# Chapter 4: Network Layer

# 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: Network Layer

# RIP ( Routing Information Protocol)

□ Distance vector algorithm

□ Included in BSD-UNIX Distribution in 1982

□ Distance metric: # of hops (max = 15 hops)
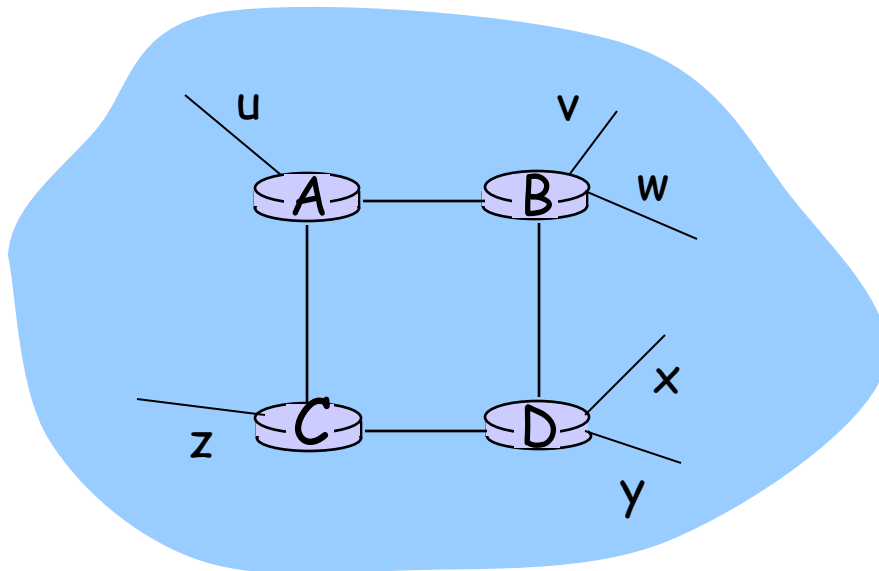


From router A to subsets:

| 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

# RIP: Example
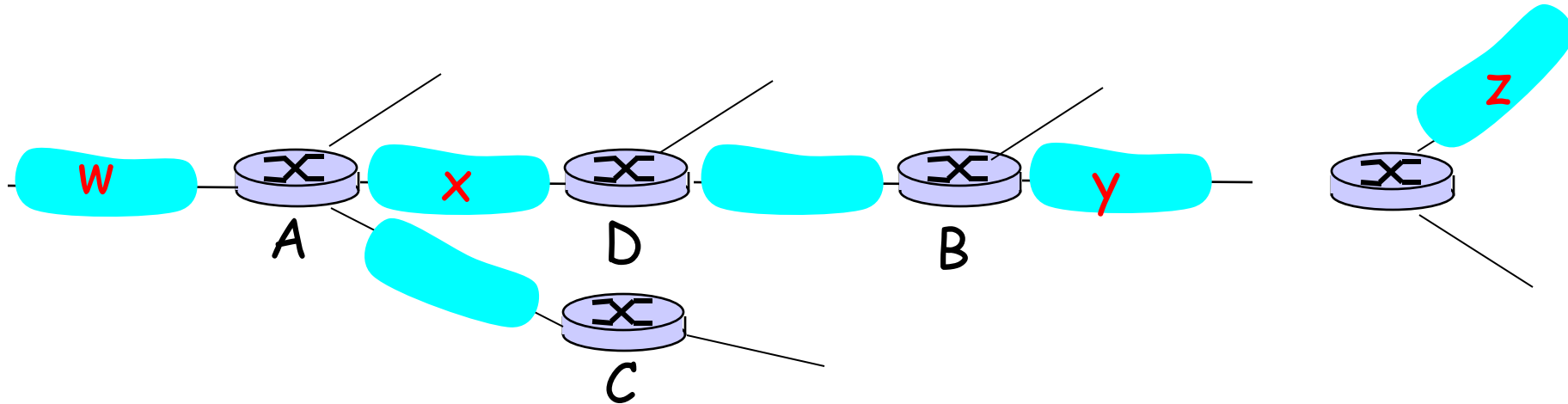


Routing table in D

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

# RIP: Example

| Dest | Next | hops |
|------|------|------|
| w | – | 1 |
| x | – | 1 |
| z | C | 4 |
| …. | … | . . . |

**Advertisement from A to D**
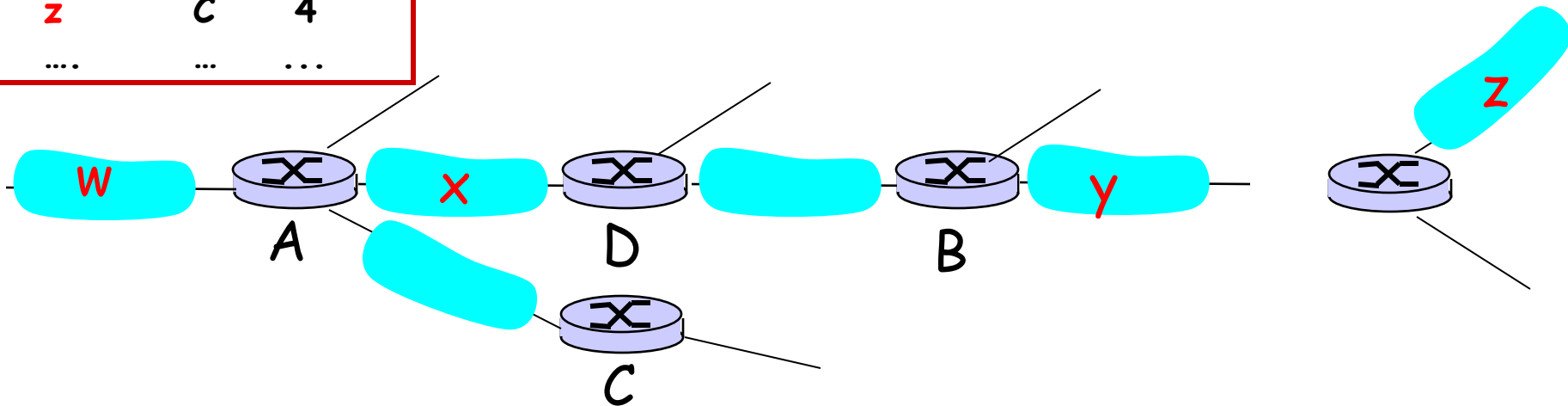


| Destination Network | Next Router | Num. of hops to dest. |
|---------------------|-------------|------------------------|
| w | A | 2 |
| y | B | 2 |
| z | ~~B~~ A | ~~7~~ 5 |
| x | – – | 1 |
| …. | …. | . . . . |

Routing table in D

# RIP: Link Failure and Recovery

If no advertisement heard after 180 sec --> neighbor/link declared dead
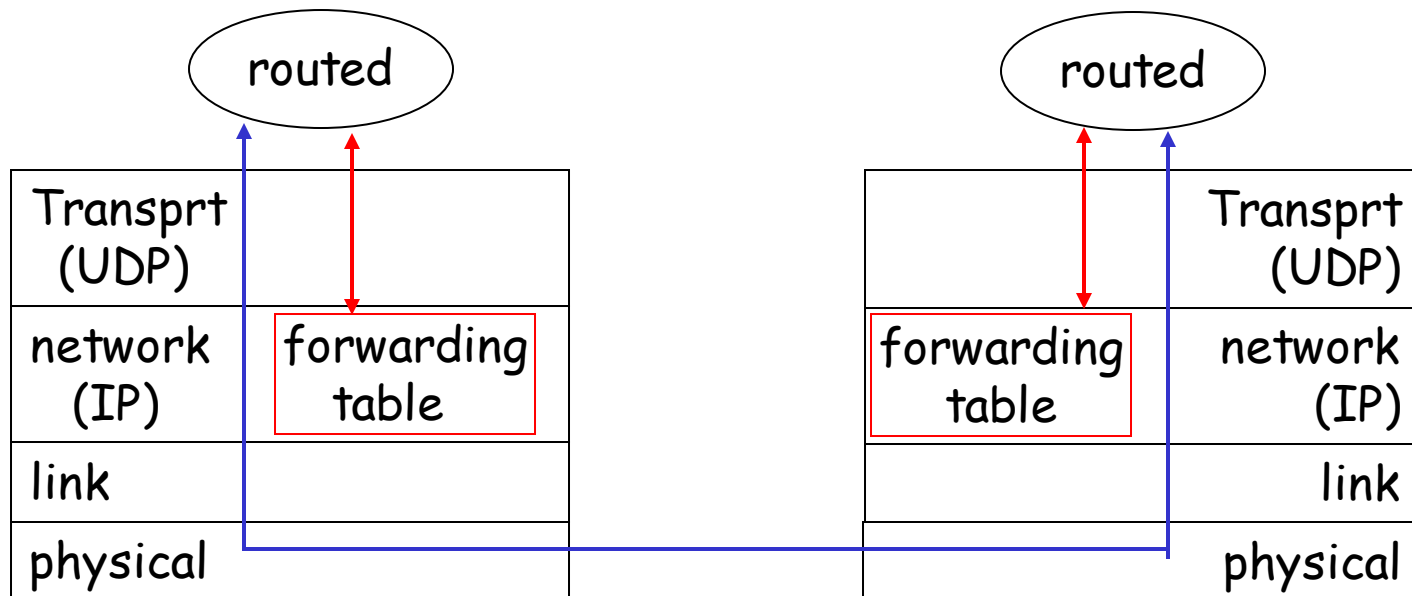
- routes via neighbor invalidated
- new advertisements sent to neighbors
- neighbors in turn send out new advertisements (if tables changed)
- link failure info quickly propagates to entire net
- poison reverse used to prevent ping-pong loops (infinite distance = 16 hops)

# RIP Table processing

□ RIP routing tables managed by **application-level** process called route-d (daemon)

□ advertisements sent in UDP packets, periodically repeated

| routed | | | routed |
|---|---|---|---|
| Transprt (UDP) | | | Transprt (UDP) |
| network (IP) | forwarding table | forwarding table | network (IP) |
| link | | | link |
| physical | | | physical |

# 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

# 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)

□ **Multi**ple same-cost **path**s 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
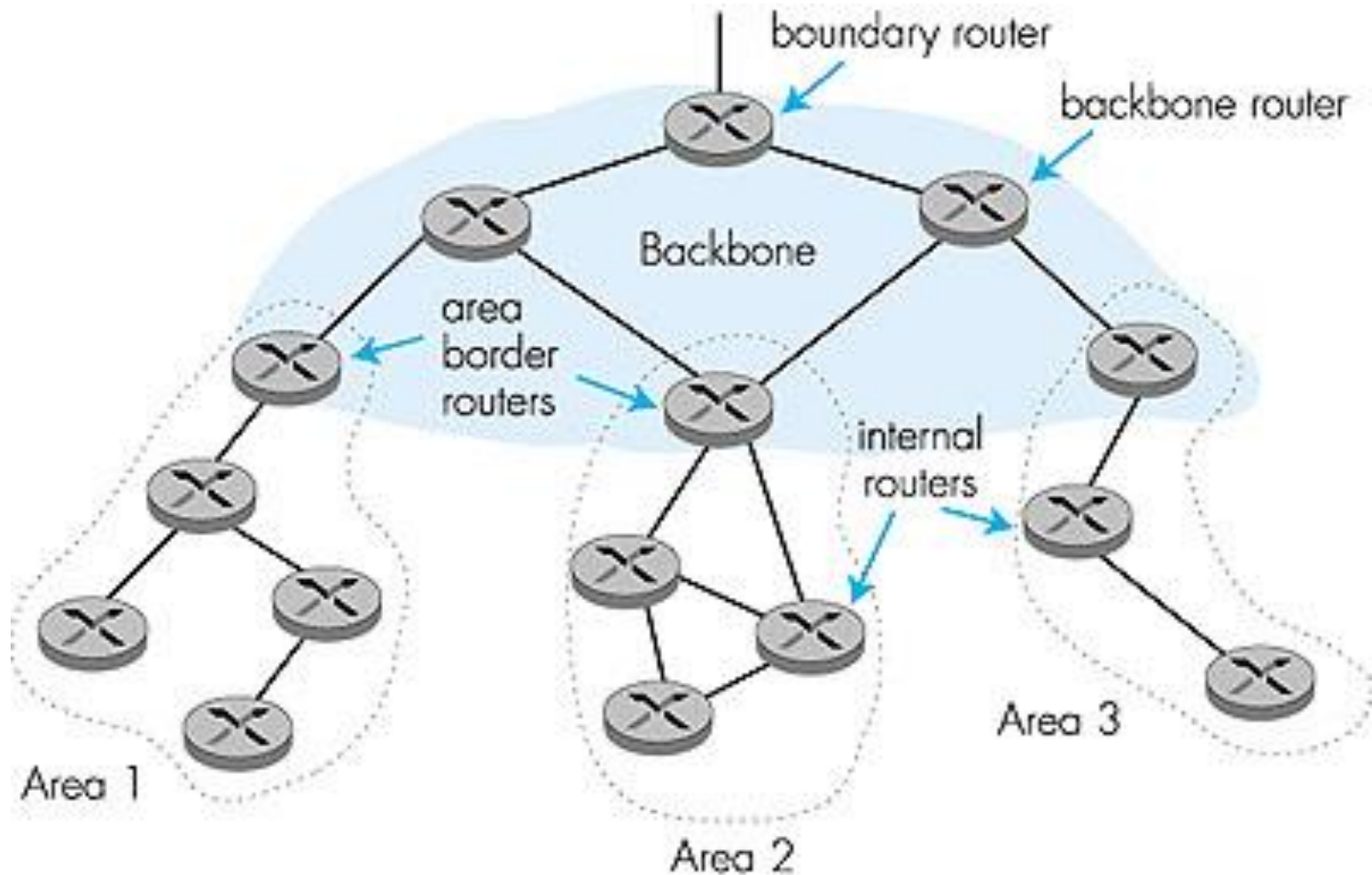
# Hierarchical OSPF

□ Two-level hierarchy: local area, backbone.
  ○ Link-state advertisements only in 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: Network Layer

# 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 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"*

# BGP basics

- Pairs of routers (BGP peers) exchange routing info over semi-permanent TCP conctns: BGP sessions
- Note that BGP sessions do not correspond to physical links.
- When AS2 advertises a prefix to AS1, AS2 is *promising* it will forward any datagrams destined to that prefix towards the prefix.
  - AS2 can aggregate prefixes in its advertisement
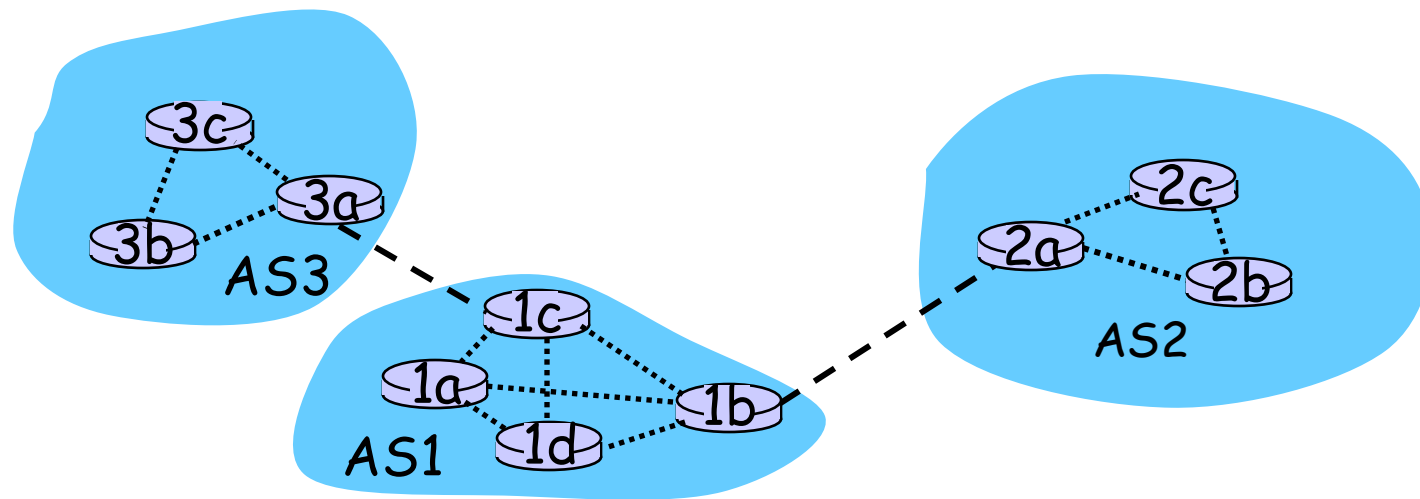


– – – – –  eBGP session

..............  iBGP session

# Distributing reachability info

□ With eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.

□ 1c can then use iBGP do 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 router learns about a new prefix, it creates an entry for the prefix in its forwarding table.



– – – – – eBGP session

·············· iBGP session
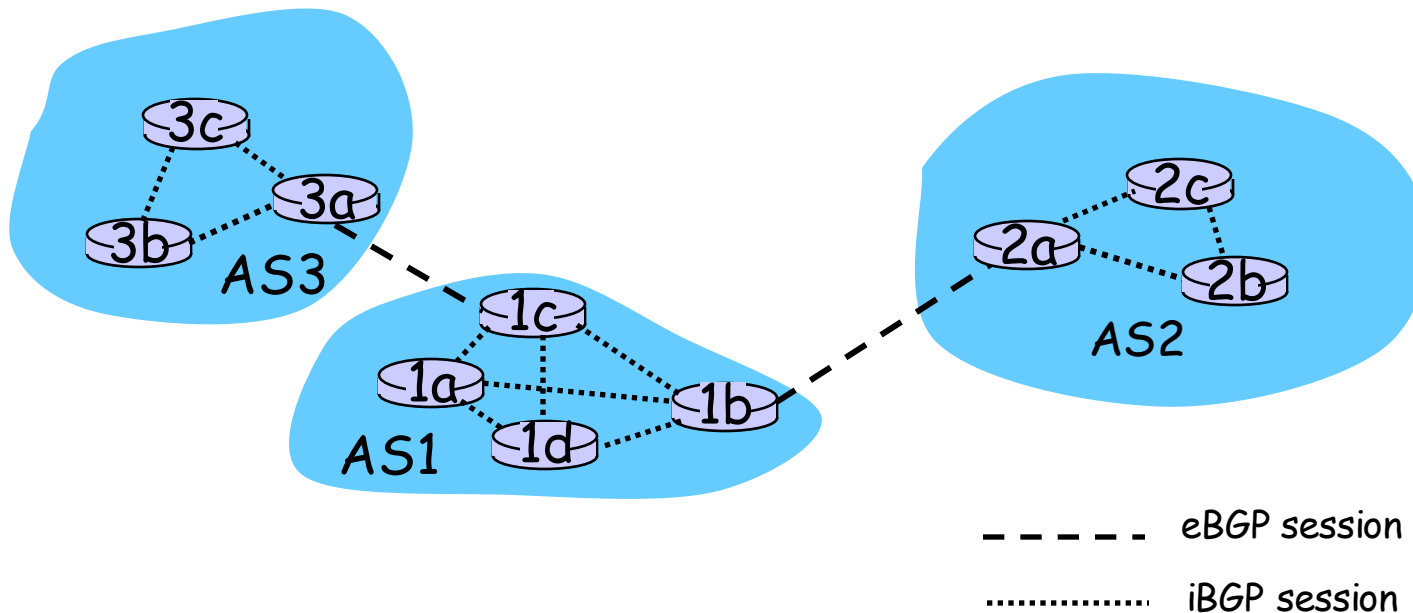
# Path attributes & BGP routes

□ When advertising a prefix, advert includes BGP attributes.

  ○ prefix + attributes = "route"

□ Two important attributes:

  ○ AS-PATH: contains the ASs through which the advert for the prefix passed: AS 67 AS 17

  ○ 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 policy to accept/decline.

# BGP route selection

□ Router may learn about more than 1 route to some prefix. Router must select route.

□ Elimination rules:

1. Local preference value attribute: policy decision
2. Shortest AS-PATH
3. Closest NEXT-HOP router: hot potato routing
4. Additional criteria

# 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 routing policy



legend:

provider network
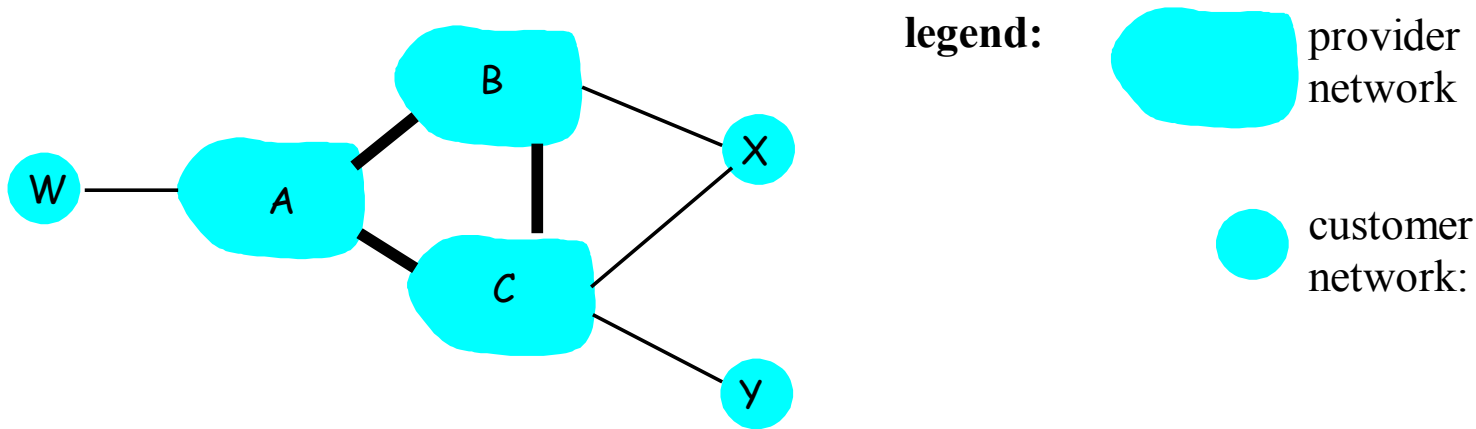
customer network:

□ A,B,C are provider networks

□ X,W,Y are customer (of provider 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 routing policy (2)



legend:

provider network

customer network:
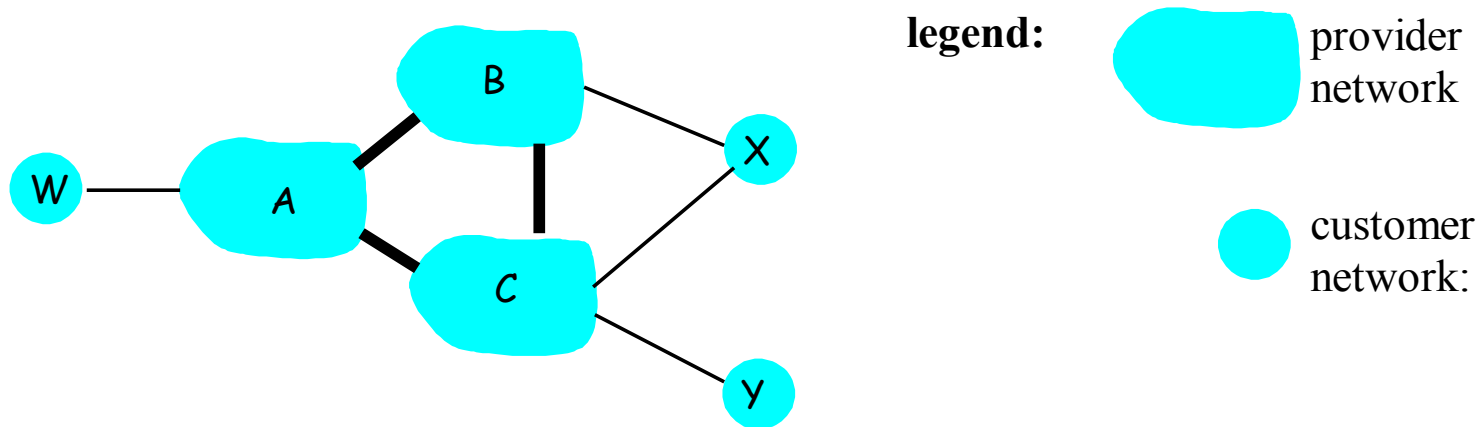
□ A advertises to B the path AW

□ B advertises to X the path BAW

□ Should B advertise to C the path BAW?
  ○ No way! 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!

# Why different Intra- and Inter-AS 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

# Summary: 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

# Questions?