

CSCI 491-01

Topics: Internet Programming

Fall 2008

## Network Layer

Derek Leonard

Hendrix College

November 17, 2008

Original slides copyright © 1996-2007 J.F Kurose and K.W. Ross

# Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

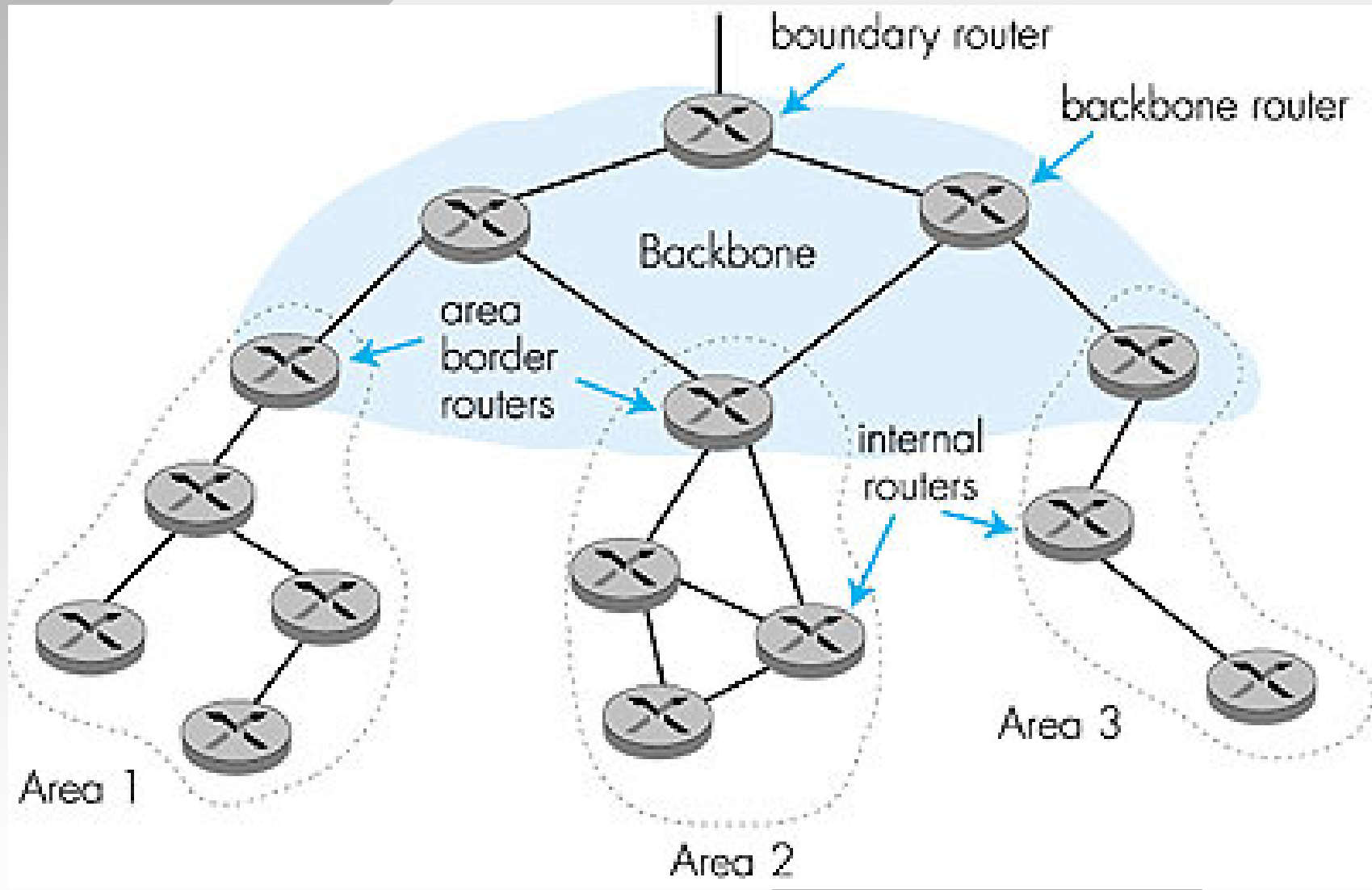
# OSPF (Open Shortest Path First)

- “Open”: publicly available
- Uses Link State algorithm
  - LS packet dissemination
  - Topology map at each node
  - Route computation using Dijkstra’s algorithm
- 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)
- **Multiple** same-cost **paths** allowed (only one path in RIP)
- Integrated uni- and **multicast** support:
  - Multicast OSPF (MOSPF) uses same topology database as OSPF
- **Hierarchical** OSPF in large domains

# Hierarchical OSPF



# Hierarchical OSPF

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

# Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

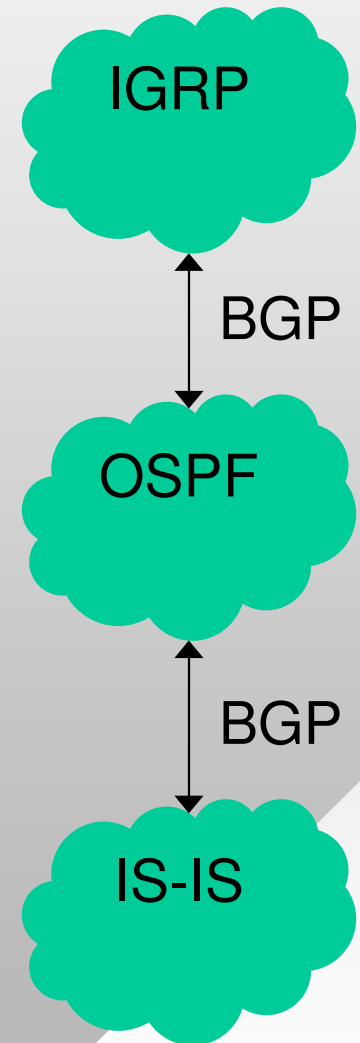
4.6 Routing in the Internet

- RIP
- OSPF
- **BGP**

4.7 Broadcast and multicast routing

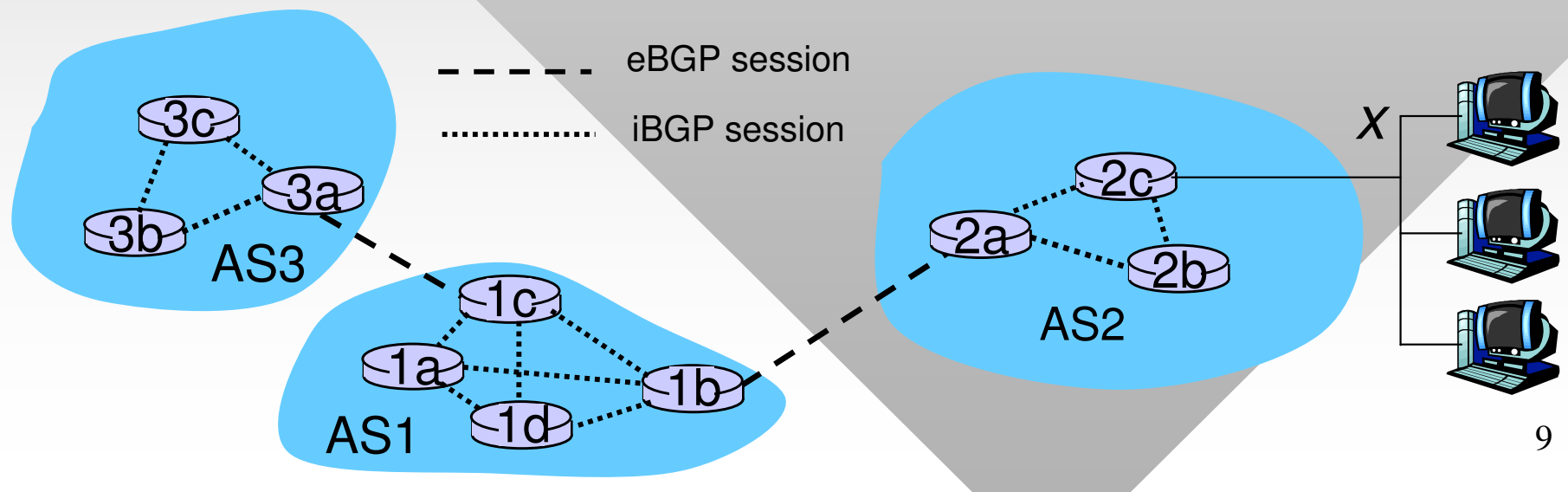
# Internet Inter-AS Routing: BGP

- **BGP (Border Gateway Protocol):** de facto standard for inter-AS (exterior) routing
- BGP provides each AS a means to:
  1. Obtain subnet reachability information from neighboring ASes
  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 the rest of the Internet: *“/ am here”*



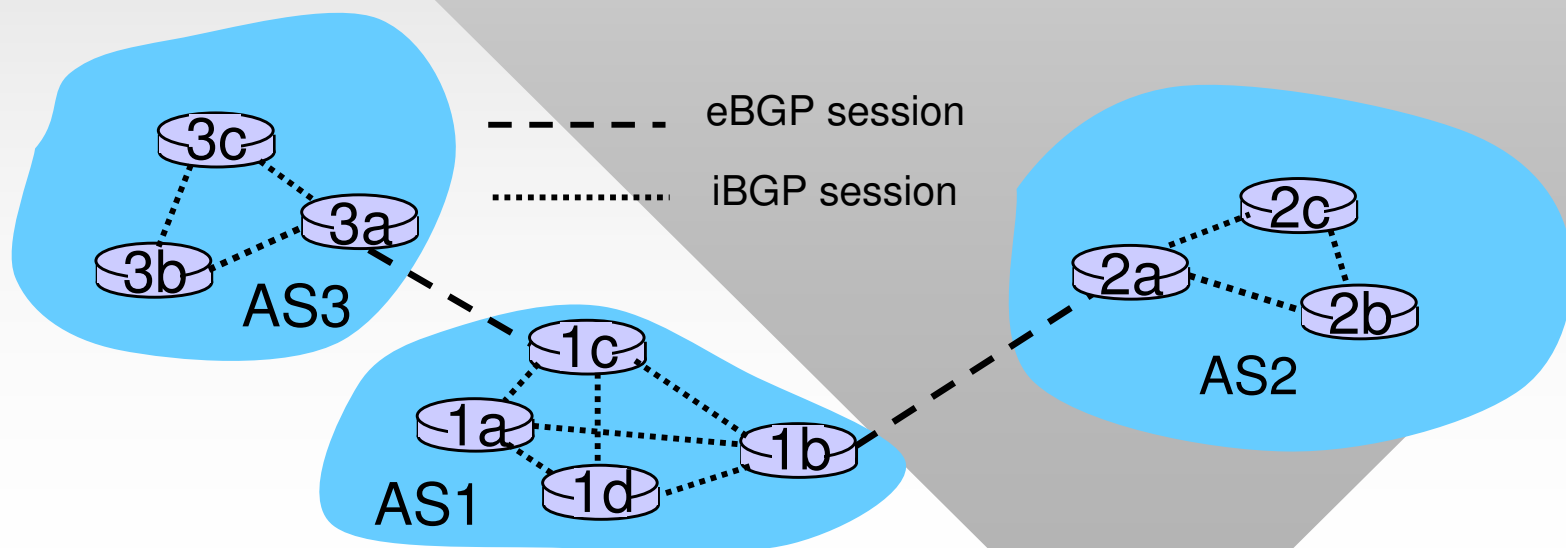
# BGP Basics

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



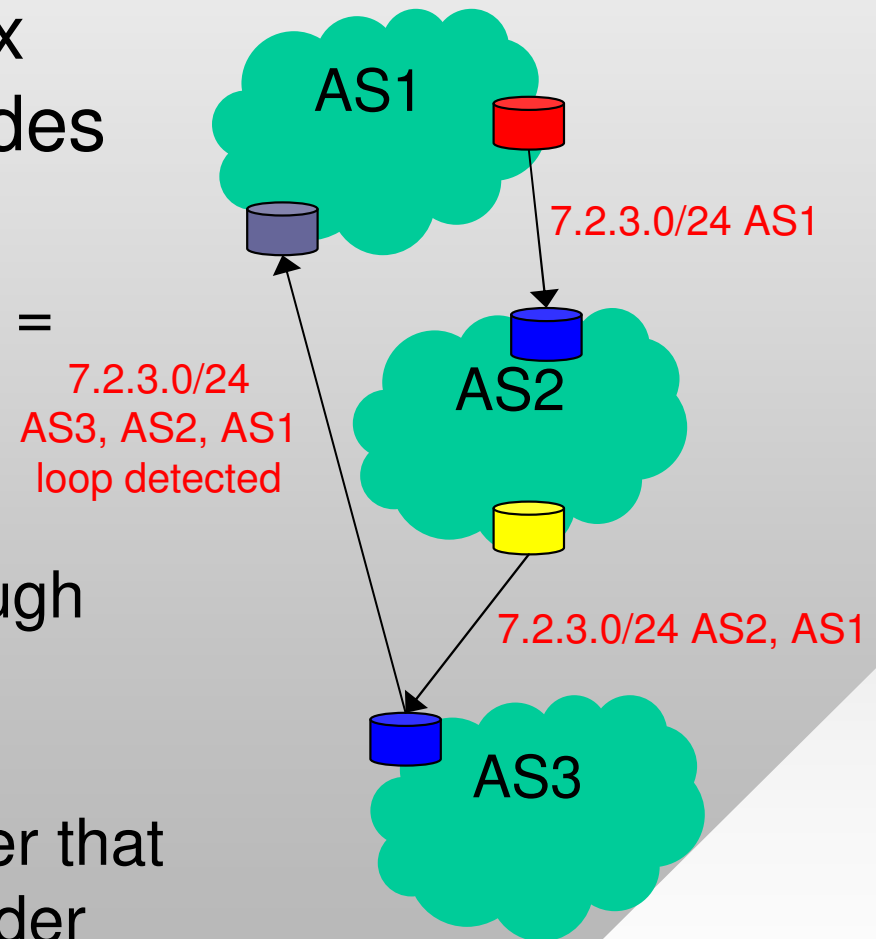
# Distributing Reachability Info

- With eBGP session between 2a and 1b, AS2 sends prefix reachability info to AS1
- 1b can then use iBGP to distribute this new prefix reachability info to all routers in AS1
- 1c may (if beneficial to AS1) re-advertise the new reachability info to AS2 over the 1c-3a 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 an IP prefix (i.e., subnet), message includes BGP **attributes**
  - Combination prefix + attributes = route
- Two important attributes:
  - **AS-PATH**: contains ASes through which the advert for the prefix passed (latest AS first)
  - **NEXT-HOP**: indicates the router that leads to this prefix (usually border router of the AS that advertised prefix; multiple values possible)

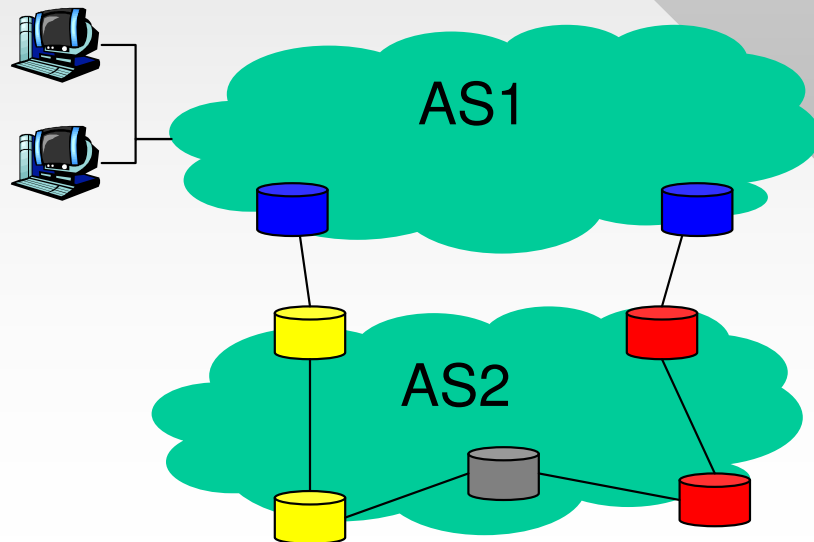
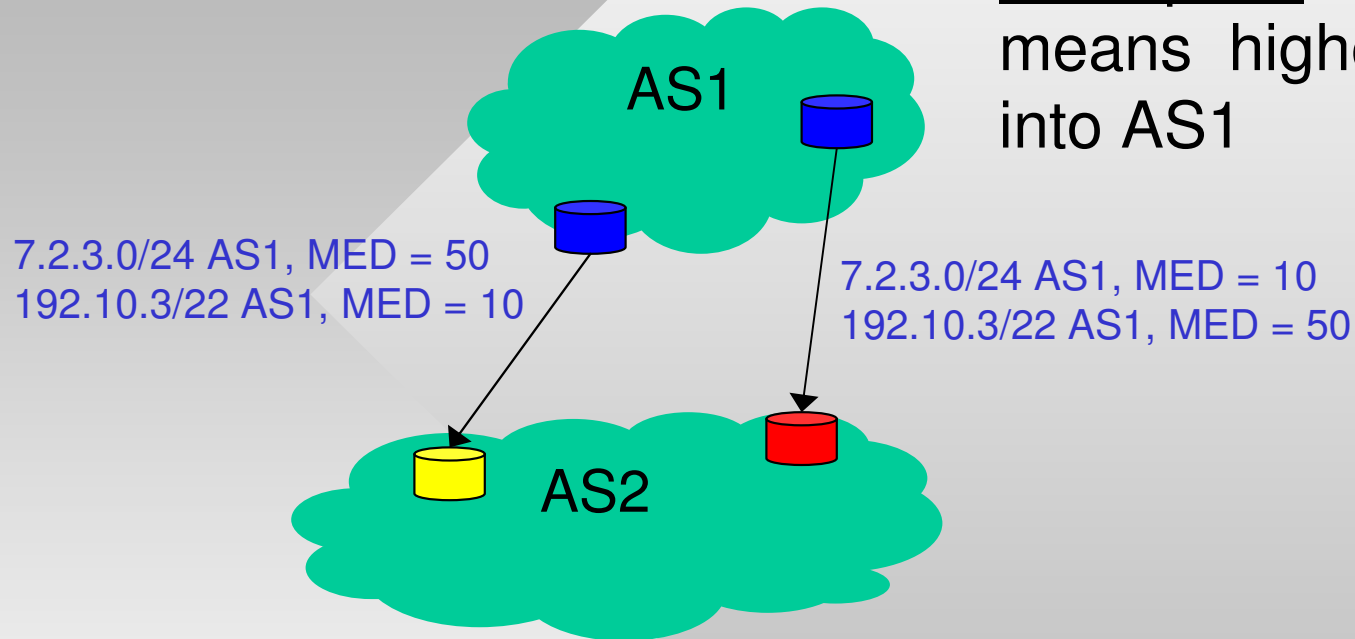


# BGP Route Selection

- When gateway router receives route advert, it uses an **import policy** to accept/decline
  - Filters and rules decide allowed/prohibited routes
- Router may learn about more than 1 route to some prefix, how do they decide which one is better?
  - **Local preference** attribute: policy decision of this AS that assigns different weight to various exit points
  - **Multi-exit discriminator (MED)** attribute: policy of foreign AS that assigns different weight to different incoming points
  - Shortest AS-PATH
  - Closest NEXT-HOP router: hot potato routing

# BGP Examples

Example 1: different MED (lower # means higher priority) for paths into AS1

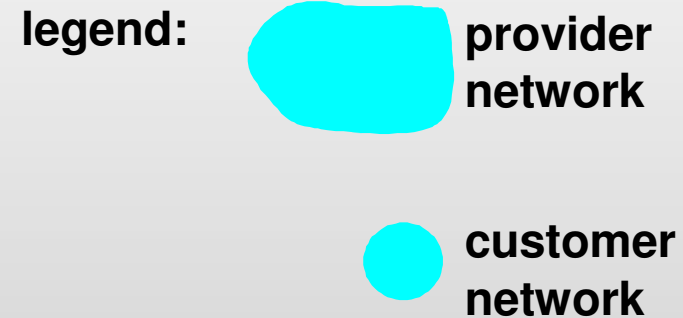
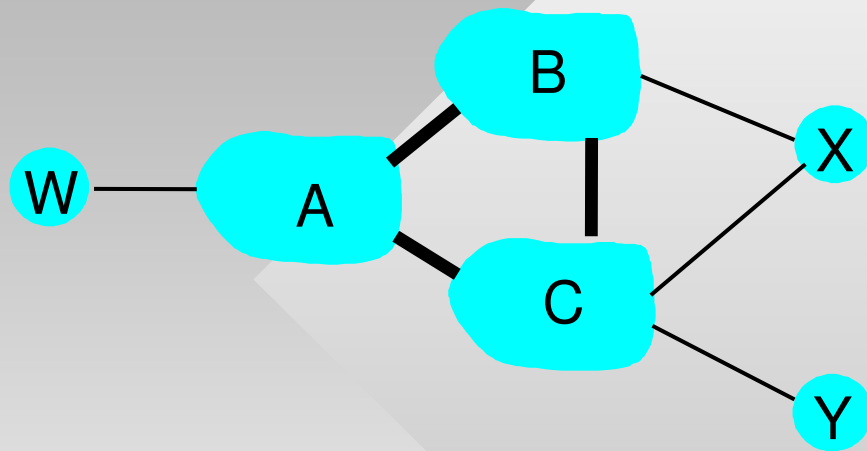


Example 2: hot-potato routing in AS2 (red routers exit on the right, yellow on the left)

# BGP Messages

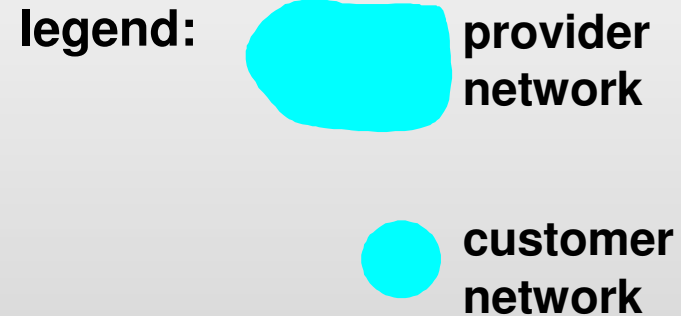
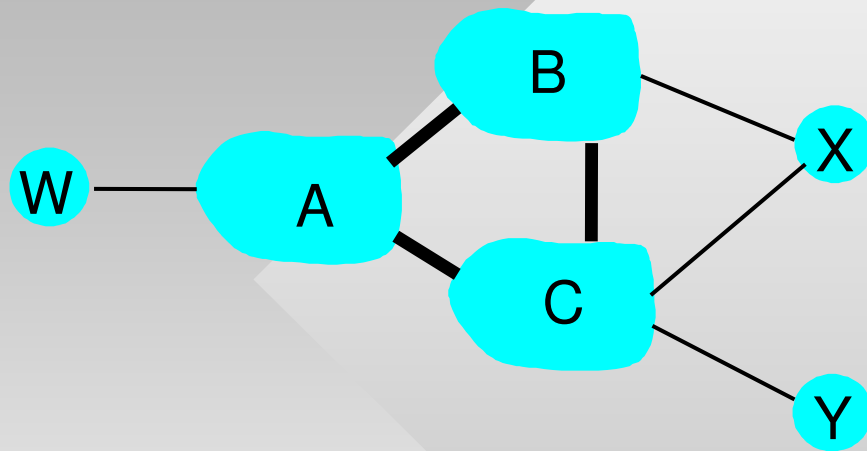
- BGP messages exchanged using TCP on port 179
  - Application-layer protocol
- BGP messages (binary header):
  - **OPEN**: opens TCP connection to peer and authenticates sender
  - **UPDATE**: advertises new path (or withdraws old)
  - **KEEPALIVE** keeps connection alive in absence of UPDATES; also 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 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 to B and C the path AW
- B advertises to X the path BAW
- Should B advertise to C the path BAW?
- Not unless B has agreed to route C's traffic!
  - 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
- ISPs want to route *only* to/from their customers!