Internet Protocols
Fall 2005

Lectures 11-12

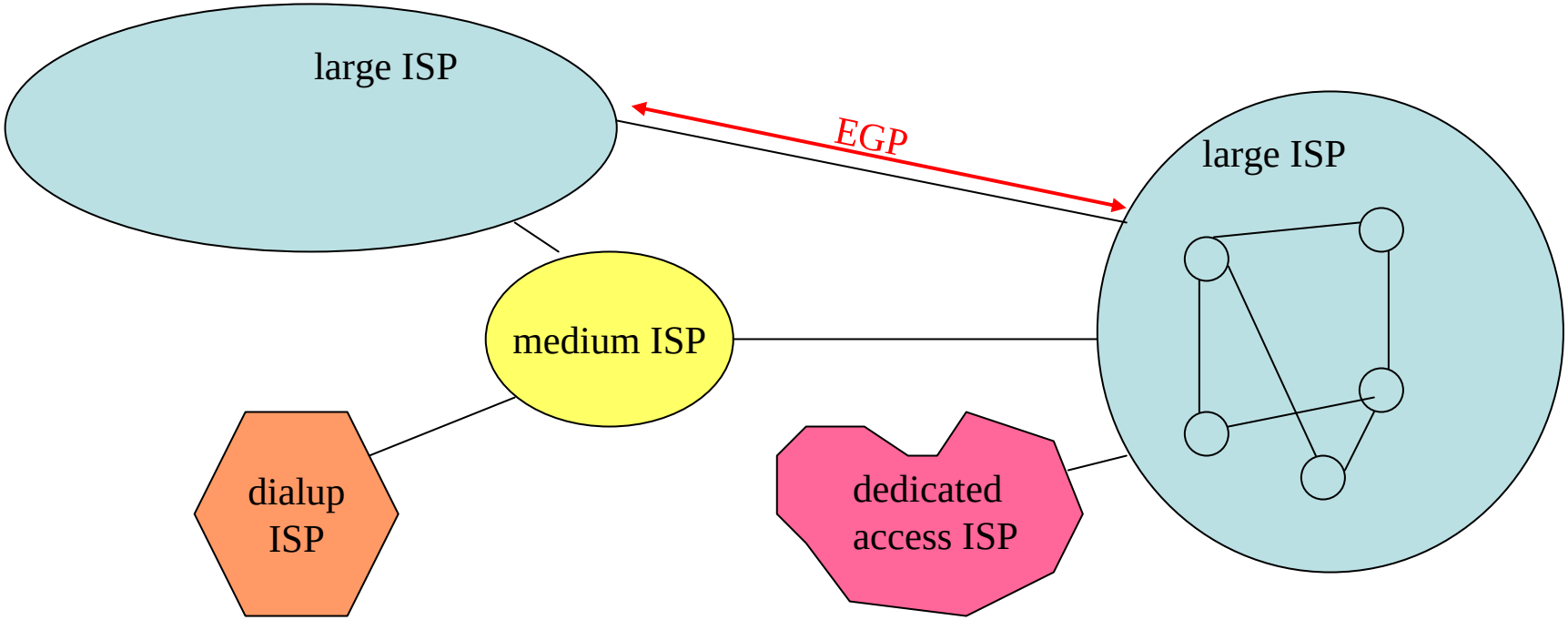
Inter-domain routing, mobility support, multicast routing

Andreas Terzis

Outline

- Inter-domain Internet Routing
  - BGP
- Routing for mobile nodes
- Multicast routing

Internet Routing Hierarchy

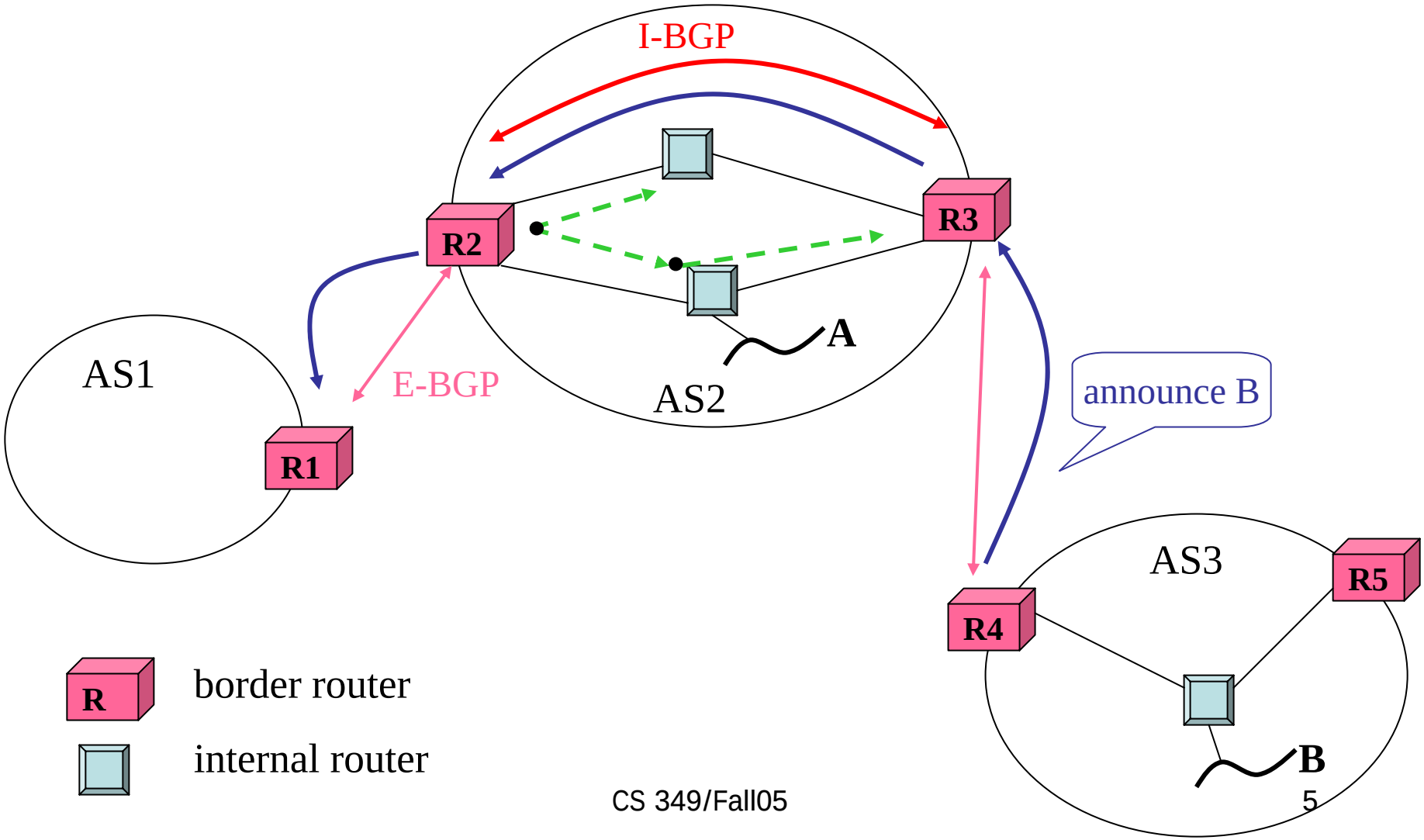


- AS (Autonomous System) - a collection of routers under the same technical and administrative domain.
- EGP (External Gateway Protocol) - used between two AS's to allow them to exchange routing information so that traffic can be forwarded across AS borders. Example: BGP

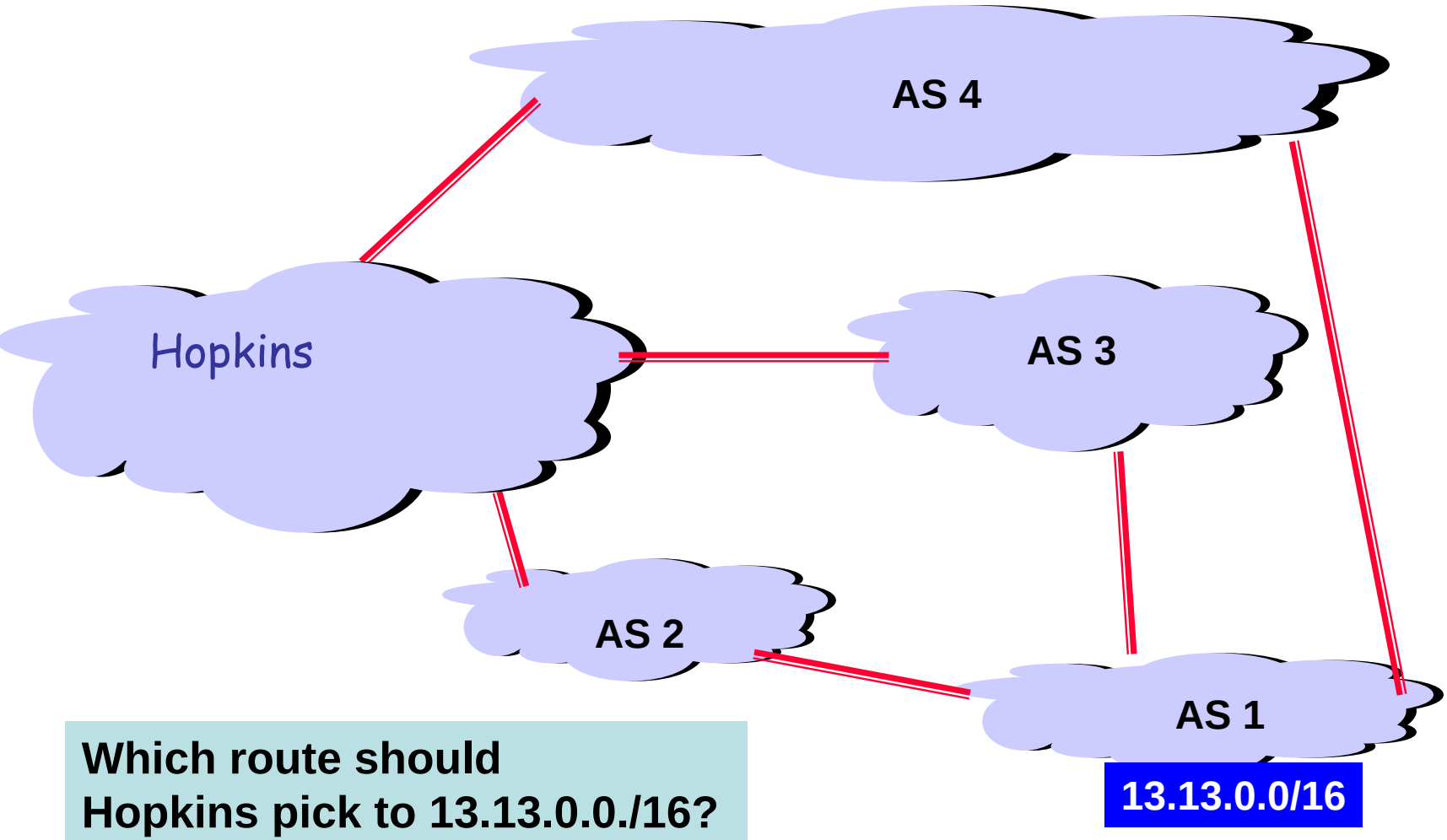
Questions

- How do we “tie” the two routing levels together?
- What are the special constraints of inter-domain routing?
- What is the routing algorithm used at the inter-domain level?

Two-Tier Routing



So Many Choices



Which route should Hopkins pick to 13.13.0.0./16?

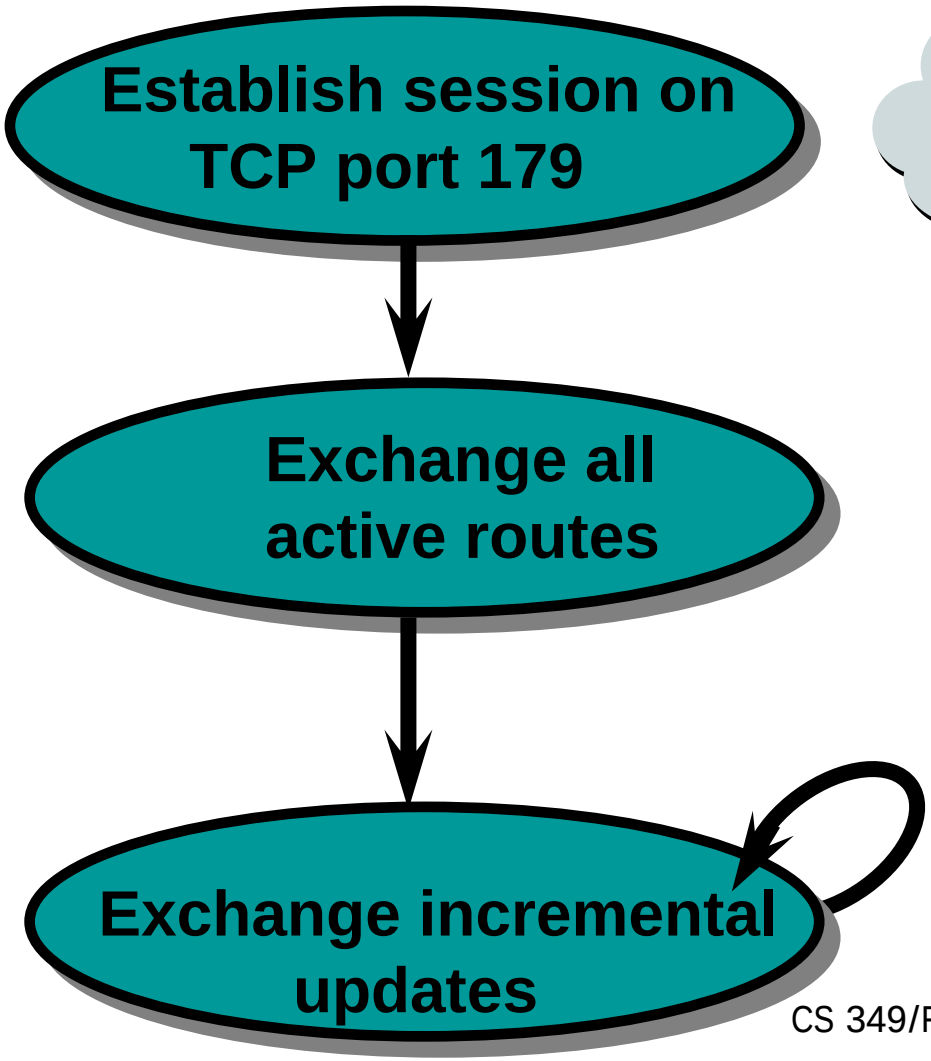
Choice of Routing Algorithm

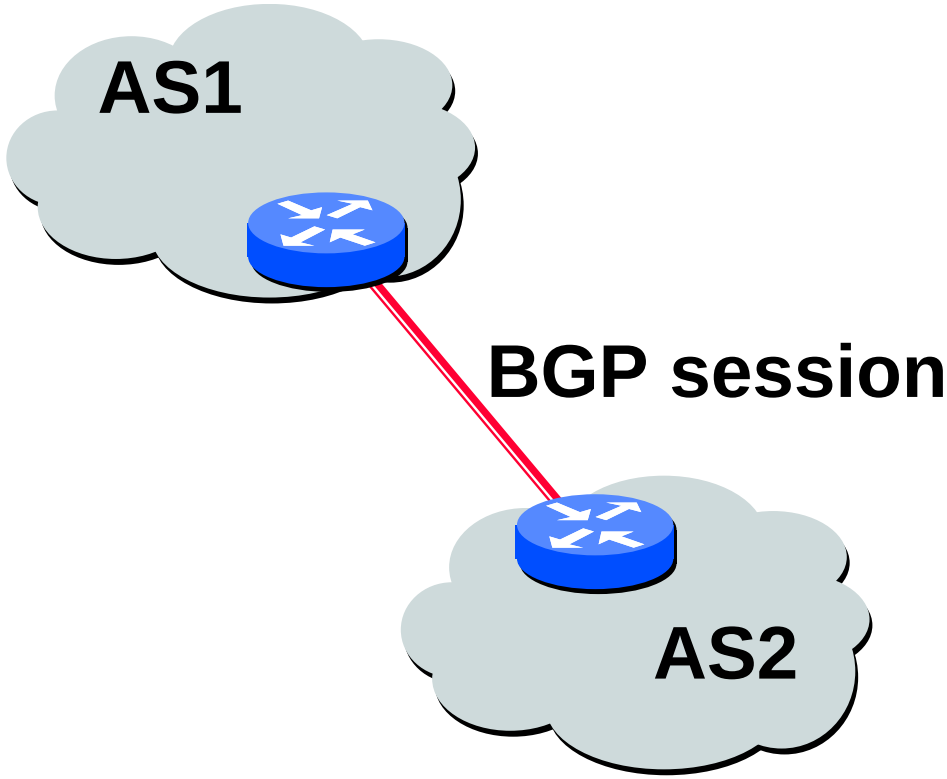
- Constraints:
  - scaling
  - Support policy-based route selection
- Link-state?
  - requires sharing of complete network informatin
  - information exchanges doesn’t scale
  - can’t express policy
- Distance Vector?
  - scales and retains privacy
  - can’t implement policy
  - can’t avoid loops if shortest paths not taken

Path Vector Protocol

- Distance vector algorithm with extra information
  - For each route, store the complete path (ASs)
  - No extra computation, just extra storage
- Advantages:
  - Can make policy choices based on set of ASs in path
  - Can easily avoid loops

BGP Operations (Simplified)





While connection is ALIVE exchange route UPDATE messages

Four Types of BGP Messages

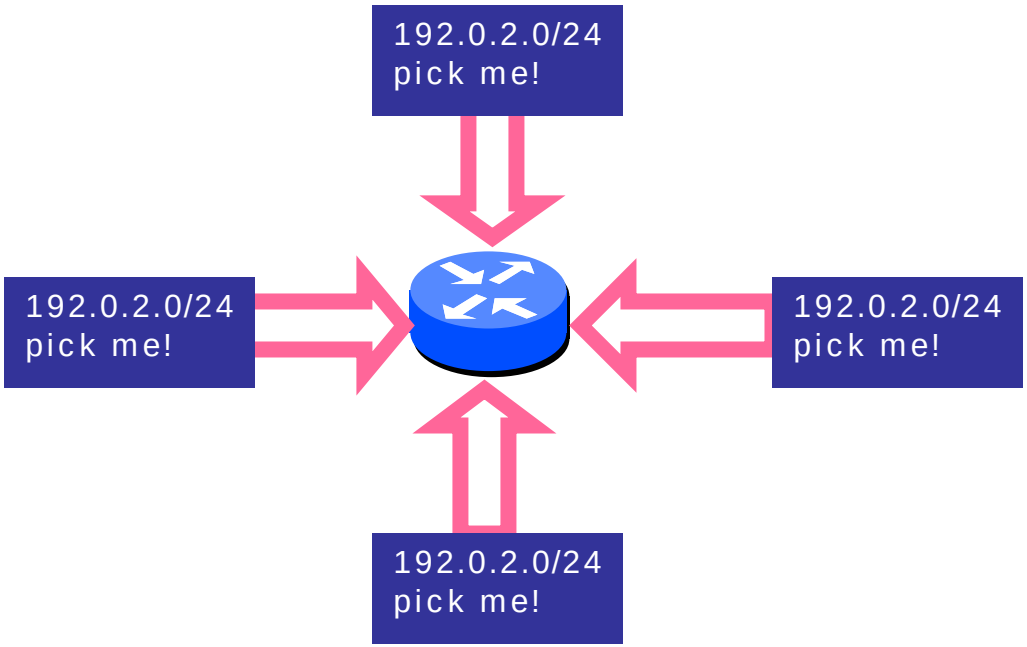
- Open : Establish a peering session.
- Keep Alive : Handshake at regular intervals.
- Notification : Shuts down a peering session.
- Update : Announcing new routes or withdrawing previously announced routes.

Announcement
=
prefix + <u>attributes values</u>

BGP Attributes

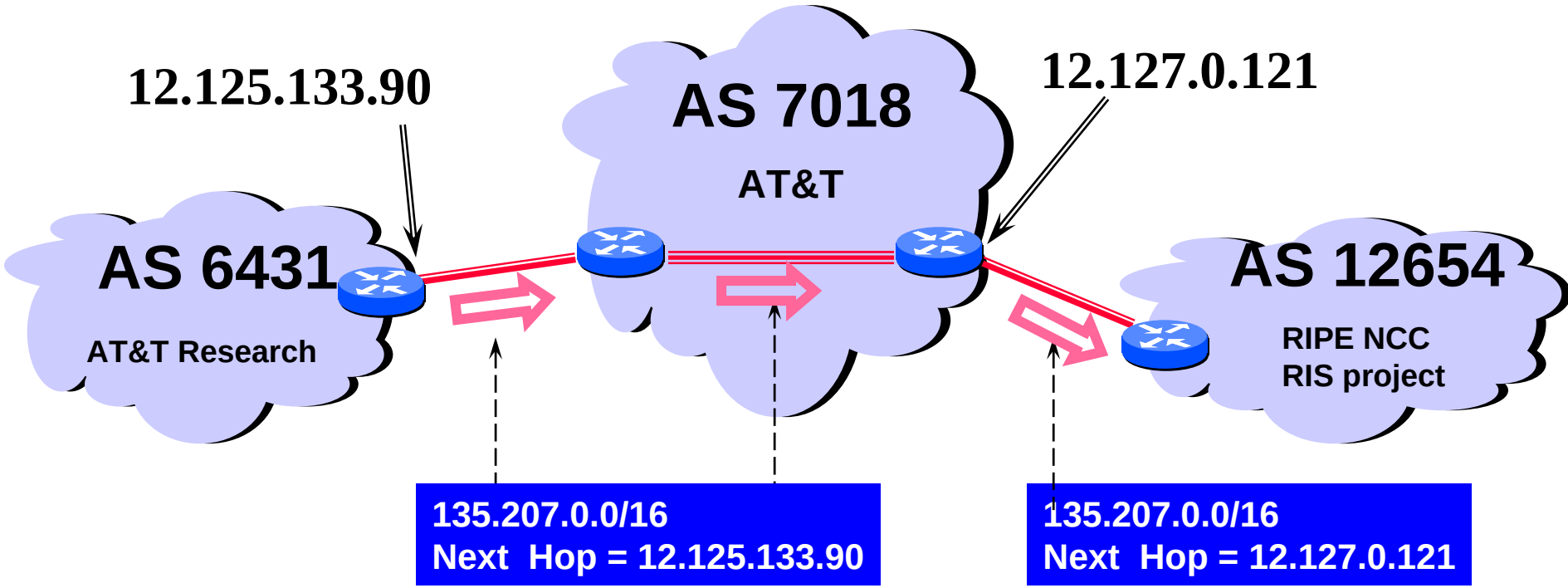
Value	Code	Reference
1	ORIGIN	[RFC1771]
2	AS_PATH	[RFC1771]
3	NEXT_HOP	[RFC1771]
4	MULTI_EXIT_DISC	[RFC1771]
5	LOCAL_PREF	[RFC1771]
6	ATOMIC_AGGREGATE	[RFC1771]
7	AGGREGATOR	[RFC1771]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC2796]
10	CLUSTER_LIST	[RFC2796]
11	DPA	[Chen]
12	ADVERTISER	[RFC1863]
13	RCID_PATH / CLUSTER_ID	[RFC1863]
14	MP_REACH_NLRI	[RFC2283]
15	MP_UNREACH_NLRI	[RFC2283]
16	EXTENDED COMMUNITIES	[Rosen]
...		
255	reserved for development	

Attributes are Used to Select Best Routes



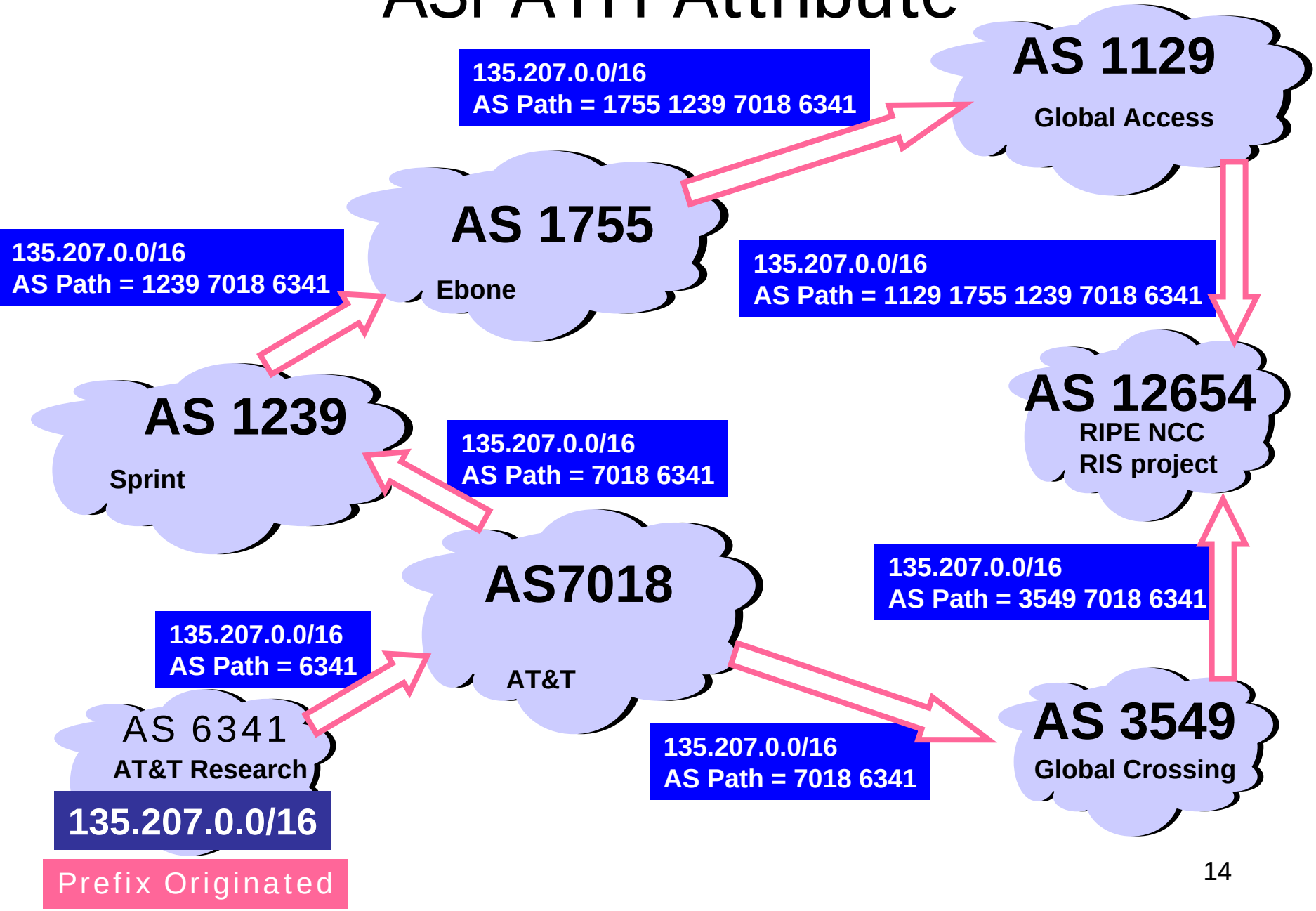
Given multiple routes to the same prefix, a BGP speaker must pick at most <u>one</u> best route
(Note: it could reject them all!)

BGP Next Hop Attribute



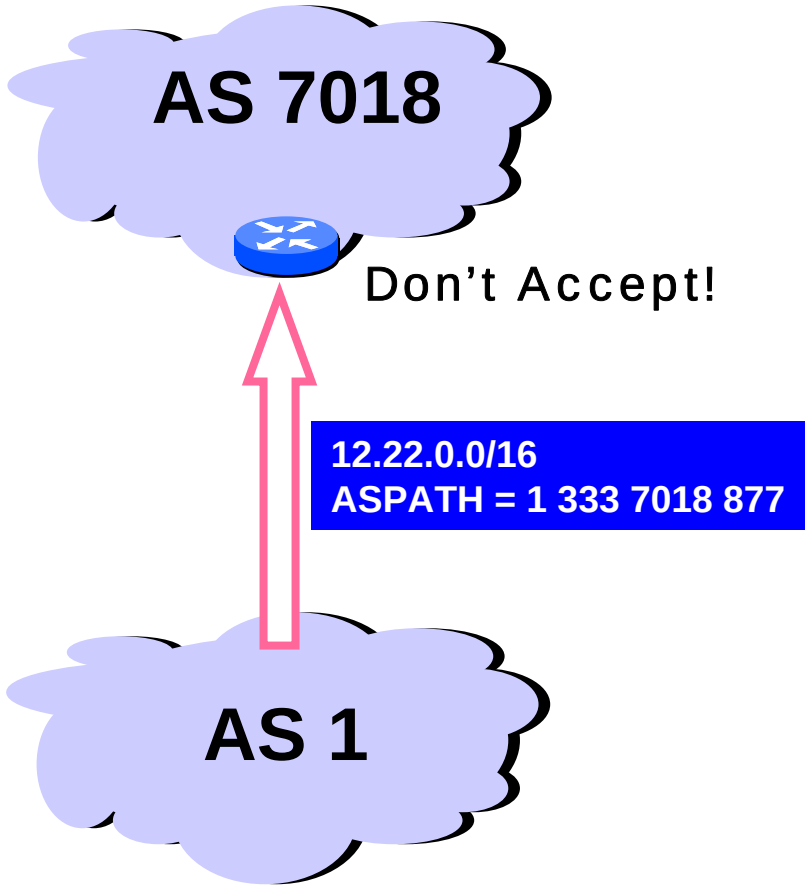
Every time a route announcement crosses an AS boundary, the Next Hop attribute is changed to the IP address of the border router that announced the route.

ASPATH Attribute

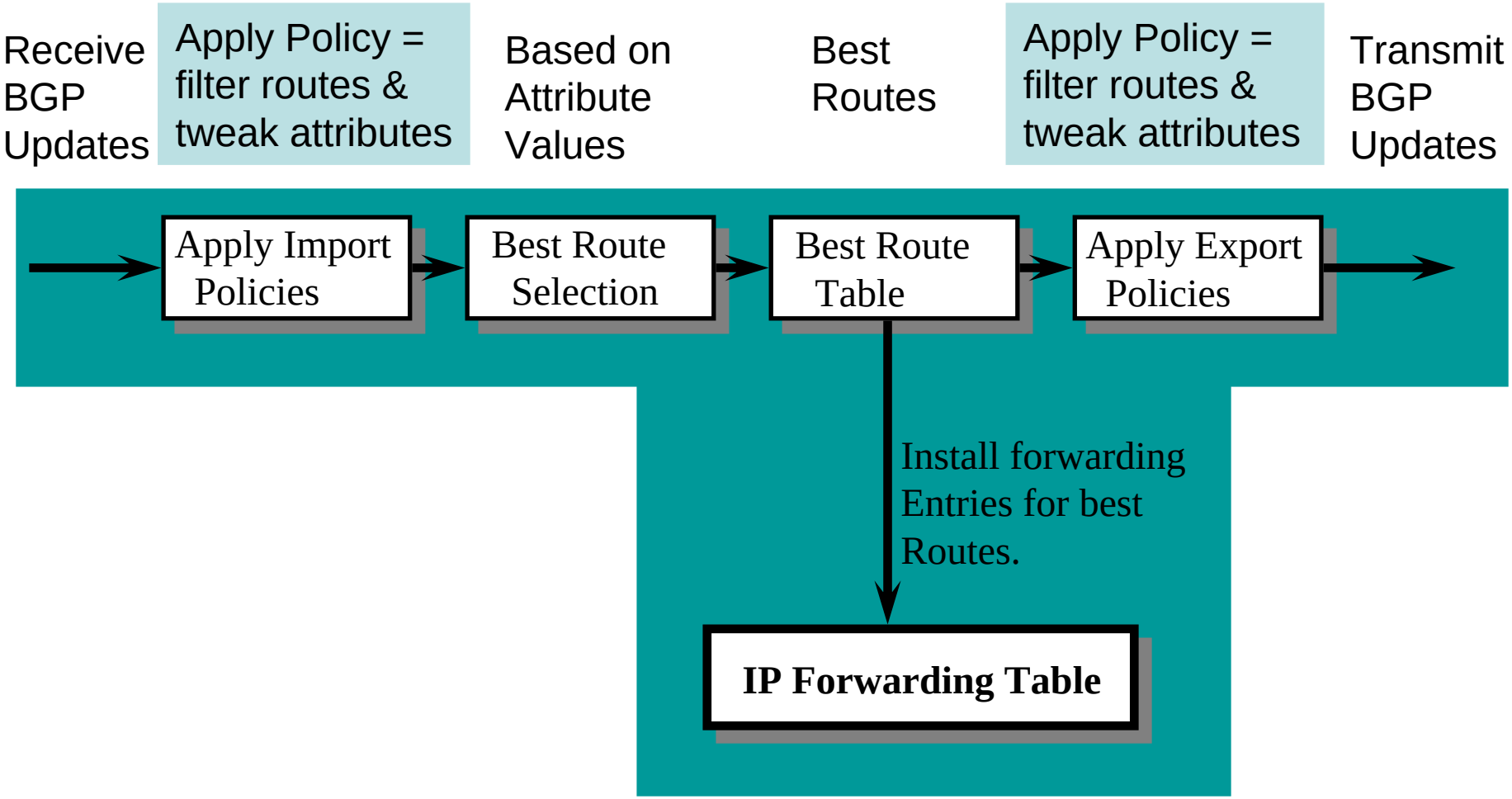


Interdomain Loop Prevention

- BGP at AS YYY will never accept a route with ASPATH containing YYY



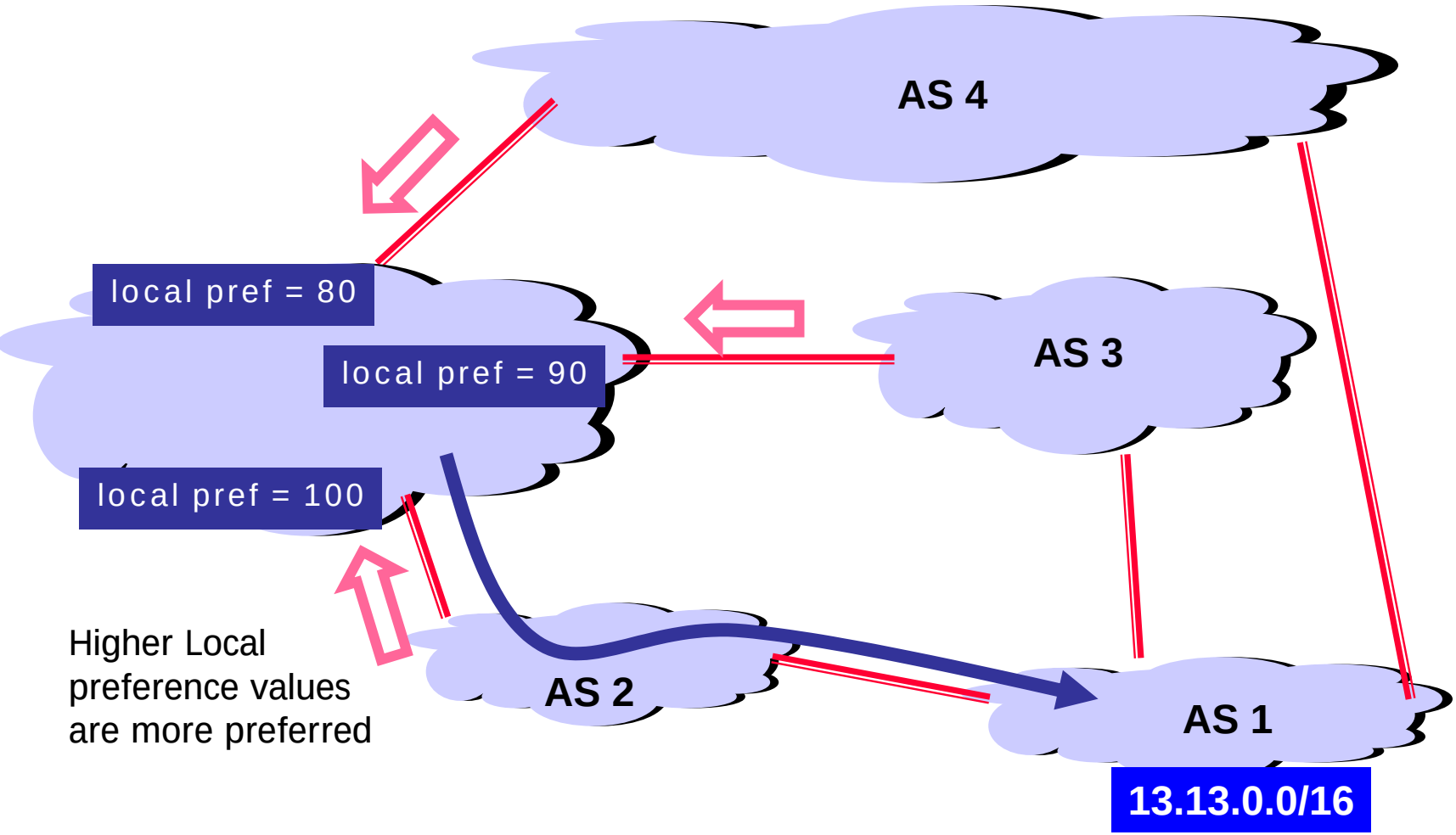
BGP Route Processing



Route Selection Summary

Highest Local Preference	Enforce relationships
Shortest ASPATH Lowest MED i-BGP < e-BGP Lowest IGP cost to BGP egress	traffic engineering
Lowest router ID	Throw up hands and break ties

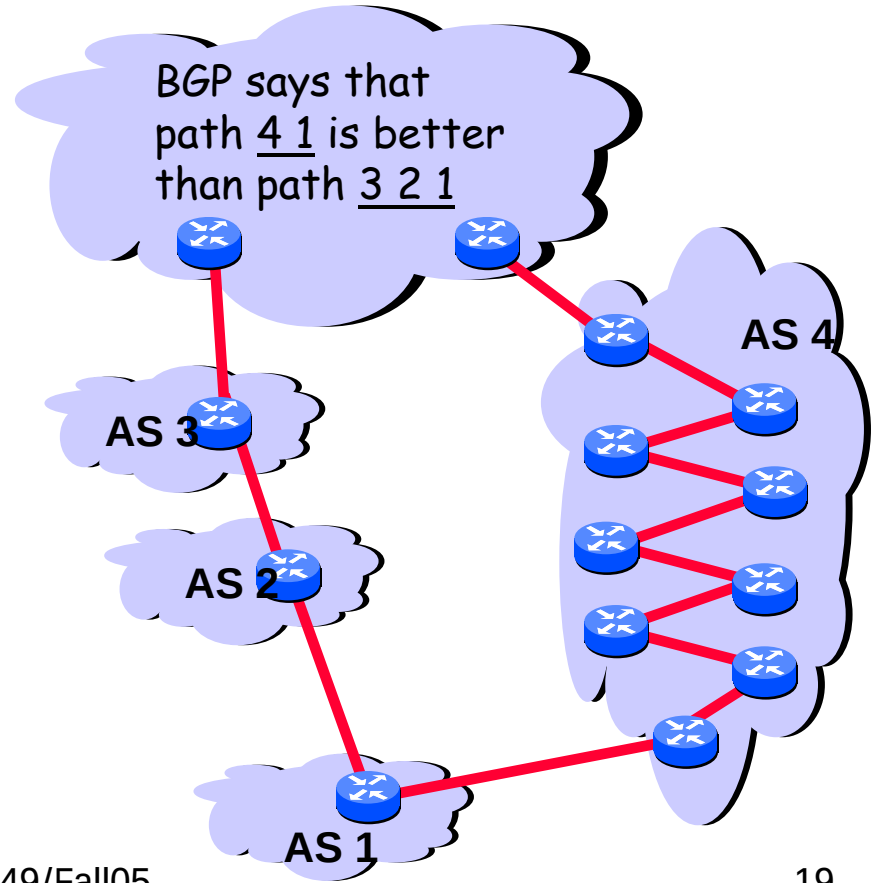
Local Preference



Higher Local preference values are more preferred

Shortest AS Path

- Route Selection Algorithm
  1. Local Pref
  2. Shortest AS Path
- Shortest AS Path != shortest path



Outline

- Inter-domain Internet Routing
  - BGP
- Routing for mobile nodes
- Multicast routing

Routing for Mobile Hosts

- Internet routing does not support mobile hosts
  - IP address is divided into network part and host part
  - Rest of the Internet knows how to forward packets to the network end-host is connected to
  - If end-host moves packets are delivered in the wrong network

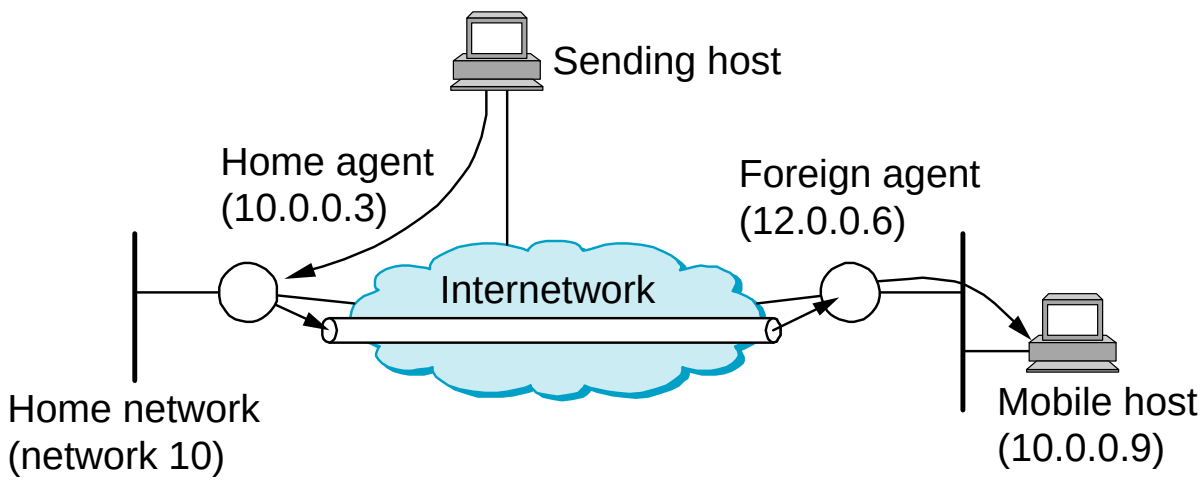
Partial Solution

- As mobile node moves it acquires different IP addresses
  - DHCP
- This is adequate for some applications
  - Web browsing, etc
- Does not work for other applications
  - Mobile host is the server
  - Long lasting connections

Mobile IP

- Solution proposed to solve the problem of mobility
- Does not require changes in most IP infrastructure
- Introduces three new entities
  - Home Agent (HA)
  - Foreign Agent (FA)
  - Mobile Host (MH)

Example


Sending host
Home agent
(10.0.0.3)
Foreign agent
(12.0.0.6)
Internetwork
Home network
(network 10)
Mobile host
(10.0.0.9)

Sub-problems

- How does the home agent intercept packets that are destined to the mobile node?
- How does the home agent deliver the packet to the foreign agent
- How does the foreign agent deliver the packet to the mobile node?

Home Agent Intercept

- Packets from Corresponding Agents (CAs) arrive at the home network
- Home agent issues proxy ARP
  - Replies with its own MAC to queries for the Mobile Hosts IP
  - Packets are delivered to the Home Agent
- When Mobile Host returns to the home network
  - Gratuitous ARP

Tunneling

- A *tunnel* exists between HA and FA
  - Home agent encapsulates IP packets for the Mobile Host using the FA's IP address
  - Rest of the Internet forwards packet as if it was destined for the FA
  - When packet arrives at the FA, it decapsulates the packet

FA Delivery

- FA bypasses regular IP routing for registered MHs and delivers directly to the MAC address
- MH can be the FA
- What about packets sourced by the MH?
  - Sent from the remote network with the MH's permanent address
  - Reverse tunnel

Route Optimization

- Packets from CA ->HA->FA follow suboptimal route
- If CA is upgraded it can create tunnel directly to the FA
  - HA sends binding update to the CA when packets arrive to it
  - FA can also send binding updates
- Security implications?

Outline

- Inter-domain Internet Routing
  - BGP
- Routing for mobile nodes
- Multicast routing

Multicast Routing

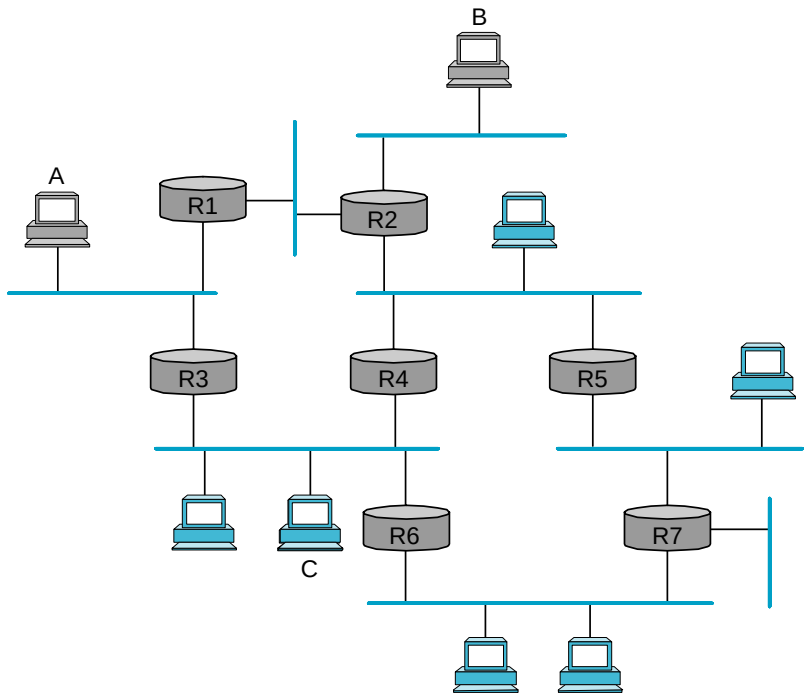
- Applications that want to send same data to multiple destinations
  - Audio, Video, Stock Updates
- Multicast model
  - Sender sends to Multicast Address
  - Receivers join Multicast Address
    - IGMP
  - Packets are delivered to all hosts that joined the multicast address
- How do the receivers know which multicast group to join?

Distance-Vector Multicast

- DVMRP uses a two step approach
  - Flooding mechanism that delivers packets to the whole network
  - *Pruning* networks that have no receivers

Reverse-Path Broadcast

- Flood packets from S to everyone
  - Router forwards packets to all interfaces (minus incoming) if packet arrives on correct interface
    - Interface R would use to send packets to S
  - Packet will be sent multiple times if multiple routers connected on same LAN
    - Designate parent for each network
    - Router with shortest distance to S



Reverse-Path Multicast

- RPB delivers packets to all networks
  - Obviously undesirable
- Leaf networks know if any interested receivers are present
- Routers propagate this information to upstream neighbors
  - Tree is pruned
- What happens if a new receiver joins at a pruned network?

Alternatives

- DVMRP not appropriate for *sparse* groups
- Protocol Independent Multicast (PIM) solves this issue
  - Rendezvous Point (RP)