The Network Layer:
IP Addressing & Routing

Smith College, CSC 249
March 25, 2008

slides mostly from J.F Kurose and K.W. Ross,
copyright 1996-2007

First half of semester recap

Concept:

- A network provides the means for applications *not* directly connected to exchange messages
- Layered model
  - Computer networking is implemented in discrete, well-defined sub-tasks

Goals:

- Understand for each layer
  - Services required from layer below
  - Services offered to layer above
  - Key principles of each layer

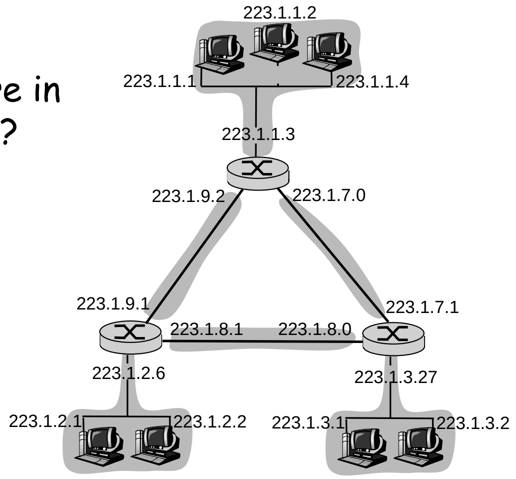
Chapter 4: Network Layer

Chapter goals:

- Understand principles behind network layer services:
  - network layer service models
  - forwarding versus routing
  - how a router works
  - routing (path selection)
- Implementation in the Internet

Subnets

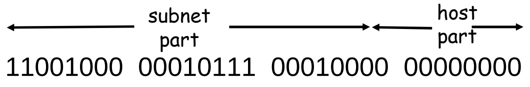
How many are in this figure?



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

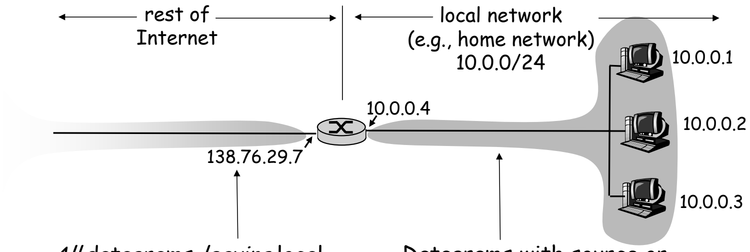


200.23.16.0/23

Chapter 4: Network Layer

- ☐ 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
- ☐ 4.7 Broadcast and multicast routing

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 view:
  - A range of addresses is not needed from the ISP: just one IP address for all devices
  - NAT can change addresses of devices in the local network without notifying outside world
  - You/your NAT can change ISPs without changing addresses of devices in the local network
  - Devices inside the local net are not explicitly addressable, visible by outside world (a security plus)

NAT Question #15 on Handout

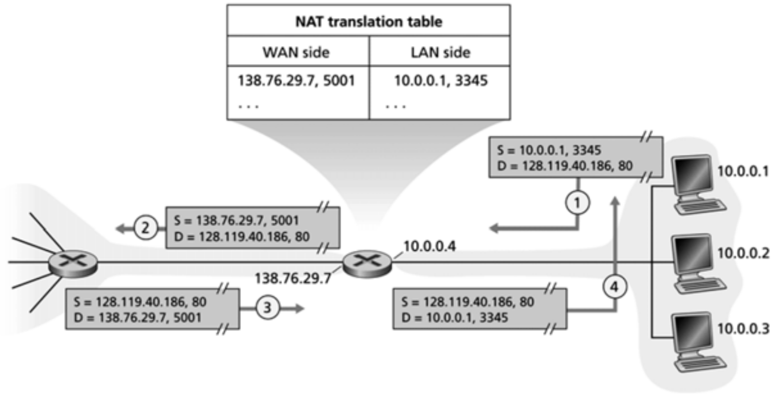


Figure 4.22 ♦ Network address translation

NAT Controversies?

- Port numbers are used by NAT to identify hosts within the local network – but ports are for addressing *processes* not *hosts*
- Routers should only process packets up to layer 3 (ports associated with app socket)
- violates end-to-end argument
  - NAT possibility must be taken into account by application designers, eg, P2P applications
  - Interfering nodes should not modify IP addresses and port numbers
- Address shortage should instead be solved by IPv6

Chapter 4: Network Layer

- 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
- 4.7 Broadcast and multicast routing

ICMP: Internet Control Message Protocol

- Used by hosts & routers to communicate network-level information
  - error reporting
  - echo request/reply
- Is in the network-layer just "above" IP:
  - ICMP msgs are carried in IP datagrams

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, Ping and ICMP

- Good discussion in text on page 357

Chapter 4: Network Layer

- 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
- 4.7 Broadcast and multicast routing

IPv6

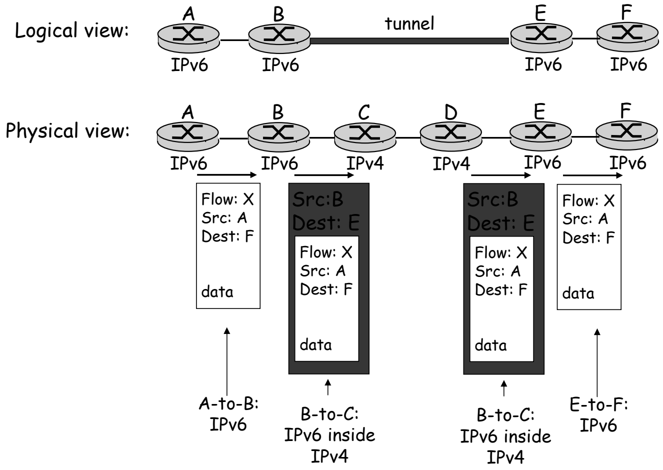
- Initial motivation: 32-bit address space soon to be completely allocated.
- Additional motivation:
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS

  IPv6 datagram format:
  - fixed-length 40 byte header
  - no fragmentation allowed
- Transition?
- Not all routers can be upgraded simultaneously
  - How will the network operate with mixed IPv4 and IPv6 routers?

Transition From IPv4 To IPv6

- Not all routers can be upgraded simultaneous
  - No ability to mandate that everyone transitions
  - How will the network operate with mixed IPv4 and IPv6 routers?
- *Tunneling:* IPv6 carried as payload in IPv4 datagram among IPv4 routers

Tunneling



Summary Section 4.4

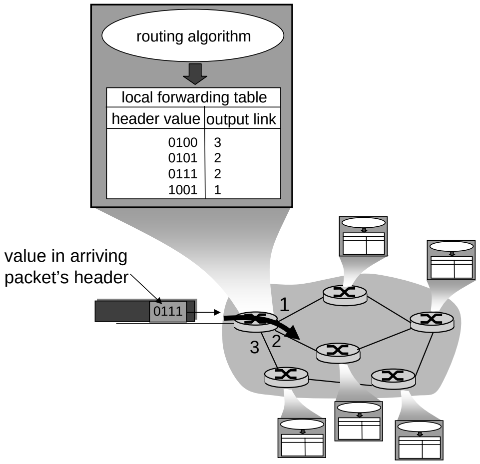
- ❑ IP Addressing
  - ❖ Class A, B…
  - ❖ Subnets
- ❑ DHCP – dynamic addressing
- ❑ NAT – network address translation
- ❑ ICMP – internet control message protocol
  - ❖ Ping and Traceroute
- ❑ IPv6
  - ❖ Overcomes IPv4 flaws, but is hard to compel transition

On to Routing and Routing Algorithms

Chapter 4: Network Layer

- ❑ 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
- ❑ 4.7 Broadcast and multicast routing

Recap: Routing v. Forwarding



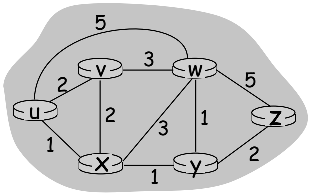
routing algorithm

local forwarding table	
header value	output link
0100	3
0101	2
0111	2
1001	1

value in arriving packet's header

Routing Algorithms

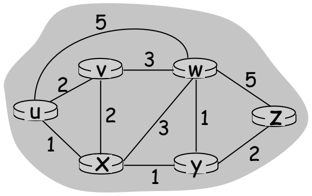
The objective of a routing algorithm is to find the least-cost path between all sources and all destinations



Routing Algorithms

- ❑ Each host is connected to a default router (the first-hop router or source router)
  - ❖ Each packet transmitted by a host is sent to this source router
- ❑ Packet needs to be routed from the source router to the destination router

Routing Notation



Graph: G = (N,E)

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

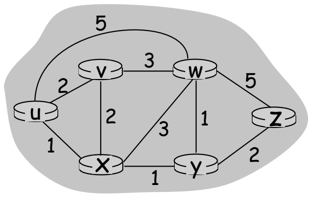
E = set of edges or links
= { (u,v), (u,x), (v,x), (v,w), (x,w), (x,y), (w,y), (w,z), (y,z) }

Routing Notation (Dijkstra)

- graph G = (N, E): is a set of nodes, N, and edges, E, where each edge is a pair of nodes from N
- 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 destination *v* (min distance from source to *v*)
- p(*v*): previous node (neighbor of *v*) along the current least cost path from the source to *v*
- N': set of nodes whose least cost path definitively known; *v* is in N' if the least cost path from source to *v* is known

Graph Notation: costs

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



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

Routing Algorithm classification

Static or dynamic?

Static:

- A program computes routes when the router boots, and they do not change
- …Routes change slowly over time

Dynamic:

- A program builds an initial forwarding table when a router boots and updates the table periodically as the network changes
- Routes change more quickly

Routing Algorithm classification

Global or decentralized information?

Global:

- All routers have complete topology & link-cost information
- "link state" algorithms

Decentralized:

- Each router knows about its physically-connected neighbors & link-costs to neighbors
- iterative process of computation, exchange of information with neighbors
- "distance vector" algorithms

Chapter 4: Network Layer

- ❑ 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
- ❑ 4.7 Broadcast and multicast routing

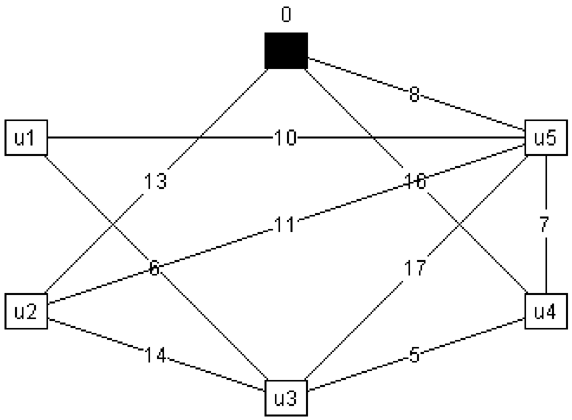
A Link-State Routing Algorithm

Dijkstra's algorithm

- ❑ computes the shortest paths in a graph by using weights on edges as a measure of distance. A path with the least number of edges may not be the path with the least weight.
- ❑ Each node has global information on network topology and edge weights

---

A Link-State Routing Algorithm



A Link-State Routing Algorithm

Dijkstra's algorithm

- ❑ computes least cost paths from one node ('source') to all other nodes
    - ❖ Determines the forwarding table for that node
- ❑ The network topology and link costs are known to all nodes
    - ❖ accomplished via "link state broadcast"
    - ❖ all nodes have the same information
- ❑ The algorithm is iterative: after k iterations, the least cost paths to k destinations are known

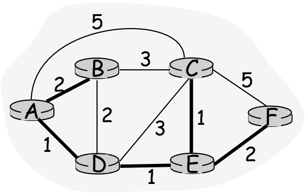
http://www-b2.is.tokushima-u.ac.jp/~ikeda/suuri/dijkstra/DijkstraApp.shtml?demo1

http://www-b2.is.tokushima-u.ac.jp/~ikeda/suuri/dijkstra/DijkstraApp.shtml?demo7

http://www-b2.is.tokushima-u.ac.jp/~ikeda/suuri/dijkstra/DijkstraApp.shtml?demo8

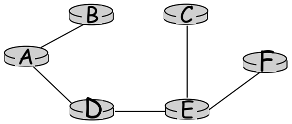
Dijkstra's algorithm: example

Step	start N'	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	infinity	infinity
1	AD	2,A	4,D		2,D	infinity
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					



Dijkstra's algorithm: example

Resulting shortest-path tree from u:



Resulting forwarding table in u:

destination	link
B	(A, B)
D	(A, D)
E	(A, D)
C	(A, D)
F	(A, D)

Chapter 4: Network Layer

- 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
- 4.7 Broadcast and multicast routing

Distance Vector Algorithm

Rather than using global information, a distance vector algorithm is:

- distributed:
  - each node communicates *only* with directly-attached neighbors
- iterative:
  - continues until no nodes exchange info.
  - *self-terminating*: no "signal" to stop
- asynchronous:
  - nodes need *not* exchange information or iterate in lock step!

Distance Vector Algorithm

Bellman-Ford Equation, an important relationship among costs of least-cost paths

Define

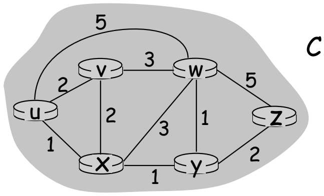
$d_x(y)$:= cost of least-cost path from x to y

Then

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

where min is taken over all neighbors v of x

Bellman-Ford Equation



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

B-F equation says:

$$\begin{aligned} d_u(z) &= \min \{ c(u,v) + d_v(z), \\ &\qquad c(u,x) + d_x(z), \\ &\qquad c(u,w) + d_w(z) \} \\ &= \min \{2 + 5, \\ &\qquad 1 + 3, \\ &\qquad 5 + 3\} = \end{aligned}$$

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

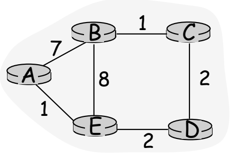
Distance Vector Routing Algorithm

Distance Table data structure

- each node has its own
  - row for each possible destination
  - column for each directly-attached neighbor
- example: in node X, for destination Y via neighbor Z:

$$\boxed{D^X(Y,Z) \begin{aligned} &= \text{distance } \textit{from} \text{ X } \textit{to} \text{ Y, } \textit{via} \text{ Z as next hop} \\ &= c(X,Z) + \min_w \{D^Z(Y,w)\} \end{aligned}}$$

Distance Table: example



$D^E(C,D) = c(E,D) + \min_w \{D^D(C,w)\}$
$=$

$D^E(A,D) = c(E,D) + \min_w \{D^D(A,w)\}$
$=$

$D^E(A,B) = c(E,B) + \min_w \{D^B(A,w)\}$
$=$

cost to destination via

destination

$D^E()$	A	B	D
A	①	15	5
B	7	8	⑤
C	6	9	④
D	4	11	②

Distance table gives routing table

cost to destination via

destination

$D^E()$	A	B	D
A	1	14	5
B	7	8	5
C	6	9	4
D	4	11	2

destination

	Outgoing link to use, cost
A	A,1
B	D,5
C	D,4
D	D,2

Distance table ⟶ Routing table

Distance vector algorithm

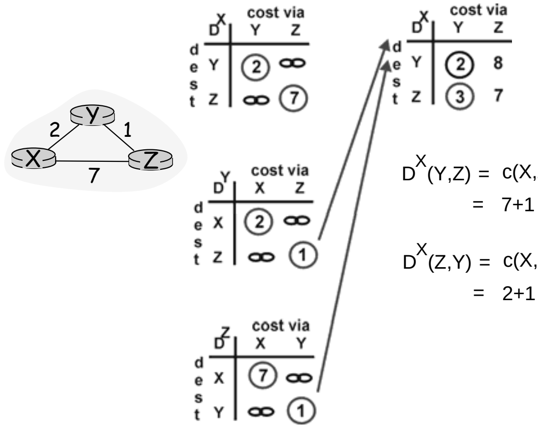
Basic idea:

- ❑ Each node begins with $D_x(y)$, an estimate of the cost of the least-cost path from itself to node y, for all nodes in N
- ❑ Each node periodically sends its own distance vector estimate to neighbors
- ❑ When a node x receives new DV estimate from neighbor, it updates its own DV using B-F equation:

$$D_x(y) \leftarrow \min_v\{c(x,v) + D_v(y)\} \quad \text{for each node } y \in N$$

- ❑ Under normal conditions, the estimate $D_x(y)$ *converge to the actual least cost* $d_x(y)$

Distance Vector Algorithm: example



D^X dest	cost via Y	cost via Z
Y	2	∞
Z	∞	7

D^Y dest	cost via X	cost via Z
X	2	∞
Z	∞	1

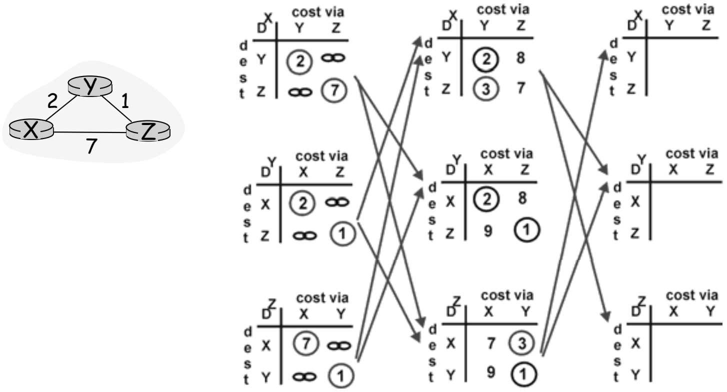
D^Z dest	cost via X	cost via Y
X	7	∞
Y	∞	1

D^X dest	cost via Y	cost via Z
Y	2	8
Z	3	7

$$D^X(Y,Z) = c(X,Z) + \min_w\{D^Z(Y,w)\}$$
$$= 7+1 = 8$$

$$D^X(Z,Y) = c(X,Y) + \min_w\{D^Y(Z,w)\}$$
$$= 2+1 = 3$$

Distance Vector Algorithm: example



D^X dest	cost via Y	cost via Z
Y	2	∞
Z	∞	7

D^Y dest	cost via X	cost via Z
X	2	∞
Z	∞	1

D^Z dest	cost via X	cost via Y
X	7	∞
Y	∞	1

D^X dest	cost via Y	cost via Z
Y	2	8
Z	3	7

D^Y dest	cost via X	cost via Z
X	2	8
Z	9	1

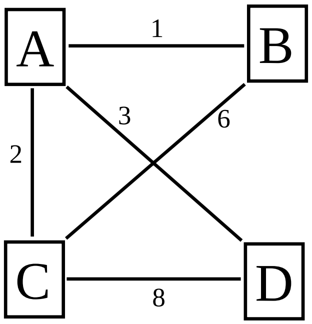
D^Z dest	cost via X	cost via Y
X	7	3
Y	9	1

D^X dest	cost via Y	cost via Z
Y		
Z		

D^Y dest	cost via X	cost via Z
X		
Z		

D^Z dest	cost via X	cost via Y
X		
Y		

Distance Vector Routing Activity



Comparison of LS and DV algorithms

- Information requirements
- Message complexity
- Convergence time varies
- Robustness: what happens if router malfunctions?
- Oscillations possible?
- Loops possible?

- To be continued...

Chapter 4: Network Layer

- 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
- 4.7 Broadcast and multicast routing

Hierarchical Routing

Our routing discussion so far – an idealization

- all routers are identical
- the network "flat"

… *not* true in practice

scale: with 200 million destinations:

- cannot store all destinations in routing tables
- routing table exchange would increase traffic

administrative autonomy

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

Hierarchical Routing

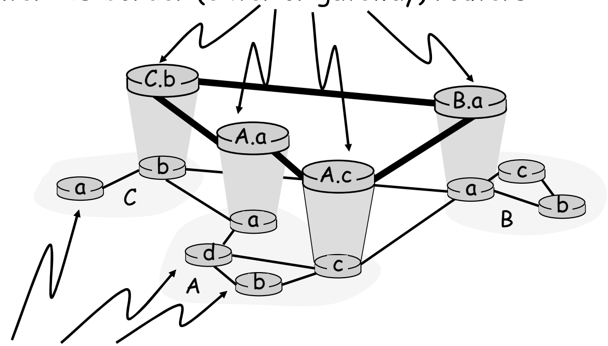
- ❑ aggregate routers into regions, "autonomous systems" (AS)
- ❑ routers in same AS run same routing protocol
  - ❖ "intra-AS" routing protocol
  - ❖ routers in different AS can run different intra-AS routing protocol

gateway 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

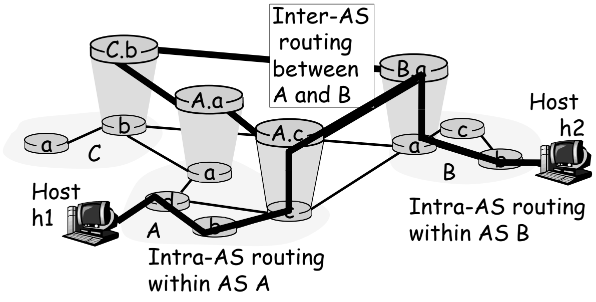
Internet AS Hierarchy

Inter-AS border (exterior gateway) routers



Intra-AS interior (gateway) routers

Intra-AS and Inter-AS routing



Inter-AS routing between A and B

Host h2

Host h1

Intra-AS routing within AS A

Intra-AS routing within AS B

- ❑ Specific inter-AS and intra-AS Internet routing protocols next class…

Summary

Forwarding:

- ❑ Leads to questions of addressing
  - ❖ Assignment of IP addresses
  - ❖ NAT, IPv6 …

Routing:

- ❑ Routing objectives
- ❑ Routing notation
- ❑ Routing classification
  - ❖ Static v. Dynamic
  - ❖ Link state v. Distance Vector
- ❑ Inter- and Intra- AS routing