

CSc 450/550
Computer Networks
Internet Routing

Jianping Pan
Summer 2007

Review

- Internet Protocol (IP)

- IP header

- addressing

- class-based, classless, hierarchical, NAT

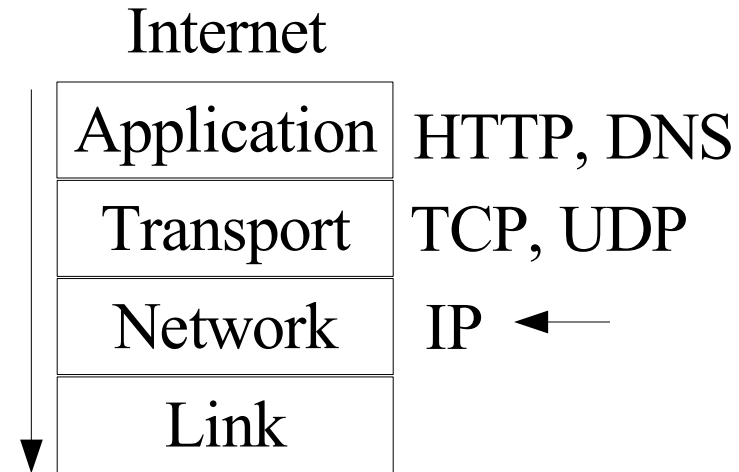
- routing algorithms

- flooding

- distance vector

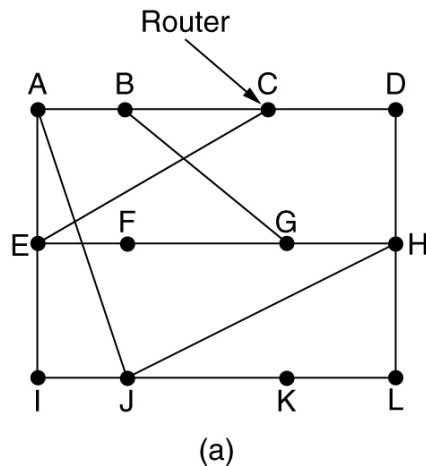
- link state

- hierarchical



Distance vector routing

- Distributed Bellman-Ford algorithm



To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28	20	A
C	25	18	19	36	28	I
D	40	27	8	24	20	H
E	14	7	30	22	17	I
F	23	20	19	40	30	I
G	18	31	6	31	18	H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	-
K	24	22	22	0	6	K
L	29	33	9	9	15	K

JA delay is 8 JI delay is 10 JH delay is 12 JK delay is 6

New routing table for J

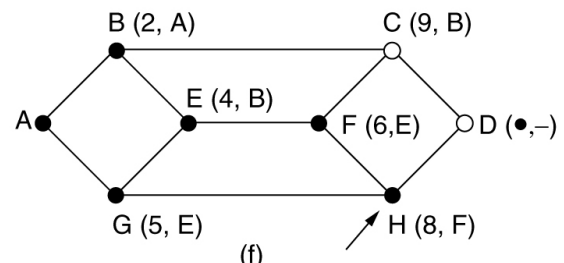
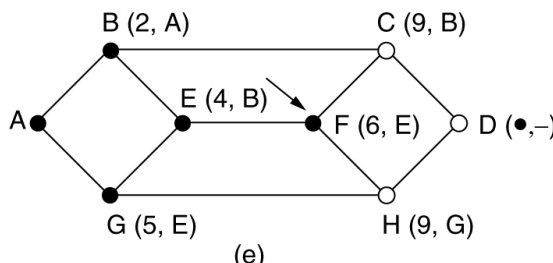
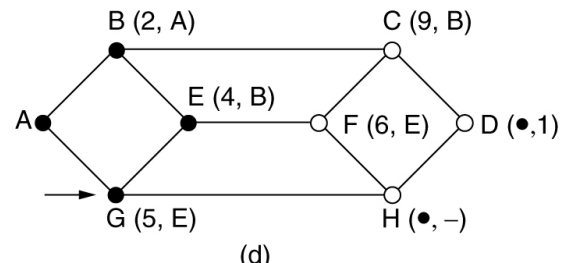
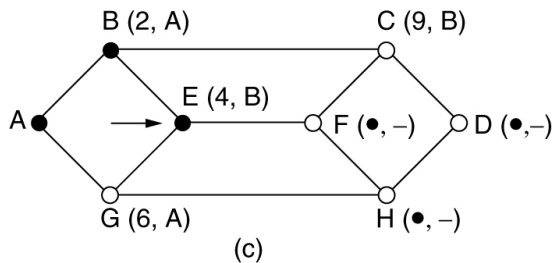
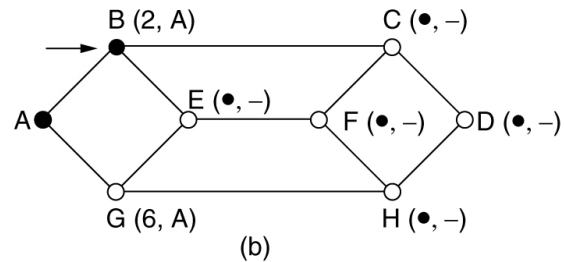
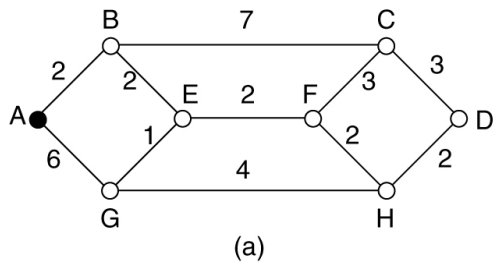
Vectors received from J's four neighbors

A B X

$$\text{Path}_{\text{new}}(A, X) = \min_B \{ \text{Path}_{\text{old}}(A, X), \text{Link}(A, B) \stackrel{(b)}{+} \text{Path}(B, X) \}$$

Link state routing

- Dijkstra algorithm



7/12/07

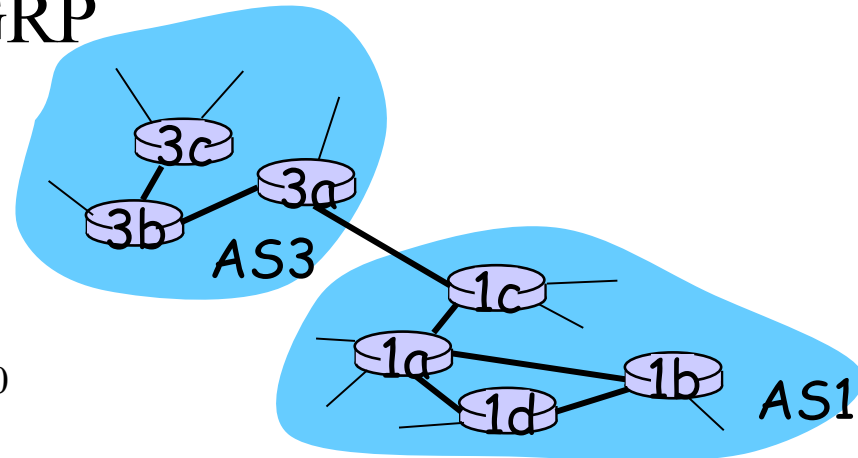
CSc 450/550

4

$$D(v) = \min(D(v), D(w) + c(w,v))$$

Today's topics

- Internet routing protocols
 - how does the Internet really route *my* packets?
 - hierarchical structures
- Intra-domain routing
 - AS: autonomous systems
 - e.g., RIP, OSPF, ISIS, IGRP
- Inter-domain routing
 - e.g., BGP



Intra-AS Routing

- Also known as **Interior Gateway Protocols (IGP)**
- **Most common Intra-AS** routing protocols:

- **RIP**: Routing Information Protocol

algorithm

DV

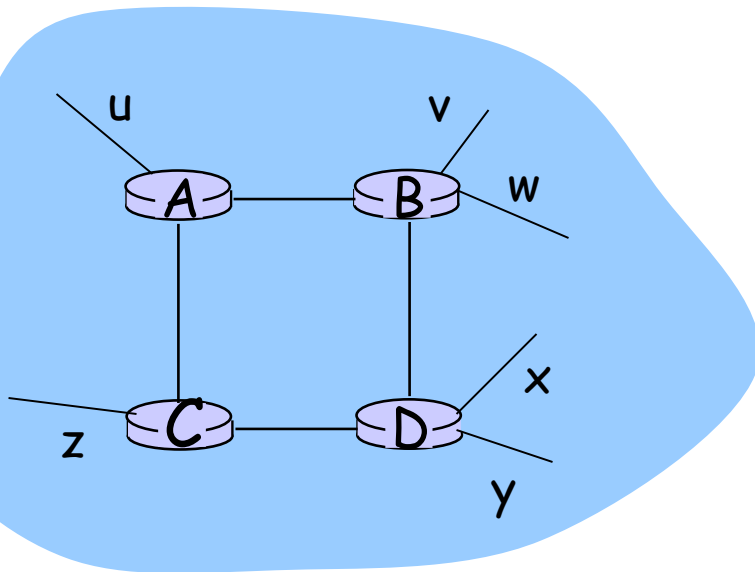
- **OSPF**: Open Shortest Path First

LS

- **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)



From router A to subsets:

<u>destination</u>	<u>hops</u>	<u>next-hop</u>
u	1	-
v	2	B
w	2	B
x	3	C
y	3	C
z	2	C

7/12/07

CSc 450/550

<number>

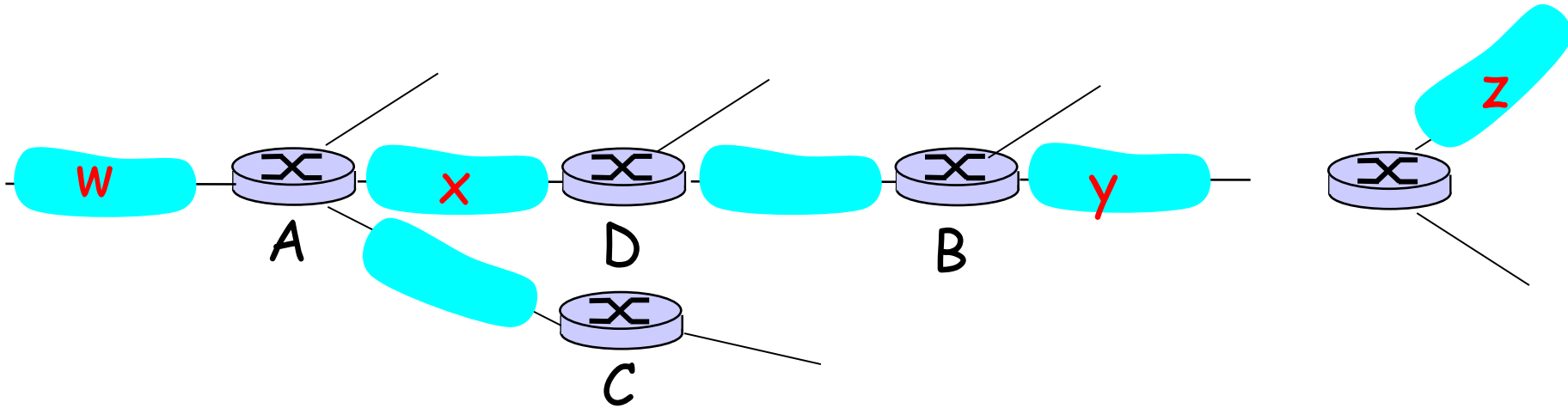
Q: why max 15 hops?

H: CTI problems in DV

RIP advertisements

- **Distance vectors: exchanged** among neighbors every **30 sec** via **RIP Response Message** (also called **advertisement**)
- Each advertisement: list of up to **25 destination subnets** **within** AS.
 - UDP port 520
 - /etc/services
 - newer implementations also use TCP
- **Routing metric: the number of hops**

RIP: Example



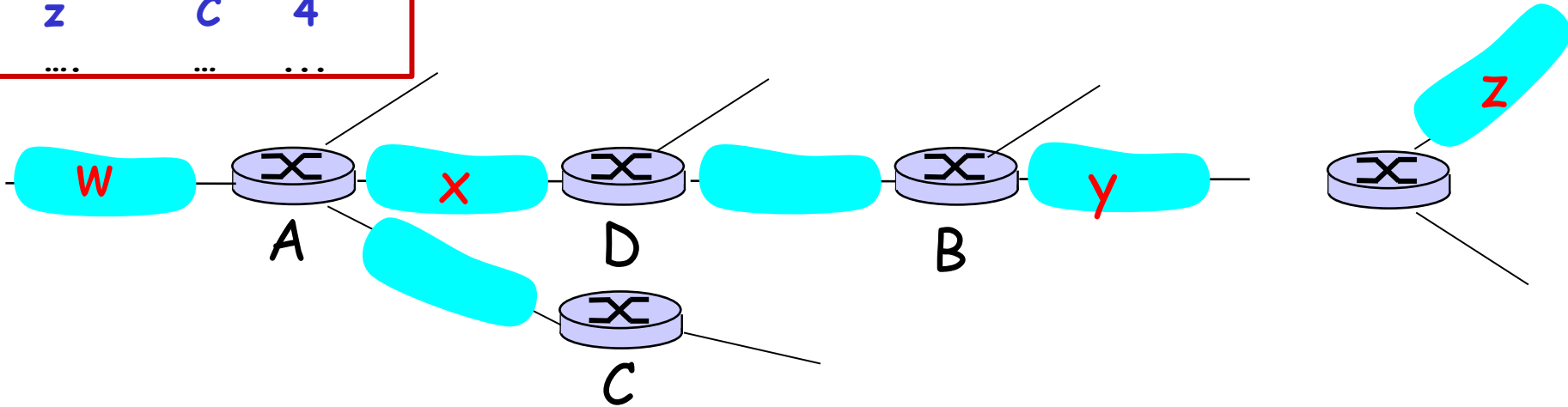
Destination Network	Next Router	Num. of hops to dest.
W	A	2
Y	B	2
Z	B	7
X	--	1
...

Routing table in router D

RIP: Route update example

Dest	Next	hops
w	-	1
x	-	1
z	C	4
...

Advertisement from A to D



Destination Network	Next Router	Num. of hops to dest.
w	A	2
y	B	2
z	B A	7 5
x	--	1
...

RIP: Link Failure and Recovery

If **no RIP advertisement** heard after **180 sec** --> neighbor/link declared **dead (not reachable)**

- routes via neighbor invalidated
- **new advertisements** sent to neighbors
- neighbors in turn send out **new advertisements** (if tables changed)
- link failure info quickly propagates to entire net
- **poison reverse** used to prevent **ping-pong loops** (infinite distance = **16** hops)

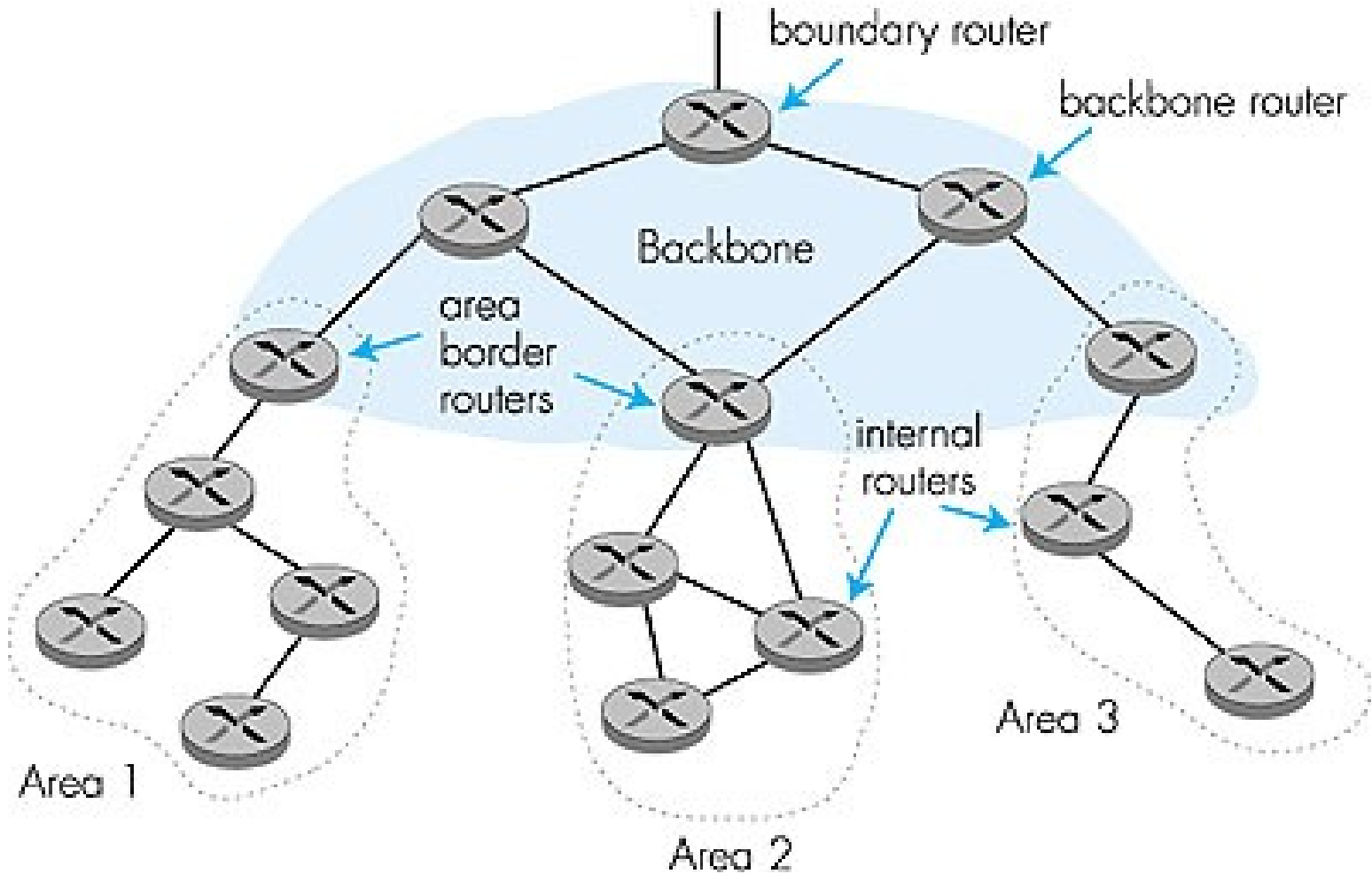
OSPF (Open Shortest Path First)

- “open”: publicly available
- Uses Link State algorithm
 - LS packet dissemination
 - A router constructs a complete topological map (i.e., a graph)
 - Route computation using Dijkstra’s algorithm to determine a shortest path tree to all subnets
- 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)
 - IP protocol ID: 89
 - /etc/protocols
- Multiple routing metrics

OSPF messages

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

Hierarchical OSPF



Hierarchical OSPF

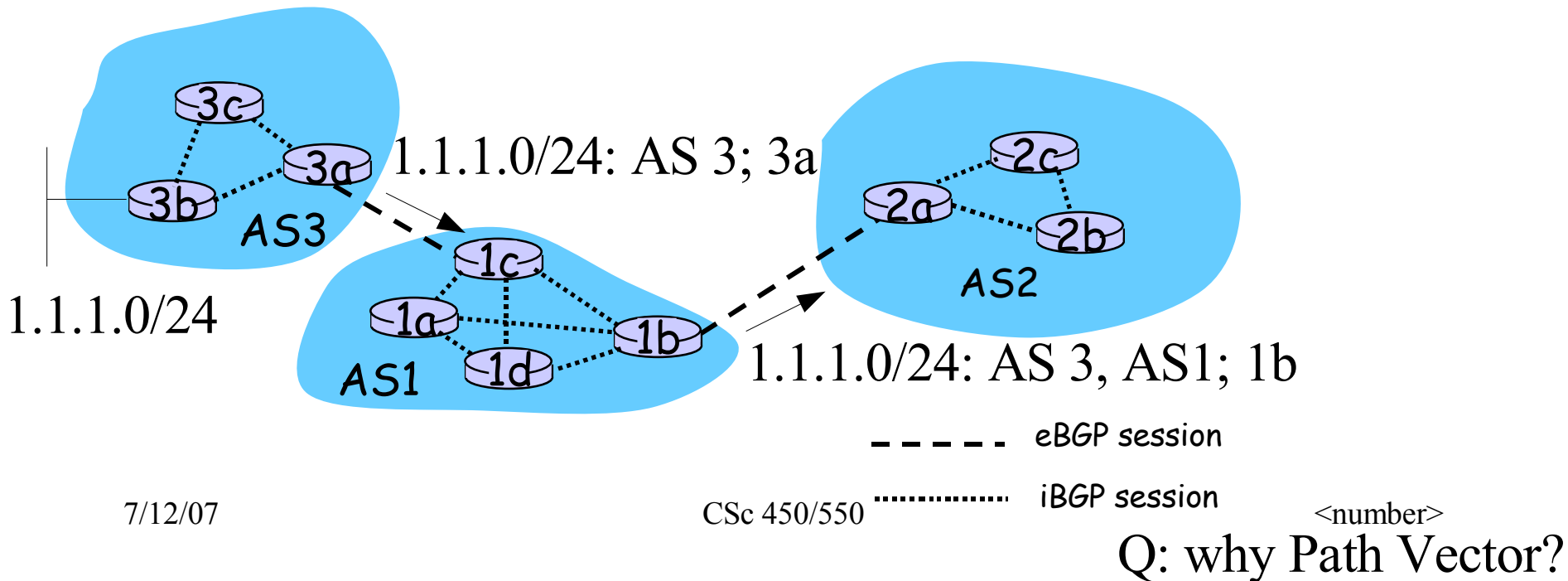
- **Two-level hierarchy:** local area, backbone.
 - Link-state advertisements only in area
 - each node has detailed area topology; only know direction (shortest path) to subnets in other areas.
- **Area border routers:** “summarize” distances to nets in own area, advertise to other Area Border routers.
- **Backbone routers:** run OSPF routing limited to backbone.
- **Boundary routers:** connect to other AS's.

Internet **inter-AS** routing: BGP

- **BGP (Border Gateway Protocol):** *the de facto exterior gateway protocol (version 4)*
- BGP provides **each AS** a means to:
 1. Obtain subnet **reachability information** from neighboring ASs.
 2. Propagate the **reachability** information to **all** routers **internal** to the AS.
 3. Determine “**good**” routes to subnets based on reachability information and policy.
- Allows a **subnet** to advertise its existence to **rest of the Internet:** *“I am here”*

Distributing reachability info

- With eBGP session between 3a and 1c, AS3 sends a **prefix reachability info**, **reachable from AS3 to the prefix**, to AS1.
- 1c can then use iBGP to distribute this new prefix reach info to **all routers in AS1**
- 1b can then **re-advertise** the new reach info to AS2 over the 1b-to-2a eBGP session
- When a **router** learns about a **new prefix**, it creates an **entry** for the **prefix** in its **forwarding table**.



Path attributes & BGP routes

- When advertising a prefix, advertisement includes **BGP attributes**.
 - **prefix + attributes = “route”**
 - path vector routing
- **Two important attributes**:
 - **AS-PATH**: contains the ASs through which the advert for the prefix passed: **AS3 AS1**
 - **NEXT-HOP**: Indicates the specific **internal-AS router** to **next-hop** AS. (There may be multiple links from current AS to next-hop-AS.)
- When **gateway** router receives route advert, uses **import routing policy** to accept/decline.

BGP messages

- BGP **messages** exchanged using **TCP**.
- 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

BGP route selection

- ❑ Router may learn about **more than 1 route** to some **prefixes**. Router must **select route**.
- ❑ Elimination rules:
 1. Local preference value attribute: **routing policy** decision
 2. Shortest **AS-PATH**
 3. Closest **NEXT-HOP** router: **hot potato routing**
 4. Additional criteria

Comparison: intra/inter-domain routing

Policy:

- **Inter-AS:** admin wants control over how its traffic routed, who routes through its net.
- **Intra-AS:** single admin, so no policy decisions needed

Scale:

- hierarchical routing saves table size, reduced update traffic

Performance:

- **Intra-AS:** can focus on performance
- **Inter-AS:** policy may dominate over performance

This lecture

- Internet routing
 - RIP: distance vector routing
 - OSPF: hierarchical, link-state routing
 - BGP basics: path vector routing
- Explore further
 - Advanced Computer Networks (Topics Course)
<http://www.cs.uvic.ca/~pan/csc485>
 - IGP routing metrics, BGP routing dynamics
 - routing security

Next lectures

- The Link Layer

- framing control
- flow/error control
- media access control

- IEEE 802 family

- IEEE 802.3: Ethernet
- IEEE 802.11: “wireless Ethernet” (WiFi)

