AMRITA
VISHWA VIDYAPEETHAM
DEEMED TO BE UNIVERSITY

15CSE312
COMPUTER NETWORKS
3-0-0 3

Chapter 4: Network Layer

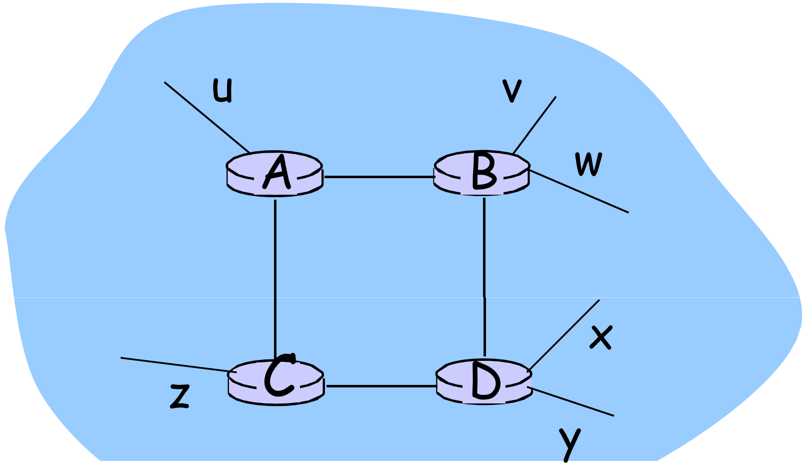
Routing Protocols in Internet

- RIP
- OSPF
- BGP

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 subnets:

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 subnets within AS

RIP: Example

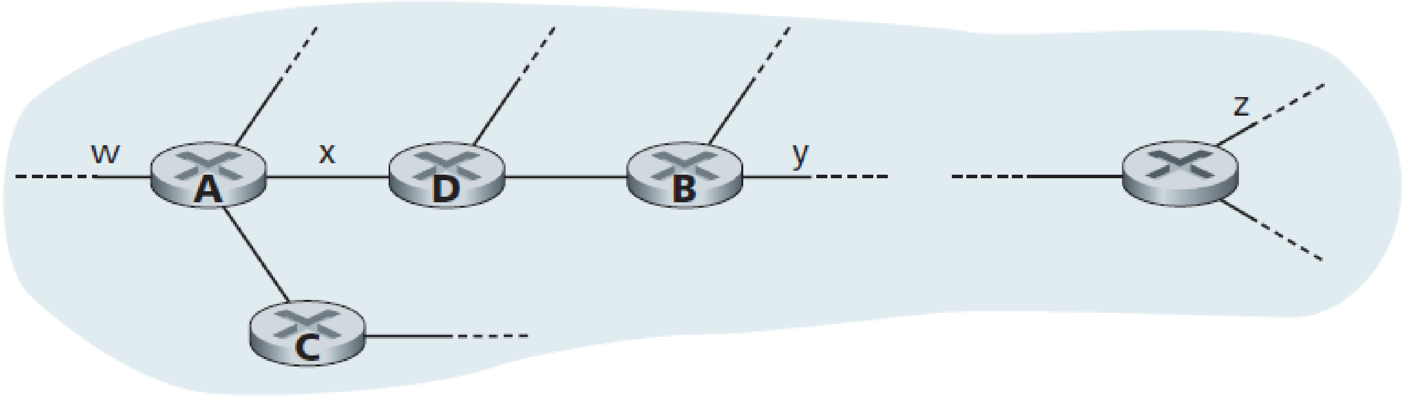
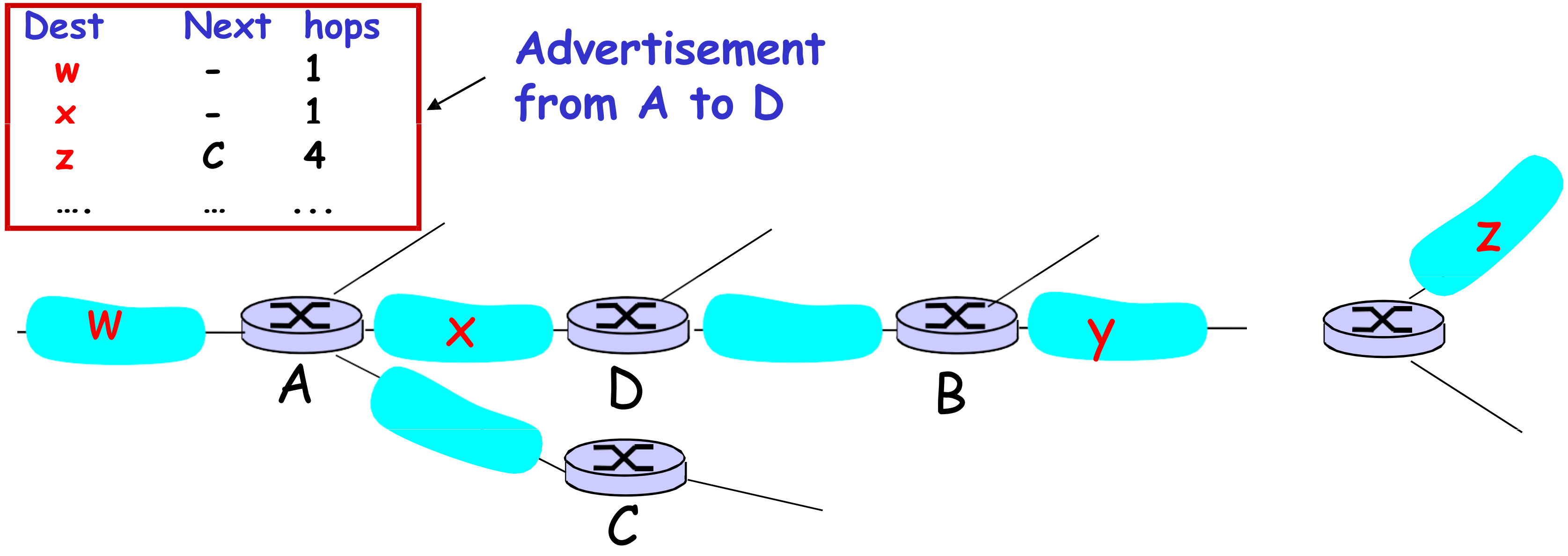


Figure 4.35 ♦ A portion of an autonomous system

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

Routing/Forwarding table in D

RIP: Example



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

Routing/Forwarding table in D

RIP: Link Failure and Recovery

If no advertisement heard after 180 sec --> neighbor/link declared dead

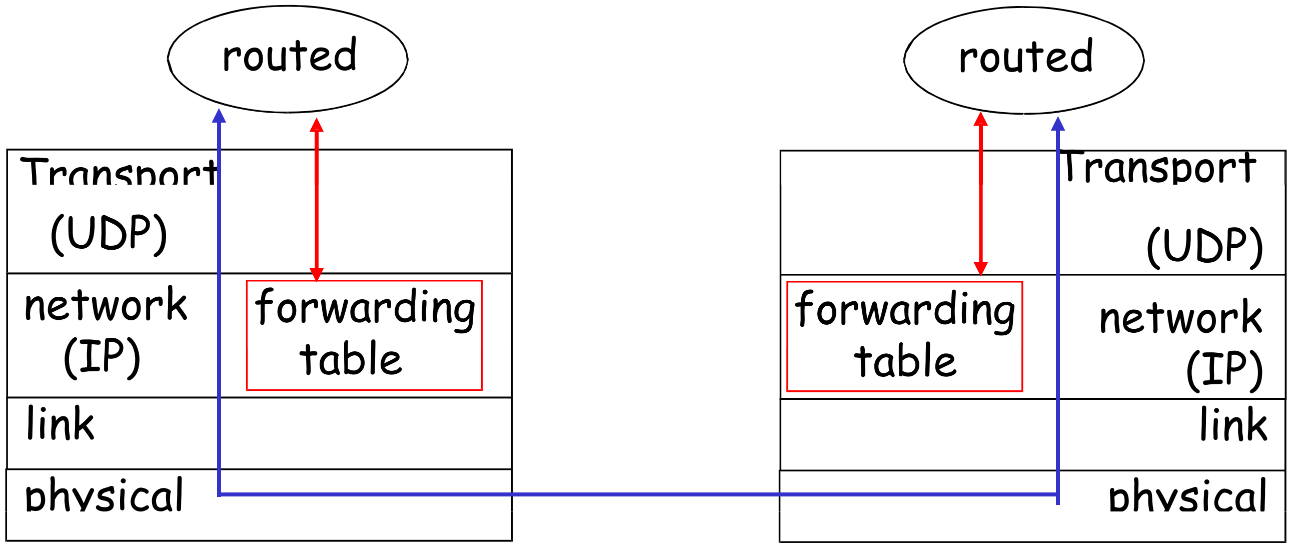
- routes via neighbor invalidated
- new advertisements sent to neighbors
- neighbors in turn send out new advertisements (if tables changed)
- link failure info quickly (?) propagates to entire net
- *poison reverse* used to prevent ping-pong loops (infinite distance = 16 hops)

  -which a router actively advertises routes as unreachable over the interface over which they were learned by setting the route metric to infinite (16 for RIP).

A router can also request information about its neighbor's cost to a given destination using RIP's request message. Routers send RIP request and response messages to each other over UDP using port number 520.

RIP Table processing

- RIP routing tables managed by **application-level** process called route-d (daemon)
- advertisements sent in UDP packets, periodically repeated



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)

OSPF was conceived as the successor to RIP and as such has a number of advanced features.
At its heart, however, OSPF is a link-state protocol that uses flooding of link-state information and a Dijkstra least-cost path algorithm

- "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
- Individual link costs are configured by the network administrator
  - The administrator might choose to set all link costs to 1, thus achieving minimum-hop routing, or might choose to set the link weights to be inversely proportional to link capacity in order to discourage traffic from using low-bandwidth links. (No mandate)

OSPF

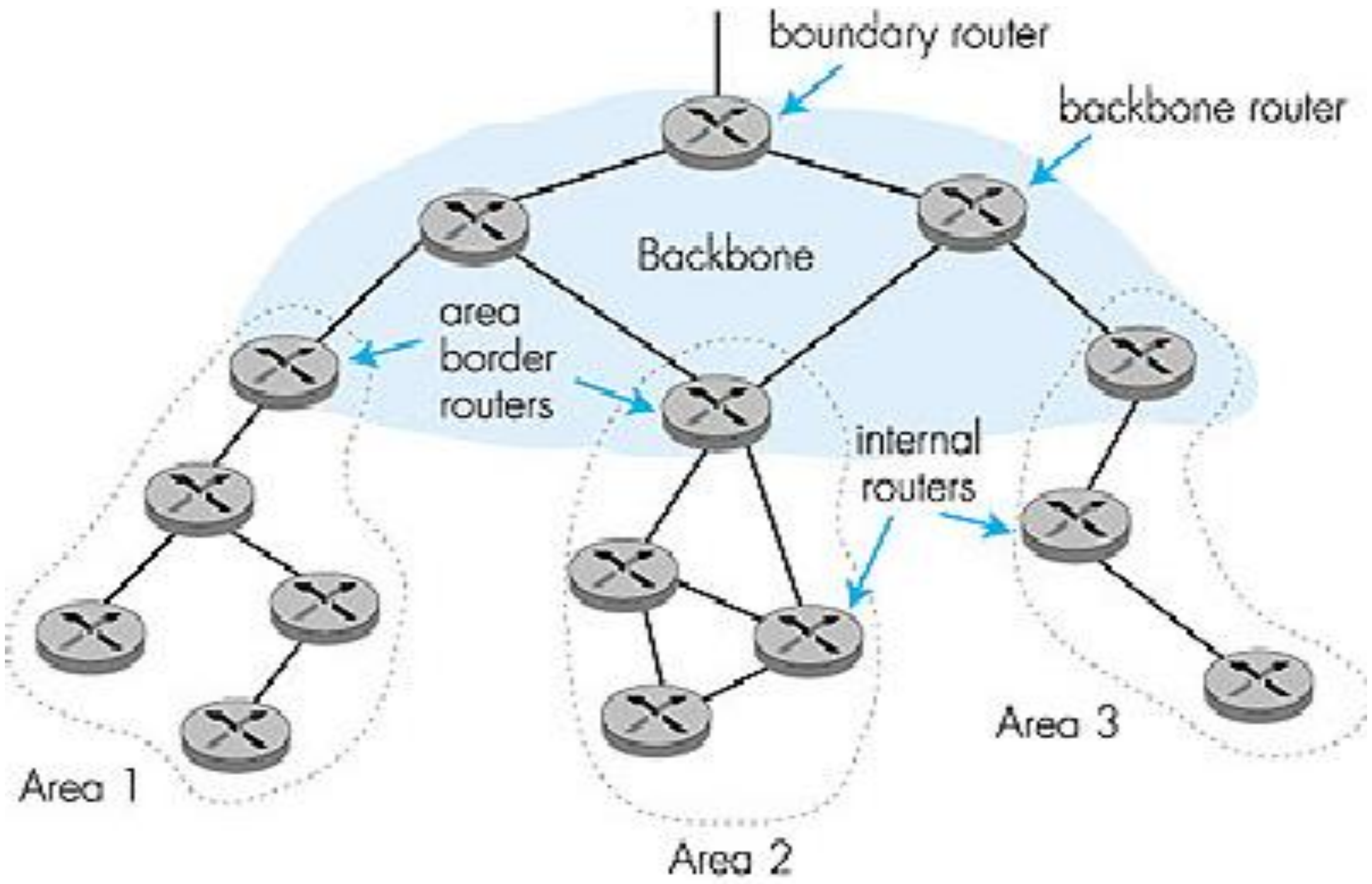
- With OSPF, a router broadcasts routing information to all other routers in the autonomous system, not just to its neighboring routers.
- router broadcasts linkstate information whenever there is a change in a link's state (for example, a change in cost or a change in up/down status).
- It also broadcasts a link's state periodically (at least once every 30 minutes), even if the link's state has not changed. RFC 2328
- "this periodic updating of link state advertisements adds robustness to the link state algorithm."

OSPF "advanced" features (not in RIP)

security: all OSPF messages authenticated (to prevent malicious intrusion). Two types of authentication can be configured—simple and MD5

- 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



An OSPF autonomous system can be configured hierarchically into areas
Each area runs its own OSPF link-state routing algorithm, with each router in an area broadcasting its link state to all other routers in that area.

Within each area, one or more **area border routers** are responsible for routing packets outside the area.

Lastly, exactly one OSPF area in the AS is configured to be the **backbone** area. The primary role of the backbone area is to route traffic between the other areas in the AS.

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

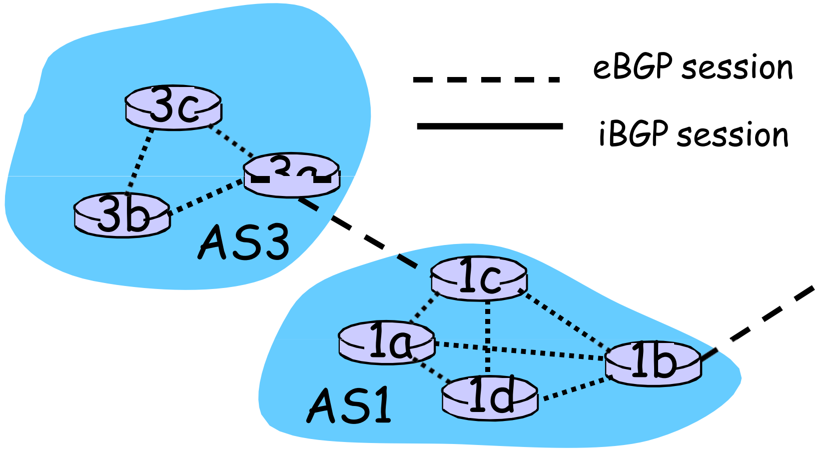
- 4. 1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Broadcast and multicast routing

Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *th* de facto standard
- BGP provides each AS a means to:
  1. Obtain subnet reachability information from neighboring ASs.
  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"*

BGP basics

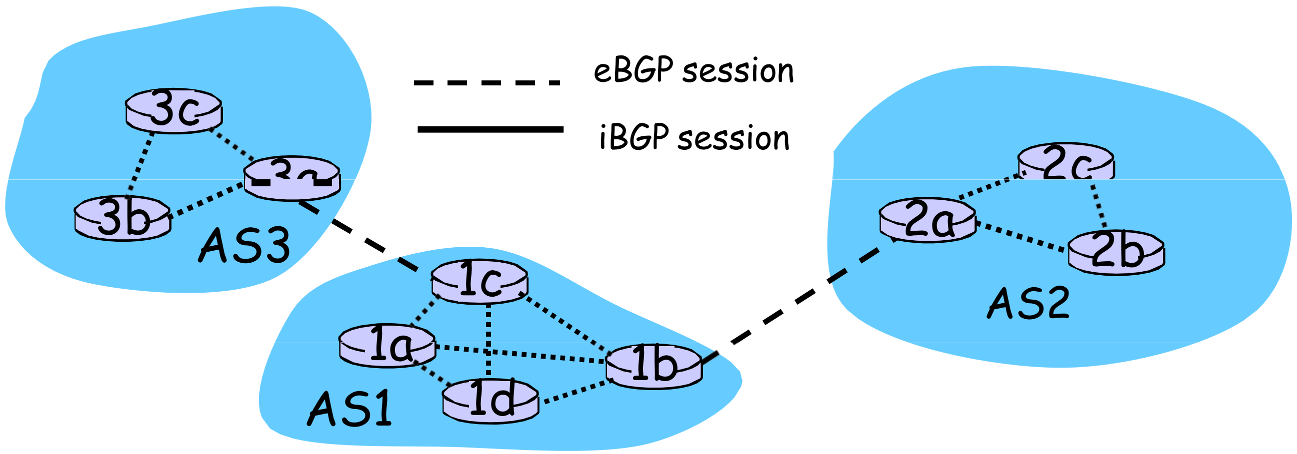
- pairs of routers (BGP peers) exchange routing info over semi-permanent TCP connections: BGP sessions
  - BGP sessions need not correspond to physical links.
- AS2 advertises a prefix to AS1:
  - AS2 *promises* it will forward datagrams towards that prefix.
  - AS2 can aggregate prefixes in its advertisement



- A BGP session that spans two Ass is called an external BGP (eBGP) session, and
- A BGP session between routers in the same AS is called an internal BGP (iBGP) session.

Distributing reachability info

- using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.
  - 1c can then use iBGP do distribute new prefix info to all routers in AS1
  - 1b can then re-advertise new reachability info to AS2 over 1b-to-2a eBGP session
- when router learns of new prefix, it creates entry for prefix in its forwarding table.



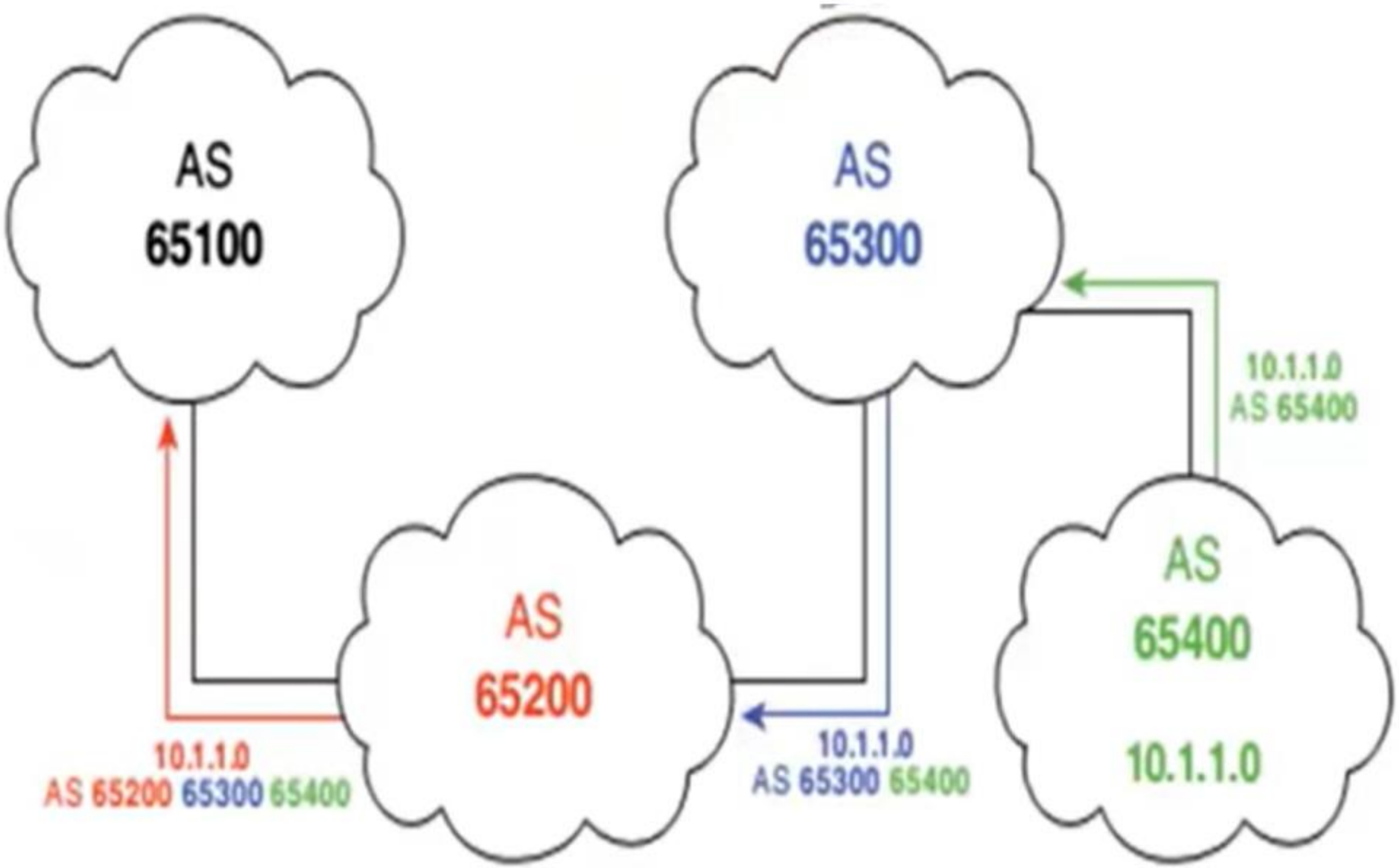
In BGP, an autonomous system is identified by its globally unique autonomous system number (ASN)

Path attributes & BGP routes

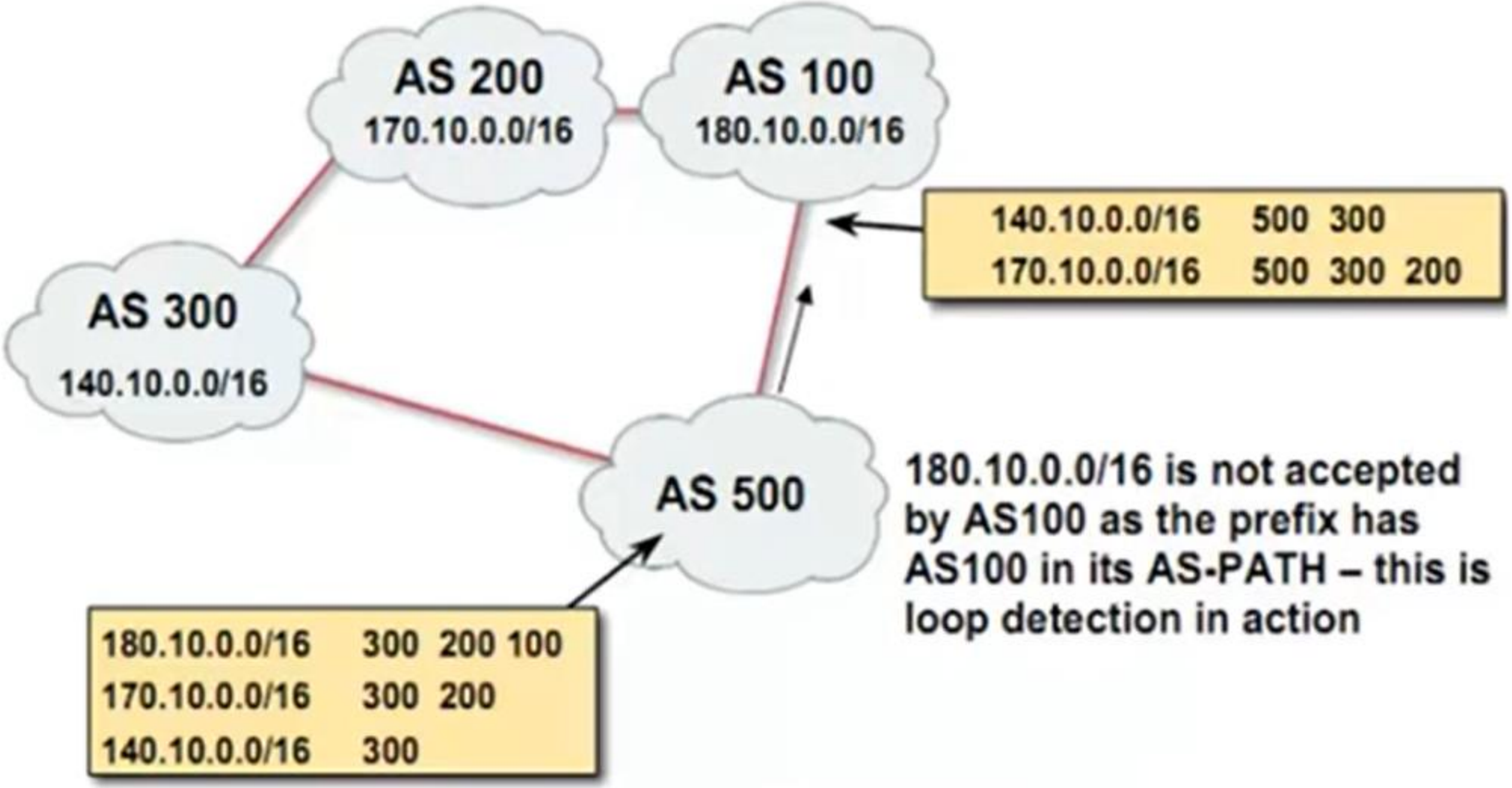
- advertised prefix includes BGP attributes.
  - prefix + attributes = "route"
- two important attributes:
  - AS-PATH: contains ASs through which prefix advertisement has passed: e.g, AS 67, AS 17
  - NEXT-HOP: indicates specific internal-AS router to next-hop AS. (may be multiple links from current AS to next-hop-AS)
- when gateway router receives route advertisement, uses import policy to accept/decline.

When a router advertises a prefix across a BGP session, it includes with the prefix a number of BGP attributes. In BGP jargon, a prefix along with its attributes is called a route. Thus, BGP peers advertise routes to each other. Two of the more important attributes are AS-PATH and NEXT-HOP:

Prefix+AS-PATH

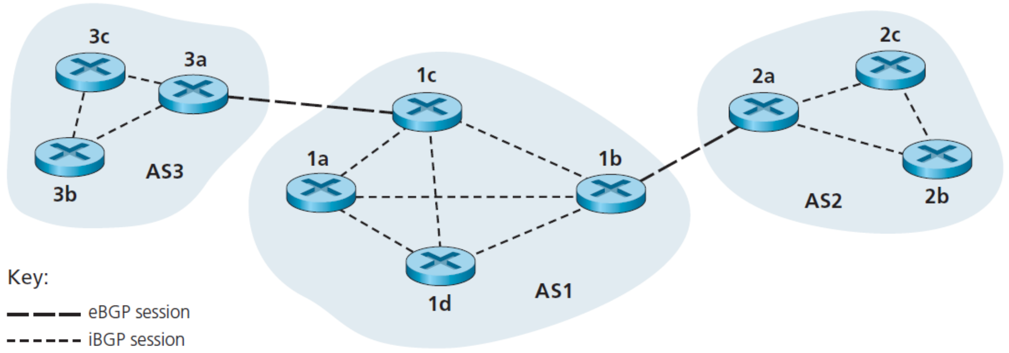


BGP loop Prevention mechanism



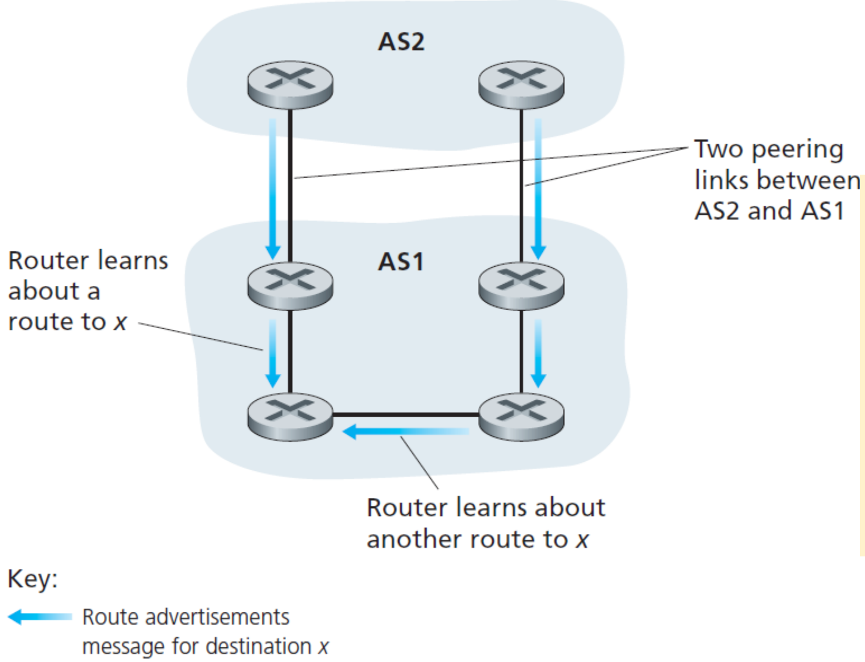
NEXT-HOP attributes in advertisements are used to determine which peering link to use

The NEXT-HOP is the router interface that begins the AS-PATH.



Consider what happens when the gateway router 3a in AS3 advertises a route to gateway router 1c in AS1 using eBGP. The route includes the advertised prefix, which we'll call x, and an AS-PATH to the prefix. This advertisement also includes the NEXT-HOP, which is the IP address of the router 3a interface that leads to 1c.

After learning about this route to x, router 1d may want to forward packets to x along the route, that is, router 1d may want to include the entry (x, I) in its forwarding table, where I is its interface that begins the least-cost path from 1d towards the gateway router 1c. To determine I, 1d provides the IP address in the NEXT-HOP attribute to its intra-AS routing module



AS1 and AS2 are connected by two peering links. A router in AS1 could learn about two different routes to the same prefix x. These two routes could have the same AS-PATH to x, but could have different NEXT-HOP values corresponding to the different peering links. Using the NEXT-HOP values and the intra-AS routing algorithm, the router can determine the cost of the path to each peering link, and then apply hot-potato routing (see Section 4.5.3) to determine the appropriate interface

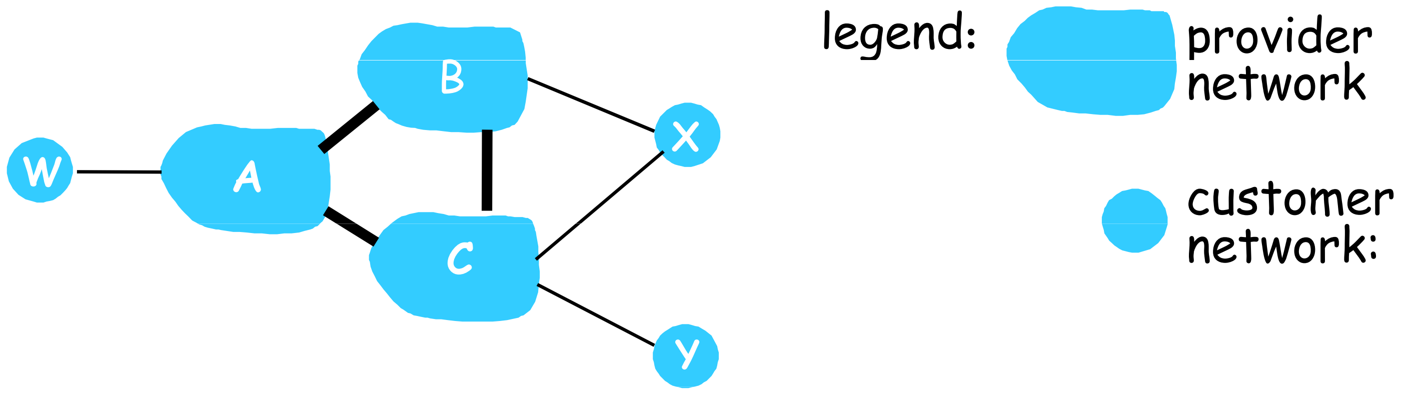
BGP route selection

- router may learn about more than 1 route to some prefix. Router must select route.
- elimination rules:
  1. local preference value attribute: policy decision- The local preference of a route could have been set by the router or could have been learned by another router in the same AS. This is a policy decision that is left up to the AS's network administrator.
  2. shortest AS-PATH (From the remaining routes (all with the same local preference value),)
  3. closest NEXT-HOP router: From the remaining routes (all with the same local preference value and the same AS-PATH length) hot potato routing Here, closest means the router for which the cost of the least-cost path, determined by the intra-AS algorithm, is the smallest.
  4. additional criteria

BGP messages

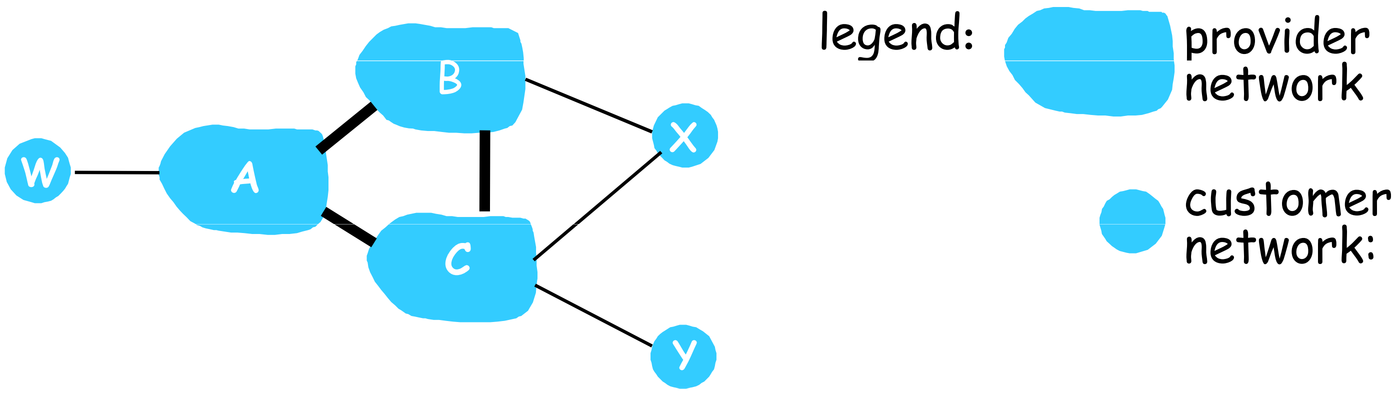
- BGP messages exchanged using TCP.
- BGP messages:
  - OPEN: opens TCP connection to peer and authenticates sender
  - UPDATE: advertises new path (or withdraws old)
  - KEEPALIVE keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - NOTIFICATION: reports errors in previous msg; also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

BGP routing policy



- 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)



- A advertises path AW to B
- B advertises path BAW to X
- Should B advertise path BAW to C?
  - 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!

A rule of thumb followed by commercial ISPs is that :

any traffic flowing across an ISP's backbone network must have either a source or a destination (or both) in a network that is a customer of that ISP;
otherwise the traffic would be getting a free ride on the ISP's network

Thank You