

# Introduction to Routing



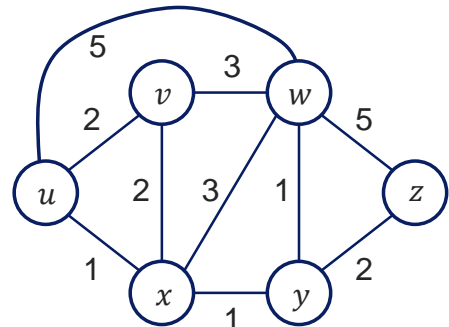
What is the least cost path between  $u$  and  $z$ ?

- Come up with a routing algorithm avoids periodic updates, or in other words discovers the route on demand
  - What is the effect of link changes in network topology?

Source Routing

- Come up with a routing algorithm based on node location?
  - Node  $x$  knows its location  $l(x)$
  - Node  $x$  can compute  $d_x(y) = \Delta(l(x), l(y))$
  - What happens when node mobility is high?

Geographic Routing



- Network model
  - Graph  $G = (V, E)$ 
    - $V = \{u, v, w, x, y, z\}$
    - $E = \{(u, v), (u, x), (v, x), \dots\}$
  - Link cost  $c(x, y)$ 
    - $c(x, y) = c(y, x)$
    - If  $(x, y) \notin E$  then  $c(x, y) = \infty$

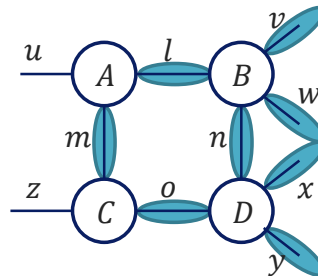
Mountains & Minds

213

## Routing Information Protocol (RIP)



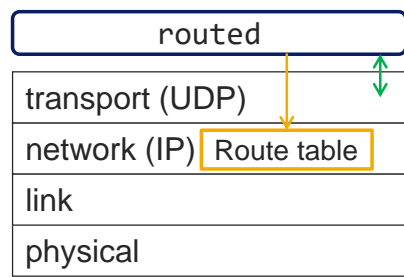
- Early inclusion in BSD-UNIX
- A **distance vector** protocol
  - Each hop cost 1, max path 15 hops
  - Poisoned reverse (infinity = 16 hops)
  - DVs exchanges with neighbors every 30 seconds
    - Up to 25 subnets advertised
    - Straight out of the routing table
  - Links removed and routes invalidated if no advertisement after 180 seconds



Dist. from router A to destination **subnets**:

subnet	hops
u	1
v	
w	
x	
y	
z	

RIP packets are encapsulated in UDP (port 530). Is that not inefficient?



Route adverts. to other routers

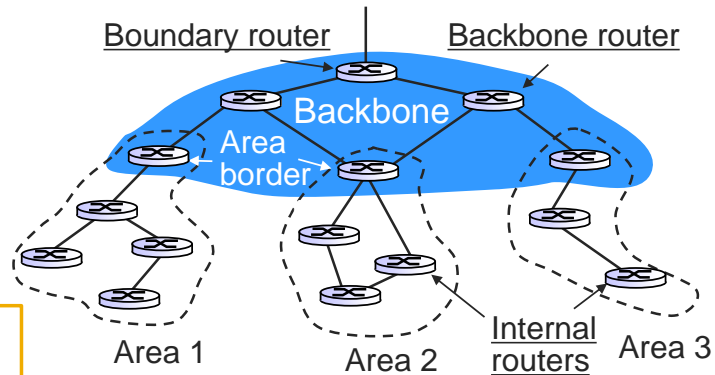
Mountains & Minds

214

# Open Shortest-Path First (OSPF)



- “Open”: publicly available
  - Cisco’s IGRP proprietary until 2015
- A **link state** protocol
  - LS information dissemination directly through IP (via raw sockets)
  - Topology map at every node
  - Route computation using Dijkstra’s
- Improvements over RIP
  - Security: OSPF routers and adverts authenticated
  - Multiple same-cost paths allowed
  - Support for multicast
  - Different cost metrics allowed
- Hierarchical OSPF
  - Local areas and backbone
  - LS adverts only within areas
  - Routers in area only know shortest path direction to other areas



How to configure weights for low latency paths? How about high bandwidth?

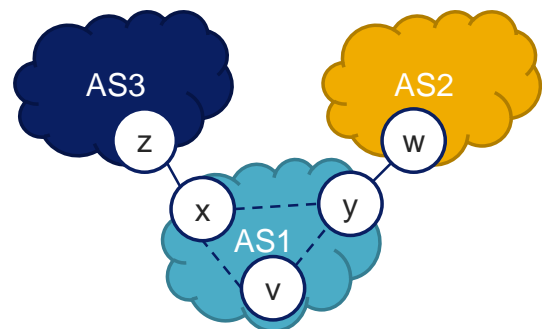
Mountains & Minds

215

## Hierarchical Routing



- Aggregate routers into regions
  - Autonomous Systems (AS)
  - One or more AS’s per ISP
- Routers in same AS run same routing protocol
  - **intra-AS** routing protocol
    - RIP, OSPF, IGRP
  - Routers in different AS can run different intra-AS routing protocol
- Gateway routers
  - At the “edge” of its own AS
  - Link to edge routers in another AS
  - Run **inter-AS** routing protocol
    - BGP



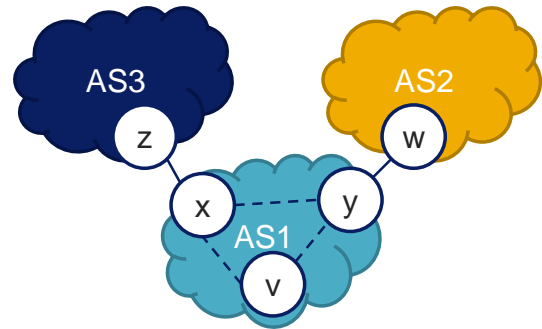
Mountains & Minds

216

# Why different Intra-AS and Inter-AS routing ?



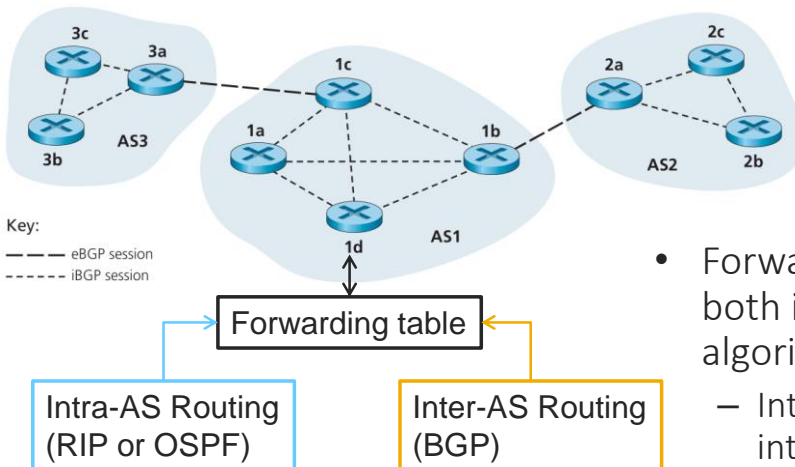
- Scale:
  - inter-AS: scalability – hierarchical addressing reduces routing table size and update traffic
  - intra-AS: flexibility – need to meet needs of different customers
- Performance:
  - inter-AS: scalability, or policy may dominate over performance
  - intra-AS: focus on performance and low cost
- Policy:
  - inter-AS: admin wants control over how its traffic is routed and who routes through its network.
  - intra-AS: single admin, so decisions can be made based on efficiency



Mountains &amp; Minds

217

## Interconnected ASes



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

Mountains &amp; Minds

218

# Interconnected ASes Example

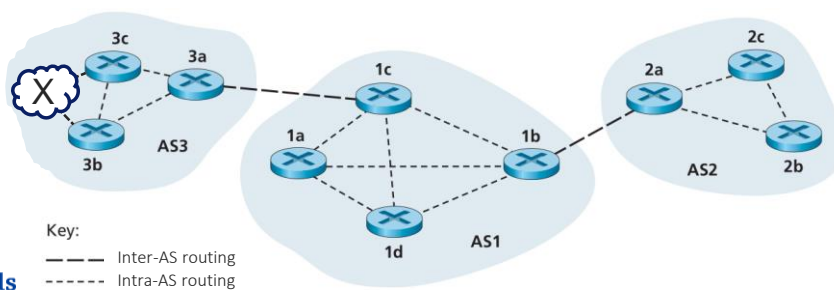


## Intra-AS routing

- Routers in AS3 exchange reachability information to subnet X
- Routing table configured at 3a
- And used by inter-AS routing

## Inter-AS routing

- Routers in AS1 receive reachability information for X from AS3 through inter-AS routing protocol
- Inter-AS routing propagates reachability from 1c to other AS1 border routers
- Inter-AS routing forwards AS3 reachability to AS2



Mountains &amp; Minds

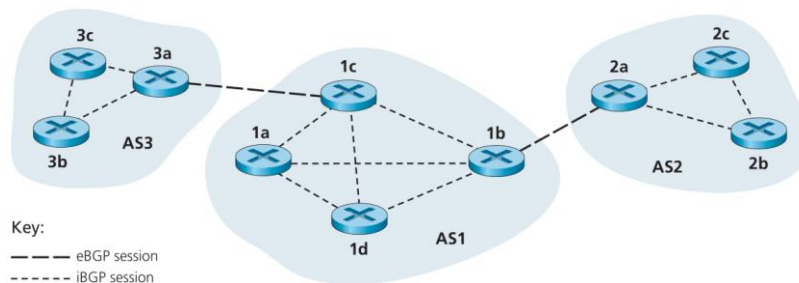
219

# Border Gateway Protocol (BGP)



- The *de facto* inter-domain routing protocol
  - “Glue that holds the Internet together”
  - Allows subnet to say: “I am here”
- BGP provides each AS a means to:
  - eBGP: obtain subnet reachability information from neighboring ASs.
  - iBGP: propagate reachability information to all AS-internal routers.
  - Determine “good” routes to other networks based on reachability information and policy.
- BGP messages
  - Over semi-permanent BGP connections between “peers”
    - May be over intra-AS paths
  - Prefix (aggregate)
    - 138.16.64/22
  - NEXT-HOP
    - IP of router interface
  - AS-PATH (path vector)
    - ASN routing path e.g., AS 67, AS 17

A bit like source routing!



Mountains &amp; Minds

220

# Border Gateway Protocol (BGP)

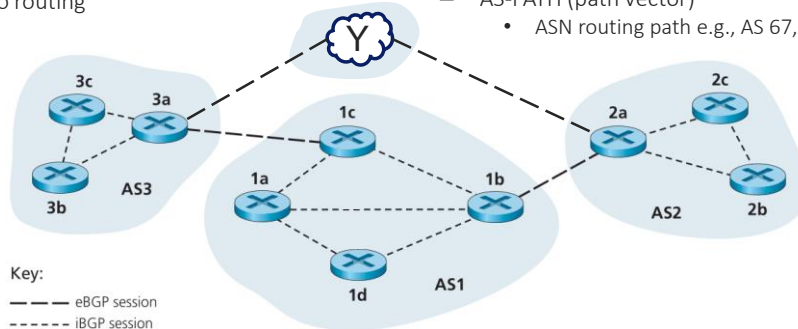


How does router 1d forward packets to subnet Y?

- Local preference value attribute (policy)
  - e.g., cheaper peering with AS2
- Shortest AS-PATH
- Closest NEXT-HOP router
  - Hot potato routing

- BGP messages
  - Over semi-permanent BGP connections between "peers"
    - May be over intra-AS paths
  - Prefix (aggregate)
    - 138.16.64/22
  - NEXT-HOP
    - IP of router interface
  - AS-PATH (path vector)
    - ASN routing path e.g., AS 67, AS 17

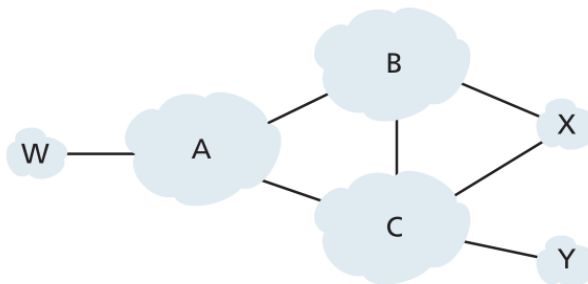
Route advertisements can be ignored by an *import policy*, e.g., never route through AS3



Mountains & Minds

221

## BGP routing policy



- A, B, C are provider networks
- X, W, Y are customers, but ISPs in their own right
- X is dual-homed: attached to two networks

- Should B advertise path BAW to C?
- Why does B not route to W through C?
- How does X prevent B routing to C through itself?

Mountains & Minds

222

# ICMP



- Internet Control Message Protocol (ICMP)
  - Used by hosts and routers to communicate network-level information
    - Error reporting: unreachable host, network, port, protocol
    - Echo request/reply (used by ping)
  - Network-layer “above” IP
    - ICMP messages carried in IP datagrams
    - One of the “upper layer” types
- 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	4	fragmentation required
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

Mountains &amp; Minds

223

## Traceroute and ICMP



- Source sends series of UDP segments to the destination
  - First set has TTL =1
  - Second set has TTL=2, etc.
  - Unlikely port number
- When nth set of datagrams arrives to nth router:
  - Router discards datagrams
  - Sends source ICMP messages (type 11, code 0)
  - ICMP messages includes name of router and IP address
- Stopping criteria:
  - when ICMP messages arrives, source records RTTs
  - UDP segment eventually arrives at destination host
  - Destination returns ICMP “port unreachable” message (type 3, code 3)
  - Source stops

Mountains &amp; Minds

224