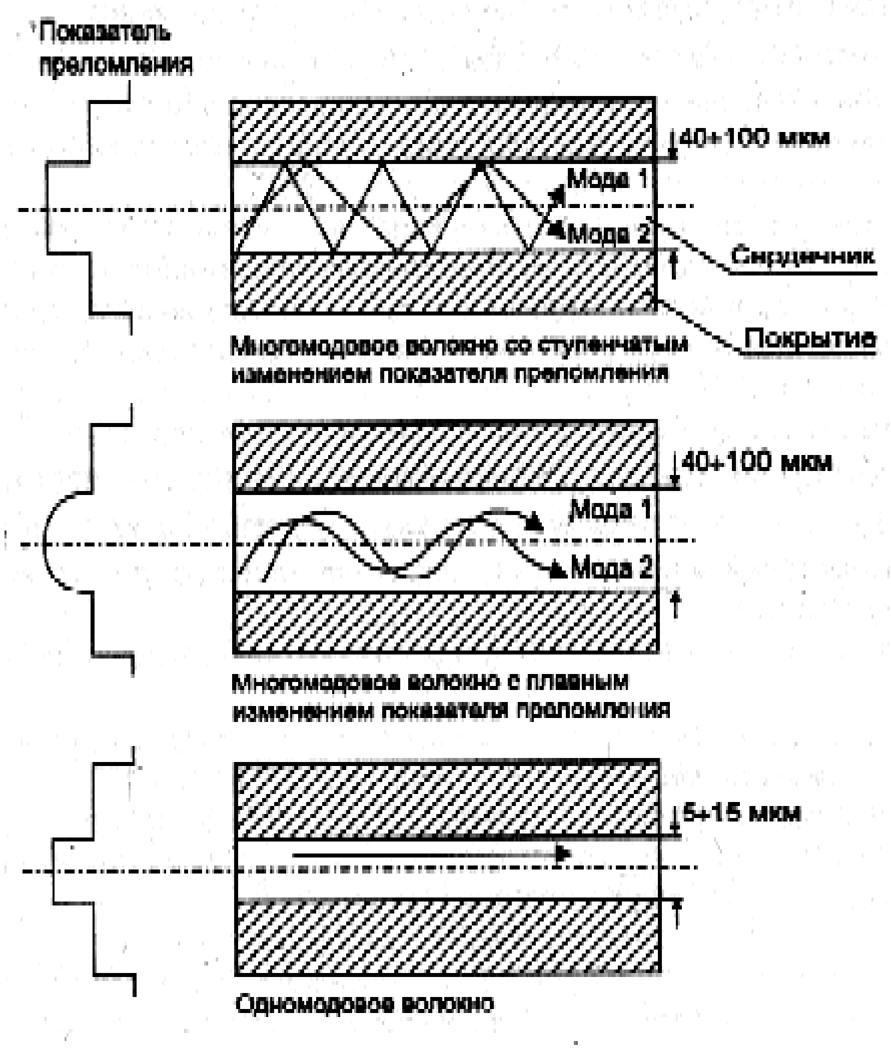

Laser beam
misses the detector
Photodetector
Region of
turbulent seeing
Laser
Heat rising
off the building

Electromagnetic spectrum

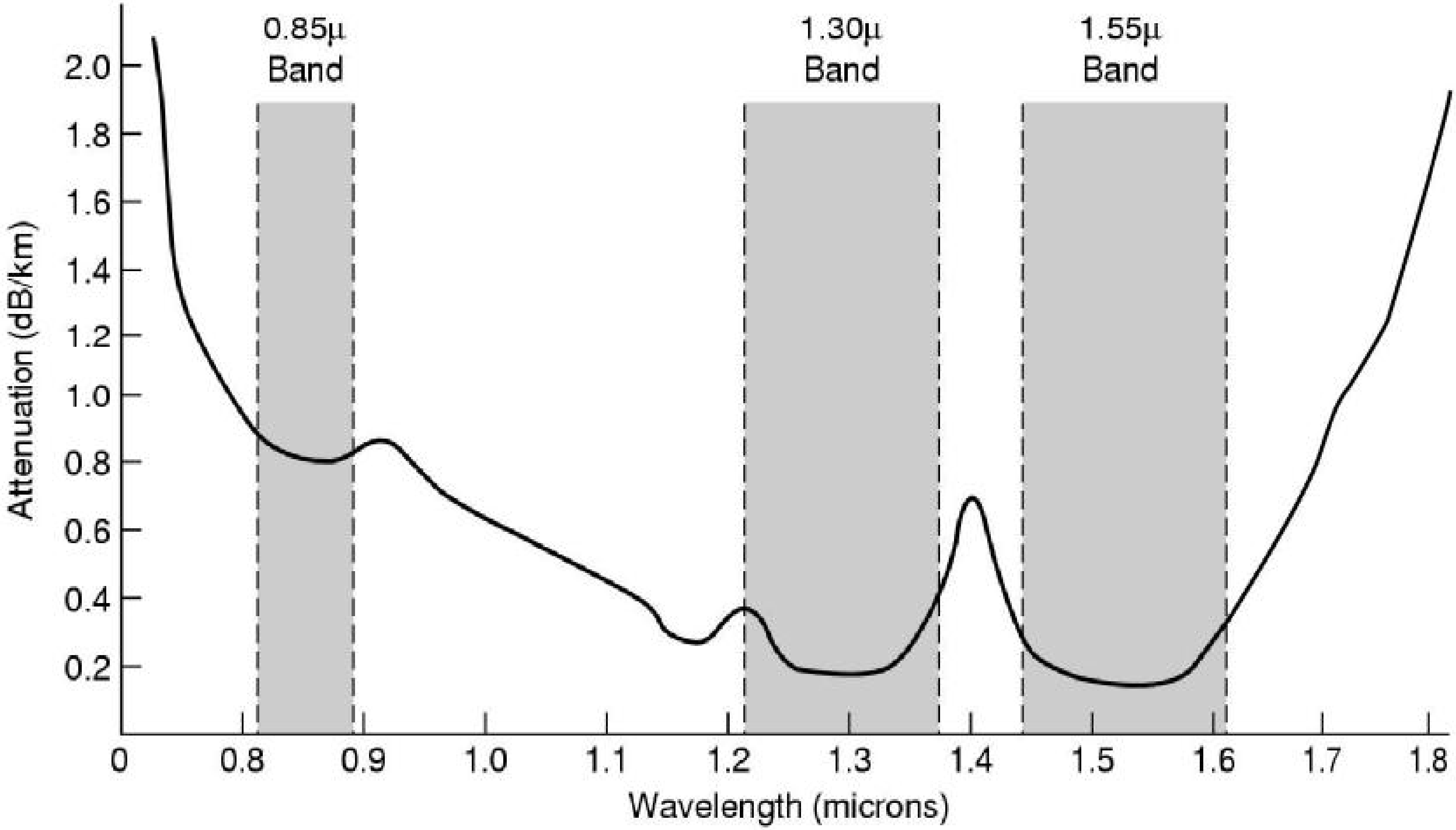


f (Hz) 10^0 10^2 10^4 10^6 10^8 10^{10} 10^{12} 10^{14} 10^{16} 10^{18} 10^{20} 10^{22} 10^{24}
Radio
Microwave
Infrared
UV
X-ray
Gamma Ray
Visible light
f (Hz) 10^4 10^5 10^6 10^7 10^8 10^9 10^{10} 10^{11} 10^{12} 10^{13} 10^{14} 10^{15} 10^{16}
Twisted pair
Satellite
Fiber optics
Coax
Terrestrial microwave
AM radio
FM radio
Maritime
TV
Band
LF
MF
HF
VHF
UHF
SHF
EHF
THF

FO cables: MMF, SMF


Показатель преломления
40÷100 мкм
Мода 1
Мода 2
Сердечник
Покрытие
Многомодовое волокно со ступенчатым изменением показателя преломления
40÷100 мкм
Мода 1
Мода 2
Многомодовое волокно с плавным изменением показателя преломления
5÷15 мкм
Одномодовое волокно

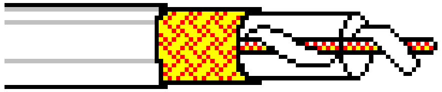
FO: attenuation in IR region



Physical Media: coax, fiber

Coaxial cable:

- two concentric copper conductors
- bidirectional
- baseband:
  - single channel on cable
  - legacy Ethernet
- broadband:
  - multiple channel on cable
  - HFC



Fiber optic cable:

- glass fiber carrying light pulses, each pulse a bit
- high-speed operation:
  - high-speed point-to-point transmission (e.g. 500! Gps)
- low attenuation: repeaters spaced far apart ; immune to electromagnetic noise

Physical media: radio

- signal carried in electromagnetic spectrum
- no physical “wire”
- bidirectional
- propagation environment effects:
  - reflection
  - obstruction by objects
  - interference

Radio link types:

- terrestrial microwave
  - e.g. up to 45 Mbps channels
- LAN (e.g., WaveLAN)
  - 2, 11, 54, 300-600 Mbps
- wide-area (e.g., cellular)
  - e.g. 3G: hundreds of kbps
- satellite
  - up to 50Mbps channel (or multiple smaller channels)
  - 270 msec end-end delay
  - geosynchronous versus LEOS

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

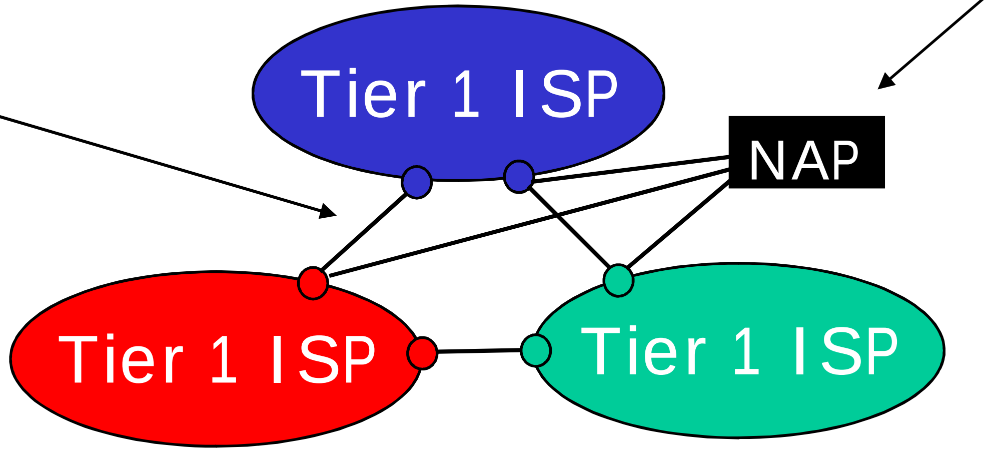
Protocol layers, service models

History

IP addressing and routing

Internet structure: network of networks

- roughly hierarchical
- at center: "tier-1" ISPs (e.g., UUNet, Sprint, AT&T, NORDUnet, C&W, TeliaSonera, Level3), national/international coverage
  - treat each other as equals

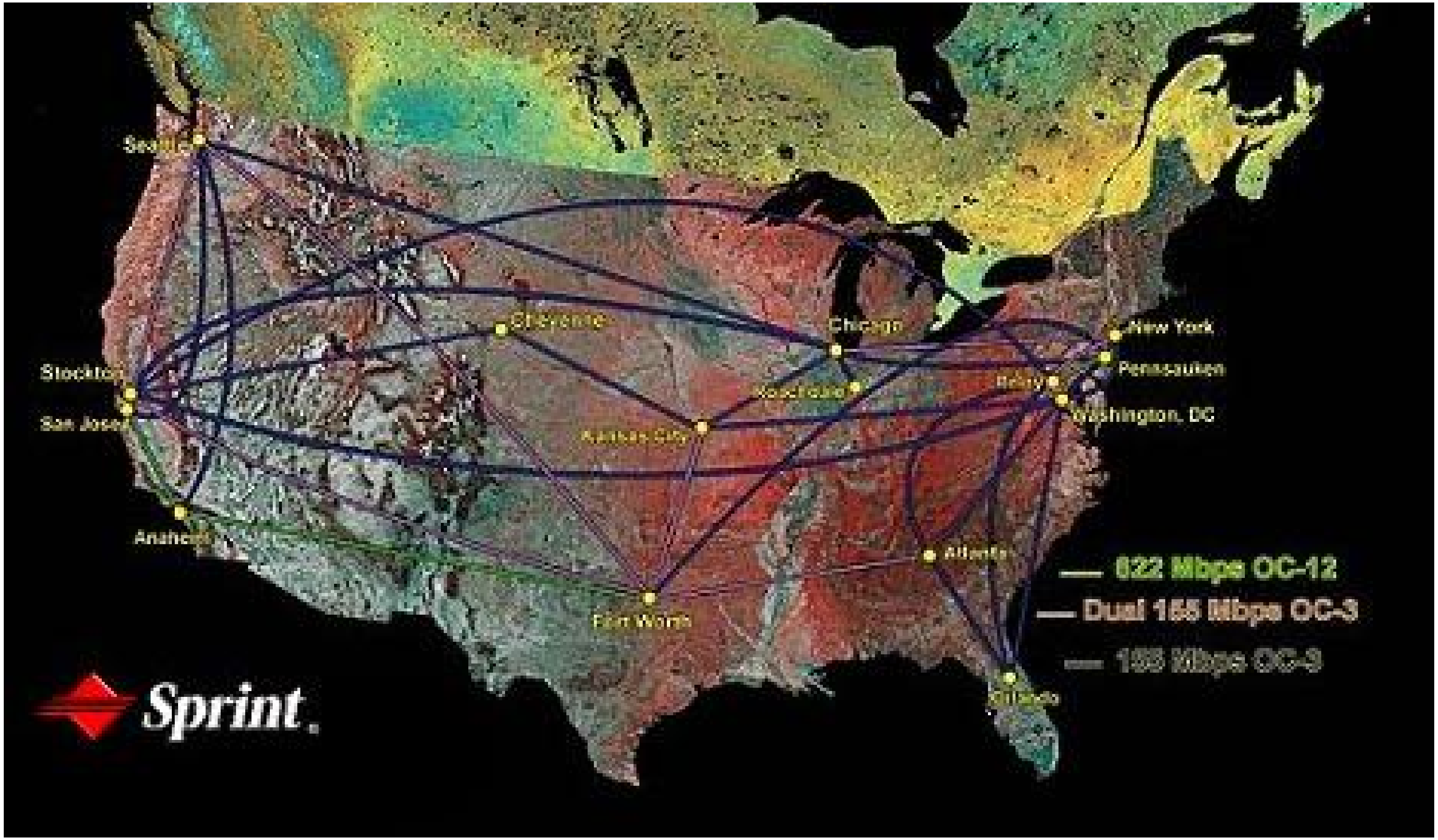


Tier-1 providers interconnect (peer) privately

Tier-1 providers also interconnect at public network access points (NAPs)

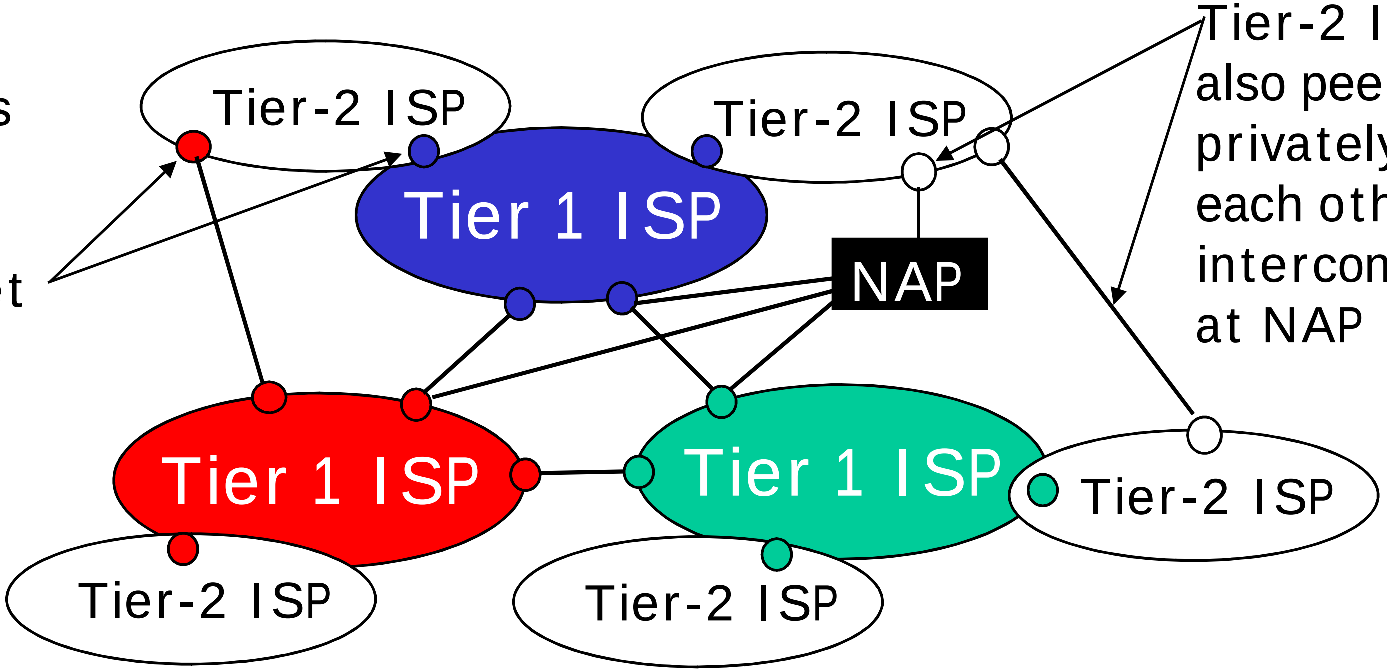
Tier-1 ISP: e.g., Sprint

Sprint US backbone network



Internet structure: network of networks

- "Tier-2" ISPs: smaller (often regional) ISPs
  - Connect to one or more tier-1 ISPs, possibly other tier-2 ISPs



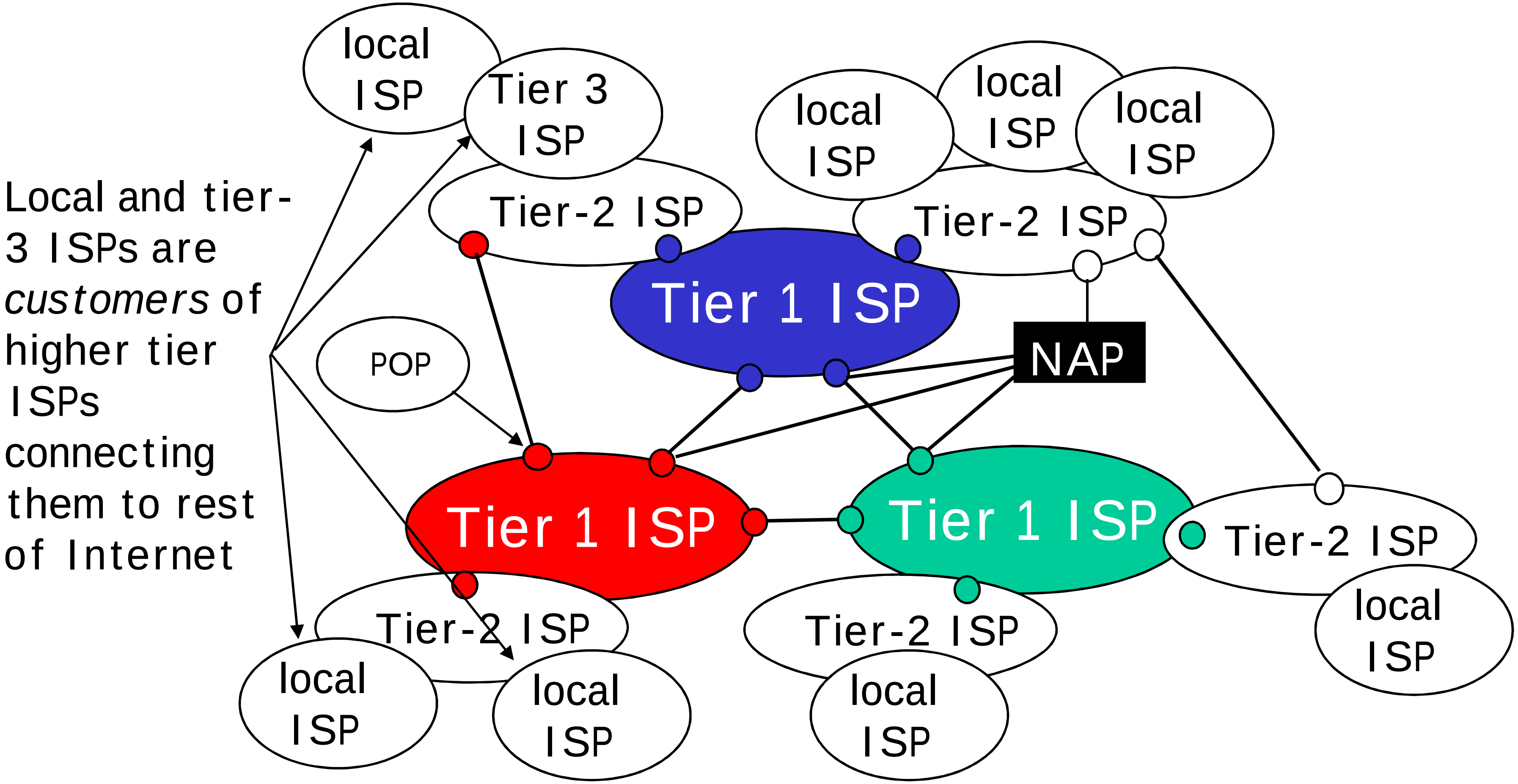
Tier-2 ISP pays tier-1 ISP for connectivity to rest of Internet

- tier-2 ISP is *customer* of tier-1 provider

Tier-2 ISPs also peer privately with each other, interconnect at NAP

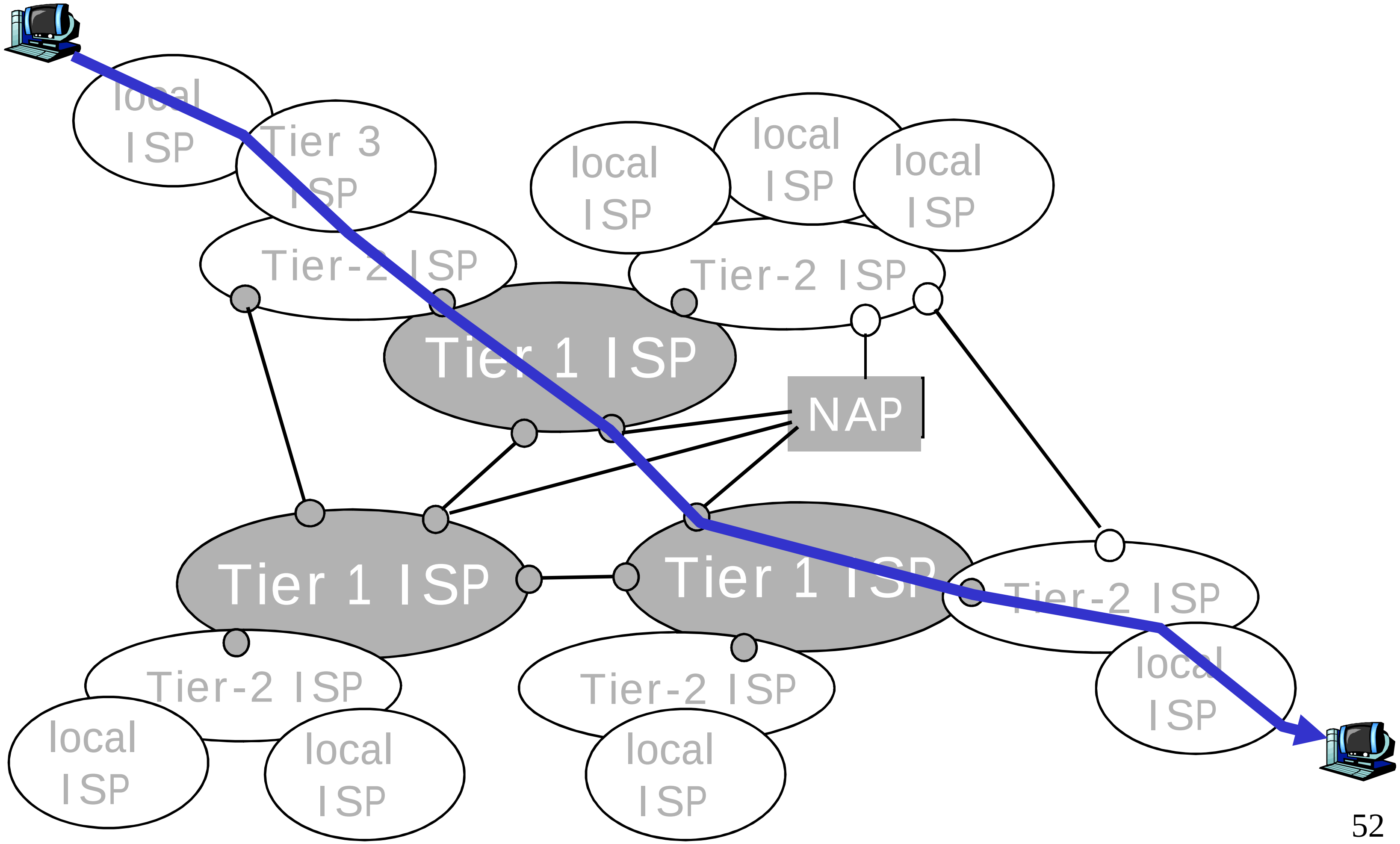
Internet structure: network of networks

- "Tier-3" ISPs and local ISPs
  - last hop ("access") network (closest to end systems)

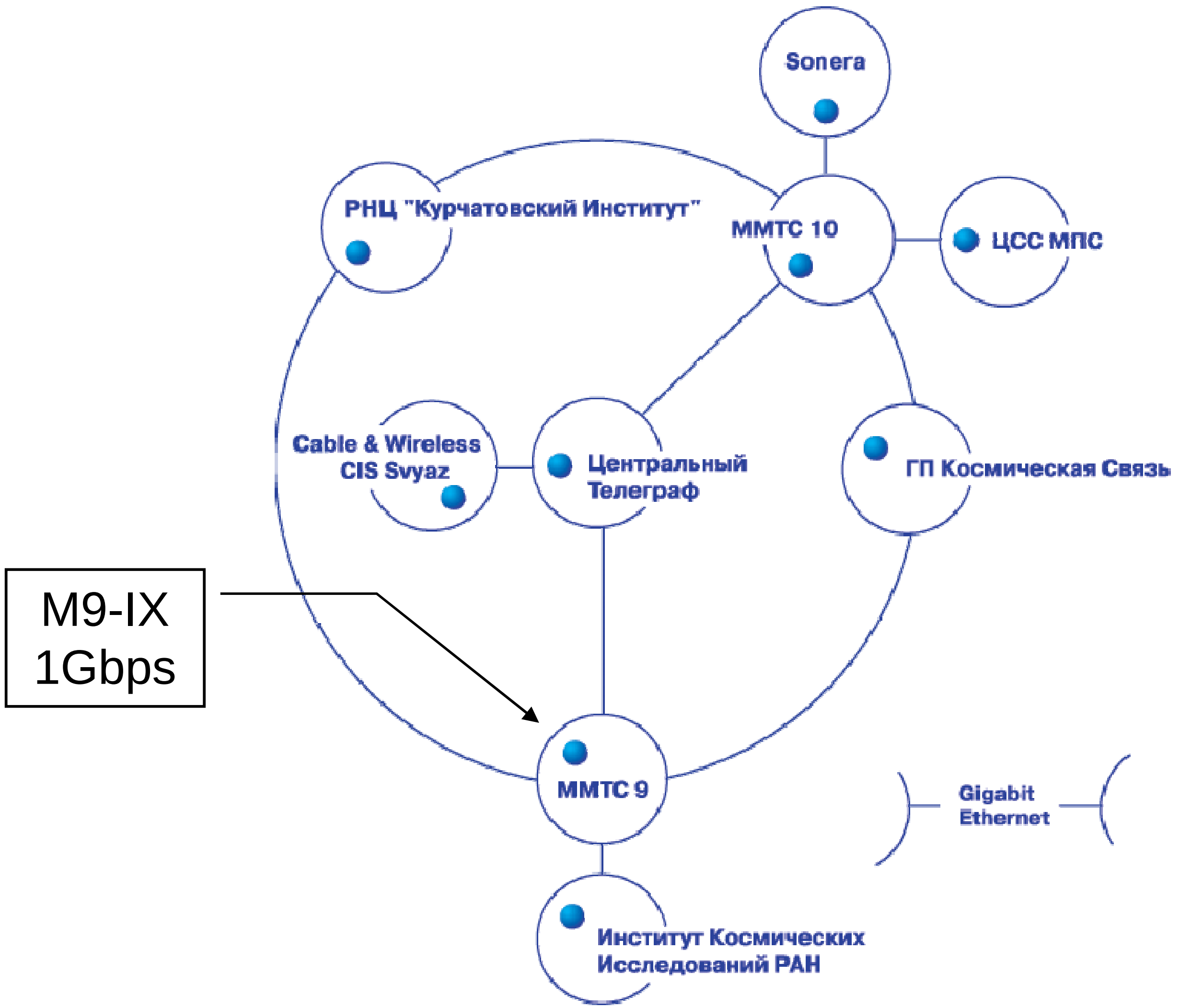


Internet structure: network of networks

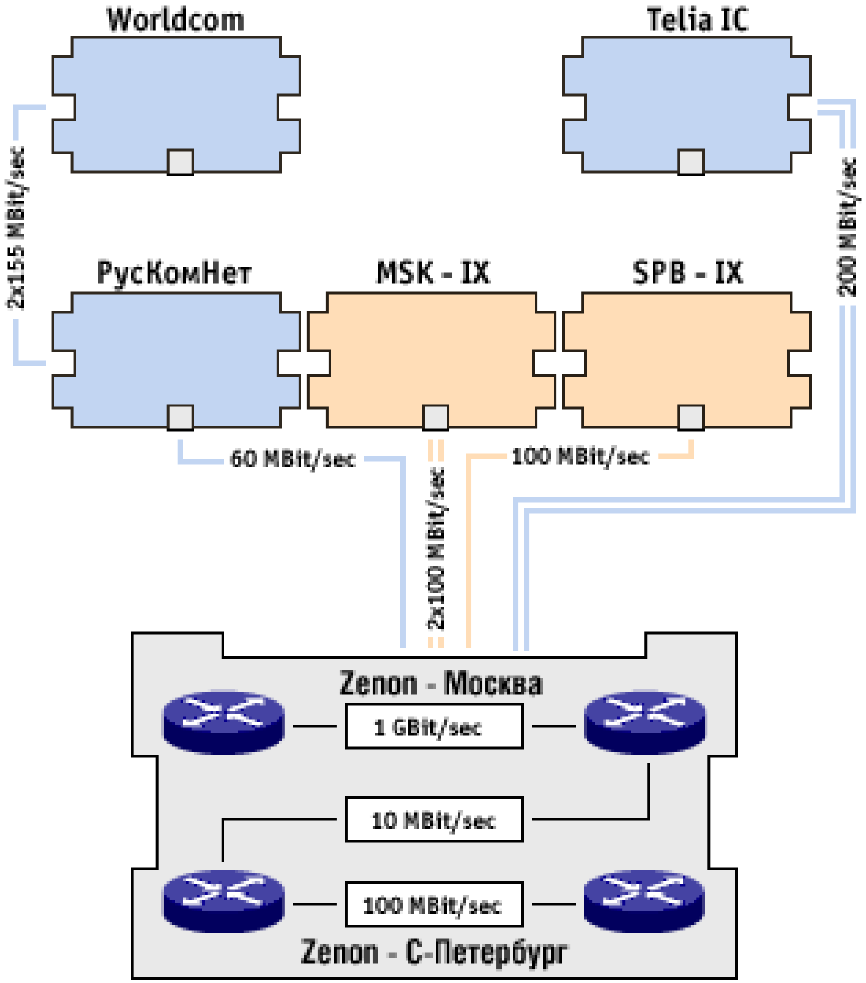
- a packet passes through many networks!



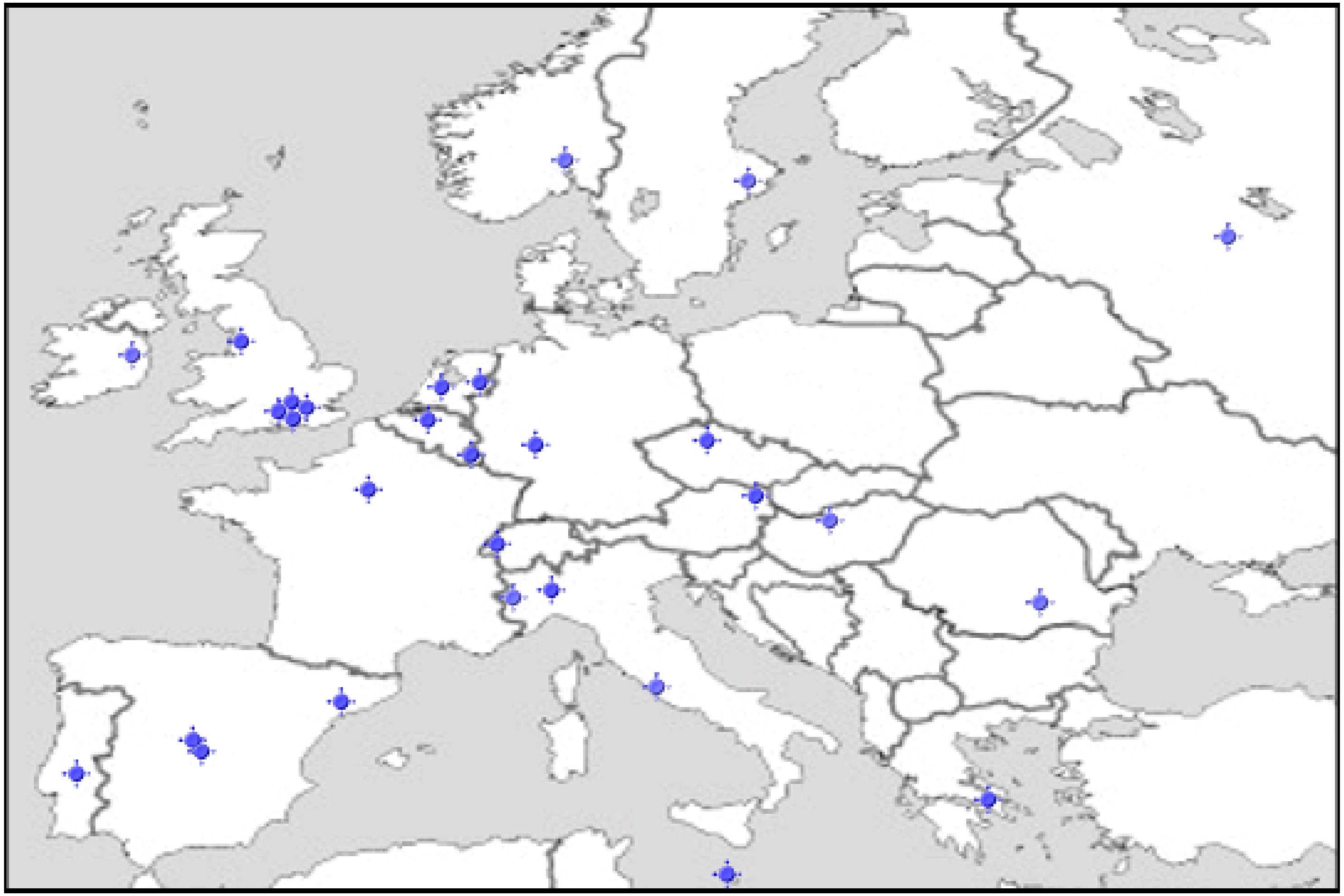
MSK-IX


Sonera
РНЦ "Курчатовский Институт"
ММТС 10
ЦСС МПС
Cable & Wireless
CIS Svyaz
Центральный
Телеграф
ГП Космическая Связь
M9-IX
1Gbps
ММТС 9
Gigabit
Ethernet
Институт Космических
Исследований РАН

Typical usage of Internet eXchange (IX)



Euro-IX



http://www.euro-ix.net/isp/choosing/map/

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

Protocol layers, service models

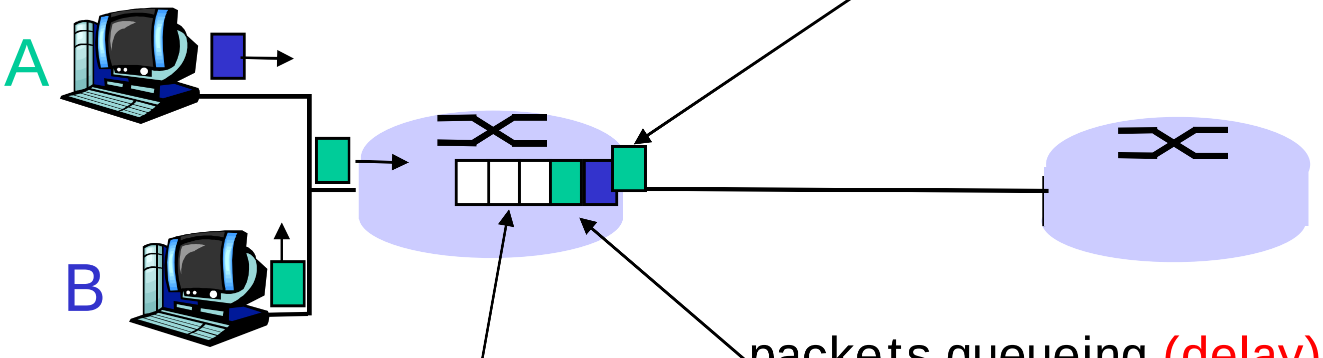
History

IP addressing and routing

How do loss and delay occur?

packets *queue* in router buffers

- packet arrival rate to link exceeds output link capacity
- packets queue, wait for turn



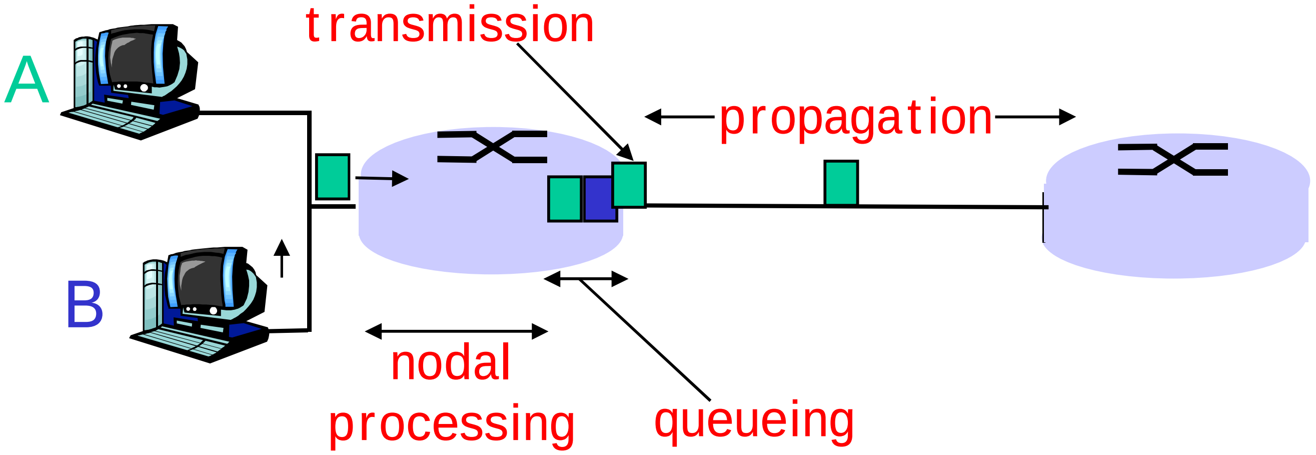
packet being transmitted (delay)

packets queueing (delay)

free (available) buffers: arriving packets dropped (loss) if no free buffers

Four sources of packet delay

- 1. nodal processing:
  - check bit errors
  - determine output link
- 2. queueing
  - time waiting at output link for transmission
  - depends on congestion level of router



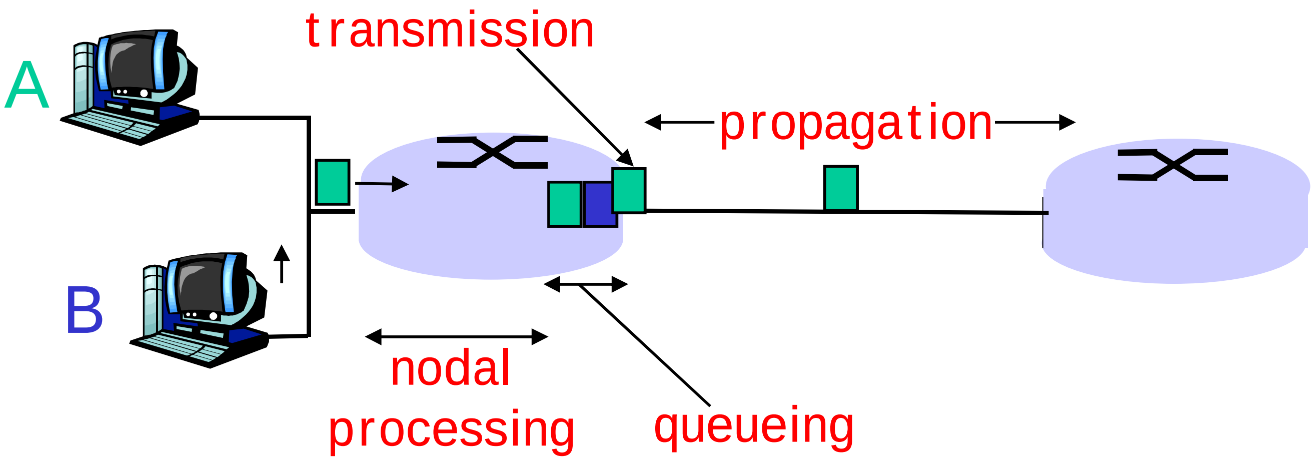
Delay in packet-switched networks

3. Transmission delay:

- q R=link bandwidth (bps)
- q L=packet length (bits)
- q time to send bits into link = L/R

4. Propagation delay:

- q d = length of physical link
- q s = propagation speed in medium (~2x10^8 m/sec)
- q propagation delay = d/s



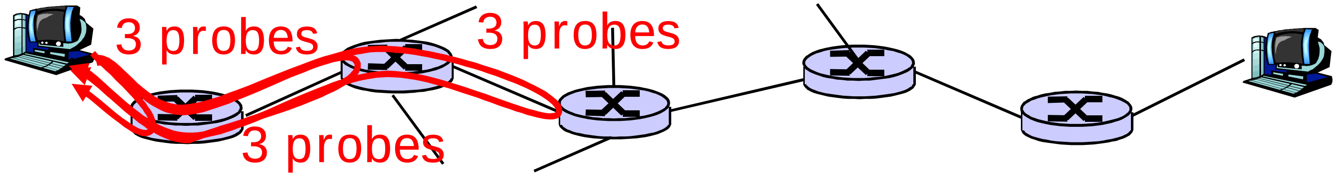
Nodal delay

$$\boxed{d_{\text{nodal}} = d_{\text{proc}} + d_{\text{queue}} + d_{\text{trans}} + d_{\text{prop}}}$$

- d_{proc} = processing delay
  - typically a few microsecs or less
- d_{queue} = queuing delay
  - depends on congestion
- d_{trans} = transmission delay
  - = L/R, significant for low-speed links
- d_{prop} = propagation delay
  - a few microsecs to hundreds of msecs

"Real" Internet delays and routes

- What do "real" Internet delay & loss look like?
- **Traceroute** program: provides delay measurement from source to router along end-end Internet path towards destination. For all *i:*
  - sends three packets that will reach router *i* on path towards destination
  - router *i* will return packets to sender
  - sender times interval between transmission and reply.



“Real” Internet delays and routes

traceroute: gaia.cs.umass.edu to www.eurecom.fr

Three delay measements from gaia.cs.umass.edu to cs-gw.cs.umass.edu

```
1  cs-gw (128.119.240.254)  1 ms  1 ms  2 ms
2  border1-rt-fa5-1-0.gw.umass.edu (128.119.3.145)  1 ms  1 ms  2 ms
3  cht-vbns.gw.umass.edu (128.119.3.130)  6 ms 5 ms 5 ms
4  jn1-at1-0-0-19.wor.vbns.net (204.147.132.129)  16 ms 11 ms 13 ms
5  jn1-so7-0-0-0.wae.vbns.net (204.147.136.136)  21 ms 18 ms 18 ms
6  abilene-vbns.abilene.ucaid.edu (198.32.11.9)  22 ms  18 ms  22 ms
7  nycm-wash.abilene.ucaid.edu (198.32.8.46)  22 ms  22 ms  22 ms
8  62.40.103.253 (62.40.103.253)  104 ms 109 ms 106 ms
9  de2-1.de1.de.geant.net (62.40.96.129)  109 ms 102 ms 104 ms
10  de.fr1.fr.geant.net (62.40.96.50)  113 ms 121 ms 114 ms
11  renater-gw.fr1.fr.geant.net (62.40.103.54)  112 ms  114 ms  112 ms
12  nio-n2.cssi.renater.fr (193.51.206.13)  111 ms  114 ms  116 ms
13  nice.cssi.renater.fr (195.220.98.102)  123 ms  125 ms  124 ms
14  r3t2-nice.cssi.renater.fr (195.220.98.110)  126 ms  126 ms  124 ms
15  eurecom-valbonne.r3t2.ft.net (193.48.50.54)  135 ms  128 ms  133 ms
16  194.214.211.25 (194.214.211.25)  126 ms  128 ms  126 ms
17  * * *
18  * * *
19  fantasia.eurecom.fr (193.55.113.142)  132 ms  128 ms  136 ms
```

trans-oceanic link

* means no reponse (probe lost, router not replying)

Packet loss

- queue (aka buffer) preceding link in buffer has finite capacity
- when packet arrives to full queue, packet is dropped (aka lost)
- lost packet may be retransmitted by previous node, by source end system, or not retransmitted at all

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

Protocol layers, service models

History

IP addressing and routing

Protocol "Layers"

Networks are complex!

- many "pieces":
  - hosts
  - routers
  - links of various media
  - applications
  - protocols
  - hardware, software

Question:

Is there any hope of *organizing* structure of network?

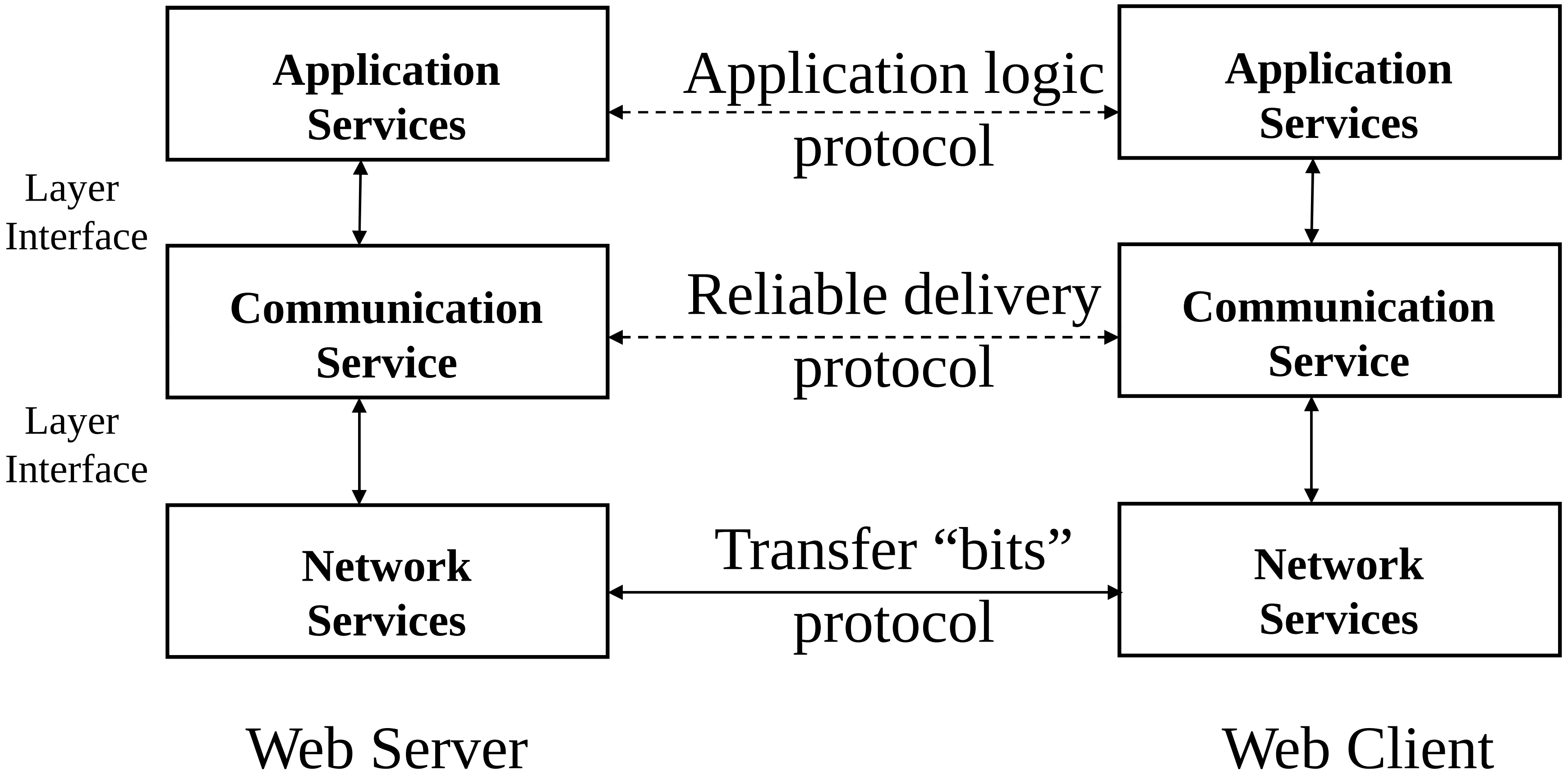
Or at least our discussion of networks?

Why layering?

Dealing with complex systems:

- explicit structure allows identification, relationship of complex system's pieces
  - layered reference model for discussion
- modularization eases maintenance, updating of system
  - change of implementation of layer's service transparent to rest of system
  - e.g., change in gate procedure doesn't affect rest of system
- layering considered harmful?

Layers, Protocols, Interfaces



Layered Architecture (Review 1/2)

- Networks organized as a stack of layers?
  - The purpose of a layer is to offer services to the layer above it using an interface (programming language analogy: libraries hide details while providing a service)
  - Reduces design complexity
- Protocols: peer-to-peer layer-n conversations
- Data Transfer: each layer passes data & control information to the layer below; eventually physical medium is reached.

Layered Architecture Review (2/2)

- A set of layers & protocols sometimes is called a Network Architecture. These specifications enable hardware/software developers to build systems compliant with a particular architecture.
  - E.g., TCP/IP, OSI

Layering: Design Issues

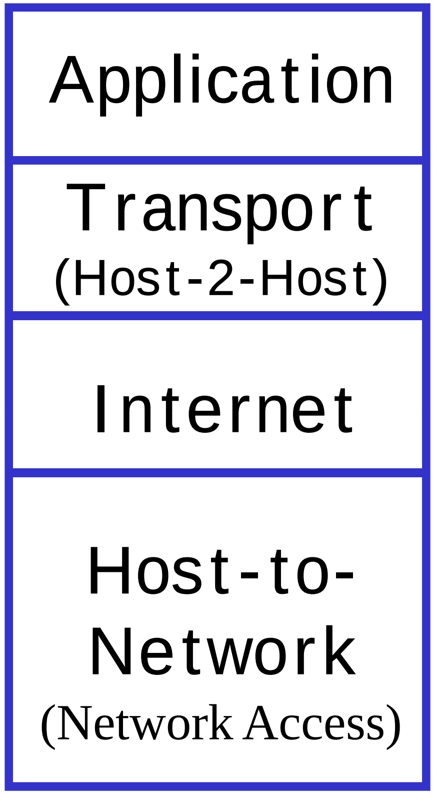
- Identify senders/receivers?
  - Addressing
- Unreliable physical communication medium?
  - Error detection
  - Error control
  - Message reordering
- Sender can swamp the receiver?
  - Flow control
- Multiplexing/Demultiplexing

Reference Models

- TCP/IP Model
- Open Systems Interconnection (OSI) Model

TCP/IP Model: History

- Originally used in the ARPANET
- ARPANET required networks using leased telephone lines & radio/satellite networks to interoperate
- Goals of the model are:
  - Seamless interoperability
  - Wide-ranging applications
  - Fault-tolerant to some extent



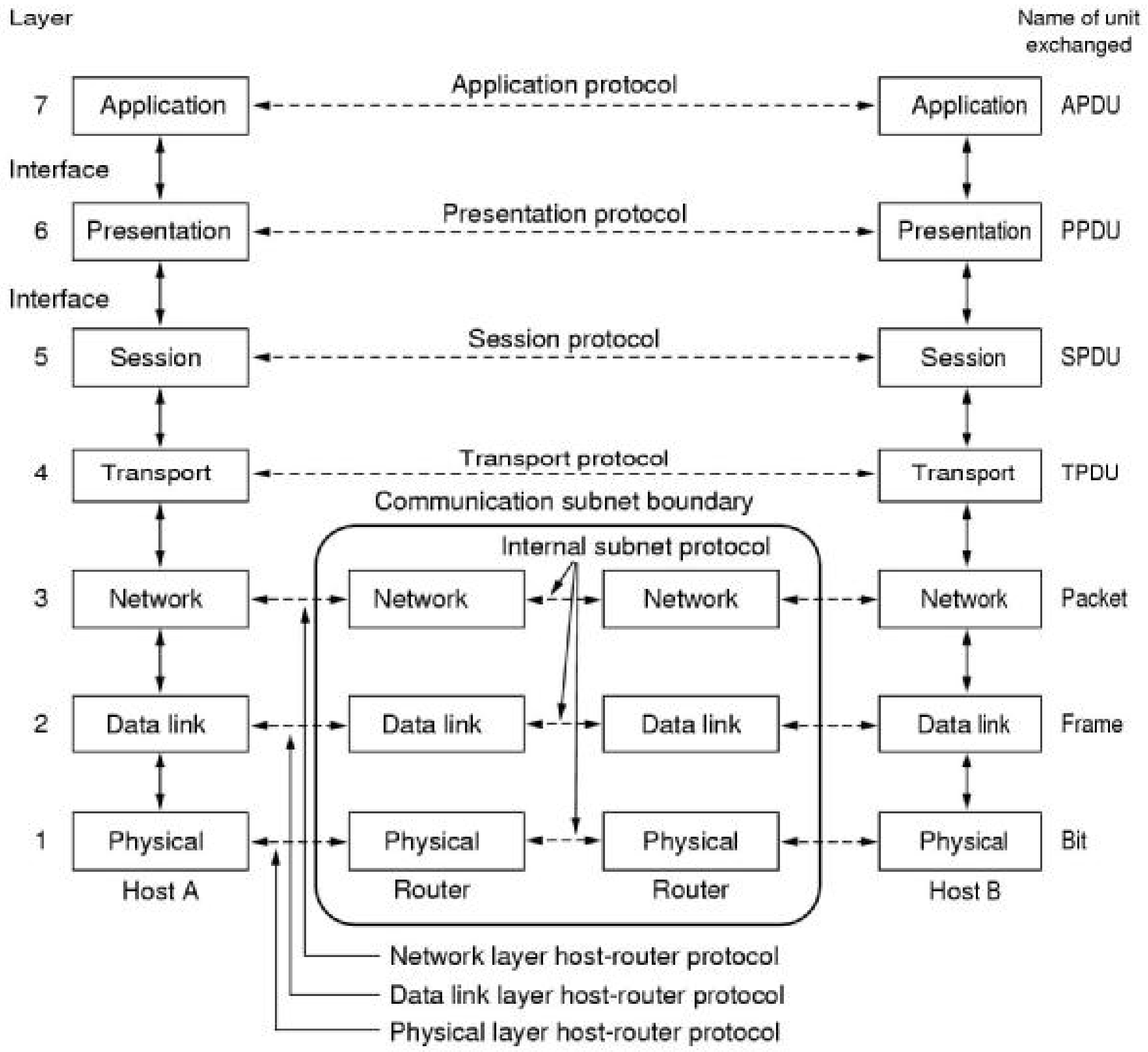
Application

Transport
(Host-2-Host)

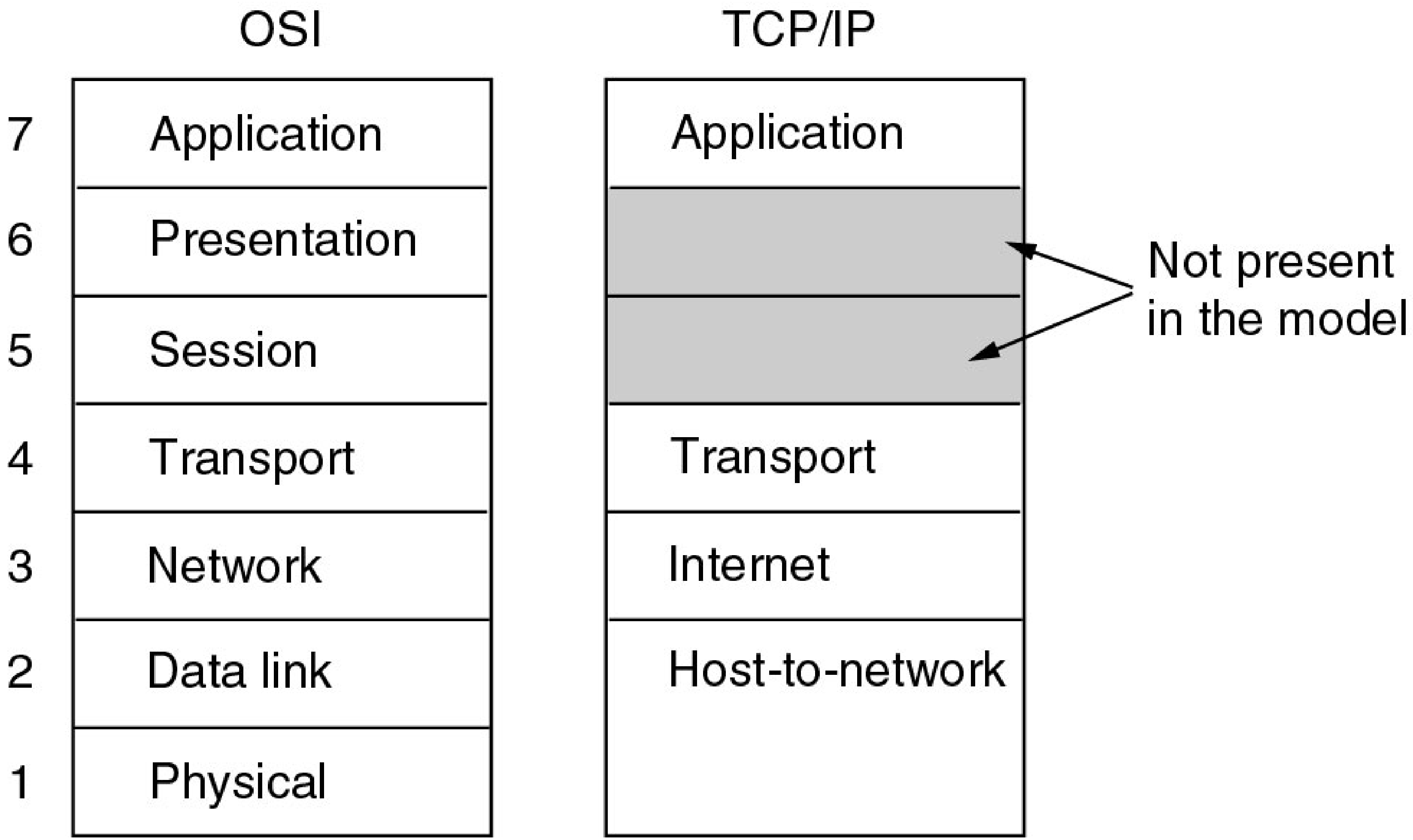
Internet

Host-to-Network
(Network Access)

OSI / ISO (1984)


Layer
Name of unit exchanged
7
Application
Application protocol
Application
APDU
Interface
6
Presentation
Presentation protocol
Presentation
PPDU
Interface
5
Session
Session protocol
Session
SPDU
4
Transport
Transport protocol
Transport
TPDU
Communication subnet boundary
Internal subnet protocol
3
Network
Network
Network
Network
Packet
2
Data link
Data link
Data link
Data link
Frame
1
Physical
Physical
Physical
Physical
Bit
Host A
Router
Router
Host B
Network layer host-router protocol
Data link layer host-router protocol
Physical layer host-router protocol

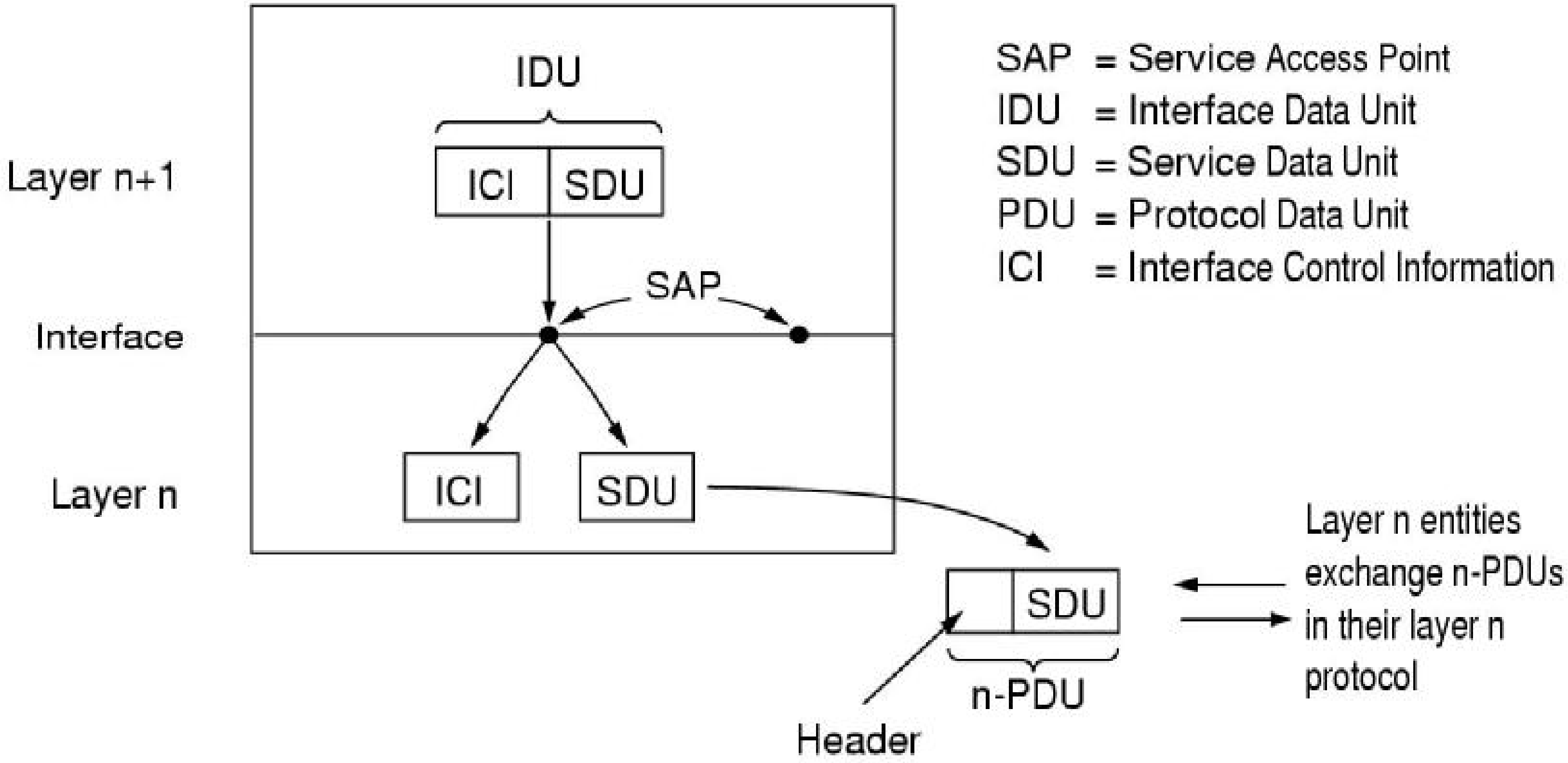
OSI vs TCP/IP



	OSI	TCP/IP
7	Application	Application
6	Presentation	
5	Session	
4	Transport	Transport
3	Network	Internet
2	Data link	Host-to-network
1	Physical	

Not present in the model

Layers interactions

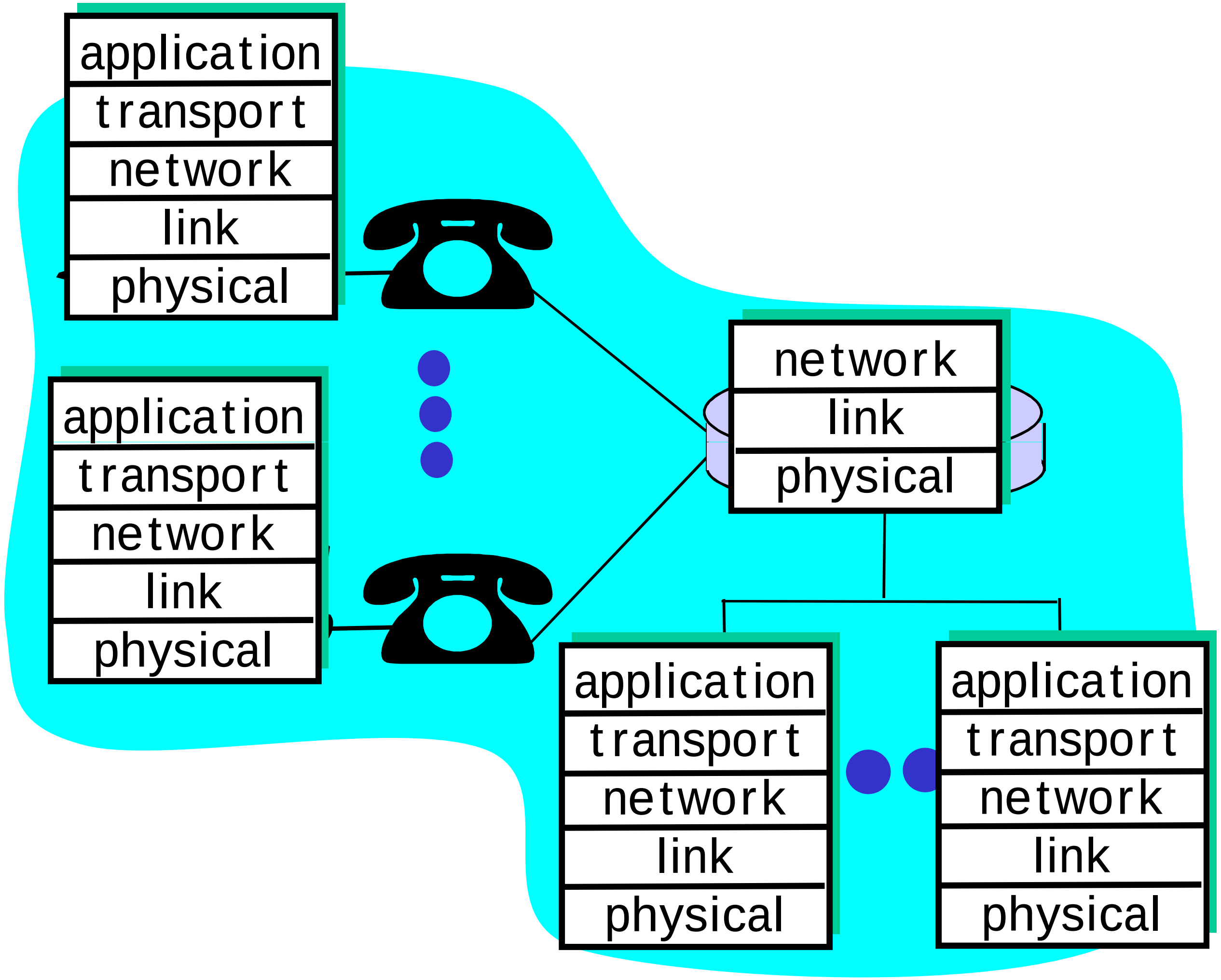


SAP = Service Access Point
IDU = Interface Data Unit
SDU = Service Data Unit
PDU = Protocol Data Unit
ICI = Interface Control Information

Layering: logical communication

Each layer:

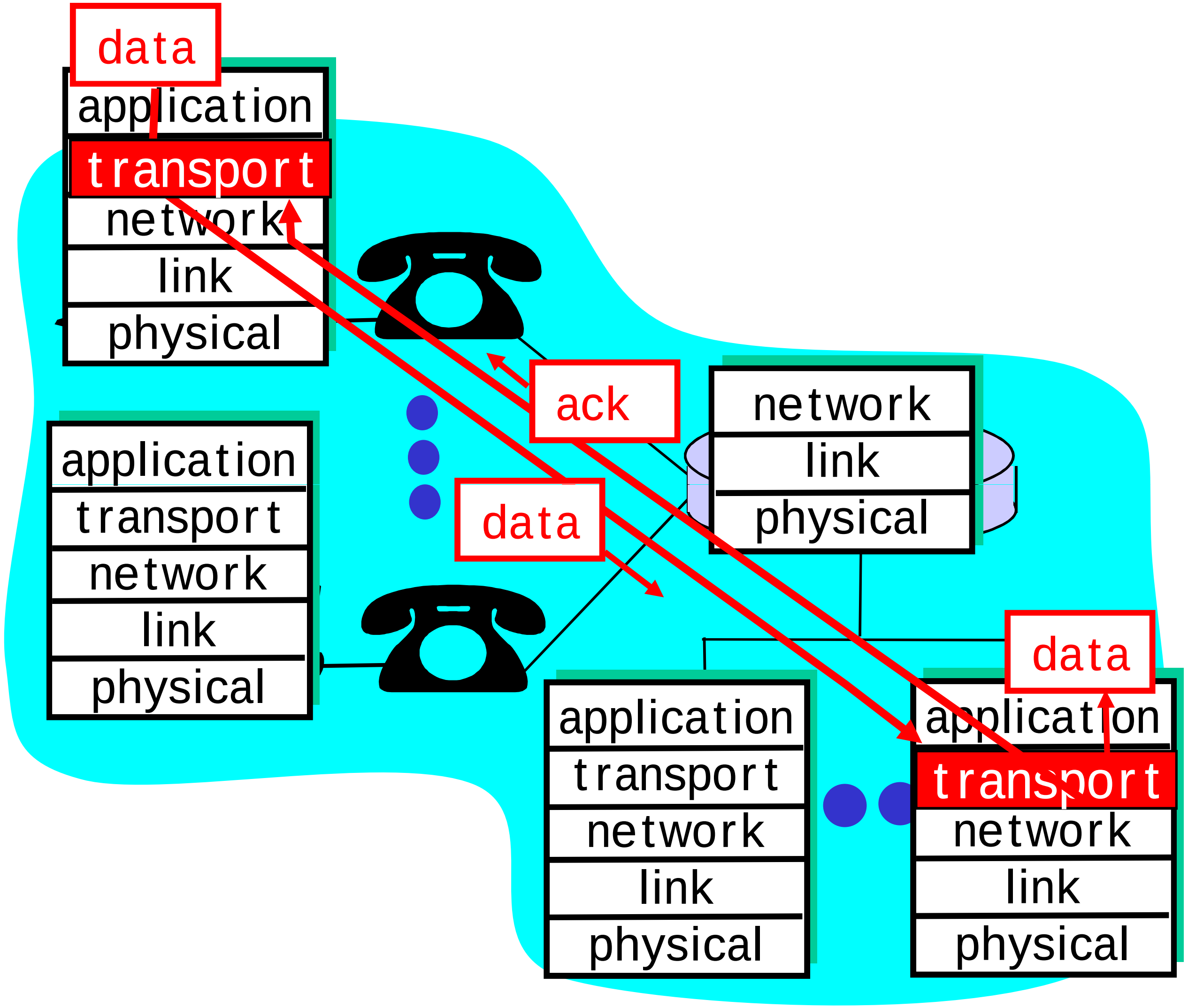
- distributed
- “entities” implement layer functions at each node
- entities perform actions, exchange messages with peers



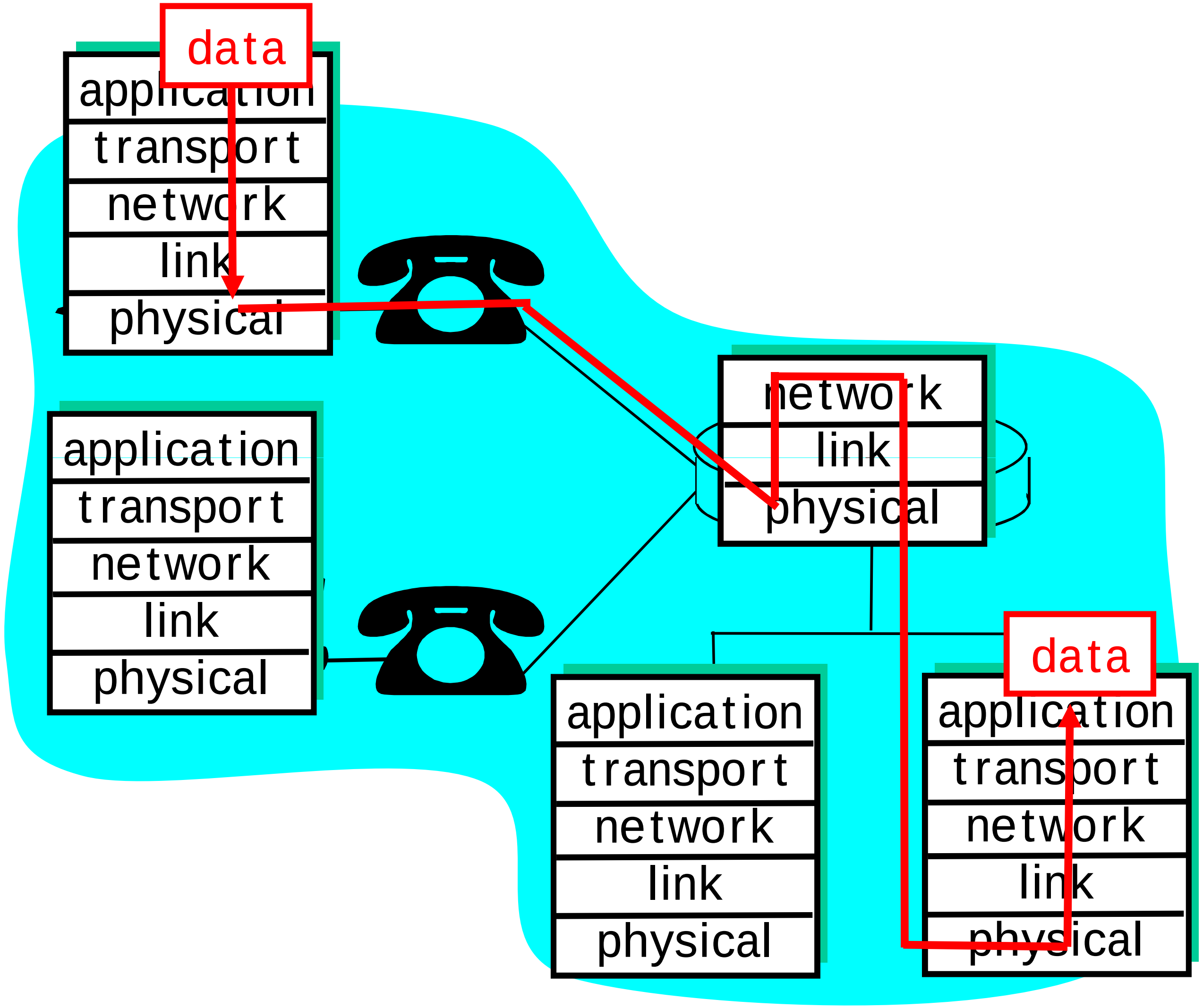
Layering: *logical* communication

E.g.: transport

- take data from app
- add addressing, reliability check info to form "datagram"
- send datagram to peer
- wait for peer to ack receipt
- analogy: post office



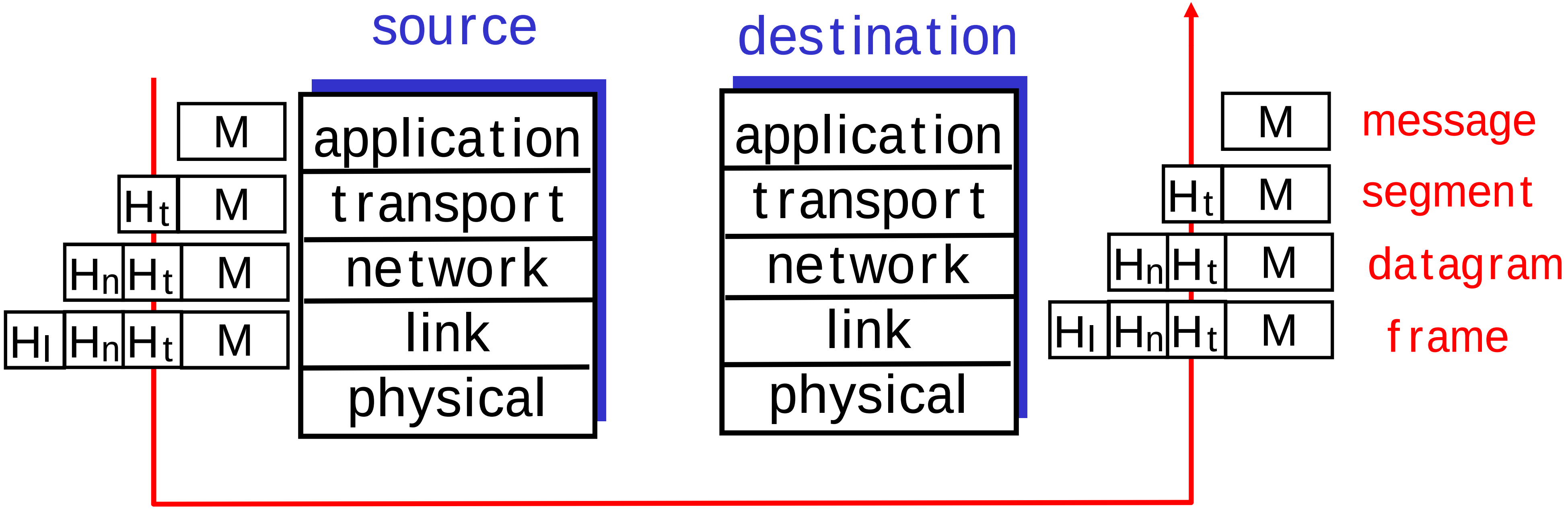
Layering: physical communication



Protocol layering and data

Each layer takes data from above

- adds header information to create new data unit
- passes new data unit to layer below



Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

Protocol layers, service models

History

IP addressing and routing

Internet History

1961-1972: Early packet-switching principles

- 1961: Kleinrock - queueing theory shows effectiveness of packet-switching
- 1964: Baran - packet-switching in military nets
- 1967: ARPAnet conceived by Advanced Research Projects Agency
- 1969: first ARPAnet node operational

- 1972:
  - ARPAnet demonstrated publicly
  - NCP (Network Control Protocol) first host-host protocol
  - first e-mail program
  - ARPAnet has 15 nodes

Internet History

1972-1980: Internetworking, new and proprietary nets

- 1970: ALOHAnet satellite network in Hawaii
- 1973: Metcalfe's PhD thesis proposes Ethernet
- 1974: Cerf and Kahn - architecture for interconnecting networks
- late70's: proprietary architectures: DECnet, SNA, XNA
- late 70's: switching fixed length packets (ATM precursor)
- 1979: ARPAnet has 200 nodes

Cerf and Kahn's internetworking principles:

- minimalism, autonomy - no internal changes required to interconnect networks
- best effort service model
- stateless routers
- decentralized control

define today's Internet architecture

Internet History

1980-1990: new protocols, a proliferation of networks

- 1983: deployment of TCP/IP
- 1982: SMTP e-mail protocol defined
- 1983: DNS defined for name-to-IP-address translation
- 1985: FTP protocol defined
- 1988: TCP congestion control

- new national networks: Csnet, BITnet, NSFnet, Minitel
- 100,000 hosts connected to confederation of networks

Internet History

1990, 2000's: commercialization, the Web, new apps

- Early 1990's: ARPAnet decommissioned
- 1991: NSF lifts restrictions on commercial use of NSFnet (decommissioned, 1995)
- early 1990s: Web
  - hypertext [Bush 1945, Nelson 1960's]
  - HTML, HTTP: Berners-Lee
  - 1994: Mosaic, later Netscape
  - late 1990's: commercialization of the Web

Late 1990's – 2000's:

- more killer apps: instant messaging, peer2peer file sharing (e.g., Naptser)
- network security to forefront
- est. 50 million host, 100 million+ users
- backbone links running at Gbps
- Internet-2

Roadmap

What *is* a Computer Network?

Network Applications

Network Taxonomies

Some Definitions

Network Structure: edge, core, access and media

Internet structure and ISPs

Delay & loss in packet-switched networks

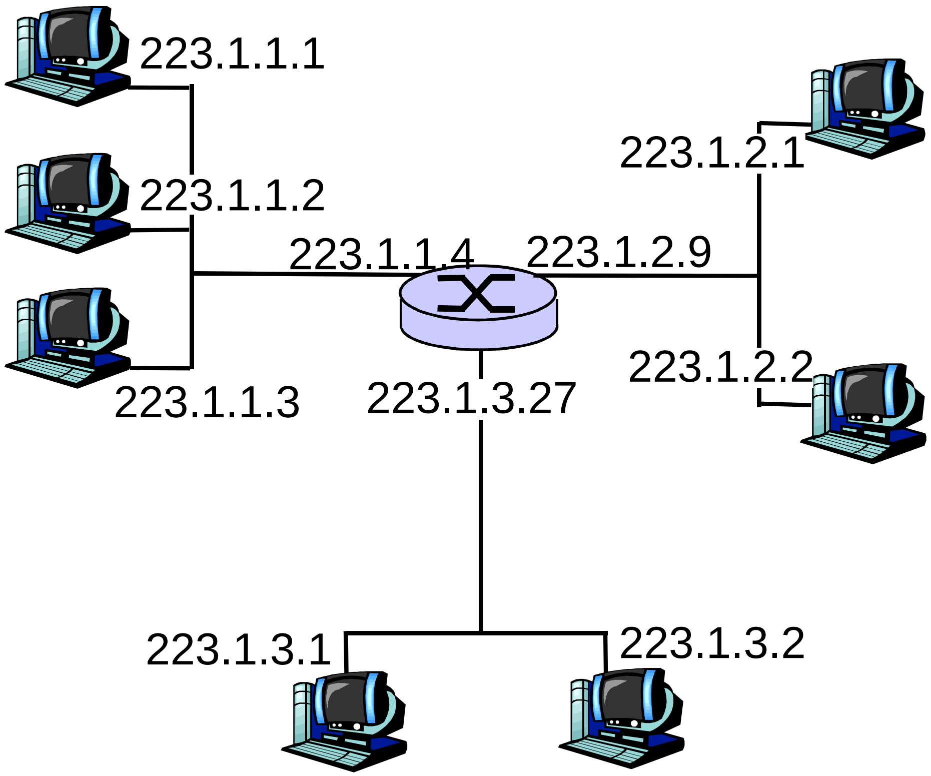
Protocol layers, service models

History

IP addressing and routing

IP Addressing: introduction

- 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 addresses associated with each interface

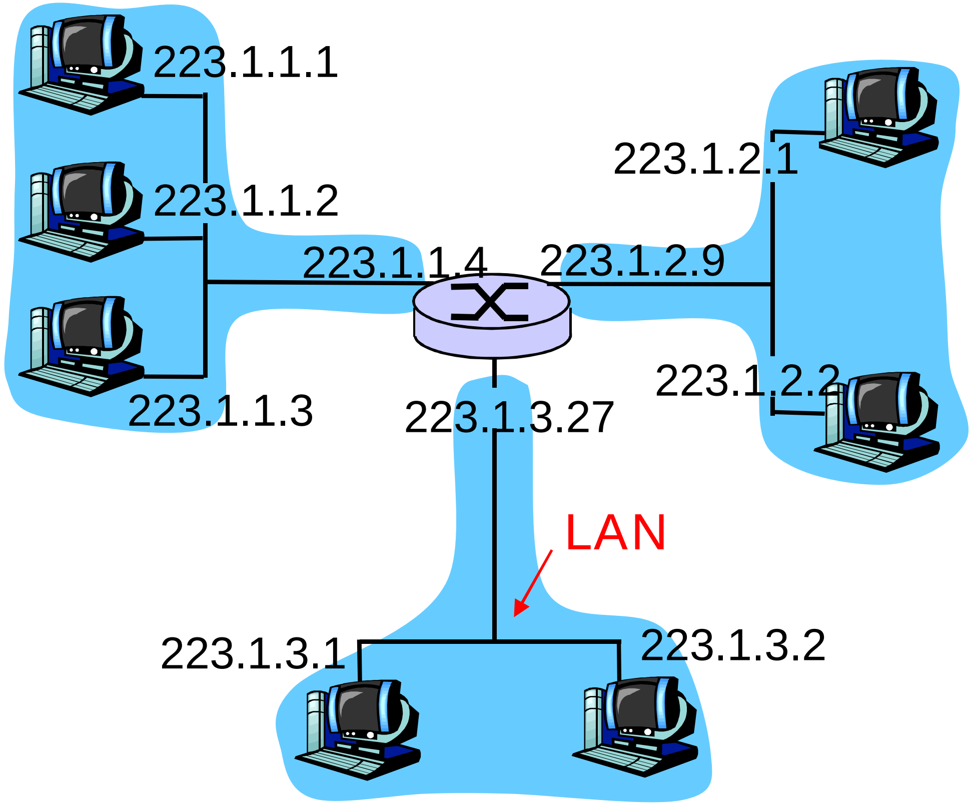


$$223.1.1.1 = \underbrace{11011111}_{223}\ \underbrace{00000001}_{1}\ \underbrace{00000001}_{1}\ \underbrace{00000001}_{1}$$

IP Addressing

- 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

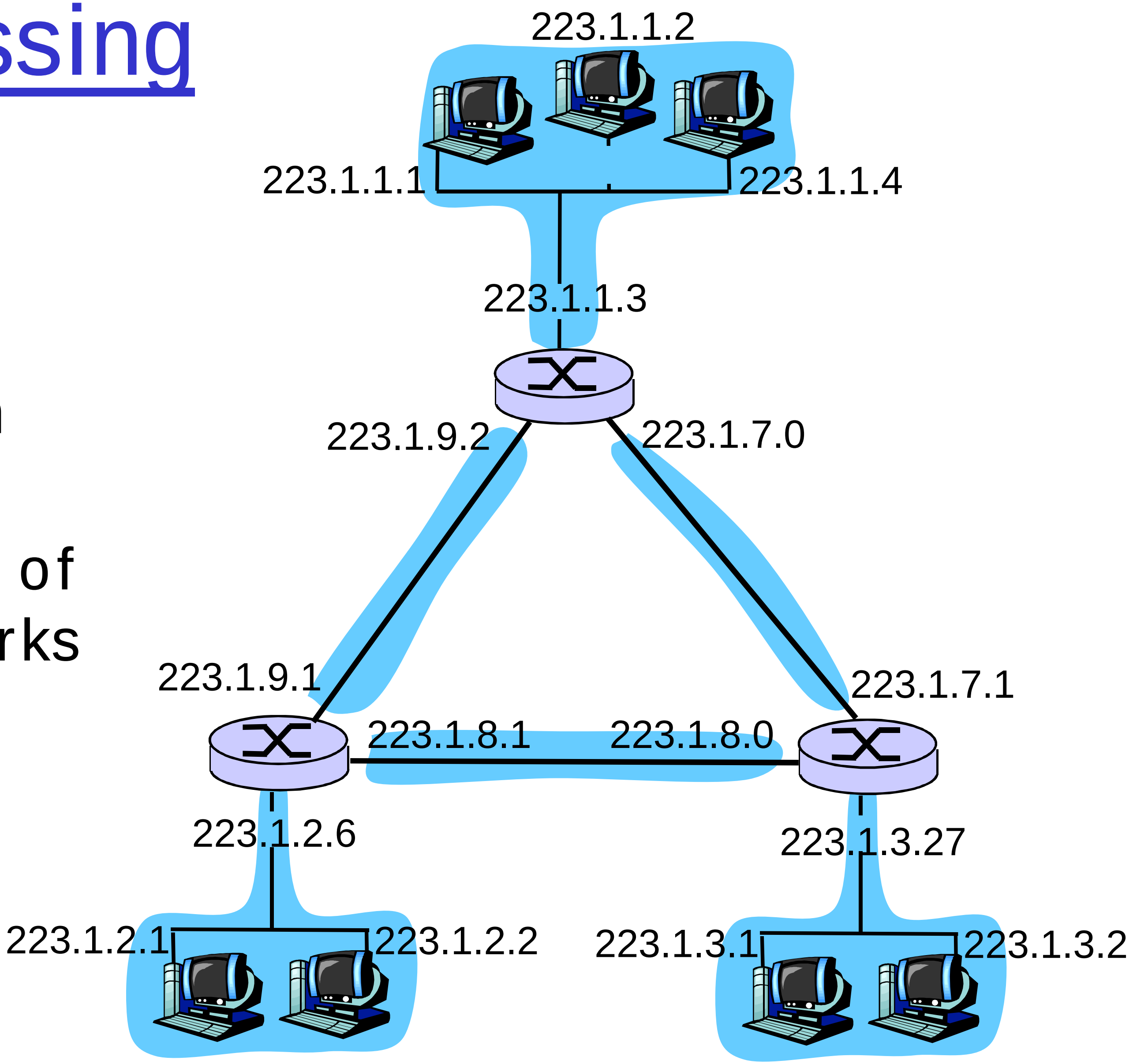


network consisting of 3 IP networks (for IP addresses starting with 223, first 24 bits are network address)

IP Addressing

How to find the networks?

- Detach each interface from router, host
- create "islands of isolated networks

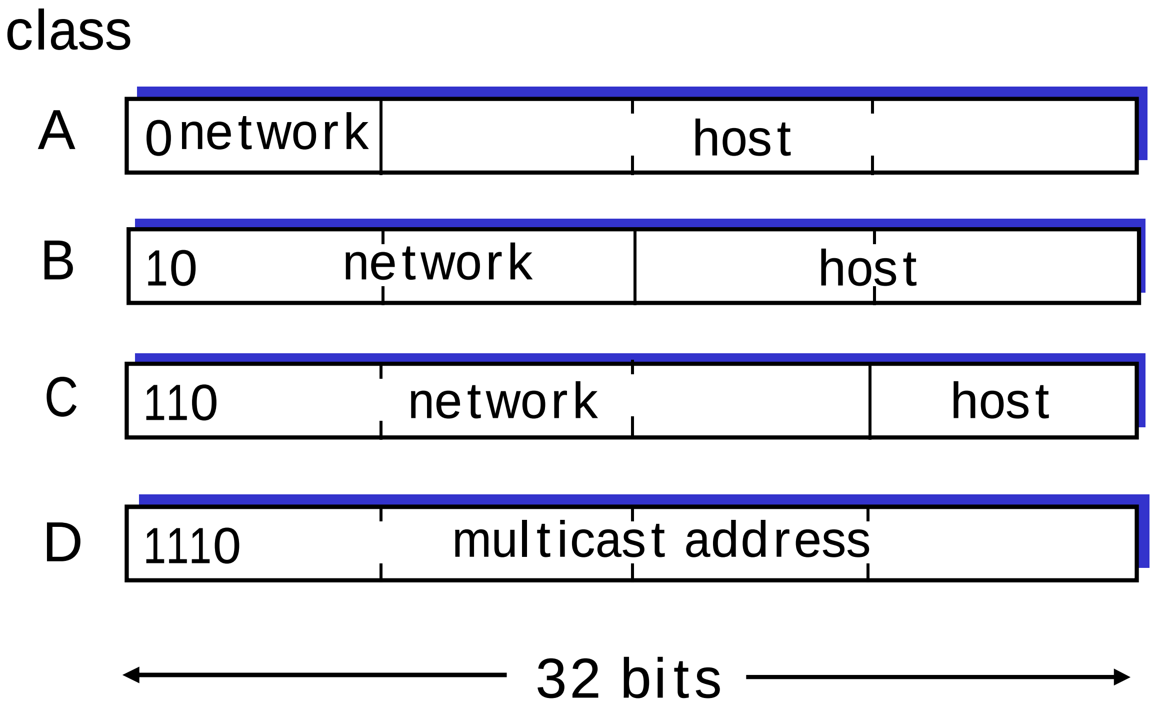


Interconnected system consisting of six networks

IP Addresses

given notion of “network”, let’s re-examine IP addresses:

“class-full” addressing:



class

A 0 network | host — 1.0.0.0 to 127.255.255.255

B 10 network | host — 128.0.0.0 to 191.255.255.255

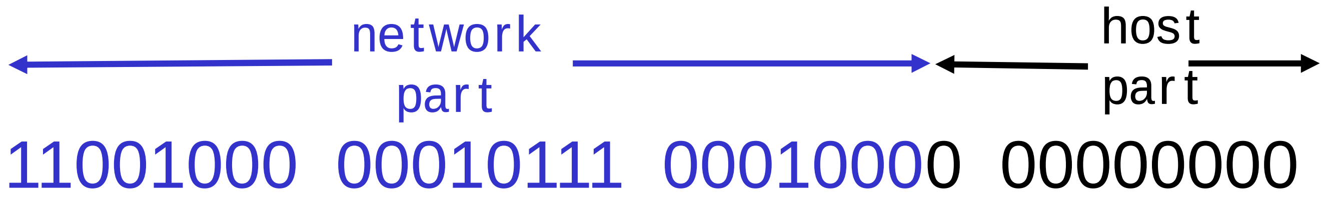
C 110 network | host — 192.0.0.0 to 223.255.255.255

D 1110 multicast address — 224.0.0.0 to 239.255.255.255

32 bits

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

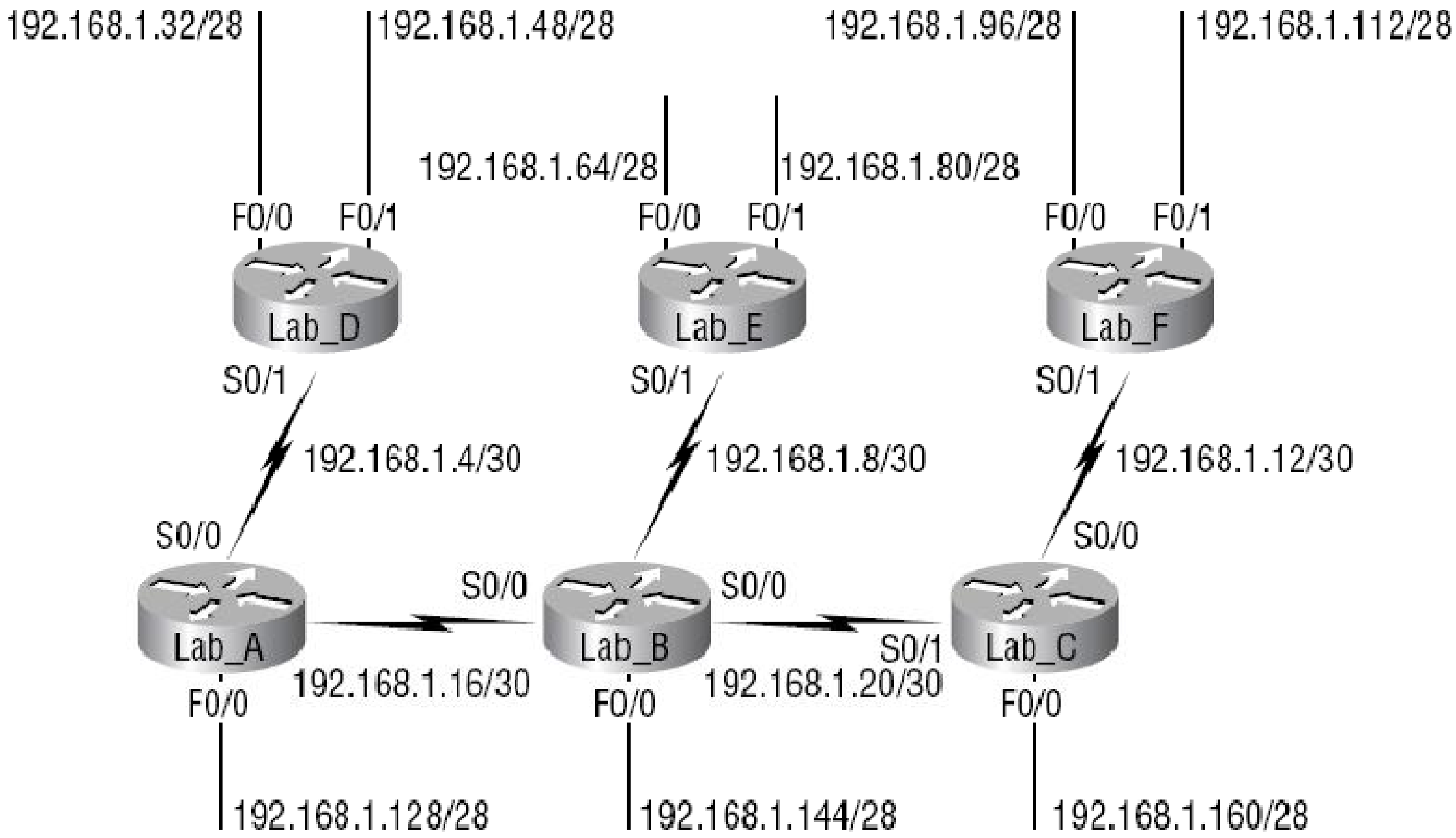


11001000 00010111 00010000 00000000

200.23.16.0/23

Variable Length Subnet Masks (VLSMs)

- Classless routing protocols should support advertisement of subnet information
- Can be used with RIPv2, EIGRP, or OSPF



IP addresses: how to get one?

Q: How does *host* get IP address?

- hard-coded by system admin in a file
  - Wintel: control-panel->network->configuration->tcp/ip->properties
  - UNIX: /etc/rc.config
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server
  - “plug-and-play”

IP addresses: how to get one?

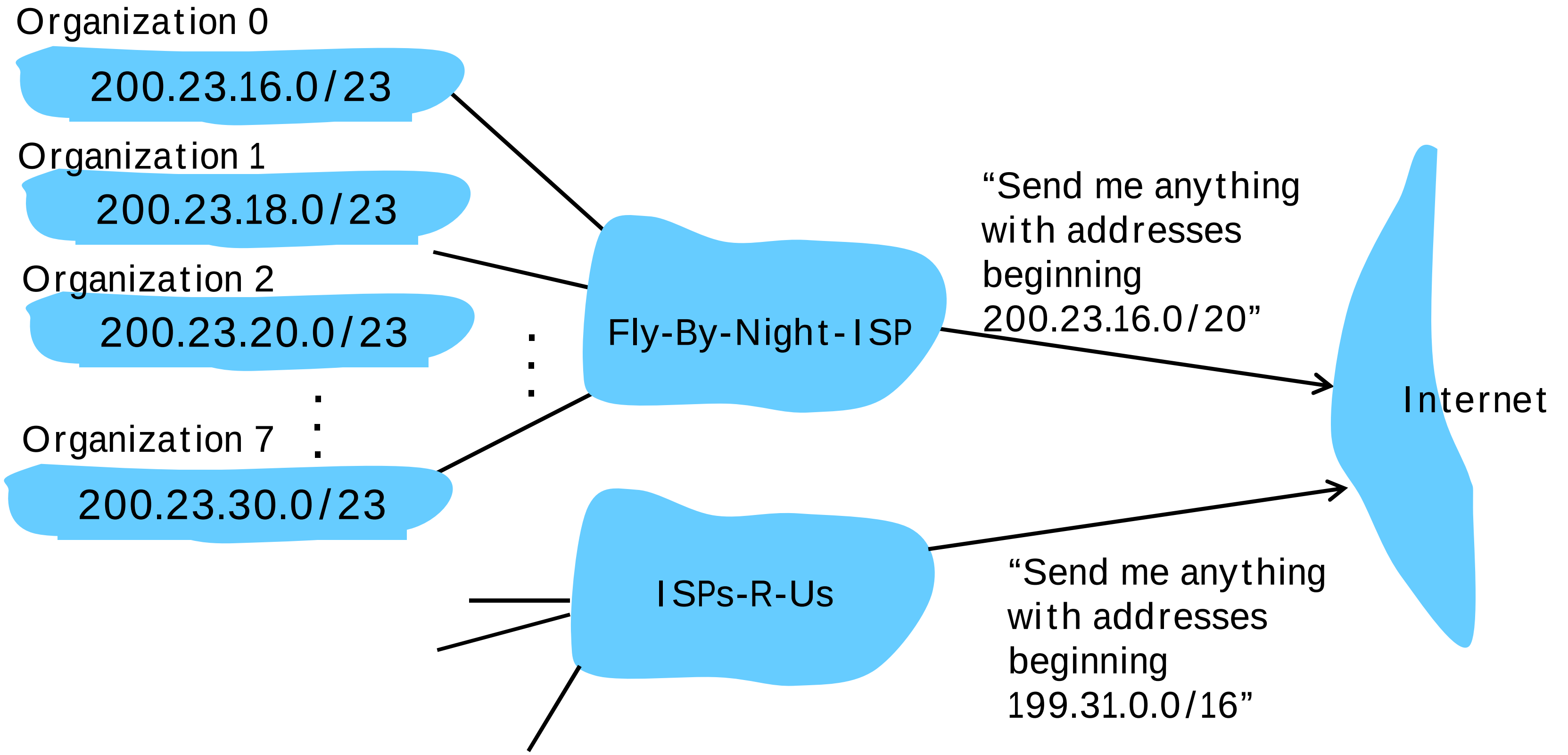
Q: How does *network* get network part of IP addr?

A: gets allocated 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

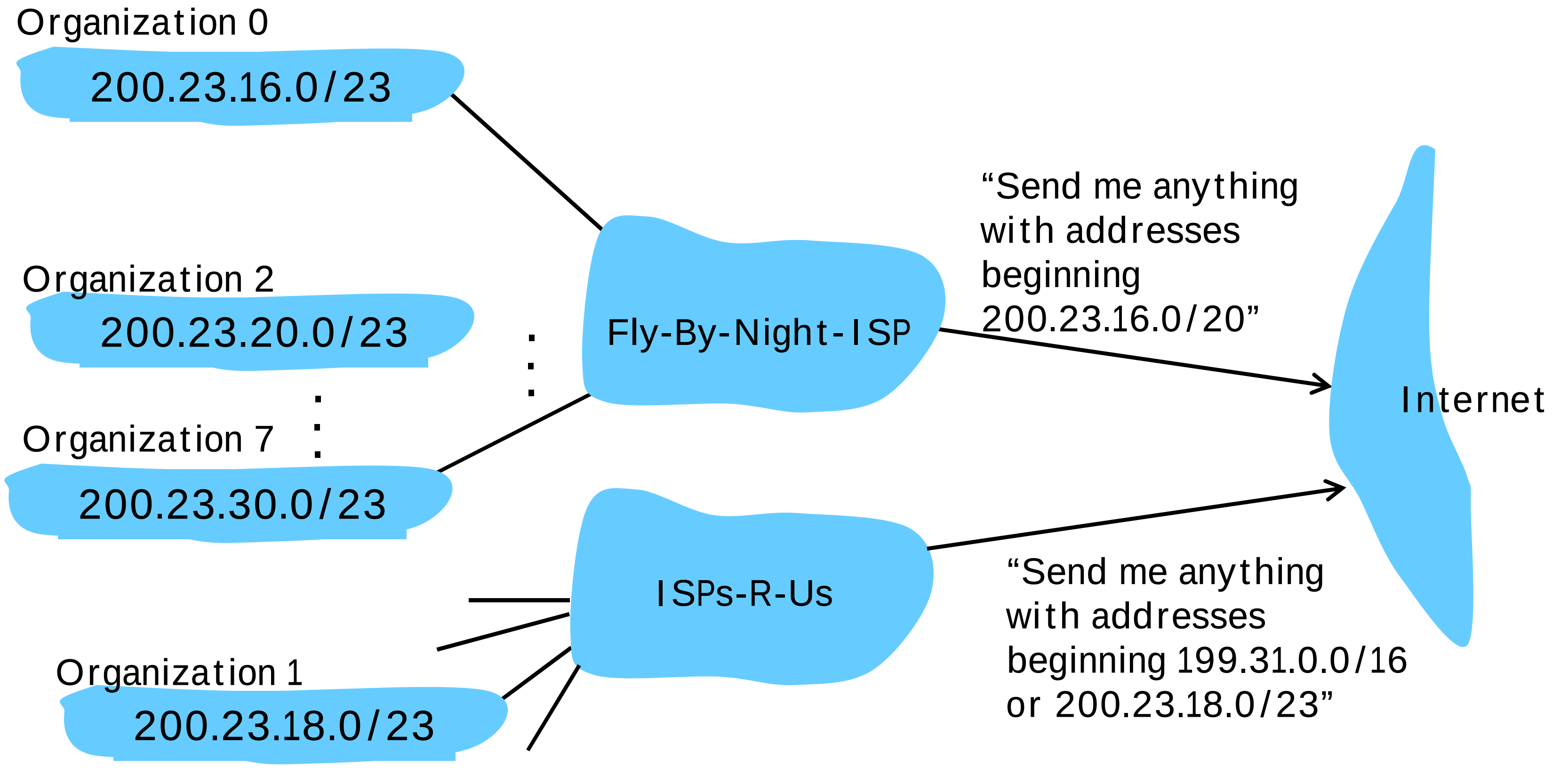
Hierarchical addressing: route aggregation (aka summarization)

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical addressing: more specific routes

ISPs-R-Us has a more specific route to Organization 1



Troubleshooting IP Addressing (CISCO’s way)

- ping 127.0.0.1
- ping NIC's address
- ping the default gateway/ router
- ping the remote server
- ping user ;) if 1-4 OK.

IP addressing: the last word...

Q: How does an ISP get block of addresses?

A: ICANN: Internet Corporation for Assigned Names and Numbers

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

Getting a datagram from source to dest.

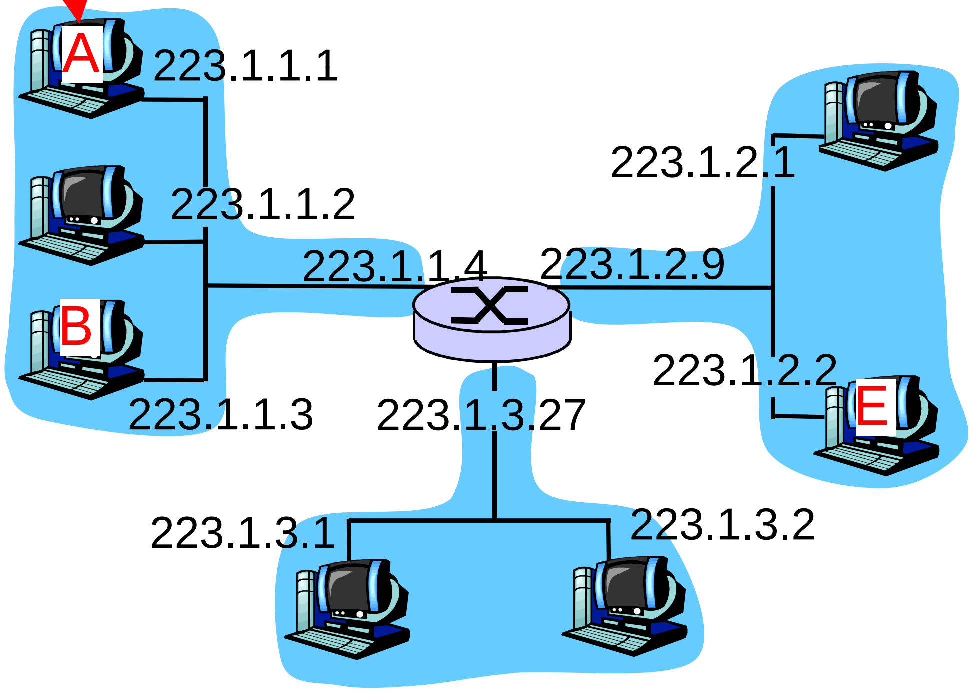
forwarding table in A

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2

IP datagram:

misc fields	source IP addr	dest IP addr	data

- datagram remains unchanged, as it travels source to destination
- addr fields of interest here



Getting a datagram from source to dest.

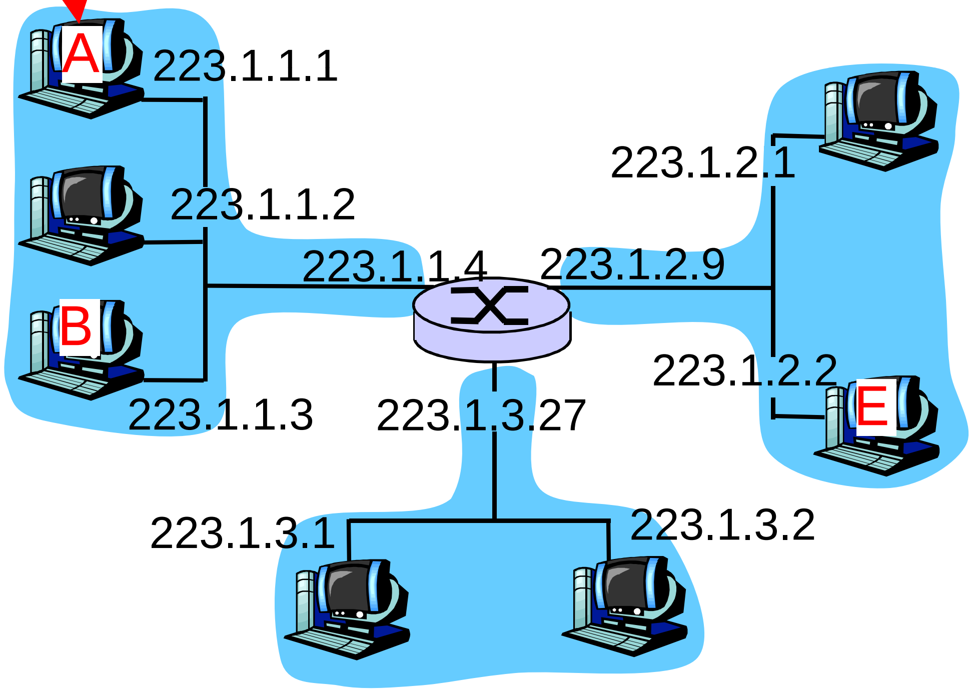
misc fields	223.1.1.1	223.1.1.3	data

Starting at A, send IP datagram addressed to B:

- look up net. address of B in forwarding table
- find B is on same net. as A
- link layer will send datagram directly to B inside link-layer frame
  - B and A are directly connected

forwarding table in A

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



Getting a datagram from source to dest.

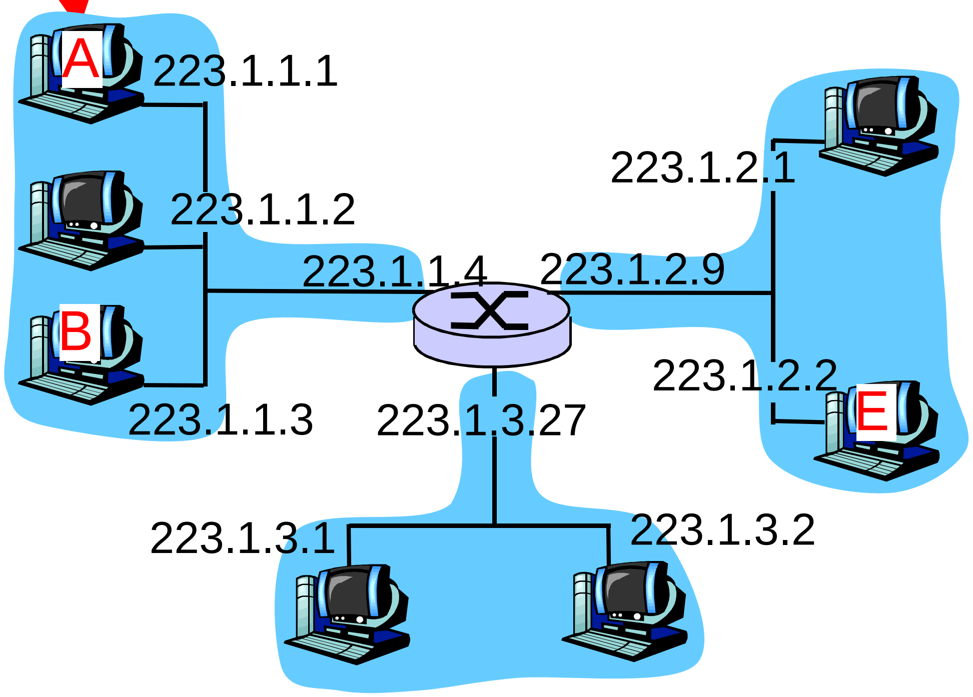
misc fields	223.1.1.1	223.1.2.3	data

forwarding table in A

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2

Starting at A, dest. E:

- look up network address of E in forwarding table
- E on *different* network
  - A, E not directly attached
- routing table: next hop router to E is 223.1.1.4
- link layer sends datagram to router 223.1.1.4 inside link-layer frame
- datagram arrives at 223.1.1.4
- continued…..



Getting a datagram from source to dest.

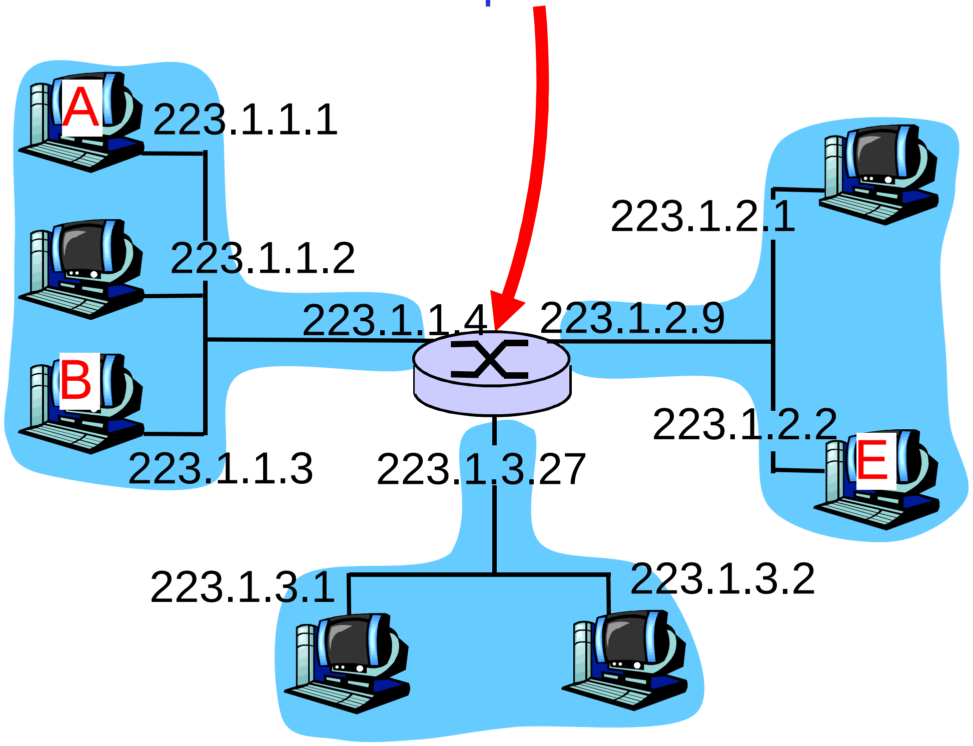
misc fields	223.1.1.1	223.1.2.3	data

forwarding table in router

Dest. Net	router	Nhops	interface
223.1.1	-	1	223.1.1.4
223.1.2	-	1	223.1.2.9
223.1.3	-	1	223.1.3.27

Arriving at 223.1.4, destined for 223.1.2.2

- look up network address of E in router's forwarding table
- E on *same* network as router's interface 223.1.2.9
  - router, E directly attached
- link layer sends datagram to 223.1.2.2 inside link-layer frame via interface 223.1.2.9
- datagram arrives at 223.1.2.2



What is routing, anyway?

Routing protocol objective:

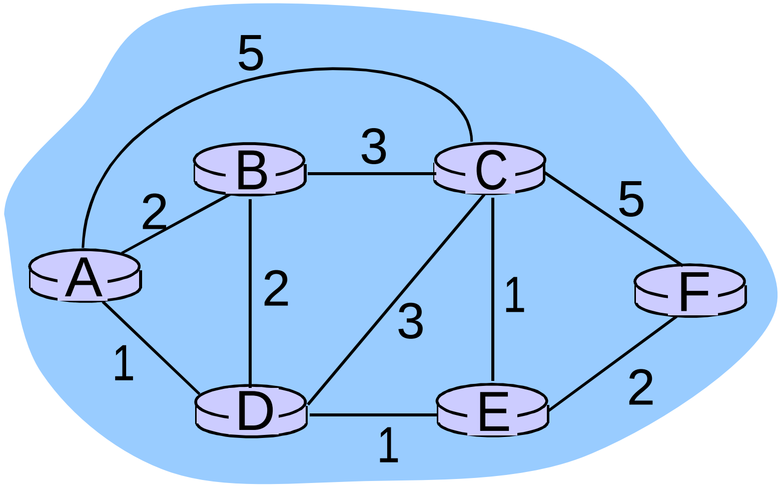
to determine "the best" path (with minimal "cost") between source and destination.

Setting routing methods:

-static

-default

-dynamic (under protocol control)



Routing Algorithm classification

Global or decentralized information?

Global:

- all routers have complete topology, link cost info
- "link state" algorithms

Decentralized:

- router knows physically-connected neighbors, link costs to neighbors
- iterative process of computation, exchange of info with neighbors
- "distance vector" algorithms

Static or dynamic?

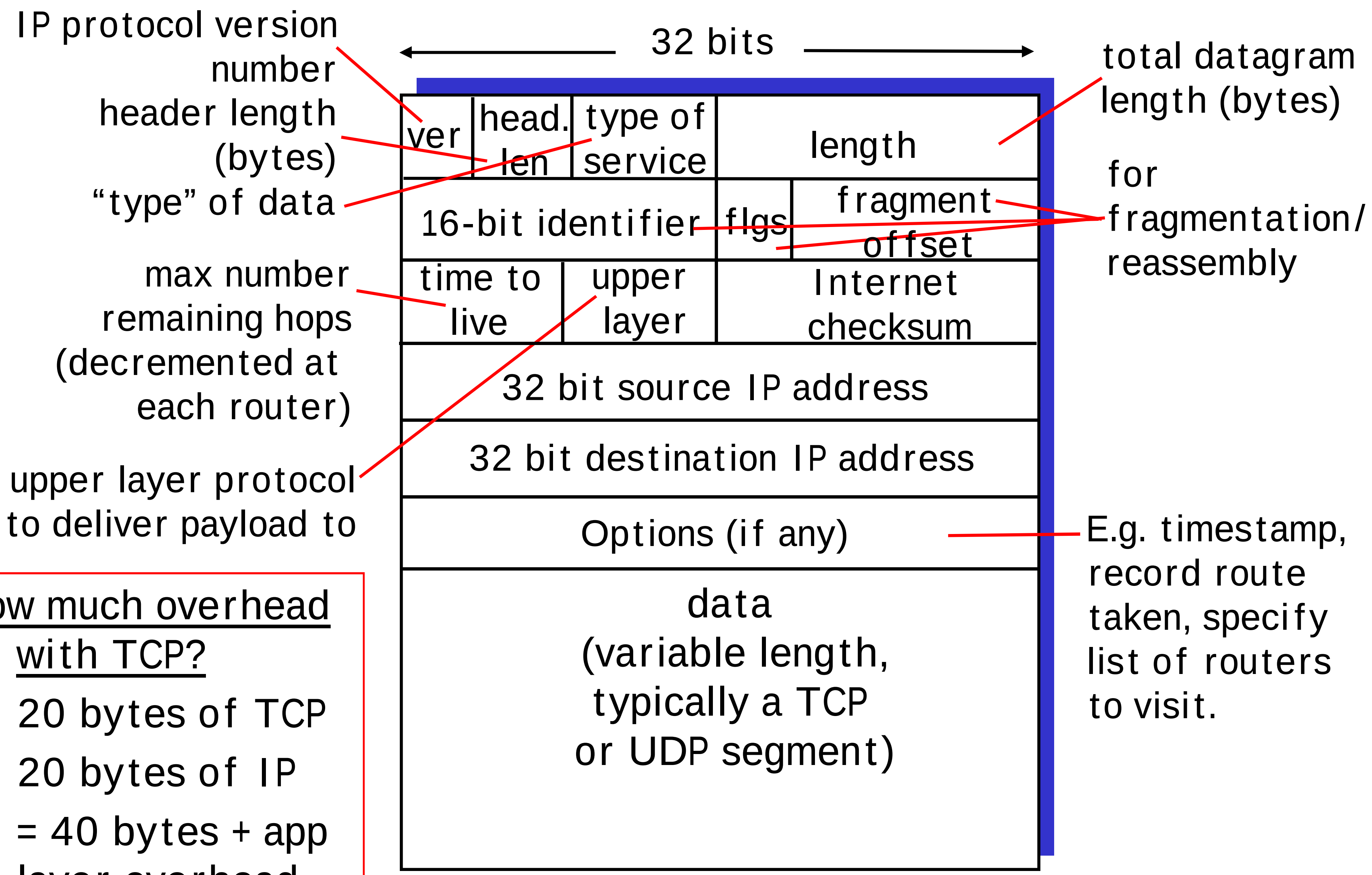
Static:

- routes change slowly over time

Dynamic:

- routes change more quickly
  - periodic update
  - in response to link cost changes

IP datagram format



IP protocol version number

header length (bytes)

“type” of data

max number remaining hops (decremented at each router)

upper layer protocol to deliver payload to

32 bits

ver	head. len	type of service	length	
16-bit identifier			flgs	fragment offset
time to live	upper layer		Internet checksum	
32 bit source IP address				
32 bit destination IP address				
Options (if any)				
data (variable length, typically a TCP or UDP segment)				

total datagram length (bytes)

for fragmentation/ reassembly

E.g. timestamp, record route taken, specify list of routers to visit.

<u>how much overhead with TCP?</u>

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

Internet Routing

scale: with 200 million destinations:

- can't store all dest's in routing tables!
- routing table exchange would swamp links!

administrative autonomy

- internet = network of networks
- each network admin may want to control routing in its own network

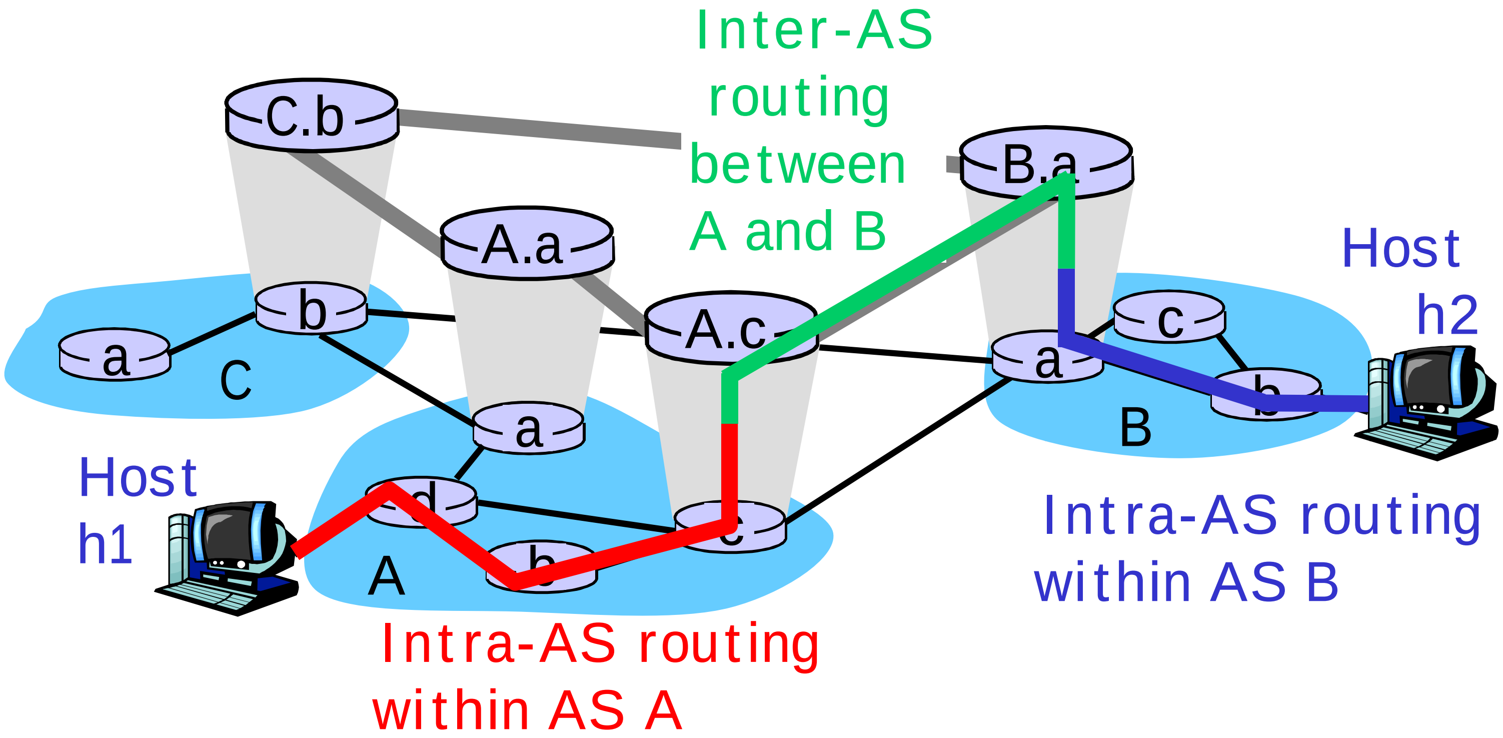
Hierarchical Routing

- aggregate routers into regions, "aut`onomous systems" (AS) – which is under sole control
- routers in same AS run same routing protocol
  - "intra-AS" routing protocol
  - routers in different AS can run different intra-AS routing protocol

gateway routers

- special routers in AS
- run intra-AS routing protocol with all other routers in AS
- *also* responsible for routing to destinations outside AS
  - run *inter-AS routing* protocol with other gateway routers

Intra-AS and Inter-AS routing


Inter-AS routing between A and B
C.b
B.a
A.a
A.c
a
b
C
c
Host h2
a
b
B
a
d
c
b
A
Host h1
Intra-AS routing within AS B
Intra-AS routing within AS A

- We'll examine specific inter-AS and intra-AS Internet routing protocols shortly

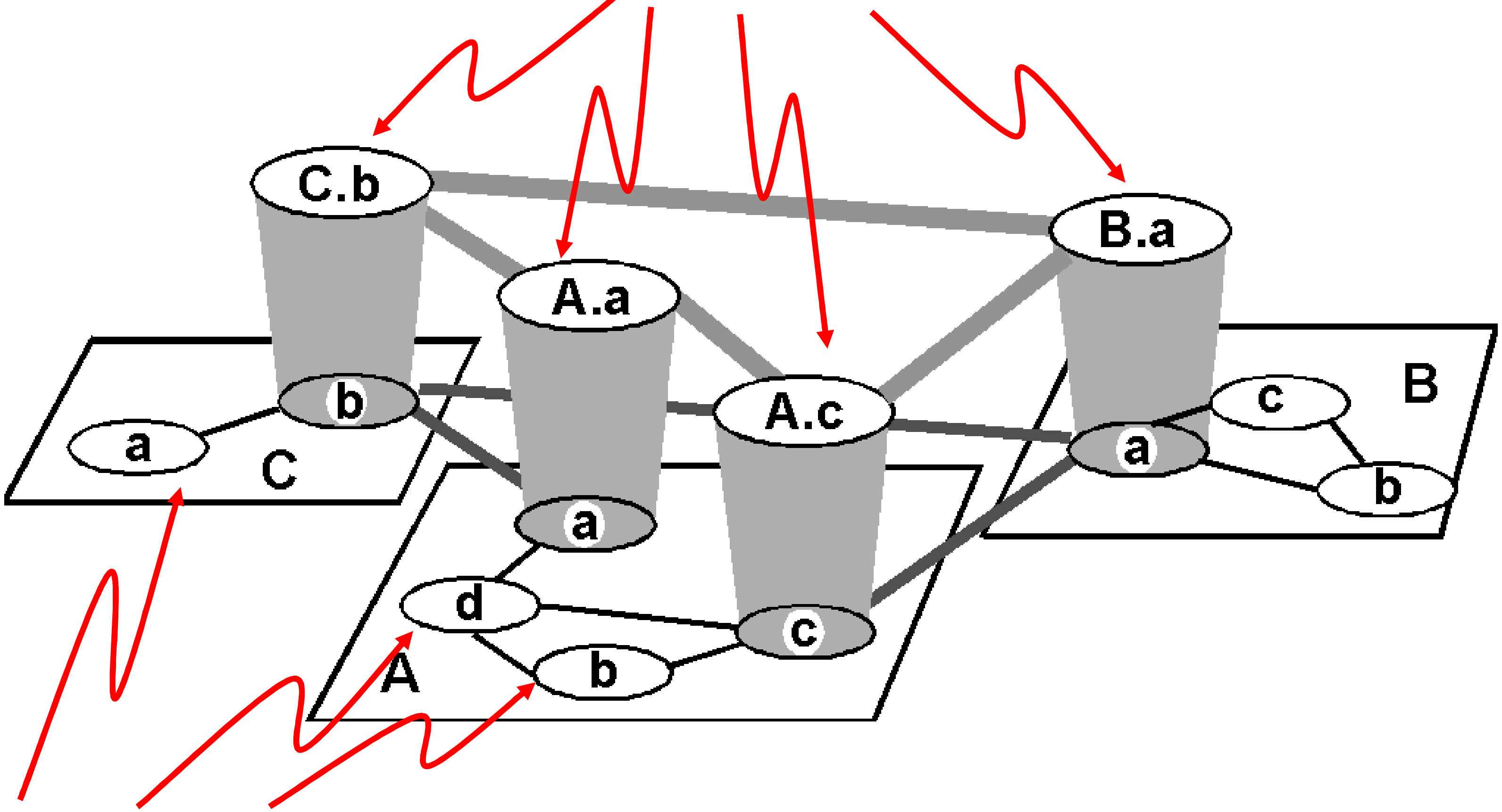
Routing in the Internet

- The Global Internet consists of Autonomous Systems (AS) interconnected with each other:
  - Stub AS: small corporation: one connection to other AS's
  - Multihomed AS: large corporation (no transit): multiple connections to other AS's
  - Transit AS: provider, hooking many AS's together

- Two-level routing:
  - Intra-AS: administrator responsible for choice of routing algorithm within network
  - Inter-AS: unique standard for inter-AS routing: BGP

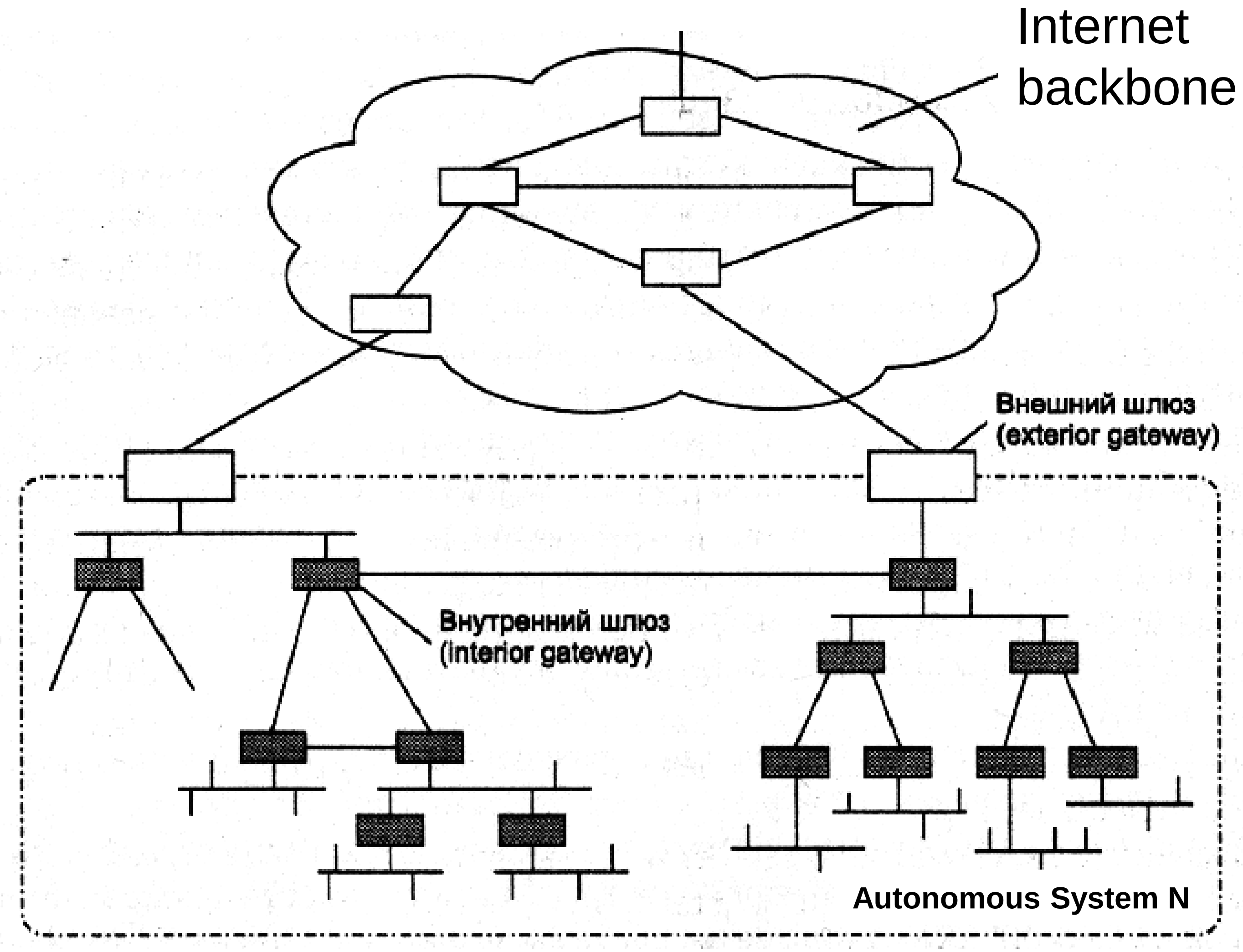
Internet AS Hierarchy

Intra-AS border (exterior gateway) routers



Inter-AS interior (gateway) routers

Internet AS Hierarchy



Internet backbone

Intra-AS Routing

- Also known as Interior Gateway Protocols (IGP)
- Most common Intra-AS routing protocols:
  - RIP(1,2): Routing Information Protocol
  - OSPF: Open Shortest Path First
  - (E)IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

RIP (Routing Information Protocol)

- Distance vector algorithm
- Included in BSD-UNIX Distribution in 1982
- Distance metric: # of hops (max = 15 hops)
- Distance vectors: exchanged among neighbors every 30 sec via Response Message (also called advertisement)
- Each advertisement: list of up to 25 destination nets within AS

OSPF (Open Shortest Path First)

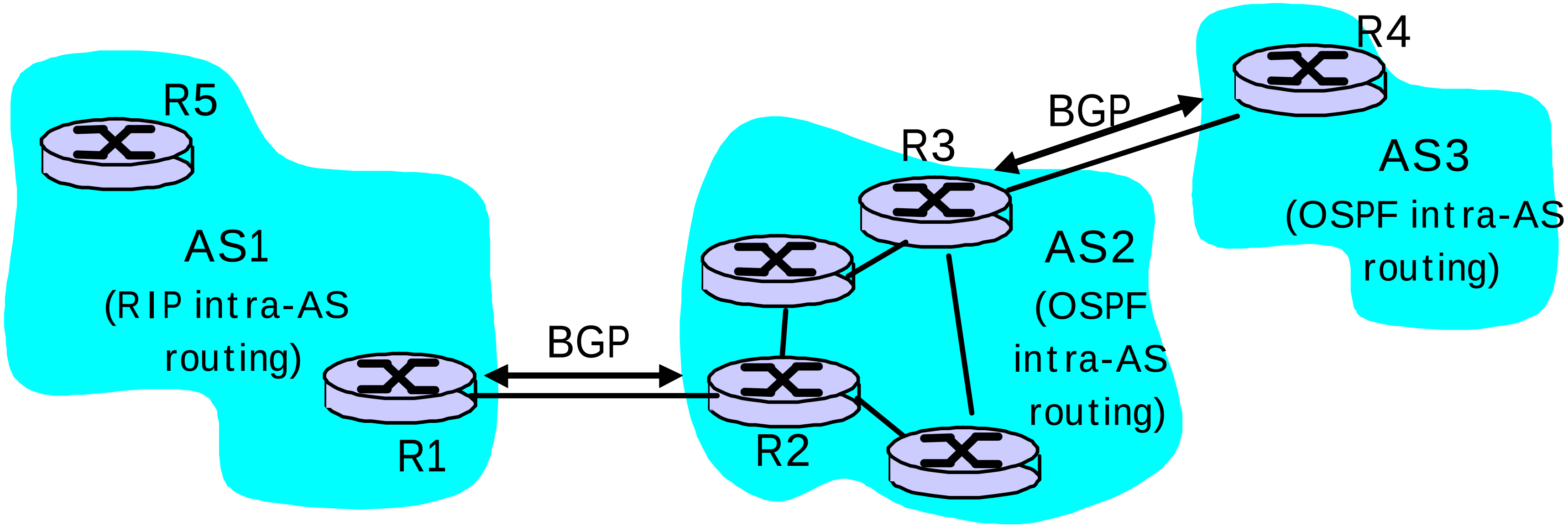
- "open": publicly available
- Uses Link State algorithm
  - LS packet dissemination
  - Topology map at each node
  - Route computation using Dijkstra's algorithm

- OSPF advertisement carries one entry per neighbor router
- Advertisements disseminated to entire AS (via flooding)
  - Carried in OSPF messages directly over IP (rather than TCP or UDP

OSPF “advanced” features (not in RIP)

- Security: all OSPF messages authenticated (to prevent malicious intrusion)
- Multiple same-cost paths allowed (only one path in RIP)
- For each link, multiple cost metrics for different TOS (e.g., satellite link cost set “low” for best effort; high for real time)
- Integrated uni- and multicast support:
  - Multicast OSPF (MOSPF) uses same topology data base as OSPF
- Hierarchical OSPF in large domains.

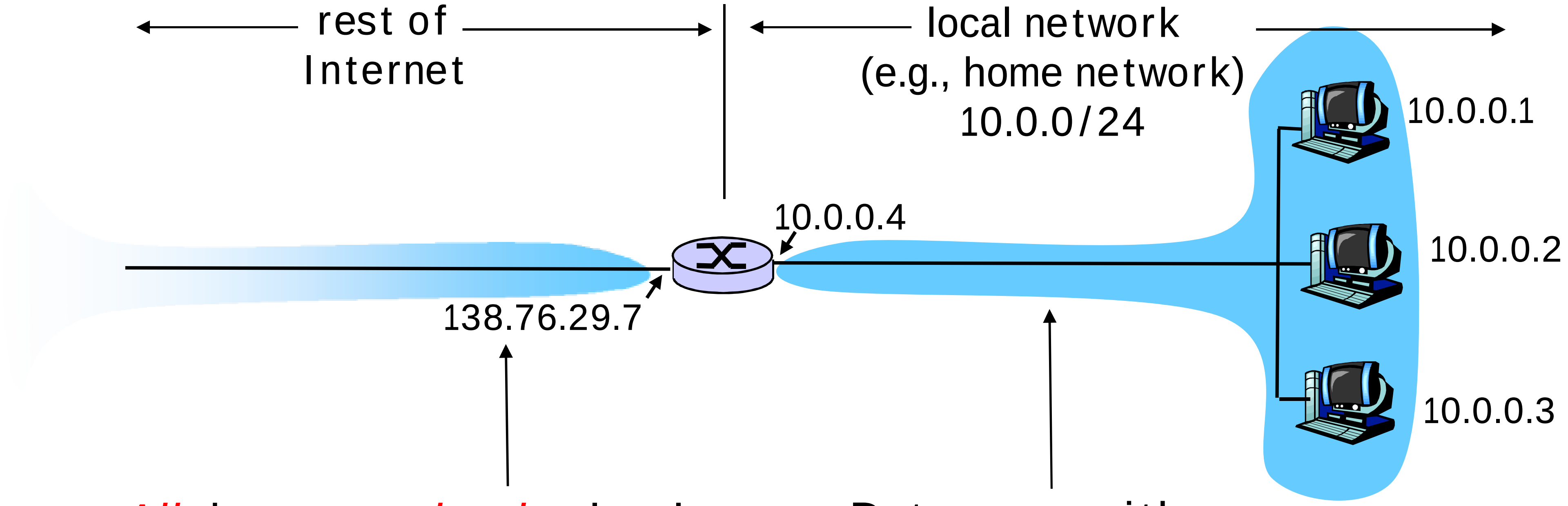
Inter-AS routing in the Internet: BGP



Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto standard
- Path Vector protocol:
  - similar to Distance Vector protocol
  - each Border Gateway broadcast to neighbors (peers) *entire path* (i.e., sequence of AS's) to destination
  - BGP routes to networks (ASs), not individual hosts

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 word is concerned:
  - no need to be allocated range of addresses from ISP: - just one IP address is used 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: Network Address Translation

- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, eg, P2P applications
  - address shortage should instead be solved by IPv6

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing error

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

IPv6

- Initial motivation: 32-bit address space soon? to be completely allocated, problems in China.
- Additional motivation:
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS

  IPv6 datagram format:
  - fixed-length 40 byte header
  - no fragmentation allowed

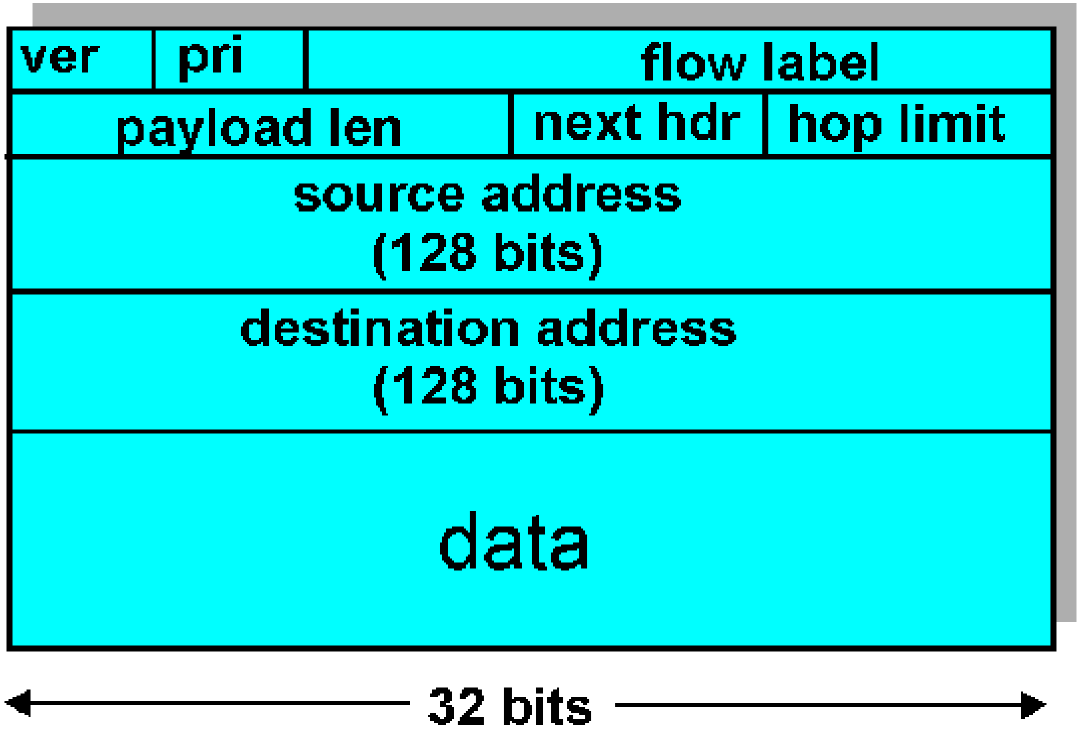
IPv6 Header (Cont)

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



ver | pri | flow label

payload len | next hdr | hop limit

source address
(128 bits)

destination address
(128 bits)

data

32 bits

Other Changes from IPv4

- *Checksum*: removed entirely to reduce processing time at each hop
- *Options:* allowed, but outside of header, indicated by "Next Header" field
- *ICMPv6:* new version of ICMP
  - additional message types, e.g. "Packet Too Big"
  - multicast group management functions

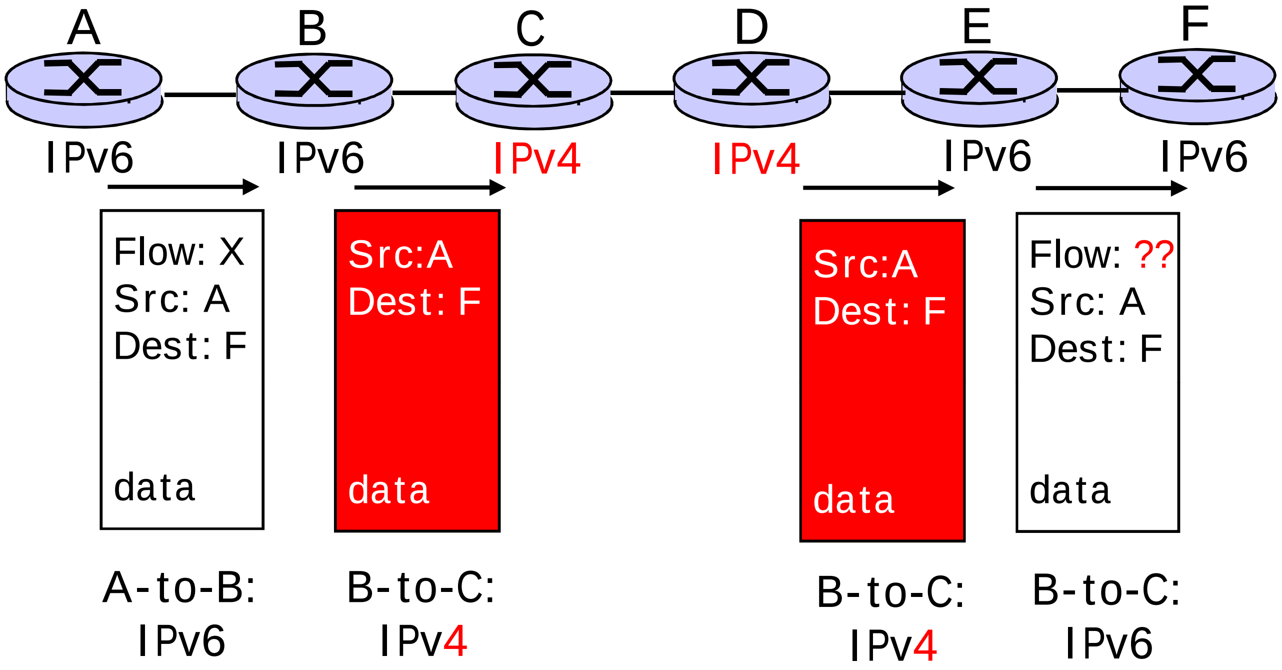
Transition From IPv4 To IPv6

- Not all routers can be upgraded simultaneous
  - no "flag days"
  - How will the network operate with mixed IPv4 and IPv6 routers?
- Two proposed approaches:
  - *Dual Stack*: some routers with dual stack (v6, v4) can "translate" between formats
  - *Tunneling:* IPv6 carried as payload in IPv4 datagram among IPv4 routers

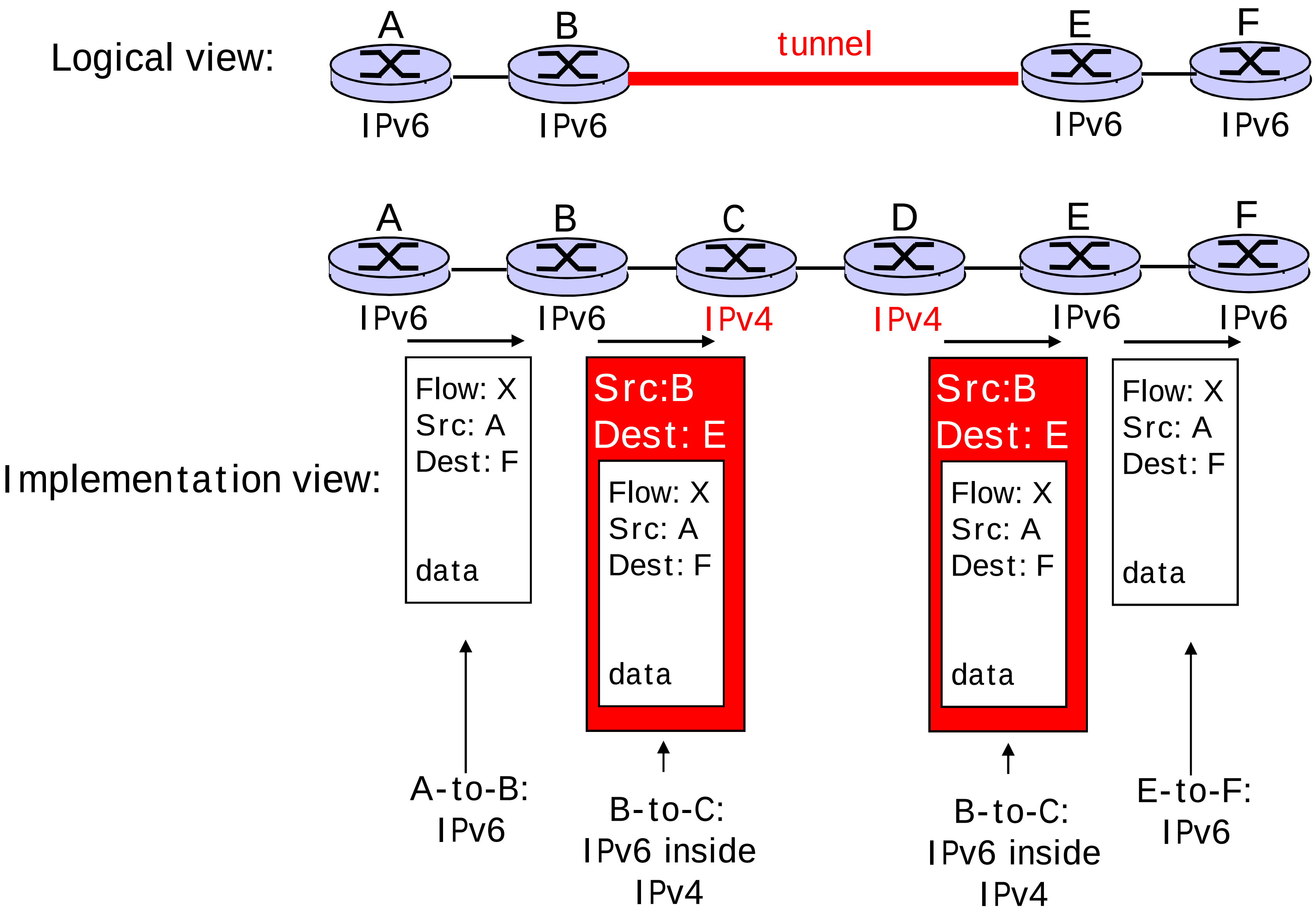
IPv6 transition plan

- An internet-draft published 08/2007 calls for an IPv6 transition plan which would require all Internet-facing servers to have IPv6 connectivity on or before January 1, 2011.
- IETF engineer John Curran proposes that migration to IPv6 happen in three stages.
  - The first stage, which would happen between now and the end of 2008, would be a preparatory stage in which organizations would start to run IPv6 servers, though these servers would not be considered by outside parties as production servers.
  - The second stage, which would take place in 2009 and 2010, would require organizations to offer IPv6 for Internet-facing servers, which could be used as production servers by outside parties.
  - Finally, in the third stage, starting in 2011, IPv6 must be in use by public-facing servers.' Then IPv4 can go away."

Dual Stack Approach



Tunneling


Logical view:
A
B
tunnel
E
F
IPv6
IPv6
IPv6
IPv6
Implementation 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

Network layer service models:

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(rt,nrt)	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

- Internet model being extended: Intserv, Diffserv
- Note: C=constant, U=unspecified, A=available, V=variable

Introduction: Summary

Covered a "ton" of material!

- Network taxonomies
- Basic definitions
- Network structure: edge, core, access
  - packet-switching versus circuit-switching
- Internet/ISP/AS structure
- performance: loss, delay
- layering and service models
- history
- IP addressing and routing

You now have:

- context, overview, "feel" of networking
- more depth, detail *to follow!*

Dynamic Routing

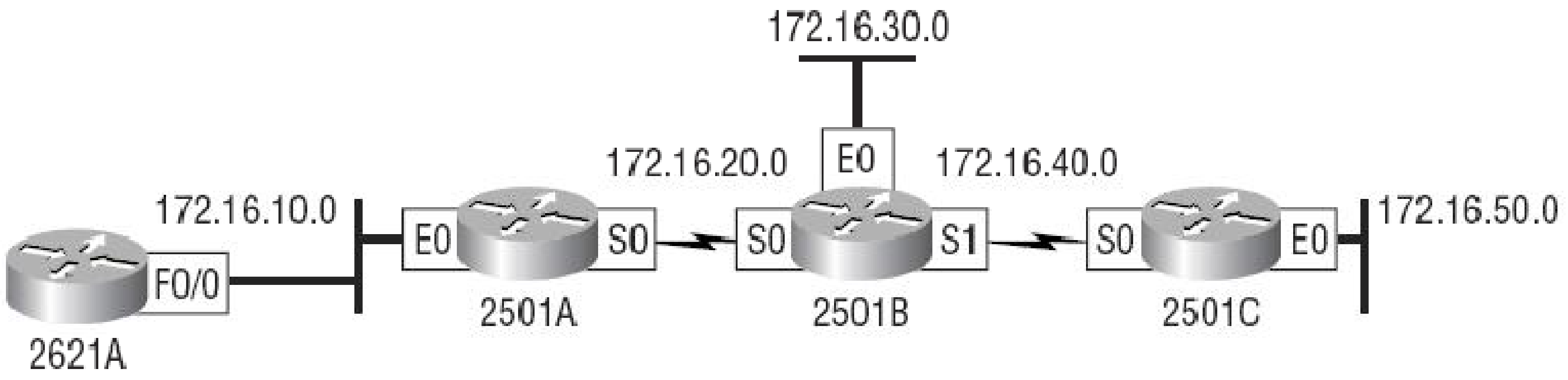
- *Routing Protocol* finds networks and updates routing table on router.
- The *Administrative Distance* (AD) is used to rate the level of *trust* of routing information received on a router from a neighbor router.
- After receiving updates, the router will place in IP-table the route with lowest AD, if ADs are equal, with lowest metric (route's quality).
- In case of equal metrics and ADs **à** router performs load ballancing.

Three classes of intra-AS/IGP routing protocols

- Distance vector
  - Each time a packet goes through a router, that’s called a *hop*.
  - Sends the entire routing table to directly connected neighbors.
  - The vector indicates the direction to the remote network.
  - “Routing by rumor”.
  - Examples – RIP, IGRP.

- Link state
  - *aka shortest-path-first protocols*
  - three separate tables:
    1. keeps track of directly attached neighbors
    2. determines the topology of the entire internetwork
    3. routing table.
  - Example – OSPF.
- Hybrid
  - use aspects of both distance vector and link state—for example, EIGRP

Distance-vector routing



Routing Table		
172.16.10.0	F0/0	0

Routing Table		
172.16.10.0	E0	0
172.16.20.0	S0	0

Routing Table		
172.16.20.0	S0	0
172.16.30.0	E0	0
172.16.40.0	S1	0

Routing Table		
172.16.40.0	S0	0
172.16.50.0	E0	0

Starting …

Routing Table		
172.16.10.0	F0/0	0
172.16.20.0	F0/0	1
172.16.30.0	F0/0	2
172.16.40.0	F0/0	2
172.16.50.0	F0/0	3

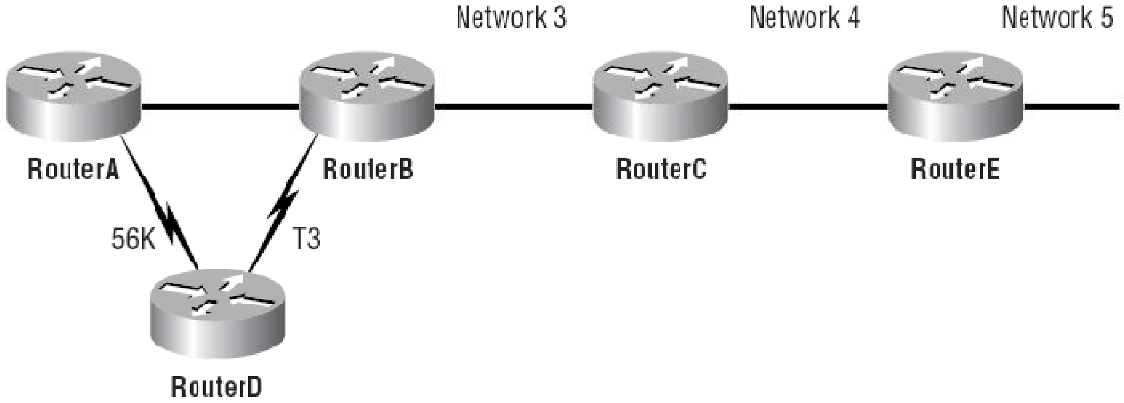
Routing Table		
172.16.10.0	E0	0
172.16.20.0	S0	0
172.16.30.0	S0	1
172.16.40.0	S0	1
172.16.50.0	S0	2

Routing Table		
172.16.20.0	S0	0
172.16.30.0	E0	0
172.16.40.0	S1	0
172.16.10.0	S0	1
172.16.50.0	S1	1

Routing Table		
172.16.40.0	S0	0
172.16.50.0	E0	0
172.16.10.0	S0	2
172.16.20.0	S0	1
172.16.30.0	S0	1

converged

Routing loop



1. Network 5 fails, Router E tells Router C
2. Router C to stop routing to Network5 through Router E
3. But Routers A, B, D don't know about Network 5 yet and keep sending updates
4. Router C will eventually send out its update and cause B to stop routing to Network 5, but Routers A and D are still not updated. To them, it appears that Network 5 is still available through Router B with a metric of 3.
5. The problem occurs when Router A sends out its regular 30-second update message, which includes the ability to reach Network 5 and now *Routers B and D receive the news that Network 5 can be reached from Router A, so Routers B and D then send out the information that Network 5 is available. Any packet destined for Network 5 will go to Router A, to Router B, and then back to Router A*.

Routing loop 2

- Maximum Hop Count
  - permits a hop count limit (e.g. 15 hops), anything that requires 16 hops is considered unreachable
- Split Horizon
  - won't advertise the route back out that same interface
- Route Poisoning
  - For example, when Network 5 goes down, Router E initiates route poisoning by advertising Network 5 as 16, or unreachable
- Holddown
  - prevents regular update messages from reinstating a route that is flapping: if route is down, starts counter.
  - if route is up again with better metric, remove holddowns

RIP

- Open standard
- maximum allowable hop count of 15
- Versions: – RIP1, RIP2, RIPng
- AD=120
- Can perform load balancing for up to six equal-cost links (four by default)
- version RIPng supports IPv6
- Version 2 similar to version 1, but in addition:
  - sends subnet mask information with the route updates – can be considered as classless routing protocol,
  - can support VLSMs and
  - summarization of network boundaries,
  - authentication,
  - uses multicast instead of ver.1 broadcasts
  - can support discontiguous networking.

- Aka R.I.P.

RIP timers

- Route update timer
  - Sets the interval (typically 30 seconds) between periodic routing updates, in which the router sends a complete copy of its routing table out to all neighbors.
- Route invalid timer
  - Determines the length of time that must elapse (180 seconds) before a router determines that a route has become invalid. It will come to this conclusion if it hasn’t heard any updates about a particular route for that period. When that happens, the router will send out updates to all its neighbors letting them know that the route is invalid.
- Holddown timer
  - This sets the amount of time during which routing information is suppressed. The default is 180 seconds.
- Route flush timer
  - Sets the time between a route becoming invalid and its removal from the routing table (240 seconds). Before it’s removed from the table, the router notifies its neighbors of that route’s impending demise. The value of the route invalid timer must be less than that of the route flush timer.
- RIP can perform load balancing for up to six equal-cost links (four by default)

Configuring RIP

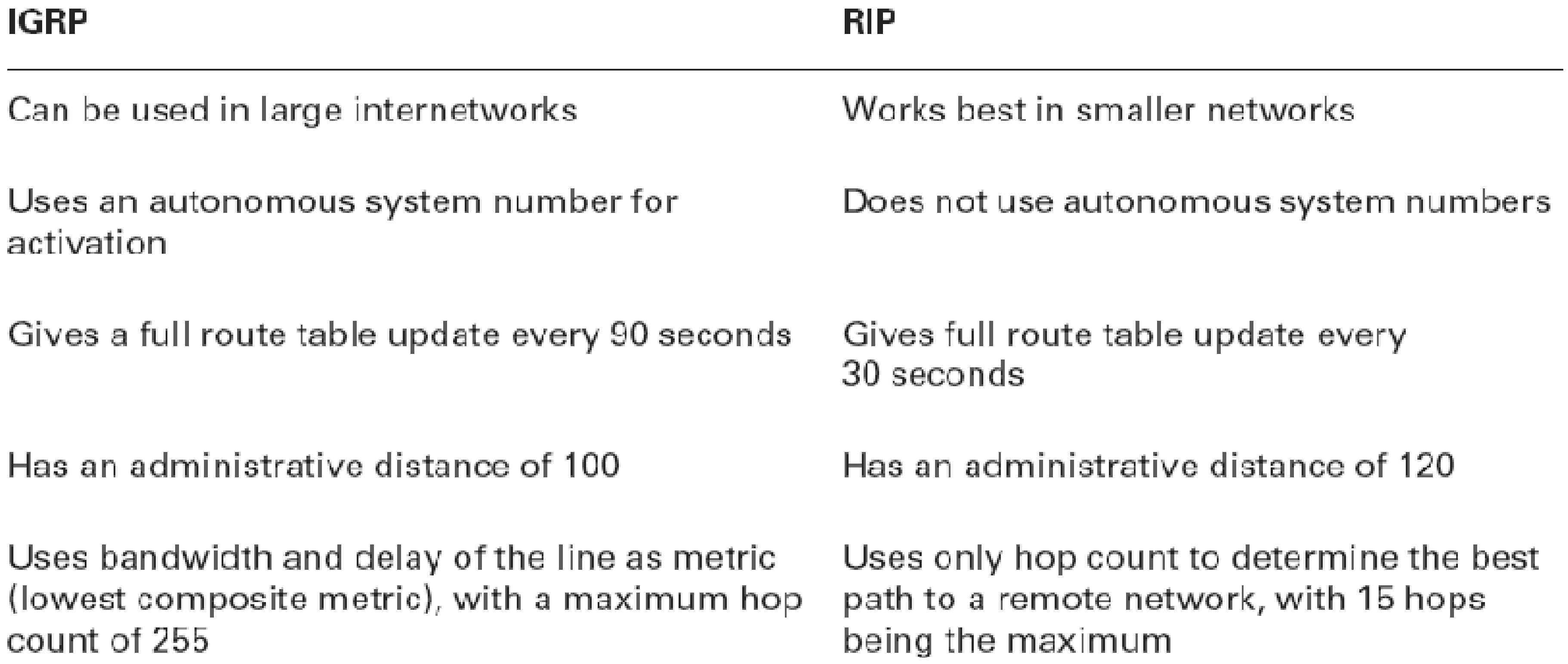
```
Lab(config)#router rip
Lab(config-router)#network 192.168.20.0
Lab(config-router)#network 192.168.30.0
Lab(config-router)#network 192.168.40.0
Lab(config-router)#network 192.168.50.0
?Lab(config-router)#version 2
?Lab(config-router)#passive-interface serial 0/1
Lab(config-router)#^Z
Lab#

Lab#sh ip route
[...]
R 192.168.30.0 [120/1] via 192.168.40.1, 00:00:04, Serial0/0
```

Q: what's wrong?

```
show ip protocol
debug ip rip
```

Interior Gateway Routing Protocol (IGRP)



IGRP	RIP
Can be used in large internetworks	Works best in smaller networks
Uses an autonomous system number for activation	Does not use autonomous system numbers
Gives a full route table update every 90 seconds	Gives full route table update every 30 seconds
Has an administrative distance of 100	Has an administrative distance of 120
Uses bandwidth and delay of the line as metric (lowest composite metric), with a maximum hop count of 255	Uses only hop count to determine the best path to a remote network, with 15 hops being the maximum

Replacement is Enhanced IGRP (EIGRP)
IGRP can load-balance up to 6 unequal links, uses bandwidth to determine how to load-balance. To load-balance over unequal-cost links, the variance command controls the load balancing between the best metric and the worst acceptable metric.
Note: Reliability, load, and maximum transmission unit (MTU) can also be used to calculate composite metric, although they are not used by default.

IGRP timers

- Update timers
  - These specify how frequently routing-update messages should be sent. The default is 90 seconds.
- Invalid timers
  - These specify how long a router should wait before declaring a route invalid if it doesn't receive a specific update about it. The default is three times the update period.
- Holddown timers
  - These specify the holddown period. The default is three times the update timer period plus 10 seconds.
- Flush timers
  - These indicate how much time should pass before a route should be flushed from the routing table. The default is seven times the routing update period. If the update timer is 90 seconds by default, then 7 x 90 = 630 seconds elapse before a route will be flushed from the route table.

Configuring IGRP

```
Lab#config t
Lab(config)#router igrp 10
Lab(config-router)#netw 192.168.10.0
Lab(config-router)#netw 192.168.20.0
Lab(config-router)#^Z
Lab#
```

```
Lab#sh ip route
[…]
C 192.168.50.0 is directly connected, FastEthernet 0/0
C 192.168.40.0 is directly connected, Serial0/0
I 192.168.30.0 [100/143723] via 192.168.40.1, 00:00:42, Serial0/0
I 192.168.20.0 [100/152365] via 192.168.40.1, 00:00;52, Serial0/0
I 192.168.10.0 [100/158350] via 192.168.20.1, 00:00:36, Serial0/0
Lab_C#
```

Note: 158350 – composite metric

```
debug ip igrp events
debug ip igrp transactions
undebug all
```

Enhanced IGRP (EIGRP)

- Support for IP, IPX, and AppleTalk via protocol-dependent modules
- Considered classless (same as RIPv2 and OSPF)
- Support for VLSM/CIDR
- Support for summaries and discontiguous networks
- Efficient neighbor discovery
- Communication via Reliable Transport Protocol (RTP): proprietary, multicast w unicast fallback
- Best path selection via Diffusing Update Algorithm (DUAL)

Neighbor Discovery

- neighborship (aka adjacencies) establishment:
  - Hello or ACK received
  - AS numbers match
  - Identical metrics (K values)
- to maintain the neighborship relationship, EIGRP routers must also continue receiving Hellos from their neighbors.
- IGRP advertises its entire routing table only when it discovers a new neighbor and forms an adjacency

EIGRP terminology

- Reported distance - the metric of a remote network, as reported by a neighbor. It is also the routing table metric of the neighbor, and is the same as the number after the slash in the topology table.
- Feasibility Condition - if, for a destination, a neighbor router advertises a distance that is strictly lower than our feasible distance, then this neighbor lies on a loop-free route to this destination.
- Feasible distance - the best metric along all paths to a remote network, including the metric to the neighbor that is advertising that remote network. This is the route that you will find in the routing table, because it is considered the best path.
- Feasible successor - provides a loop-free path whose reported distance is less than the feasible distance, and it is considered a backup route. EIGRP will keep up to 4(6) feasible successors in the topology table. Only the one with the best metric (the *successor*) is placed in the routing table. The show ip eigrp topology command will display all the EIGRP feasible successor routes known to a router.
- Successor - (successful!) provides the best route to a remote network. A successor route is used by EIGRP to forward traffic to a destination and is stored in the routing table. It is backed up by a feasible successor route that is stored in the topology table—if one is available.

By having feasible successors in the topology table as backup links, the network can converge instantly, and updates to any neighbor are the only traffic sent from EIGRP.

If there isn't alternative in the local topology table, EIGRP router asks neighbors (diffusing).

EIGRP tables

- Neighbor Table: stores data about the neighboring routers, i.e. those accessible through directly connected interfaces. There is one neighbor table for each protocol-dependent module. Sequence numbers are used to match acknowledgments with update packets. The last sequence number received from the neighbor is recorded so that out-of-order packets can be detected.

- Topology Table: contains the aggregation of the routing tables gathered from all the neighbors. This table contains a list of destination networks in the EIGRP-routed network together with their respective metrics. Also for every destination, a *successor* and a *feasible successor* are identified and stored in the table if they exist. Every destination in the topology table can be marked either as "Passive", which is the state when the routing has stabilized and the router knows the route to the destination, or "Active" when the topology has changed and the router is in the process of (actively) updating its route to that destination.
- Routing table: Stores the actual routes to all destinations; the routing table is populated from the topology table with every destination network that has its successor and optionally feasible successor identified (if unequal-cost load-balancing is enabled using the variance command). The successors and feasible successors serve as the next hop routers for these destinations

- Unlike most other distance vector protocols, EIGRP does not rely on periodic route dumps in order to maintain its topology table. *Routing information is exchanged once at a router's startup*, after which only changes are sent.
- EIGRP *doesn't send link-state packets* as OSPF does; instead, it sends traditional distance-vector updates containing information about networks plus the cost of reaching them from the perspective of the advertising router

- http://www.cisco.com/warp/public/103/eigrp-toc.html

EIGRP metric

- bandwidth = 10000000/bandwidth(Kbps))
- delay = delay(mks)

$$\left[\left(K_1 \cdot \text{Bandwidth} + \frac{K_2 \cdot \text{Bandwidth}}{256 - \text{Load}} + K_3 \cdot \text{Delay}\right) \cdot \frac{K_5}{K_4 + \text{Reliability}}\right] \cdot 256$$

- Default K values: K1=1; K2=0; K3=1; K4=0; K5=0, hence metric = 256*(bandwidth + delay)
- Reliability’s range [1..255], 255 – most reliable
- Load belongs to range [1..255], 255 - saturated

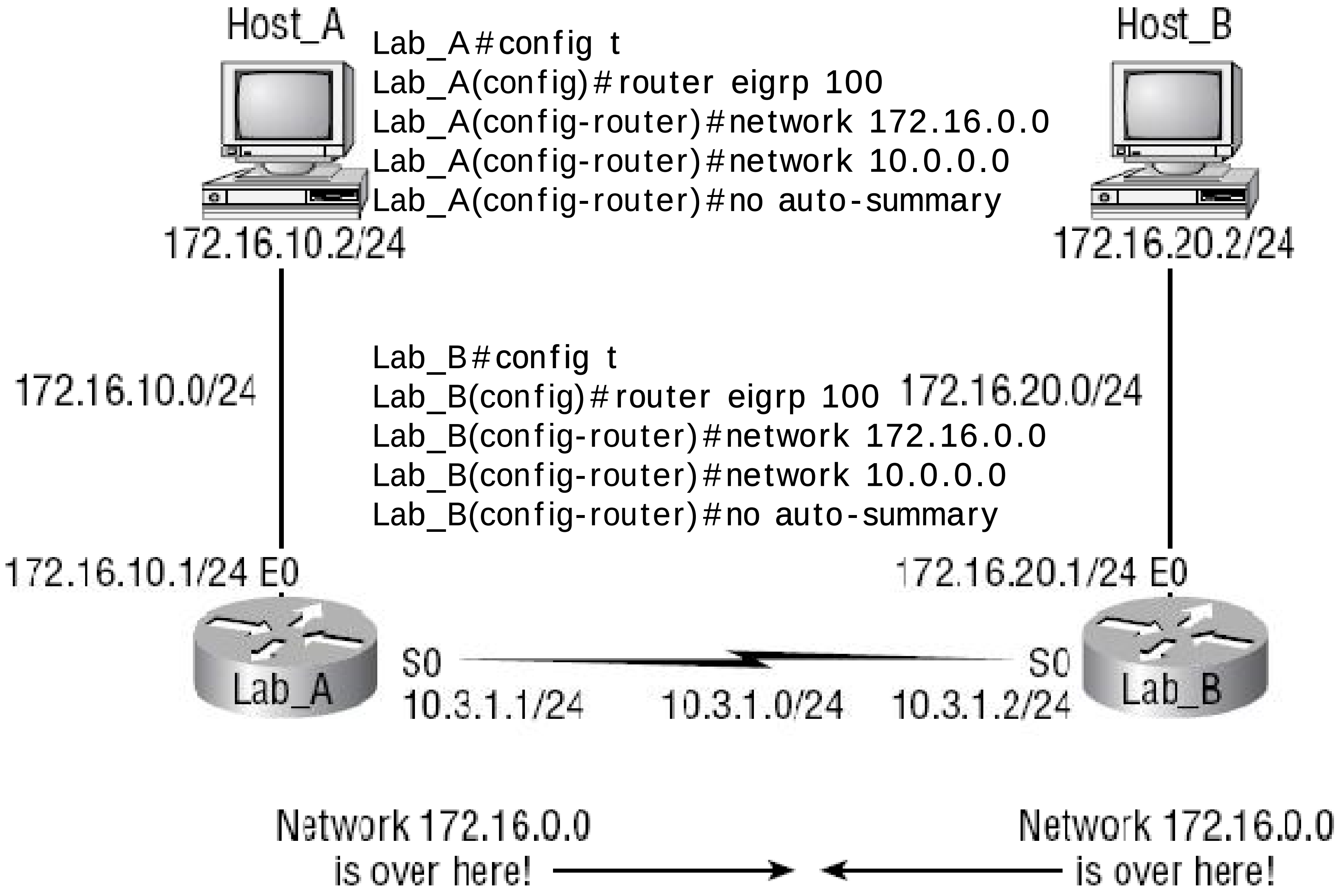
Configuring EIGRP

```
Lab#config t
Lab(config)#router eigrp 100
Lab(config-router)#network 172.16.0.0
Lab(config-router)#network 10.0.0.0
Lab(config-router)#no auto-summary
```

```
show ip route eigrp
show ip eigrp neighbors
show ip eigrp topology
```

```
Lab#sh ip route
[...]
D 192.168.50.0/24 [90/2707456] via 192.168.20.2,00:04:35, Serial0/0
```

Discontiguous networks



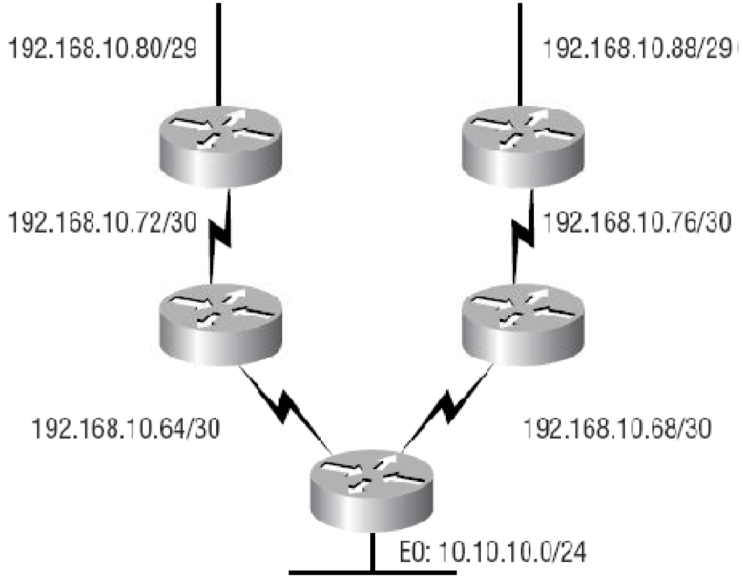
```
Lab_A#config t
Lab_A(config)#router eigrp 100
Lab_A(config-router)#network 172.16.0.0
Lab_A(config-router)#network 10.0.0.0
Lab_A(config-router)#no auto-summary
```

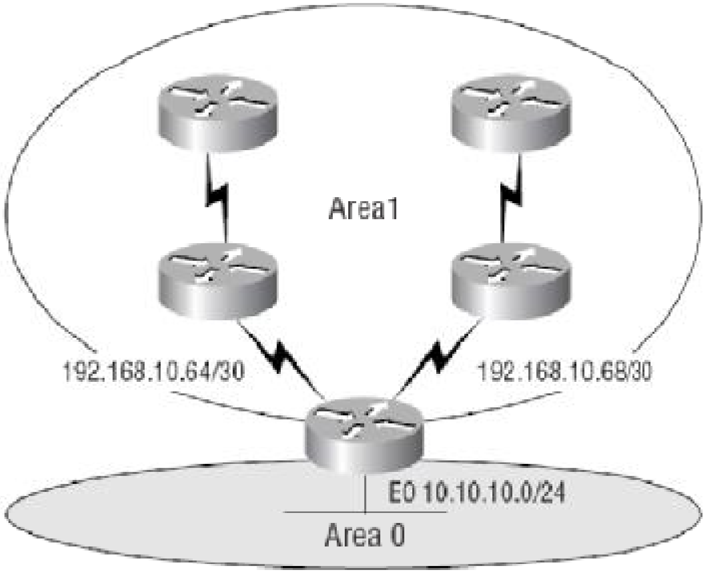
```
Lab_B#config t
Lab_B(config)#router eigrp 100
Lab_B(config-router)#network 172.16.0.0
Lab_B(config-router)#network 10.0.0.0
Lab_B(config-router)#no auto-summary
```

Redistribution

- Routes originated within a specific AS in IGRP are called internal EIGRP route
- External EIGRP routes (AD=170) appear within EIGRP route tables courtesy of either manual or automatic redistribution, and they represent networks that originated outside of the EIGRP autonomous system
- If the same AS number set for EIGRP that used for IGRP, EIGRP will automatically redistribute the routes from IGRP into EIGRP with AD=170

Configuring EIGRP and OSPF Summary Routes





```
Core#config t
Core(config)#router eigrp 10
Core(config-router)#network 192.168.10.0
Core(config-router)#network 10.0.0.0
Core(config-router)#no auto-summary
Core(config-router)#interface ethernet 0
Core(config-if)#ip summary-address eigrp 10
192.168.10.64 255.255.255.224
```

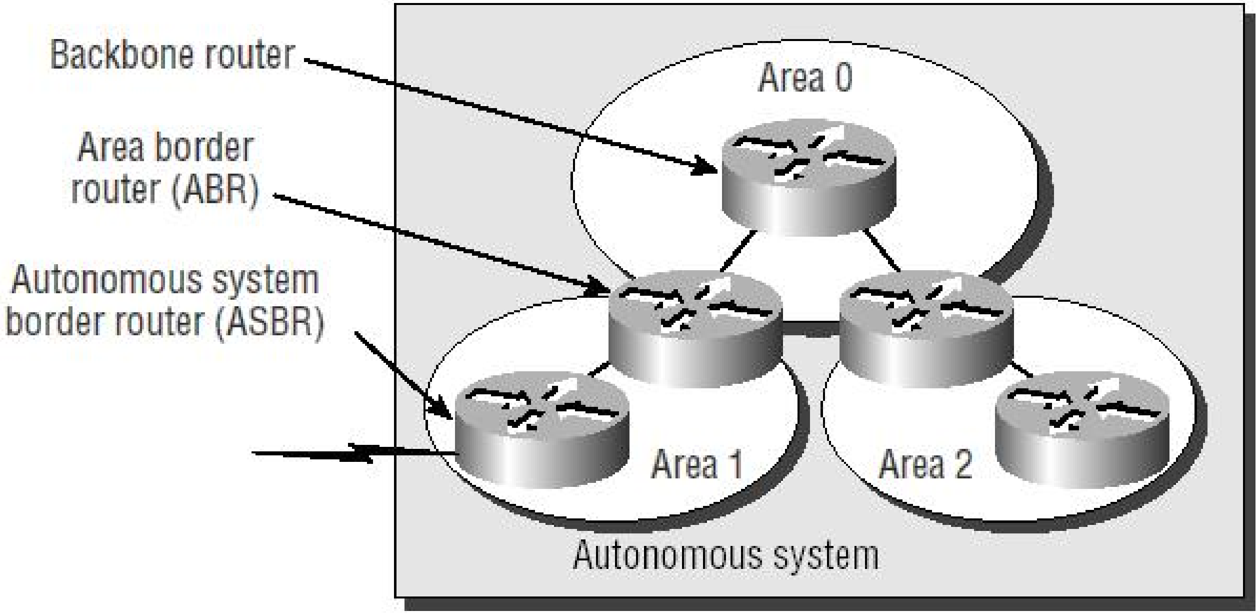
```
Core#config t
Core(config)#router ospf 1
Core(config-router)#network 192.168.10.64 0.0.0.3 area 1
Core(config-router)#network 192.168.10.68 0.0.0.3 area 1
Core(config-router)#network 10.10.10.0 0.0.0.255 area 0
Core(config-router)#area 1 range 192.168.10.64
255.255.255.224
```

Open Shortest Path First (OSPF)

- Open standard
- Minimizes routing update traffic
- Allows scalability
- Supports VLSM/CIDR
- Has unlimited hop count

Why creating OSPF in a hierarchical design?

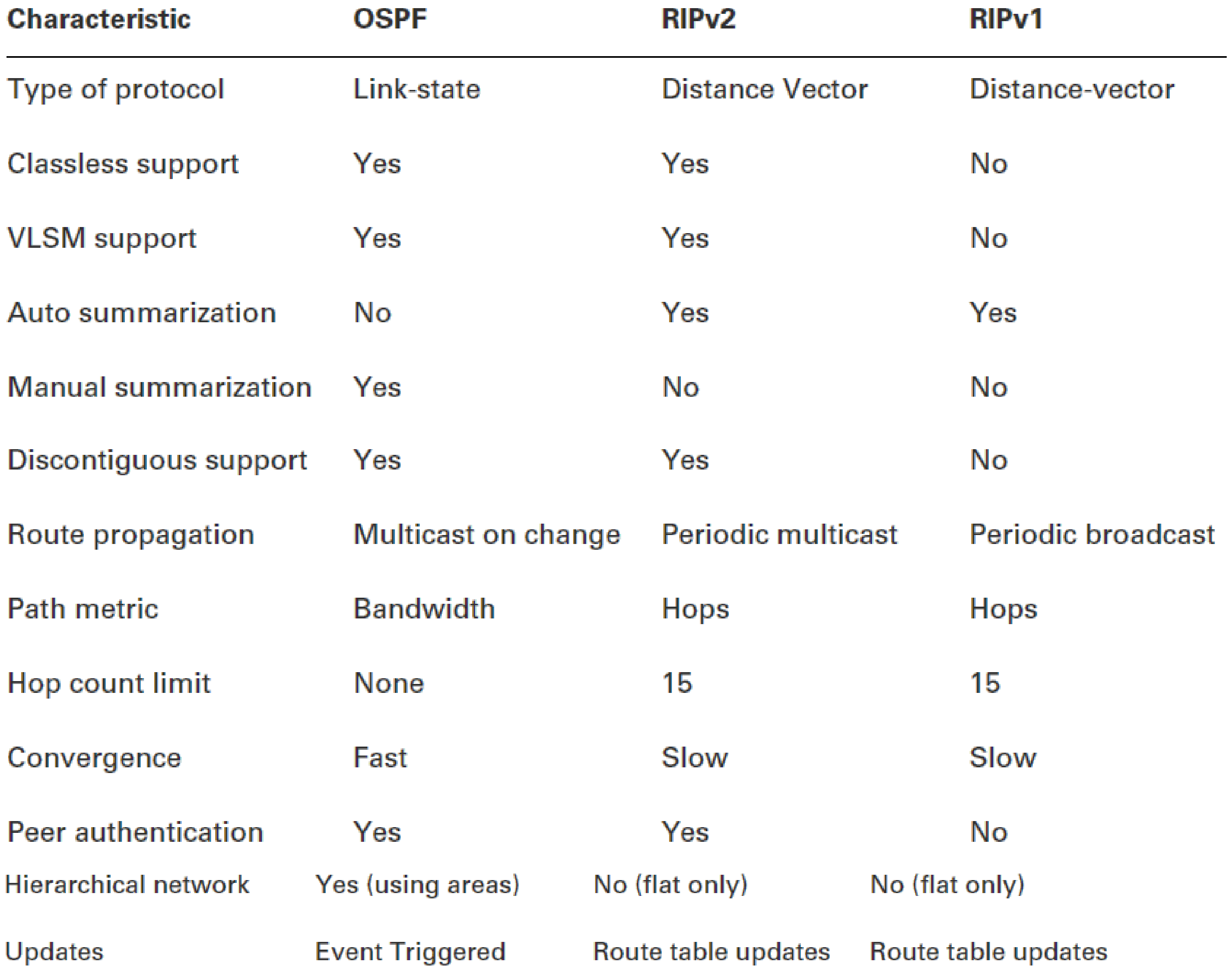
- to decrease routing overhead
- to speed up convergence
- to confine network instability to single areas of the network



Open Shortest Path First (OSPF) 2

- Routers in the same broadcast domain or at each end of a point-to-point telecommunications link form *adjacencies* when they have detected each other, when "see" itselfs in a hello packets.
- OSPF uses both unicast and multicast to send "hello packets" and link state updates. Multicast addresses 224.0.0.5 (all SPF/link state routers) and 224.0.0.6 (all Designated Routers) are reserved for OSPF. OSPF does not use TCP or UDP but uses IP directly
- The routers elect a *designated router* (DR) and a *backup designated router* (BDR) which act as a hub to reduce traffic between routers in case of broadcast network. BDR, DR are not used in P2P.

OSPF vs R I P



Characteristic	OSPF	RIPv2	RIPv1
Type of protocol	Link-state	Distance Vector	Distance-vector
Classless support	Yes	Yes	No
VLSM support	Yes	Yes	No
Auto summarization	No	Yes	Yes
Manual summarization	Yes	No	No
Discontiguous support	Yes	Yes	No
Route propagation	Multicast on change	Periodic multicast	Periodic broadcast
Path metric	Bandwidth	Hops	Hops
Hop count limit	None	15	15
Convergence	Fast	Slow	Slow
Peer authentication	Yes	Yes	No
Hierarchical network	Yes (using areas)	No (flat only)	No (flat only)
Updates	Event Triggered	Route table updates	Route table updates

OSPF terminology

- Router ID (RID) - is an IP address used to identify the router. Cisco chooses the RID by using the highest IP address of all configured loopback interfaces. If no loopback interfaces are configured with addresses, OSPF will choose the highest IP address of all active physical interfaces.
- Adjacency - is a relationship between two routers that permits the direct exchange of route updates. And not all neighbors will become adjacent—this depends upon both the type of network and the configuration of the routers.
- Link State Advertisement (LSA) - is an OSPF data packet containing link-state and routing information that's shared among OSPF routers.
- Hello protocol - provides dynamic neighbor discovery and maintains neighbor relationships. Hello packets and Link State Advertisements (LSAs) build and maintain the topological database. Hello packets are addressed to 224.0.0.5.
- Area - is a grouping of contiguous networks and routers. All routers in the same area share a common Area ID. Because a router can be a member of more than one area at a time, the Area ID is associated with specific interfaces on the router.

OSPF databases

- Neighborship database - is a list of all OSPF routers for which Hello packets have been seen. A variety of details, including the RID and state, are maintained on each router in the neighborship database.
- Topology database - contains information from all of the Link State Advertisement packets that have been received for an area. The router uses the information from the topology database as input into the Dijkstra algorithm that computes the shortest path to every network.

OSPF

- each router calculates the best/shortest path to every network in that same area forming a tree. This calculation is based upon the information collected in the topology database and an algorithm called shortest path first (SPF).
- OSPF uses a metric referred to as cost - which is associated with every outgoing interface included in an SPF tree.
- CISCO's OSPF cost=10^8/bandwidth

Configuring OSPF

- Disabling protocols with higher AD
  - Lab(config)#no router eigrp 10
  - Lab(config)#no router igrp 10
- Enabling OSPF (PID is locally significant)
  - Lab(config)#router ospf ?
  - <1-65535>
  - Lab(config)#router ospf 1
- Configuring OSPF areas
  - network 10.0.0.0   0.255.255.255   area 0

Configuring OSPF 2

```
Lab#config t
Lab(config)#router ospf 64999
Lab(config-router)#network 192.168.40.0 0.0.0.255 area 0
Lab(config-router)#network 192.168.50.0 0.0.0.255 area 0
Lab(config-router)#^Z
Lab#
```

```
show ip ospf
show ip ospf database
show ip ospf interface
```

shows: IP, Area assignment, Process ID, RID, Network type, Cost, Priority, DR/BDR election information (if applicable), Hello and Dead timer intervals, Adjacent neighbor information

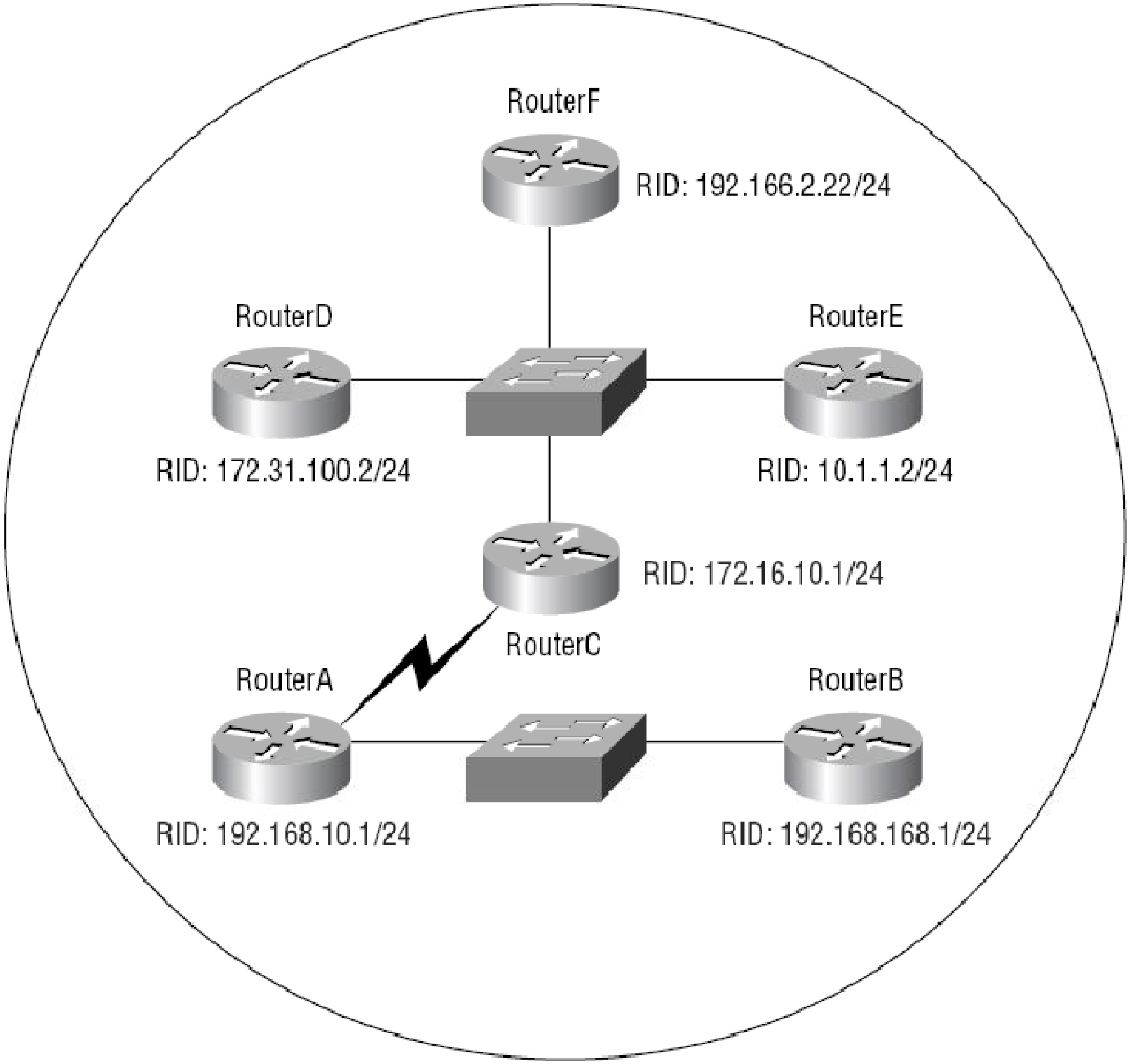
```
show ip ospf neighbor
show ip protocols
```

```
Lab#sh ip route
[...]
O 192.168.30.0/24 [110/65] via 192.168.20.2, 00:01:07, Serial0/0
```

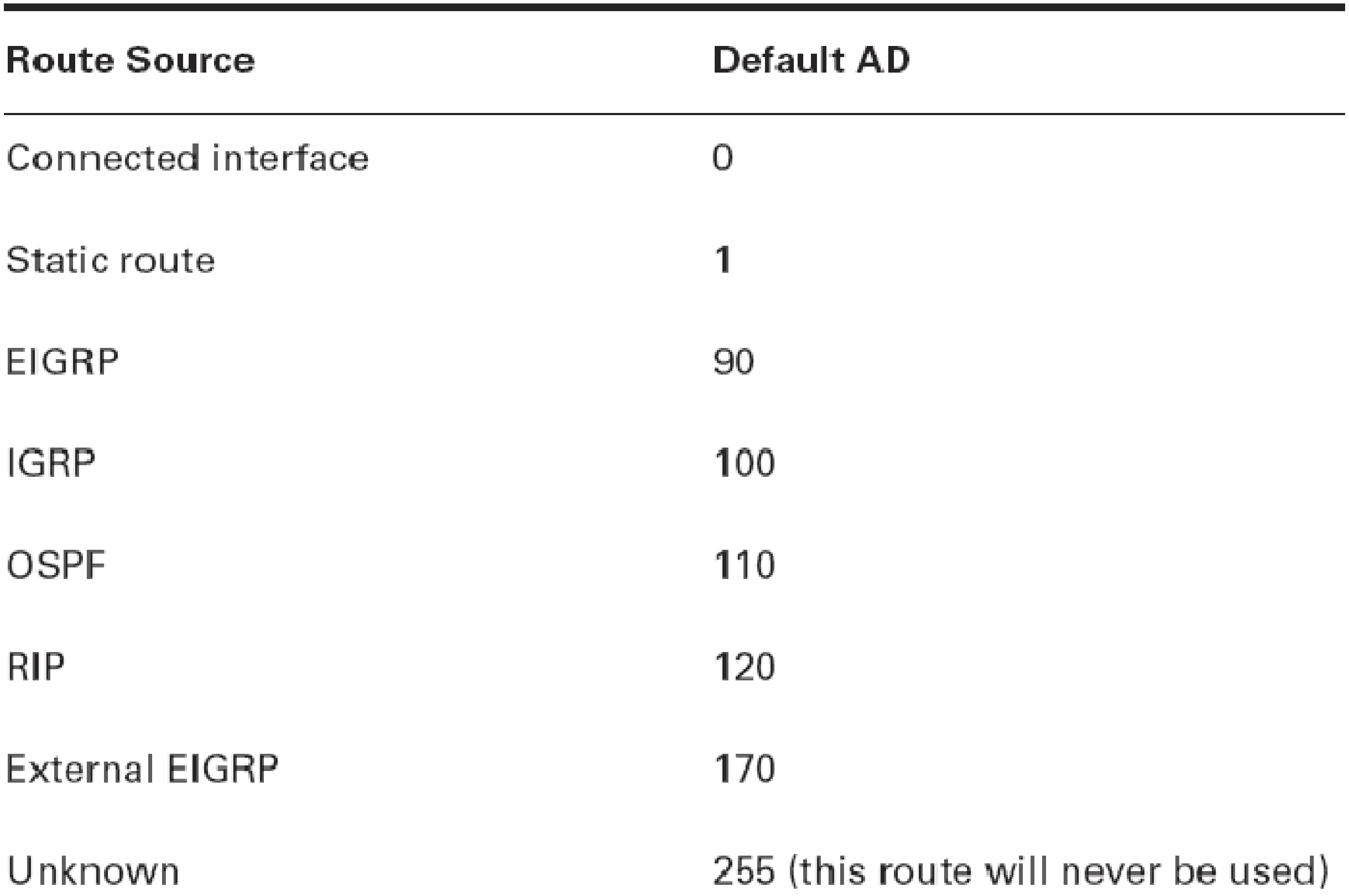
Configuring Loopback Interfaces

```
Lab#config t
Lab(config)#int loopback 0
Lab(config-if)#ip address 172.16.10.1 255.255.255.255
Lab(config-if)#no shut
Lab(config-if)#^Z
Lab#
```

Designated router example

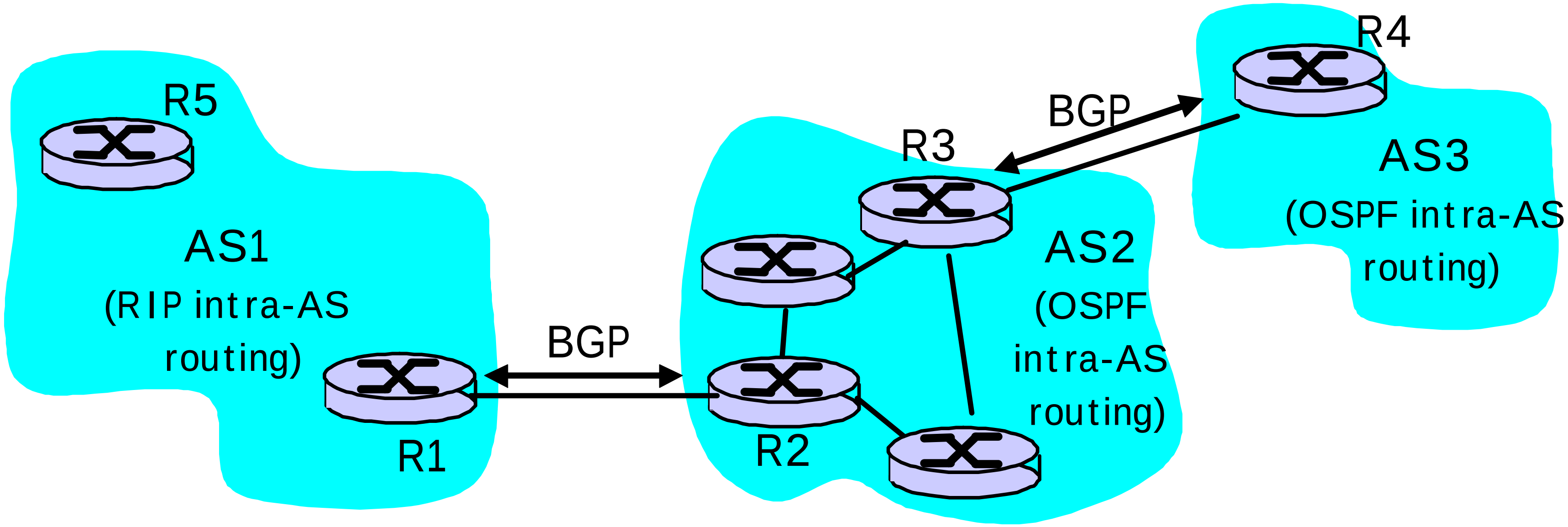


AD’s values



Route Source	Default AD
Connected interface	0
Static route	1
EIGRP	90
IGRP	100
OSPF	110
RIP	120
External EIGRP	170
Unknown	255 (this route will never be used)

Inter-AS routing in the Internet: BGP



Substitution of EGP

Since 1994, version 4 of the protocol has been in use on the Internet

Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto standard
- Path Vector protocol:
  - similar to Distance Vector protocol
  - each Border Gateway broadcast to neighbors (peers) *entire path* (i.e., sequence of AS's) to destination
  - BGP routes to ASs networks , not individual hosts
  - E.g., Gateway X may send its path to dest. Z:

Path (X,Z) = X,Y1,Y2,Y3,…,Z

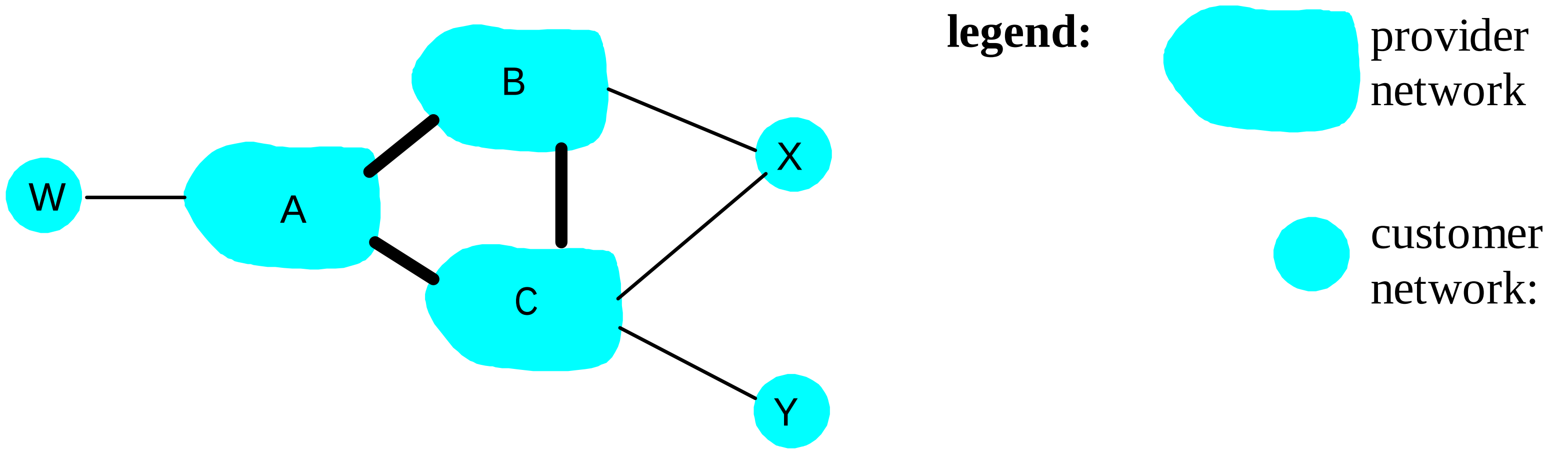
Internet inter-AS routing: BGP

Suppose: gateway X send its path to peer gateway W

- W may or may not select path offered by X
  - cost, policy (don’t route via competitors AS), loop prevention reasons.
- If W selects path advertised by X, then:

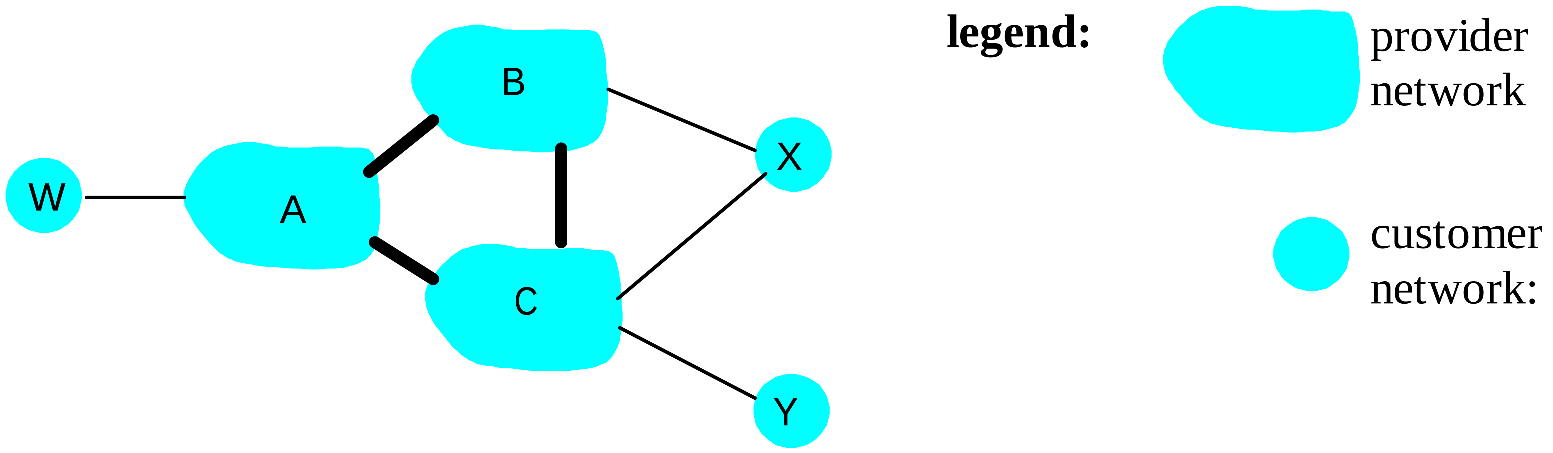
  Path (W,Z) = x, Path (X,Z)
- Note: X can control incoming traffic by controlling it route advertisements to peers:
  - e.g., don’t want to route traffic to Z -> don’t advertise any routes to Z

BGP: controlling who routes to you



- A,B,C are backbone provider networks
- X,W,Y are customers networks
- X is dual-homed: attached to two networks
  - X does not want to route from B via X to C
  - .. so X will not advertise to B a route to C

BGP: controlling who routes to you



- A advertises to B the path AW
- B advertises to X the path BAW
- Should B advertise to C the path BAW?
  - No. B gets no “revenue” for routing CBAW since neither W nor C are B’s customers
  - B wants to force C to route to w via A
  - B wants to route *only* to/from its customers!

BGP operation

Q: What does a BGP router do?

- Receiving and filtering route advertisements from directly attached neighbor(s).
- Route selection.
  - To route to destination X, which path)of several advertised) will be taken?
- Sending route advertisements to neighbors.

BGP messages

- BGP messages exchanged using TCP port=179.
- BGP messages:
  - OPEN: opens TCP connection to peer and authenticates sender
  - UPDATE: advertises new path (or withdraws old)
  - KEEPALIVE keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - NOTIFICATION: reports errors in previous msg; also used to close connection

Route parameters (BGP's attributes):

Weight

Local preference

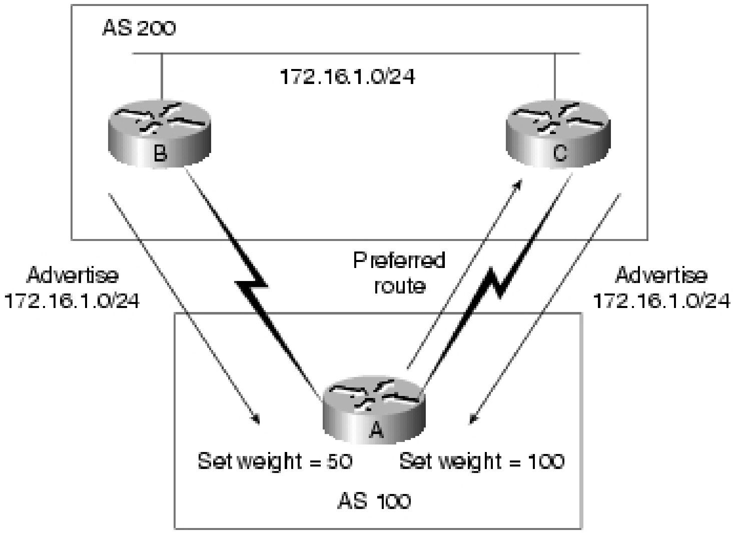
Multi-exit discriminator

Origin

AS_path

Next hop

Community



http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

Configuring BGP

- router BGP 7777
  - where 7777 - AS is under your control
- neighbor 1.2.3.4 remote-as 8888
- network 9.10.11.0 mask 255.255.255.0

- show ip bgp summary
- show ip bgp

Access Lists

Managing Traffic with Access Lists

- *access list* - is essentially a list of conditions that categorize packets
- Rules:

1. always starts with the first line of the access list, then go to line 2, and so on.
2. packet is compared with lines of the access list only until a match is made.
3. if a packet doesn’t match the condition on any of the lines in the access list, the packet will be discarded (implicit "deny").

IOS commands related to access lists

q	access-list	Creates a list of tests to filter the networks.
q		
q	host	Specifies a single host address.
q		
q	any	same as the 0.0.0.0 255.255.255.255.
q		
q	0.0.0.0 255.255.255.255	Specifies any host or any network
q		
q	ip access-group	Applies an IP access list to an interface.
q		
q	access-class	Applies a standard IP access list to a VTY line.
q		
q	show access-list	Shows all the access lists configured on the router.
q		
q	show access-list 110	Shows only access-list 110.
q		
q	show ip access-list	Shows only the IP access lists.
q		
q	show ip interface	Shows which interfaces have IP access lists applied.

Two main types of access lists

- Standard access lists
  - `Router(config)# access-list 10 deny 172.16.10.0 0.0.0.255`
  - `Router(config)# access-list 10 deny host 172.16.30.2`
  - `Router(config)# access-list 10 permit any`
  - `or Router(config)# access-list 10 permit 0.0.0.0 255.255.255.255`
  - `or Router(config)# access-list 10 permit 0.0.0.0 255.255.255.255 log`
  - `Router(config)# int s0/0`
  - `Router(config-if)# ip access-group 10 out`
- Extended access lists
  - `Router(config)#access-list 110 deny tcp any 172.16.50.0 0.0.0.255 eq 23`
  - `Router(config)#access-list 110 permit ip any any`
  - `Router(config)#interface s0/0`
  - `Router(config-if)#ip access-group 110 out`
- Named access lists are created and referred to using words vs numbers

Controlling VTY

```
Router(config)#access-list 50 permit host 172.16.10.3
Router(config)#line vty 0 4
Router(config-line)#access-class 50 in
```

Verifying access lists

- Router#show running-config
- Router#show access-list
- Router#show access-list 110
- Router#show ip access-list
- Router#show ip interface

- Labs 10.1-5 (load 8.11 configuration)