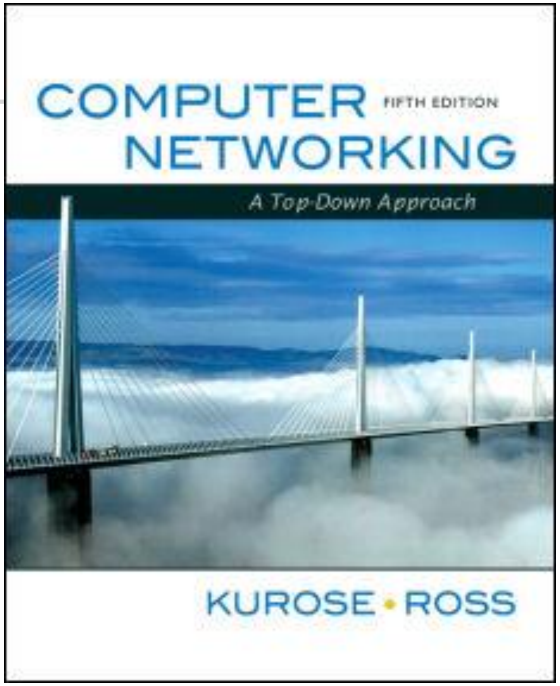
Chapter 4
Network Layer



A note on the use of these ppt slides:

We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part. In return for use, we only ask the following:

- If you use these slides (e.g., in a class) in substantially unaltered form, that you mention their source (after all, we'd like people to use our book!)
- If you post any slides in substantially unaltered form on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

Computer Networking: A Top Down Approach Featuring the Internet, 5th edition.
Jim Kurose, Keith Ross
Pearson Addison-Wesley, 2009.

Chapter 4: Network layer

Chapter goals:

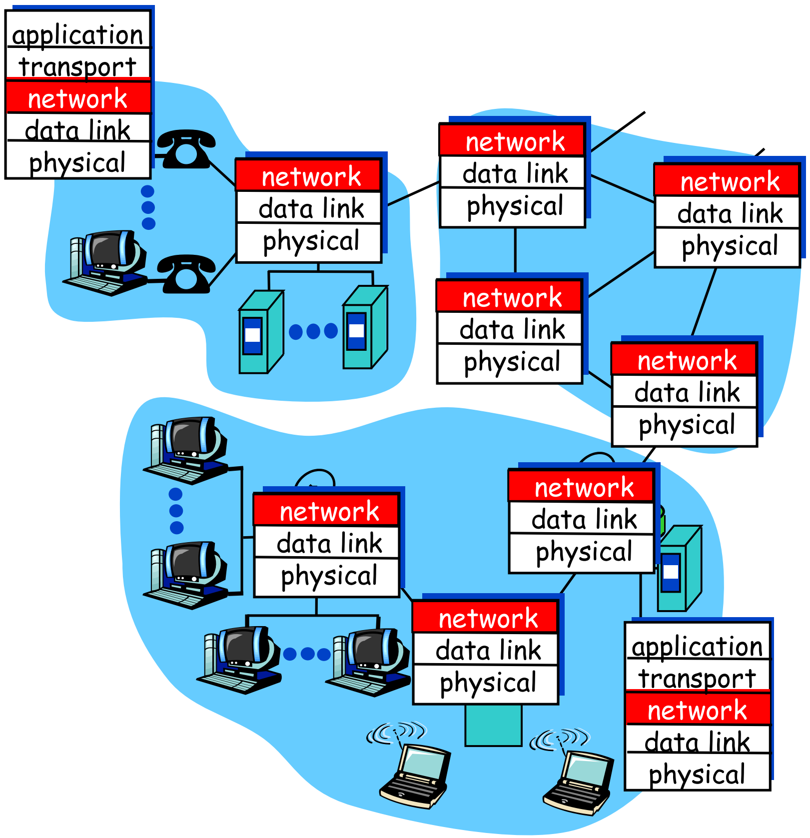
- Understand principles behind network layer services
  - Network layer service models
  - Forwarding versus routing
  - How a router works
  - Routing (path selection)
  - Dealing with scale
  - Advanced topics: IPv6, mobility
- Instantiation, implementation in the Internet

Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Network layer

- Transport segment from sending to receiving host
- On sending side encapsulates segments into datagrams
- On rcving side, delivers segments to transport layer
- Network layer protocols in *every* host, router
- Router examines header fields in all IP datagrams passing through it



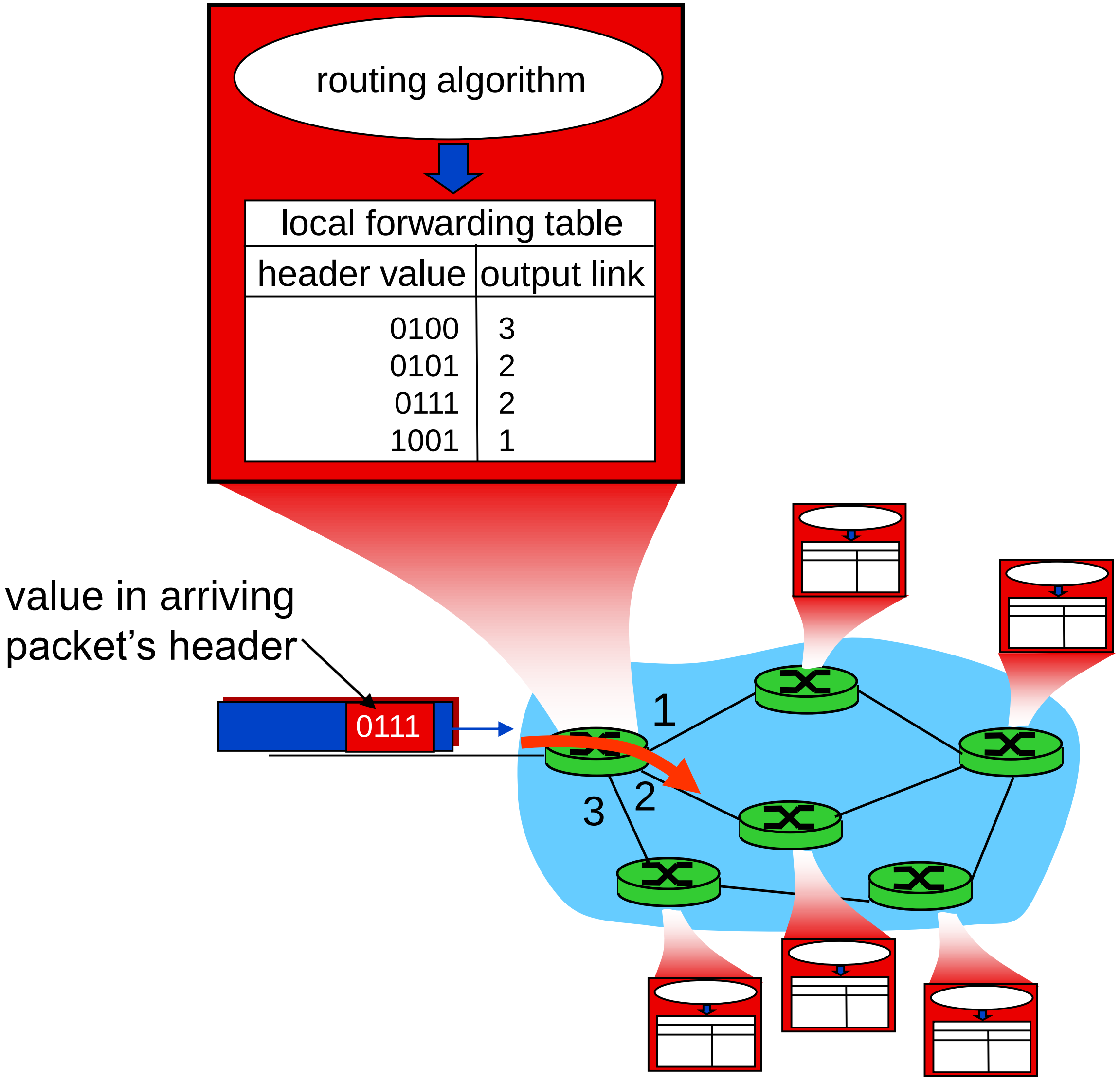
Two key network layer functions

- *Forwarding:* move packets from router's input to appropriate router output
- *Routing:* determine route taken by packets from source to destination
  - *Routing algorithms*

<u>Analogy:</u>

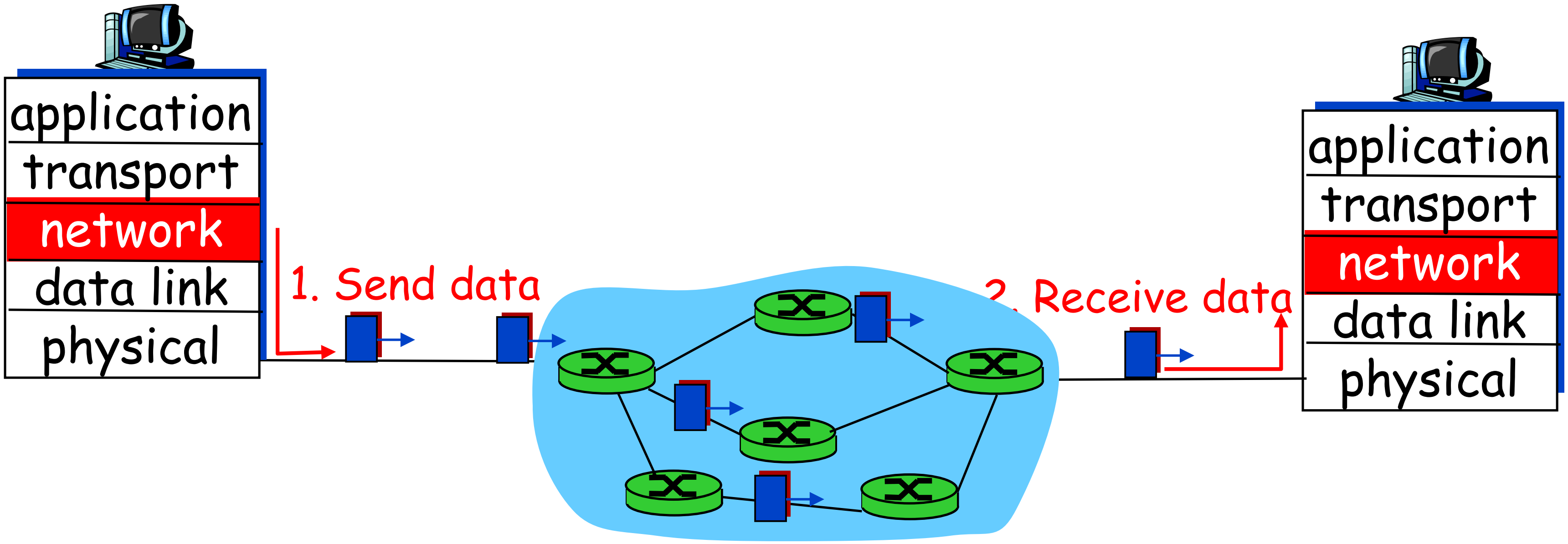
- Routing: process of planning trip from source to destination
- Forwarding: process of getting through single interchange

Interplay between routing and forwarding



Datagram networks

- No call setup at network layer
- Routers: no state about end-to-end connections
  - No network-level concept of "connection"
- Packets forwarded using destination host address
  - Packets between same source-dest pair may take different paths



Forwarding table

4 billion possible entries

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
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
otherwise	3

Longest prefix matching

Prefix Match	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?

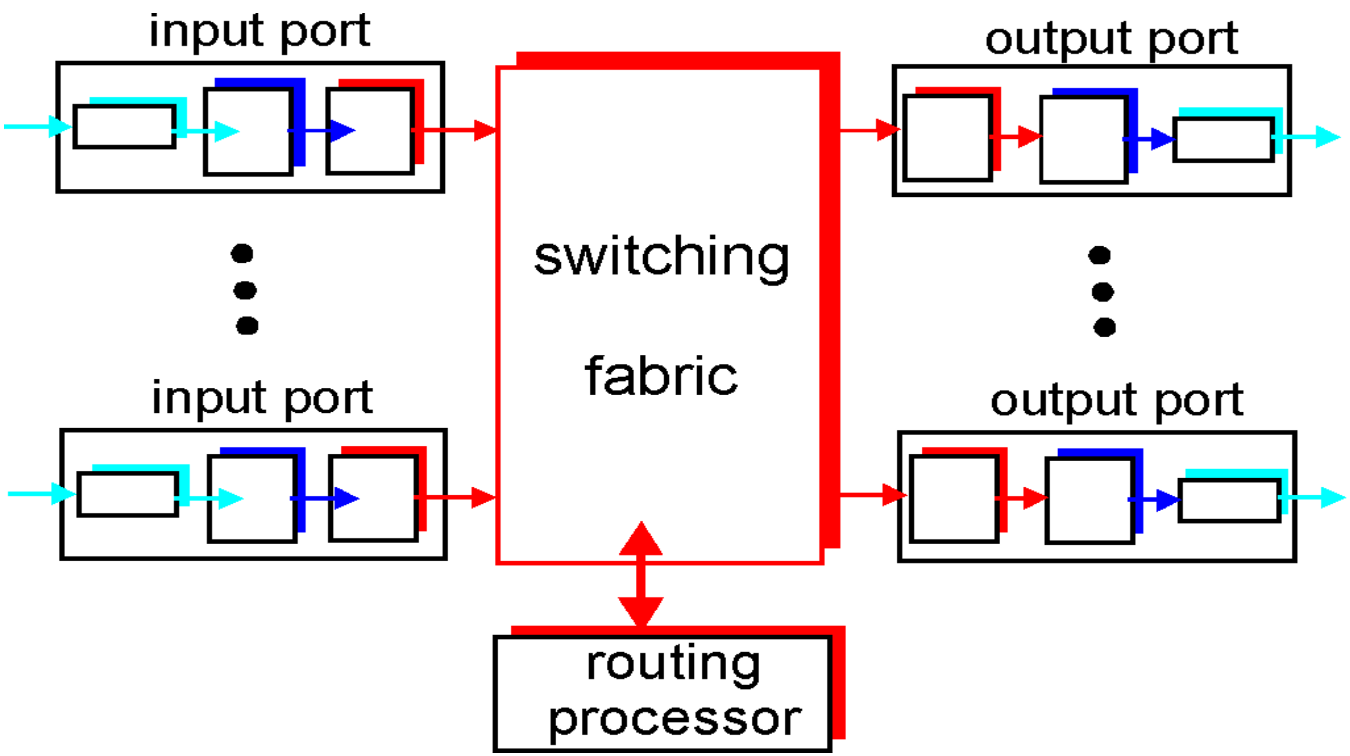
Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

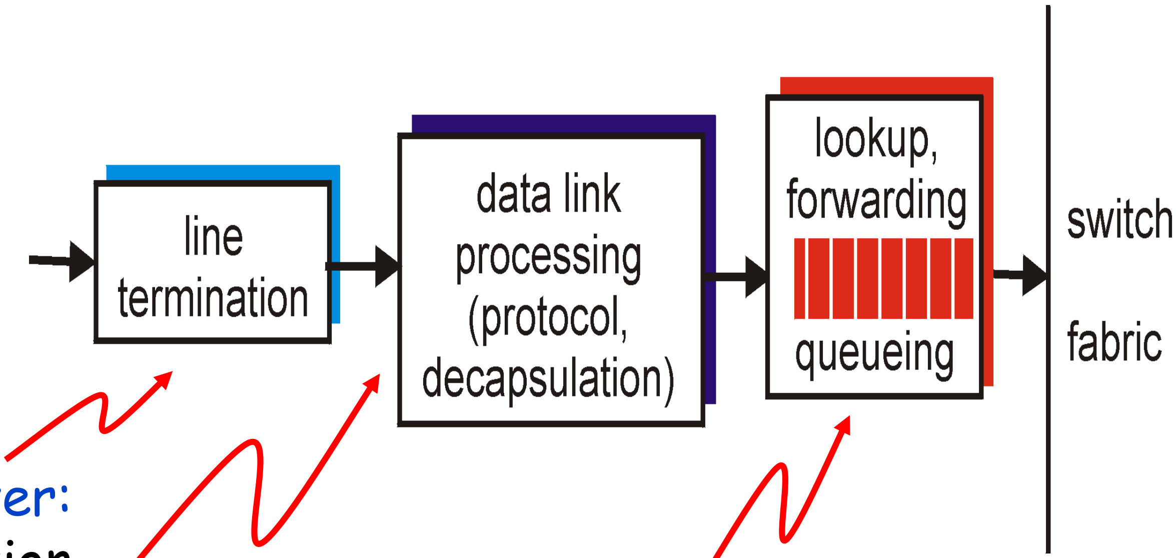
Router architecture overview

Two key router functions:

- Run routing algorithms/protocol (RIP, OSPF, BGP)
- *Forward* datagrams from incoming to outgoing link



Input port functions



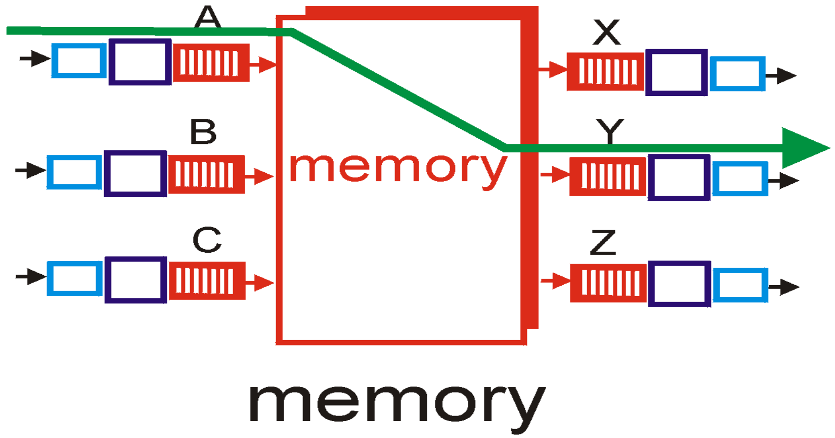
Physical layer:
bit-level reception

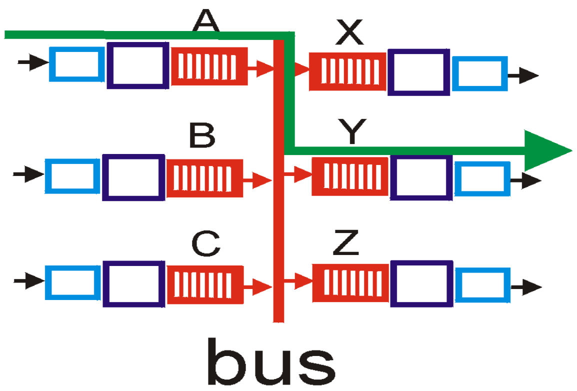
Data link layer:
e.g., Ethernet
see chapter 5

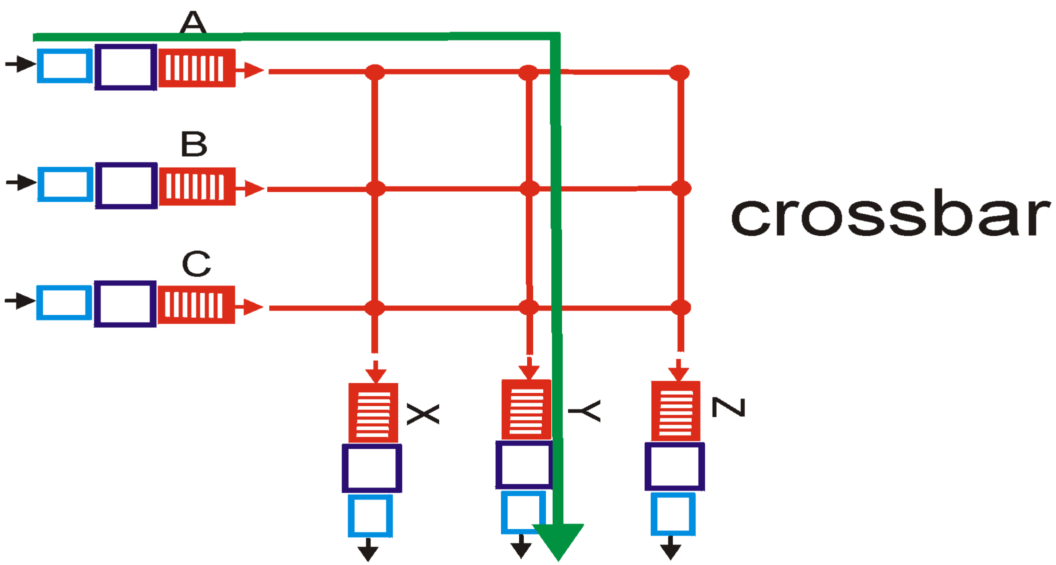
Decentralized switching:

- Given datagram dest., lookup output port using forwarding table in input port memory
- Goal: complete input port processing at ‘line speed’
- Queuing: if datagrams arrive faster than the forwarding rate into switch fabric

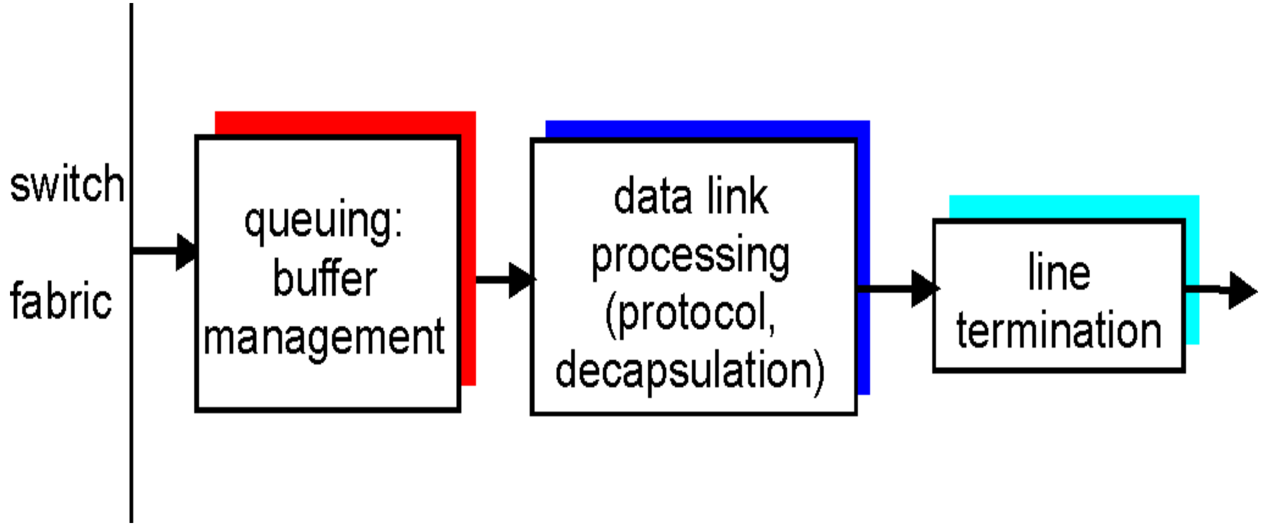
Three types of switching fabrics







Output ports



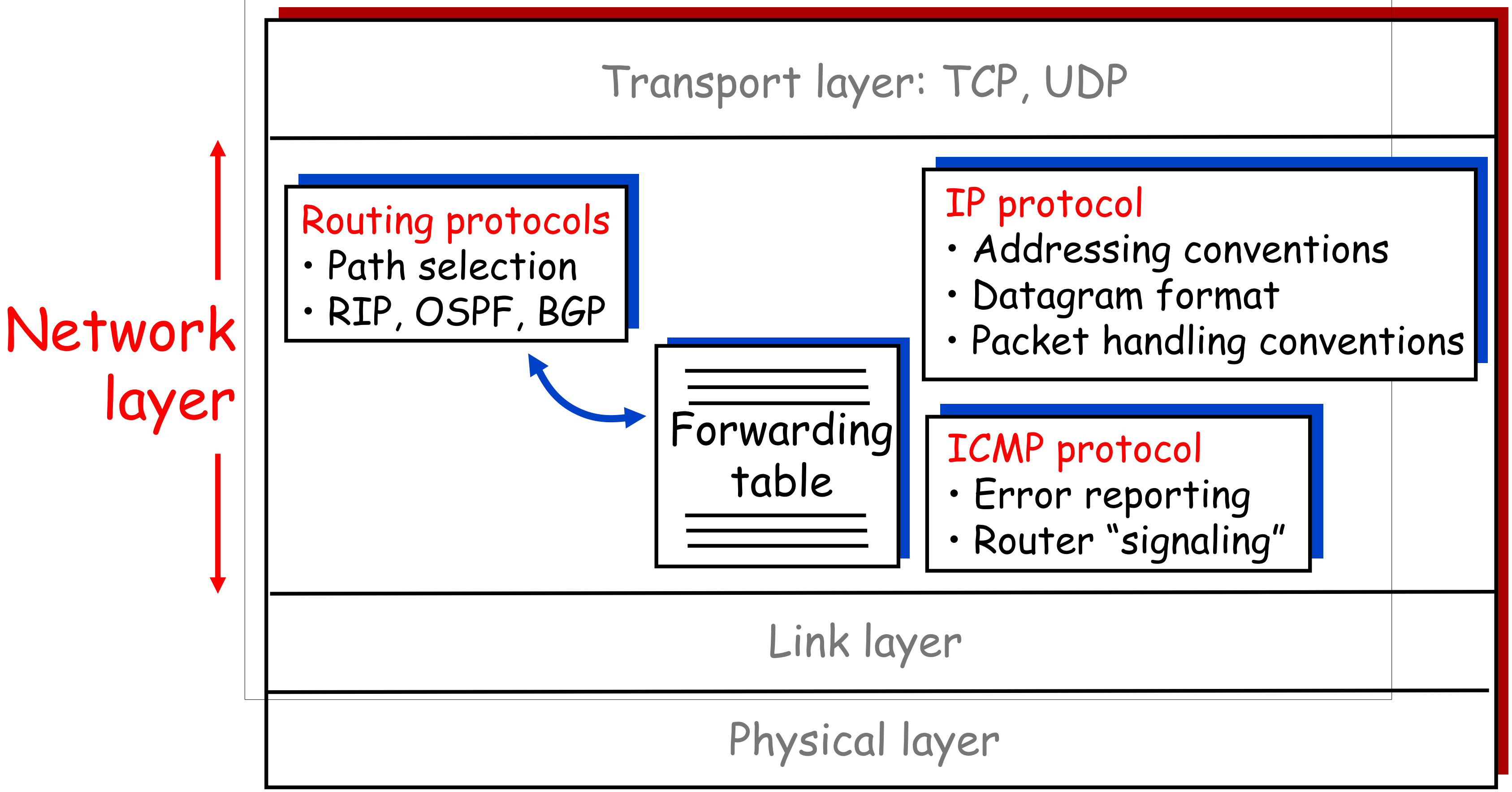
- *Buffering* required when datagrams arrive from fabric faster than the transmission rate
- *Scheduling discipline* chooses among queued datagrams for transmission

Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

The Internet network layer

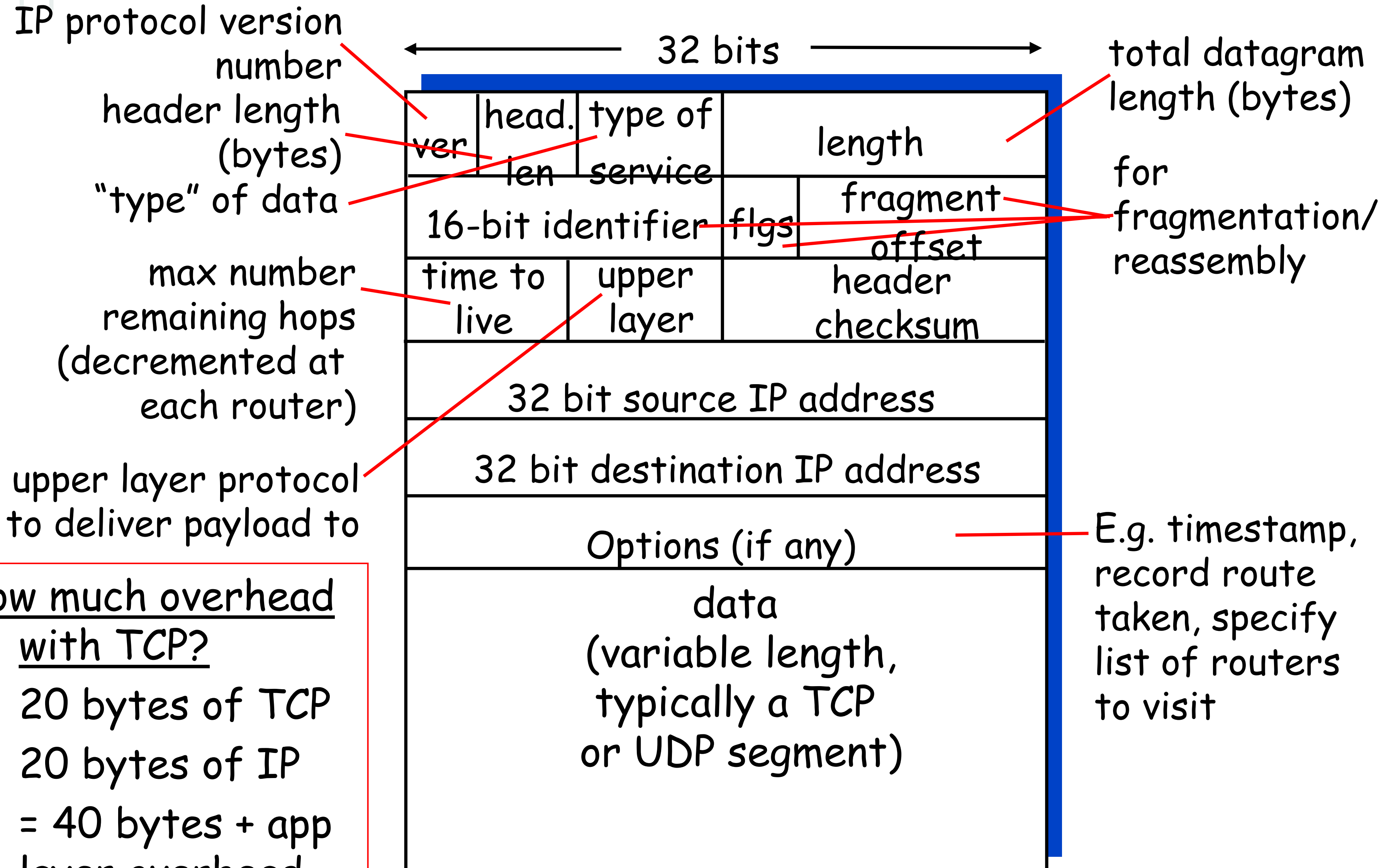
Host, router network layer functions



Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

IP datagram format



how much overhead with TCP?

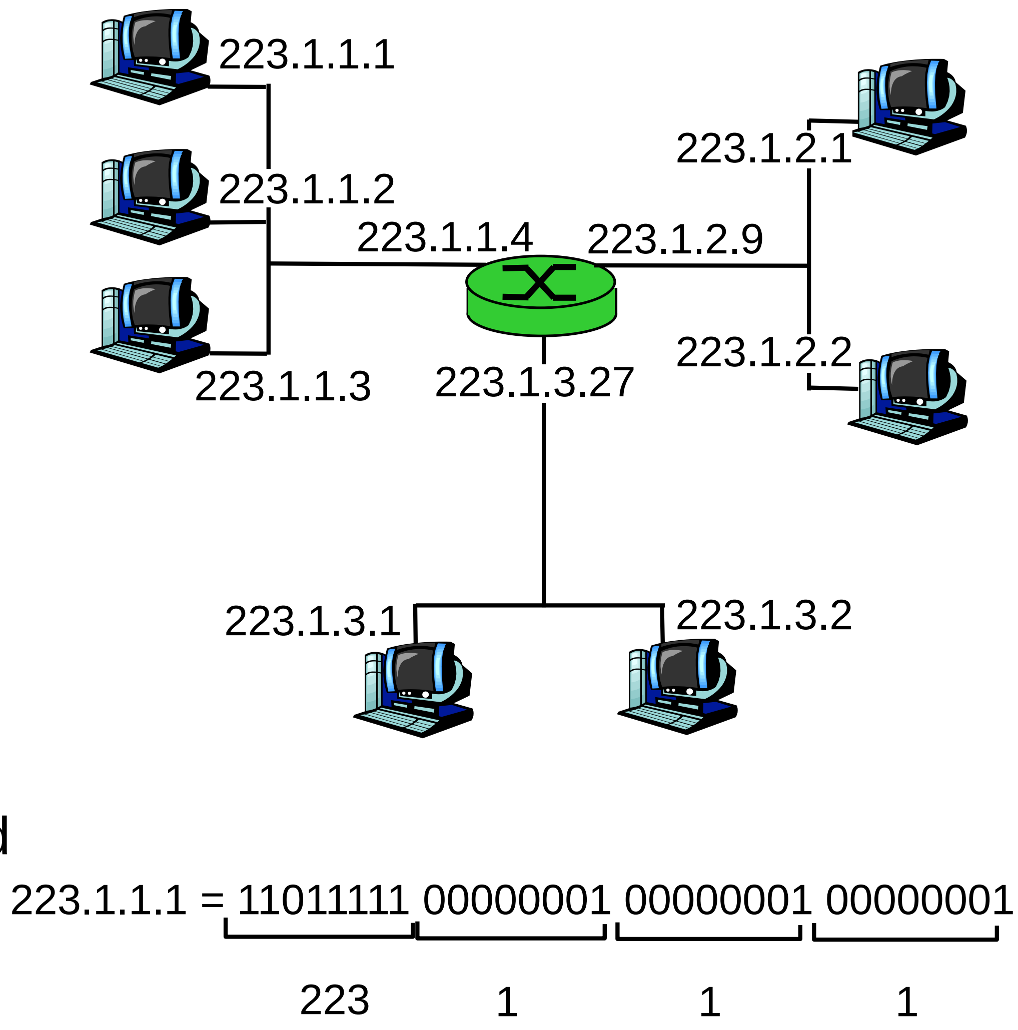
- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

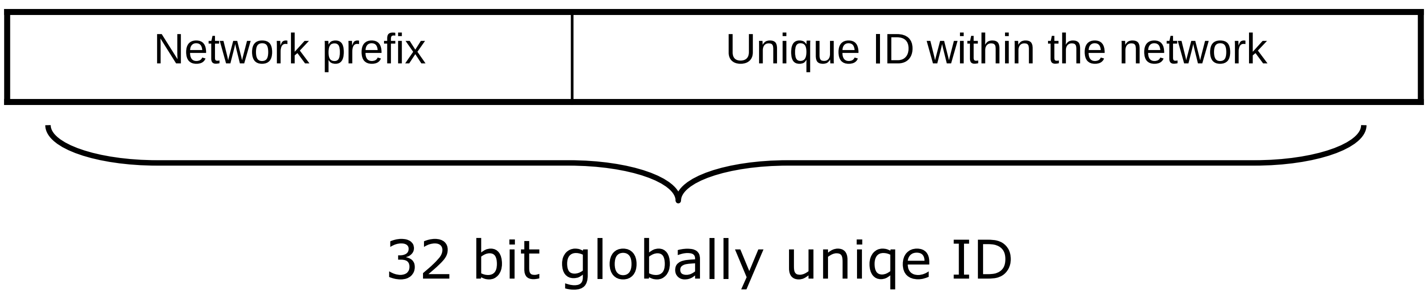
IP addressing: Introduction

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



The IPv4 address

- 4 byte length(32 bit)
  - $2^{32} \approx 4 \cdot 10^9 \Rightarrow$ running out!
- Notation
  - Binary: 10110000 10010011 00111110 11100001
  - „Dotted decimal”: 176.147.62.225
- Address = Network prefix + unique ID



IPv4 address classes

- At the beginning: 8 bit prefixen, then they introduced classes (1981)
- Different address ranges for different sized networks

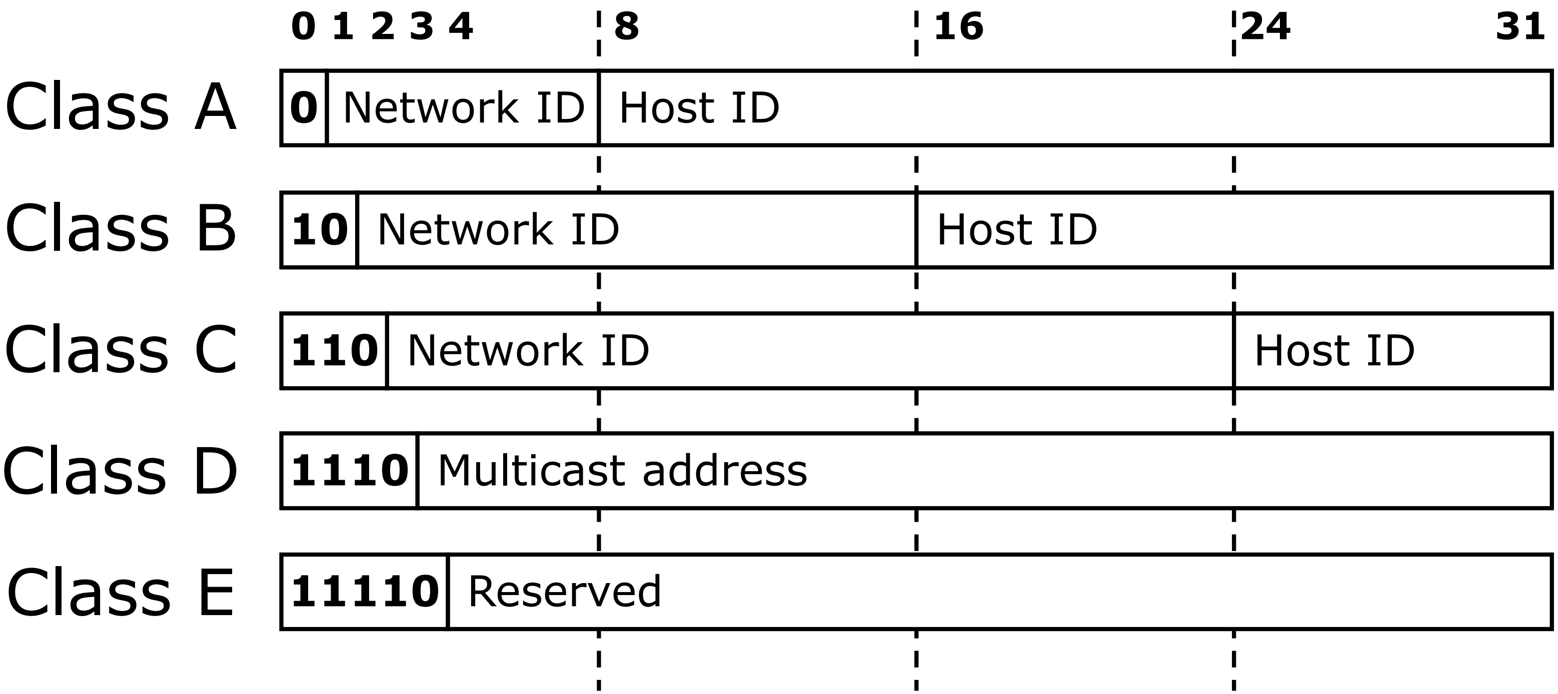
Class	Bits of prefixes	Number of networks	Number of hotsts within the networks
A	8	126	16777214
B	16	16382	65534
C	24	2097150	254

- Assignments of IP addresses (IANA: Internet Assigned Numbers Authority)
- 5 Regional Internet Registry (RIR)
  - A and B classes typically for
    - Countries
    - ISP: Internet Service Provider
    - Universities

IPv4 address classes (cont'd)

- A, B and C classes
  - For unique addressing (unicast)
- Class D
  - multicasting
  - Reserved range:
    from 224.0.0.0 to 239.255.255.255
- Class E
  - Reserved class and range
    from 240.0.0.0 to 255.255.255.255

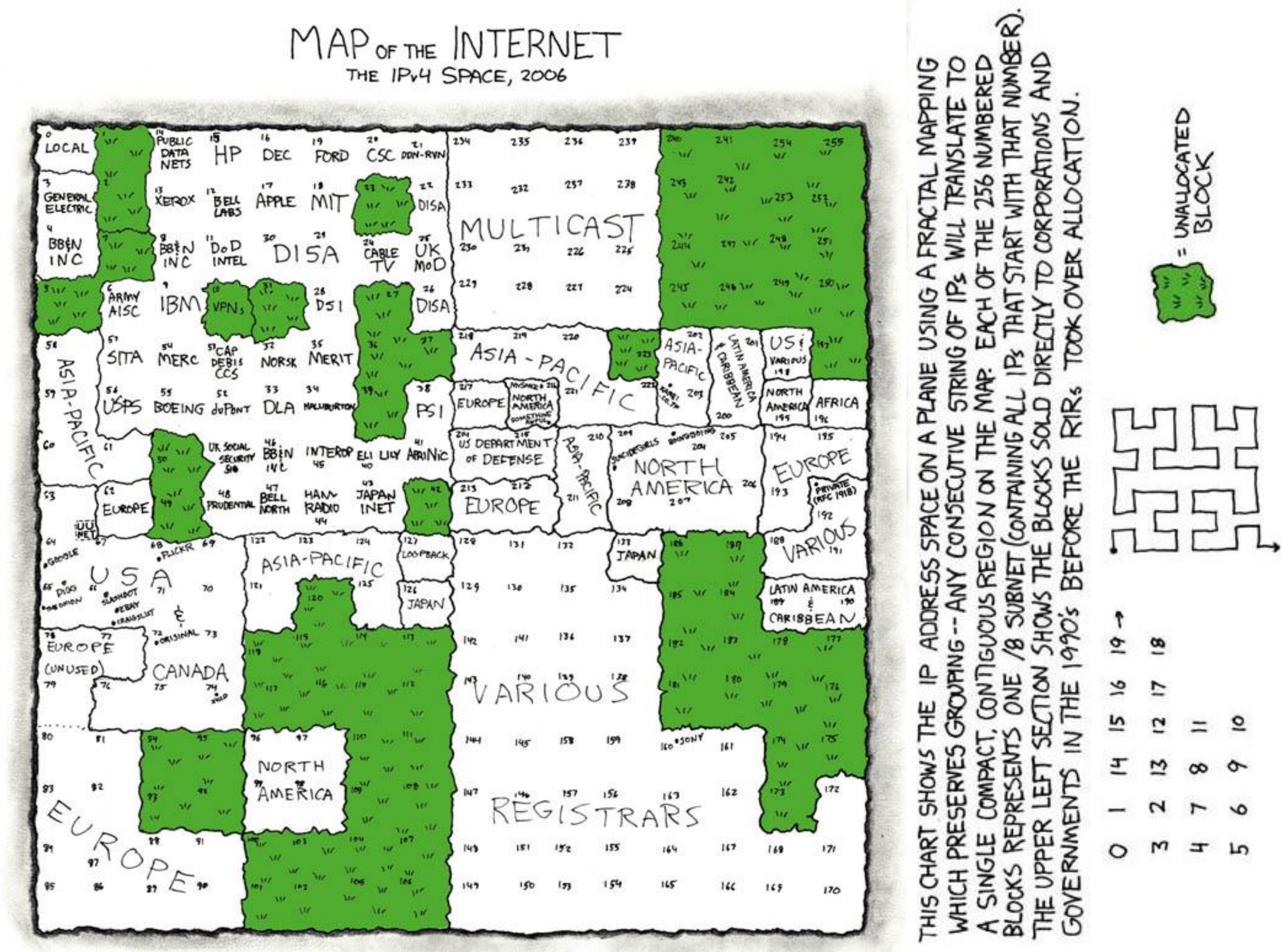
IPv4 address classes (cont'd)



- Class identifier prefix
- Disjunct address ranges

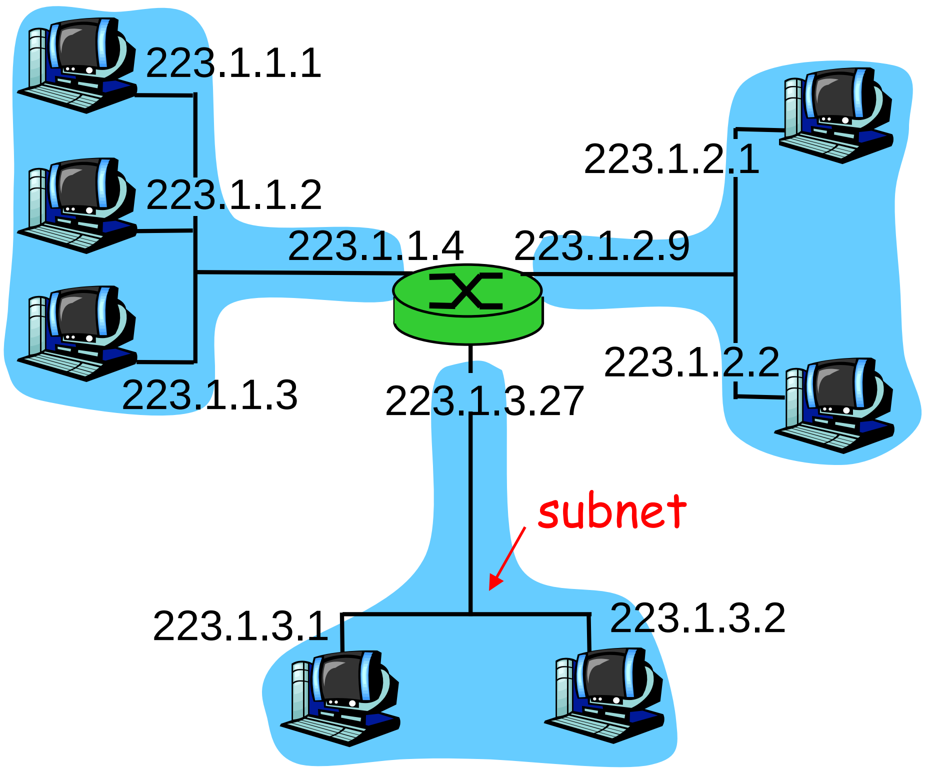
Problems with the IPv4 address classes

- Several big companies, universities received class A addresses (16 million addresses)
  - Class A too big
  - Class B too small
  - Lack of flexibility!!!
- The same with SME-s:
  - Class *C* too small
  - Class *B* too big
- Problems with routing aggregation
- The class-based system is not scalable!
  - CIDR
  - IPv6


MAP OF THE INTERNET
THE IPv4 SPACE, 2006
LOCAL
PUBLIC DATA NETS
HP
DEC
FORD
CSC
DDN-RVN
GENERAL ELECTRIC
XEROX
BELL LABS
APPLE
MIT
DISA
BBN INC
BBN INC
DoD INTEL
DISA
CABLE TV
UK MoD
ARMY AISC
IBM
VPNs
DSI
DISA
SITA
MERC
CAP DEBIS CCS
NORSK
MERIT
USPS
BOEING
duPONT
DLA
HALLIBURTON
PSI
ASIA-PACIFIC
UK SOCIAL SECURITY
BBN INC
INTEROP
ELI LILY
AfriNIC
EUROPE
PRUDENTIAL
BELL NORTH
HAM RADIO
JAPAN INET
UU NET
GOOGLE
FLICKR
USA
DIGG
SLASHDOT
EBAY
CRAIGSLIST
ONION
&
ORIGINAL
EUROPE (UNUSED)
CANADA
ASIA-PACIFIC
LOOPBACK
JAPAN
NORTH AMERICA
EUROPE
MULTICAST
ASIA-PACIFIC
EUROPE
NORTH AMERICA
MYSPACE
SOMETHING AWFUL
US DEPARTMENT OF DEFENSE
EUROPE
ASIA-PACIFIC
NORTH AMERICA
SUICIDEGIRLS
BOINGBOING
ASIA-PACIFIC
KOREA TELECOM
LATIN AMERICA & CARIBBEAN
US & VARIOUS
NORTH AMERICA
AFRICA
EUROPE
PRIVATE (RFC 1918)
JAPAN
VARIOUS
LATIN AMERICA & CARIBBEAN
VARIOUS
SONY
REGISTRARS
THIS CHART SHOWS THE IP ADDRESS SPACE ON A PLANE USING A FRACTAL MAPPING WHICH PRESERVES GROUPING -- ANY CONSECUTIVE STRING OF IPs WILL TRANSLATE TO A SINGLE COMPACT, CONTIGUOUS REGION ON THE MAP. EACH OF THE 256 NUMBERED BLOCKS REPRESENTS ONE /8 SUBNET (CONTAINING ALL IPs THAT START WITH THAT NUMBER). THE UPPER LEFT SECTION SHOWS THE BLOCKS SOLD DIRECTLY TO CORPORATIONS AND GOVERNMENTS IN THE 1990's BEFORE THE RIRs TOOK OVER ALLOCATION.
= UNALLOCATED BLOCK

Subnets

- IP address
  - Subnet part (high order bits)
  - Host part (low order bits)
- *What's a subnet?*
  - Device interfaces with same subnet part of IP address
  - Can physically reach each other without intervening router

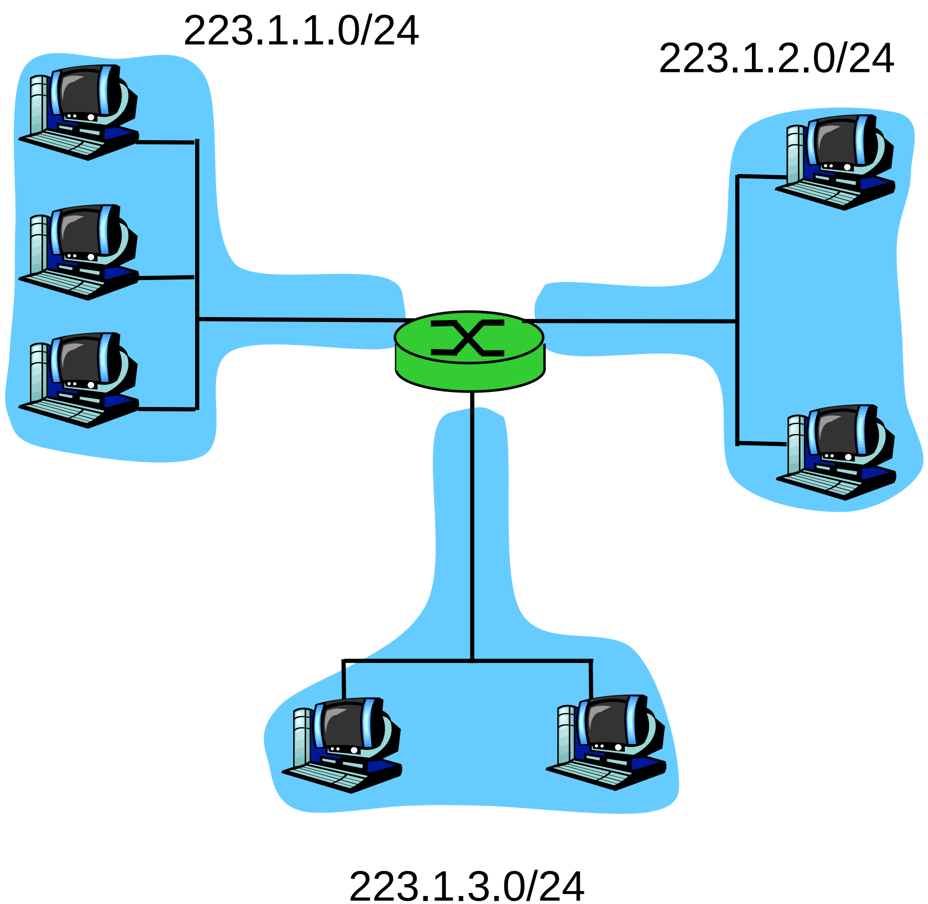


Network consisting of 3 subnets

Subnets

Recipe

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

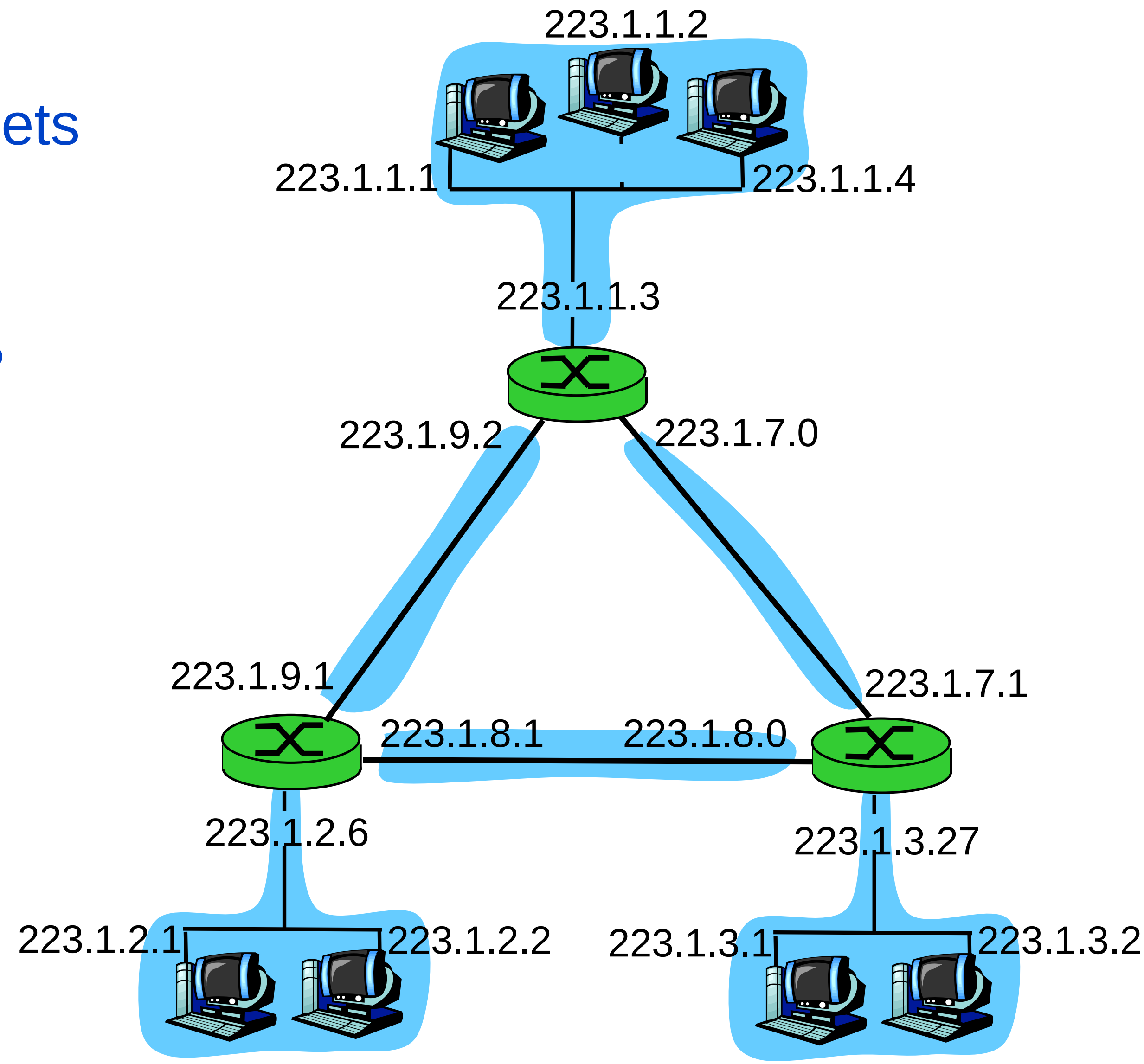


Subnet mask: /24

Subnets

How many subnets do we have?

Which are they?



IP addressing: CIDR

CIDR: Classless InterDomain Routing

- Subnet portion of address of arbitrary length
- Address format: a.b.c.d/x, where x is # bits in subnet portion of address

subnet part | host part

11001000 00010111 00010000 00000000

200.23.16.0/23

IP addresses: How to get one?

Q: How does *host* get IP address?

- Hard-coded by system admin in a file
  - Windows: control-panel->network->configuration->tcp/ip->properties
  - UNIX: /etc/rc.config
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from a server (more in next chapter)
  - “plug-and-play”

IP addresses: How to get one?

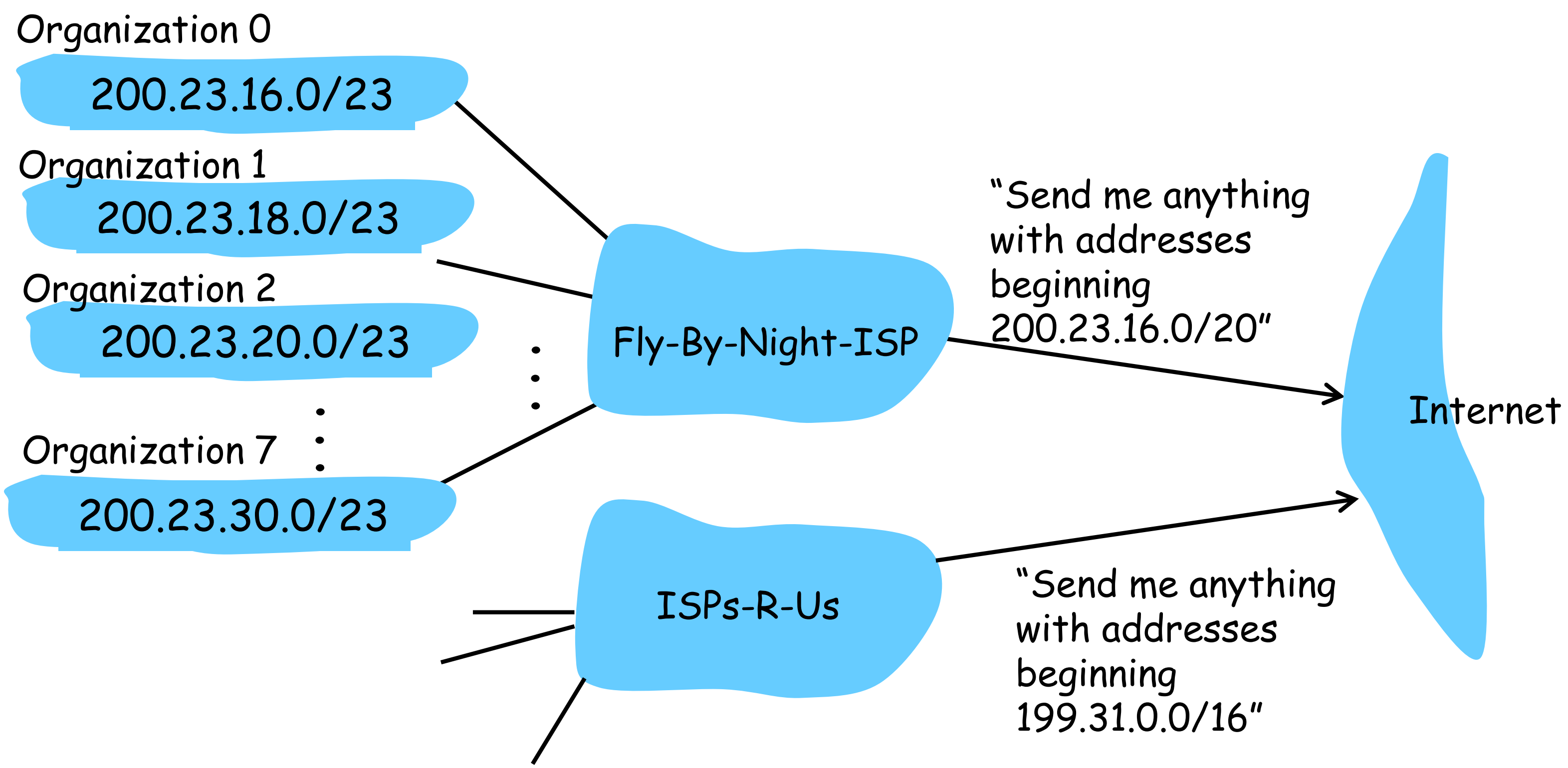
Q: How does *network* get subnet part of IP addr?

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

Hierarchical addressing allows efficient advertisement of routing information



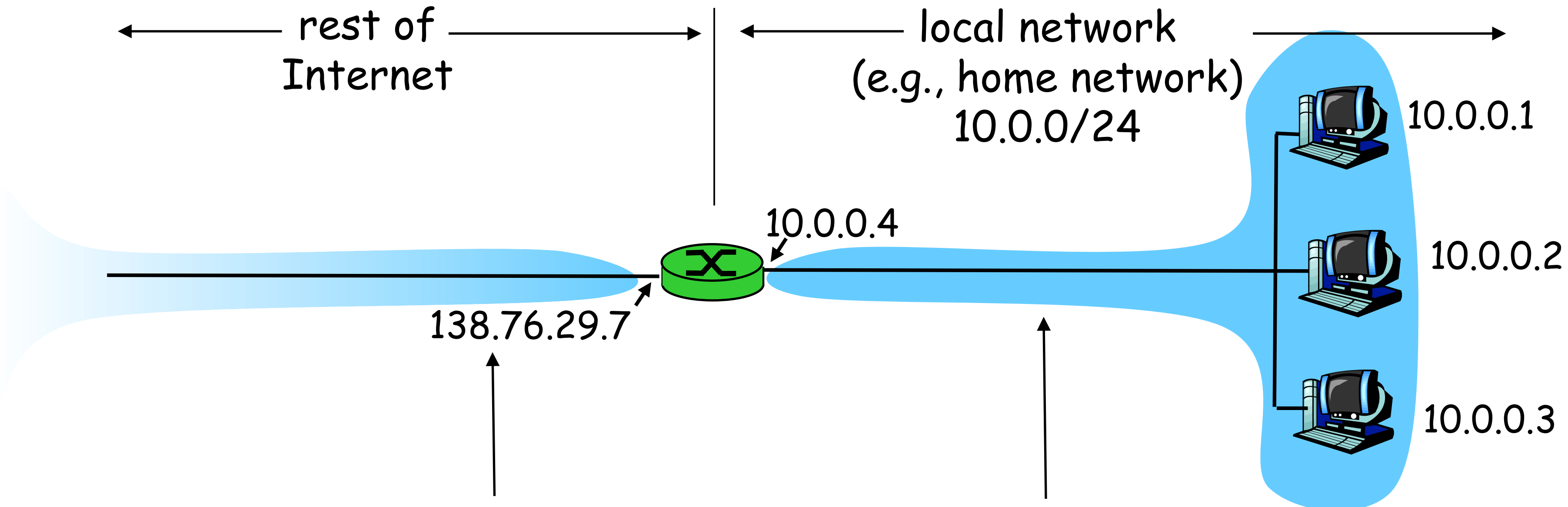
IP addressing: The last word...

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

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: 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

Chapter 4 outline

- 4.1 Introduction
- 4.2 What's inside a router
- 4.3 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.4 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.5 Routing in the Internet
  - RIP
  - OSPF
  - BGP

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

Traceroute and ICMP

- Source sends series of UDP segments to dest
  - First has TTL =1
  - Second has TTL=2, etc.
  - Unlikely port number
- When n^{th} datagram arrives to n^{th} router
  - Router discards datagram
  - And sends to source an ICMP message (type 11, code 0)
  - Message includes name of router & IP address
- When ICMP message arrives, source calculates RTT
- Traceroute does this 3 times

Stopping criterion

- UDP segment eventually arrives at destination host
- Destination returns ICMP "host unreachable" packet (type 3, code 3)
- When source gets this ICMP, stops

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

IPv6

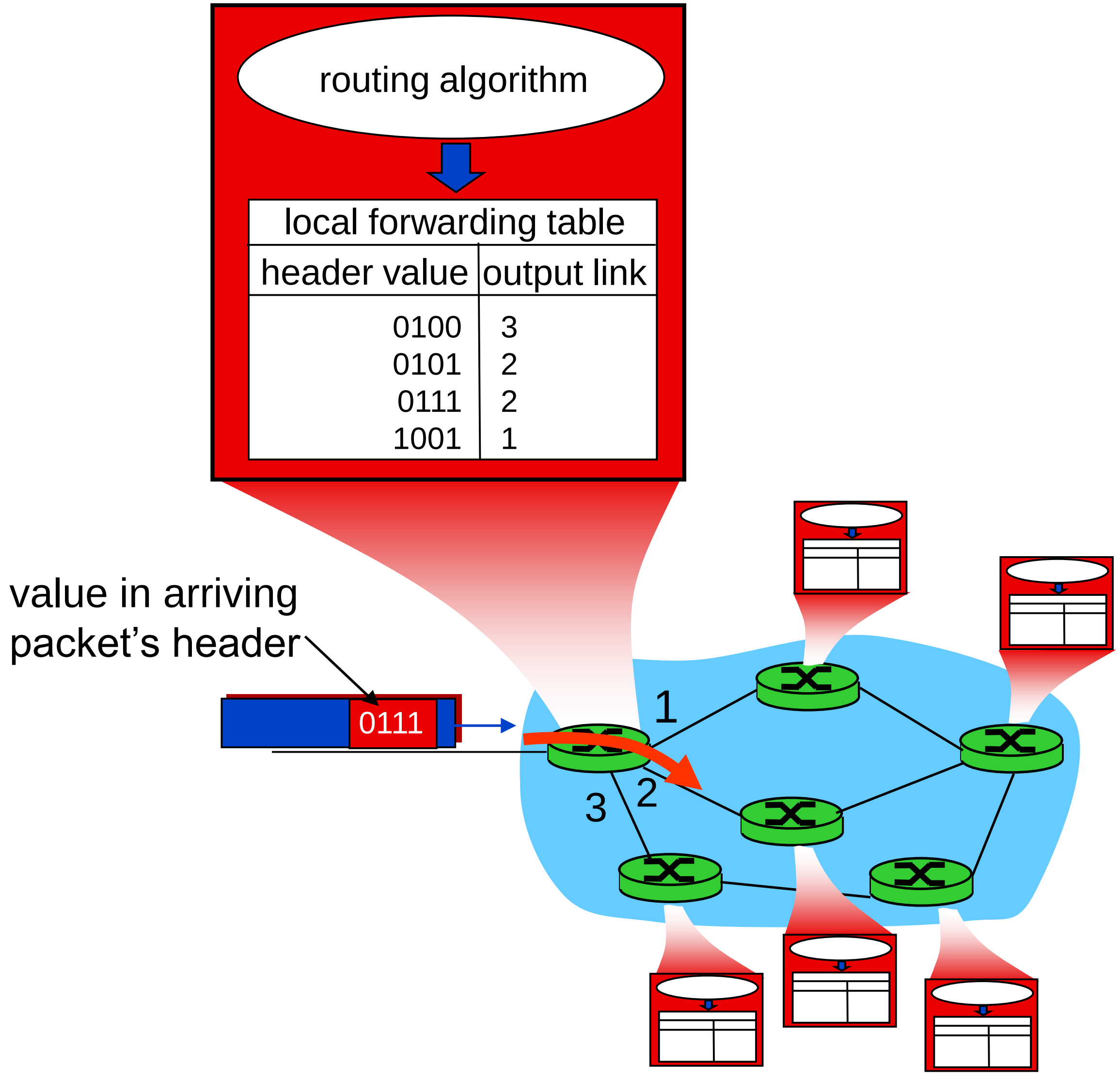
- Initial motivation: 32-bit address space soon to be completely allocated
- Additional motivation
  - Header format helps speed processing/forwarding
  - Header changes to facilitate QoS

  IPv6 datagram format
  - Fixed-length 40 byte header
  - No fragmentation allowed

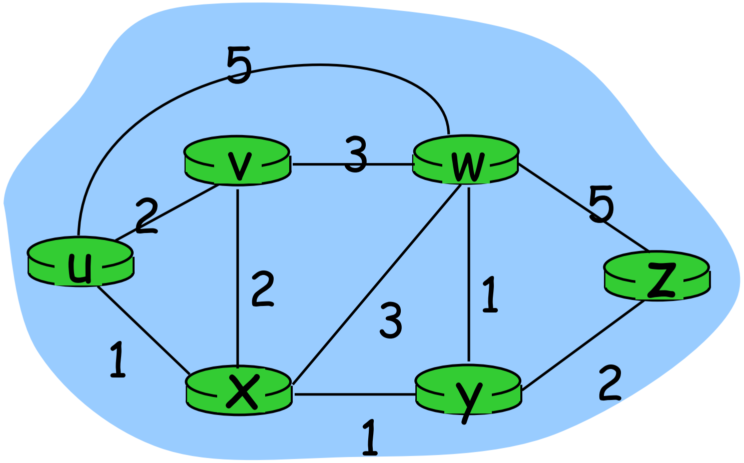
Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Interplay between routing and forwarding



Graph abstraction



Graph: G = (N,E)

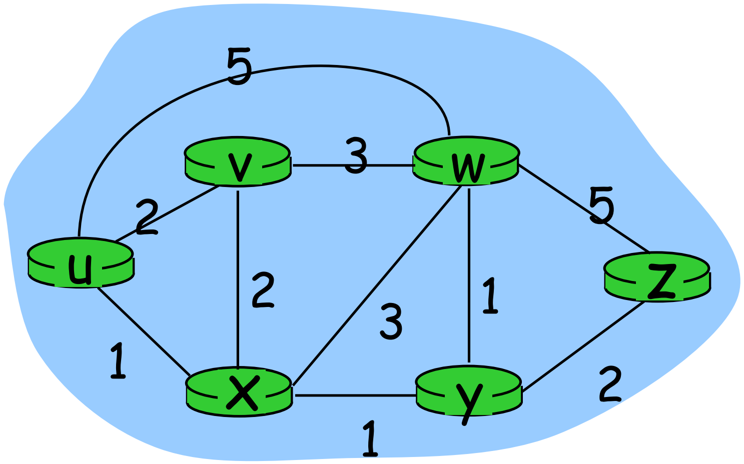
N = set of routers = { u, v, w, x, y, z }

E = set of links ={ (u,v), (u,x), (v,x), (v,w), (x,w), (x,y), (w,y), (w,z), (y,z) }

Remark: Graph abstraction is useful in other network contexts

Example: P2P, where N is set of peers and E is set of TCP connections

Graph abstraction: Costs



- $c(x,x')$ = cost of link (x,x')
  - e.g., $c(w,z) = 5$
- cost could always be 1, or inversely related to bandwidth, or inversely related to congestion

Cost of path $(x_1, x_2, x_3, \ldots, x_p) = c(x_1,x_2) + c(x_2,x_3) + \ldots + c(x_{p-1},x_p)$

Question: What's the least-cost path between u and z ?

Routing algorithm: Algorithm that finds least-cost path

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

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

A link state routing algorithm

Dijkstra's algorithm

- Net topology, link costs known to all nodes
  - accomplished via "link state broadcast"
  - all nodes have same info
- Computes least cost paths from one node ('source') to all other nodes
  - gives forwarding table for that node
- Iterative: after k iterations, know least cost path to k dest's

Notation

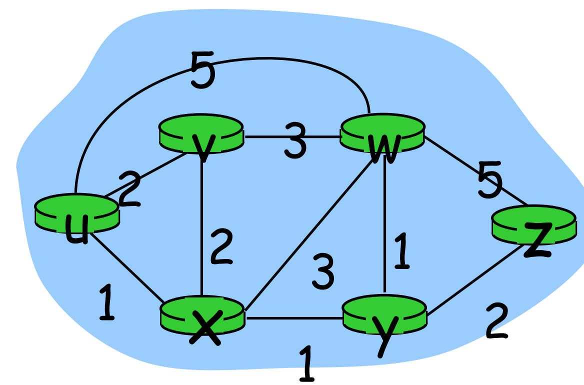
- c(x,y): link cost from node x to y; = ∞ if not direct neighbors
- D(v): current value of cost of path from source to dest. v
- p(v): predecessor node along path from source to v
- N': set of nodes whose least cost path definitively known

Dijkstra's algorithm

- c(x,y): link cost from node x to y; = ∞ if not direct neighbors
- D(v): current value of cost of path from source to dest. v
- p(v): predecessor node along path from source to v
- N': set of nodes whose least cost path definitively known

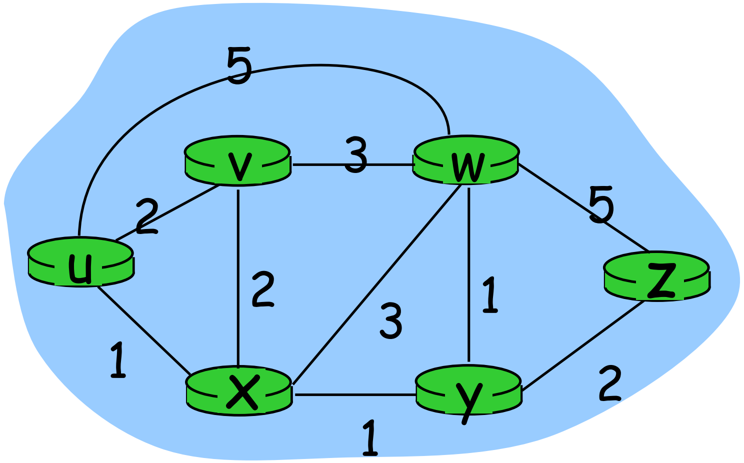
```
Initialization
  N' = {u}
  for all nodes v
    if v adjacent to u
        then D(v) = c(u,v)
    else D(v) = ∞

 Loop
   find w not in N' such that D(w) is a minimum
   add w to N'
   update D(v) for all v adjacent to w and not in N':
      D(v) = min( D(v), D(w) + c(w,v) )
   /* new cost to v is either old cost to v or known
    shortest path cost to w plus cost from w to v */
 until all nodes in N'
```



Dijkstra's algorithm: Example

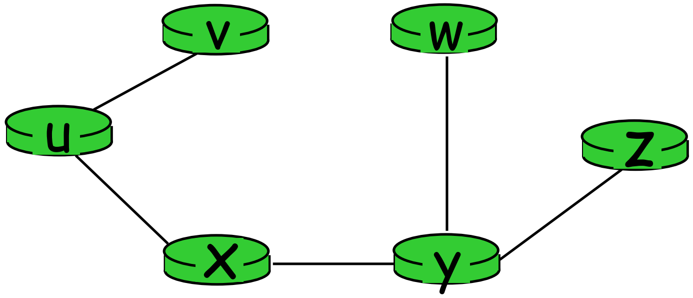
Step	N'	D(v),p(v)	D(w),p(w)	D(x),p(x)	D(y),p(y)	D(z),p(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxyvw					4,y
5	uxyvwz					



- c(x,y): link cost from node x to y; = ∞ if not direct neighbors
- D(v): current value of cost of path from source to dest. v
- p(v): predecessor node along path from source to v
- N': set of nodes whose least cost path definitively known

Dijkstra's algorithm: Example (2)

Resulting shortest-path tree from u:



Resulting forwarding table in u:

Destination	Link
v	(u,v)
x	(u,x)
y	(u,x)
w	(u,x)
z	(u,x)

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Distance vector algorithm

Bellman-Ford Equation (dynamic programming)

Define

$d_x(y)$:= cost of least-cost path from x to y

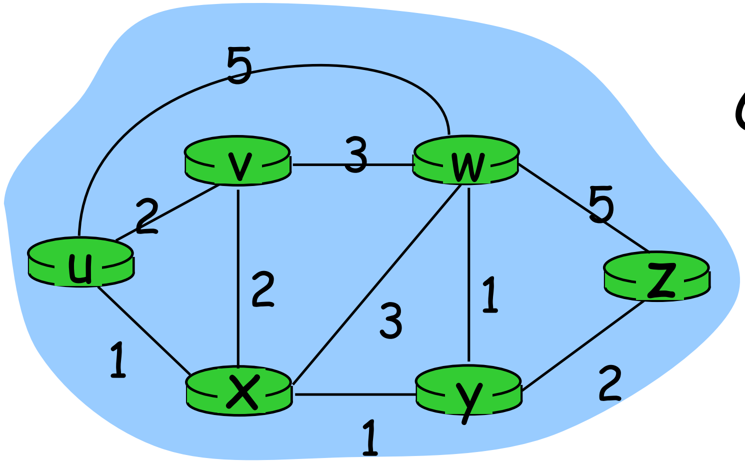
Then

$$d_x(y) = \min_v \{c(x,v) + d_v(y)\}$$

where min is taken over all neighbors v of x

- c(x,y): link cost from node x to y; = ∞ if not direct neighbors
- D(v): current value of cost of path from source to dest. v
- p(v): predecessor node along path from source to v
- N': set of nodes whose least cost path definitively known

Bellman-Ford example



Clearly, $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$

B-F equation says:

$$d_u(z) = \min \{ c(u,v) + d_v(z),\ c(u,x) + d_x(z),\ c(u,w) + d_w(z) \}$$

$$= \min \{2 + 5,\ 1 + 3,\ 5 + 3\} = 4$$

Node that achieves minimum is next hop in shortest path ➜ forwarding table

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Hierarchical routing

Our routing study thus far - idealization

- all routers identical
- network "flat"

... *not* true in practice

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

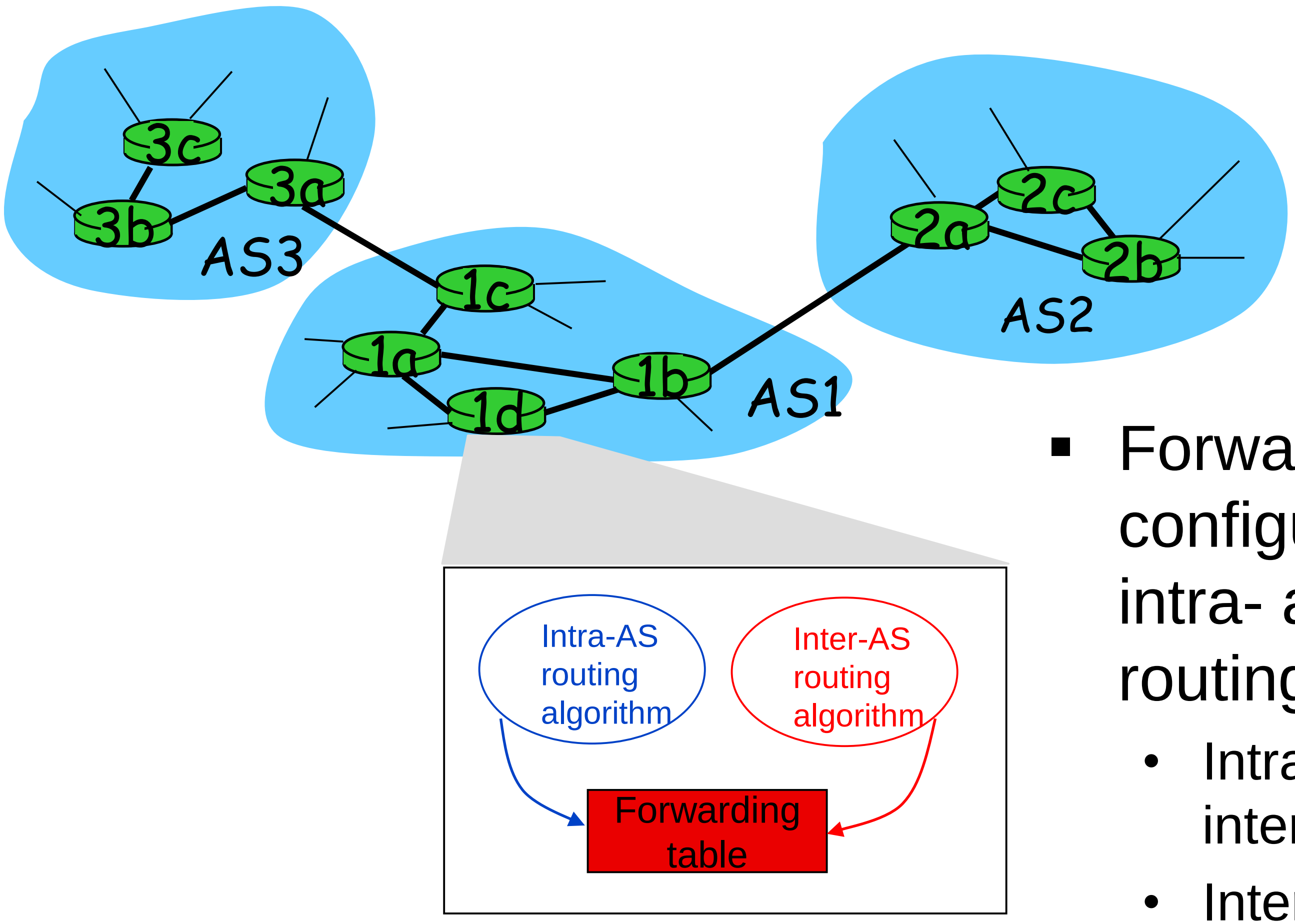
Hierarchical routing

- Aggregate routers into regions, "autonomous systems" (AS)
- Routers in same AS run same routing protocol
  - "Intra-AS" routing protocol
  - Routers in different AS can run different intra-AS routing protocol

Gateway router

- Direct link to router in another AS

Interconnected ASes



- Forwarding table is configured by both intra- and inter-AS routing algorithm
  - Intra-AS sets entries for internal dests
  - Inter-AS & Intra-AS sets entries for external dests

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Intra-AS routing

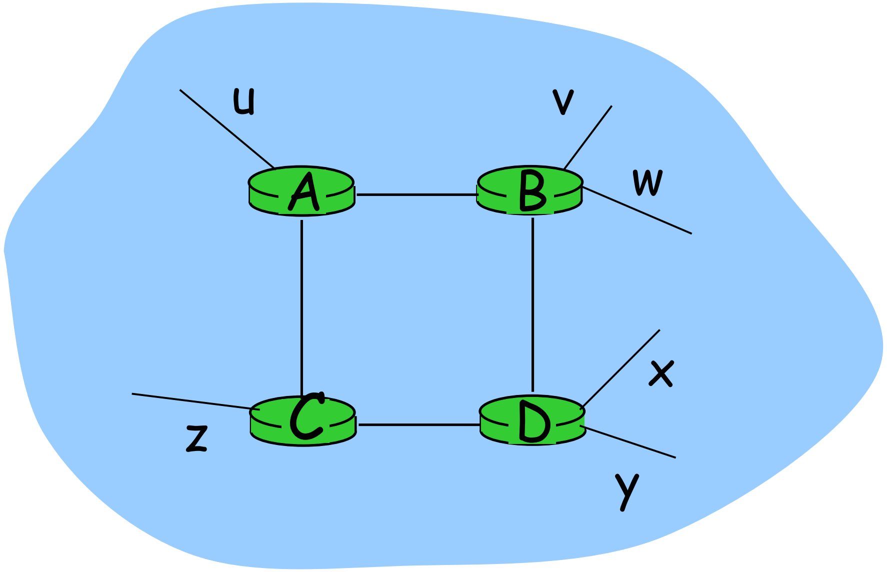
- Also known as Interior Gateway Protocols (IGP)
- Most common Intra-AS routing protocols
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First
  - IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

RIP – Routing Information Protocol

- Distance vector algorithm
- Included in BSD-UNIX Distribution in 1982
- Distance metric: # of hops (max = 15 hops)



From router A to subsets:

destination	hops
u	1
v	2
w	2
x	3
y	3
z	2

RIP advertisements

- 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

Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

OSPF – Open Shortest Path First

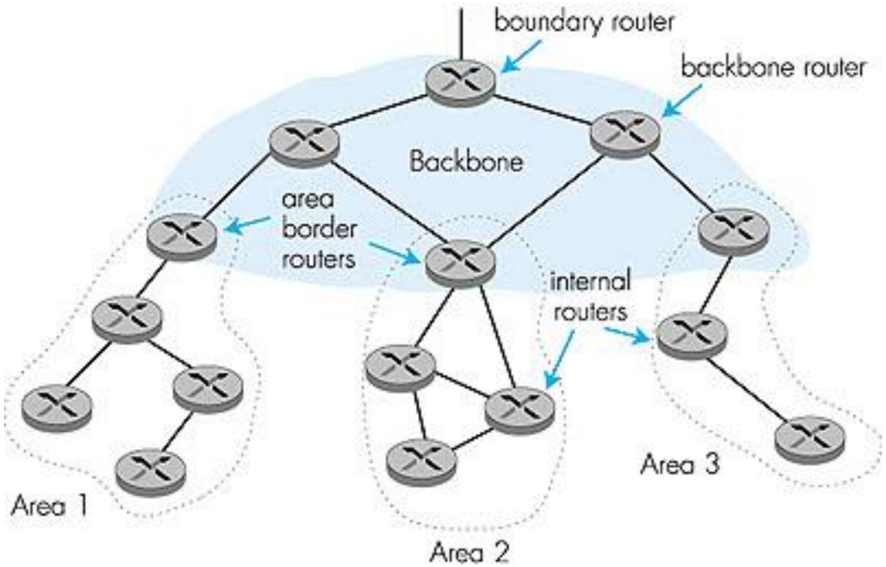
- "Open": publicly available
- Uses Link State algorithm
  - LS packet dissemination
  - Topology map at each node
  - Route computation using Dijkstra's algorithm
- OSPF advertisement carries one entry per neighbor router
- Advertisements disseminated to entire AS (via flooding)
  - Carried in OSPF 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

Hierarchical OSPF

- Two-level hierarchy: local area, backbone
  - Link-state advertisements only in the area
  - Each nodes has detailed area topology; only know direction (shortest path) to nets in other areas
- Area border routers
  - “Summarize” distances to nets in own area, advertise to other Area Border routers
- Backbone routers
  - Run OSPF routing limited to backbone
- Boundary routers
  - Connect to other AS’s



Chapter 4 outline

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP

Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto standard
- BGP provides each AS a means to
  1. Obtain subnet reachability information from neighboring ASes
  2. Propagate reachability information to all AS-internal routers
  3. Determine “good” routes to subnets based on reachability information and policy
- Allows subnet to advertise its existence to rest of Internet: *“I am here”*

Chapter 4: Summary

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP