



# PART III

## **Implementing Inter-Network Relationships with BGP**

# Routing Protocols

## ■ Autonomous System



# BGP-4

- **BGP** = Border Gateway Protocol
  - Is a Policy-Based routing protocol
  - Is the de facto EGP of today's global Internet
  - Relatively simple protocol, but **configuration is complex** and the entire world can see, and be impacted by, your mistakes.

- **1989 : BGP-1 [RFC 1105]**
  - Replacement for EGP (1984, RFC 904)
- **1990 : BGP-2 [RFC 1163]**
- **1991 : BGP-3 [RFC 1267]**
- **1995 : BGP-4 [RFC 1771]**
  - Support for Classless Interdomain Routing (CIDR)

# Routing problems

## YouTube outage underscores big Internet problem

The screenshot shows a web browser window with the address bar displaying "YouTube outage underscores big Internet problem | InfoWorld | News | 2008-02-26 | By R...". The browser's menu bar includes "File", "Edit", "View", "History", "Bookmarks", "Tools", and "Help". The page content features a navigation bar with links: "HOME", "NEWS", "TEST CENTER", "TECHNOLOGIES", "BLOGS", "AUDIO/VIDEO", "EVENTS", "AWARDS", "NEWSLETTERS", and "RESOURCES". The main headline is "YouTube outage underscores big Internet problem", with a sub-headline: "BGP data intended to block access to YouTube within Pakistan was accidentally broadcast to other service providers, causing a widespread YouTube outage". The byline reads "By Robert McMillan, IDG News Service February 26, 2008". Below the byline are links for "Talkback", "E-mail", "Printer Friendly", and "Reprints", along with a "Text Size" selector. The article text begins with "Sunday's inadvertent disruption of Google's YouTube video service underscores a flaw in the Internet's design that could some day lead to a serious security problem, according to networking experts." A sidebar on the left contains a section titled "Developers on steroids" with a link "Save months building custom Web apps with application generation". The main text continues: "The issue lies in the way ISPs share BGP (Border Gateway Protocol) routing information. BGP is the standard protocol used by routers to find computers on the Internet, but there is a lot of BGP routing data available. To simplify things, ISPs share this kind of information among each other." At the bottom of the browser window, a status bar shows "Done".

HOME NEWS TEST CENTER TECHNOLOGIES BLOGS AUDIO/VIDEO EVENTS AWARDS NEWSLETTERS RESOURCES

### YouTube outage underscores big Internet problem

**BGP data intended to block access to YouTube within Pakistan was accidentally broadcast to other service providers, causing a widespread YouTube outage**

By Robert McMillan, IDG News Service  
February 26, 2008

Talkback E-mail Printer Friendly Reprints Text Size A A

Sunday's inadvertent disruption of Google's YouTube video service underscores a flaw in the Internet's design that could some day lead to a serious security problem, according to networking experts.

**Developers on steroids**  
Save months building custom Web apps with application generation

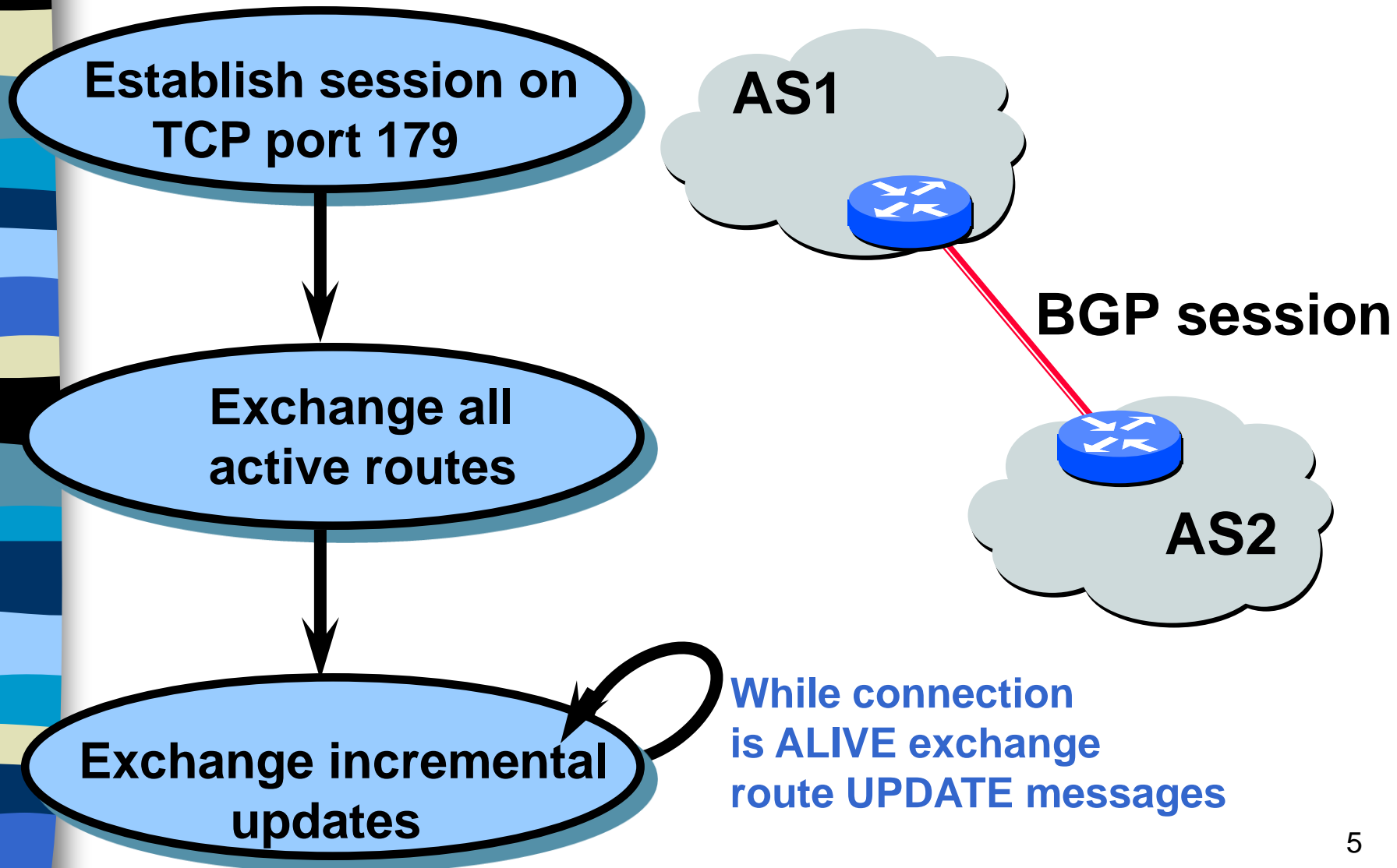
The issue lies in the way ISPs share BGP (Border Gateway Protocol) routing information. BGP is the standard protocol used by routers to find computers on the Internet, but there is a lot of BGP routing data available. To simplify things, ISPs share this kind of information among each other.

videos for some countries outage.

Done

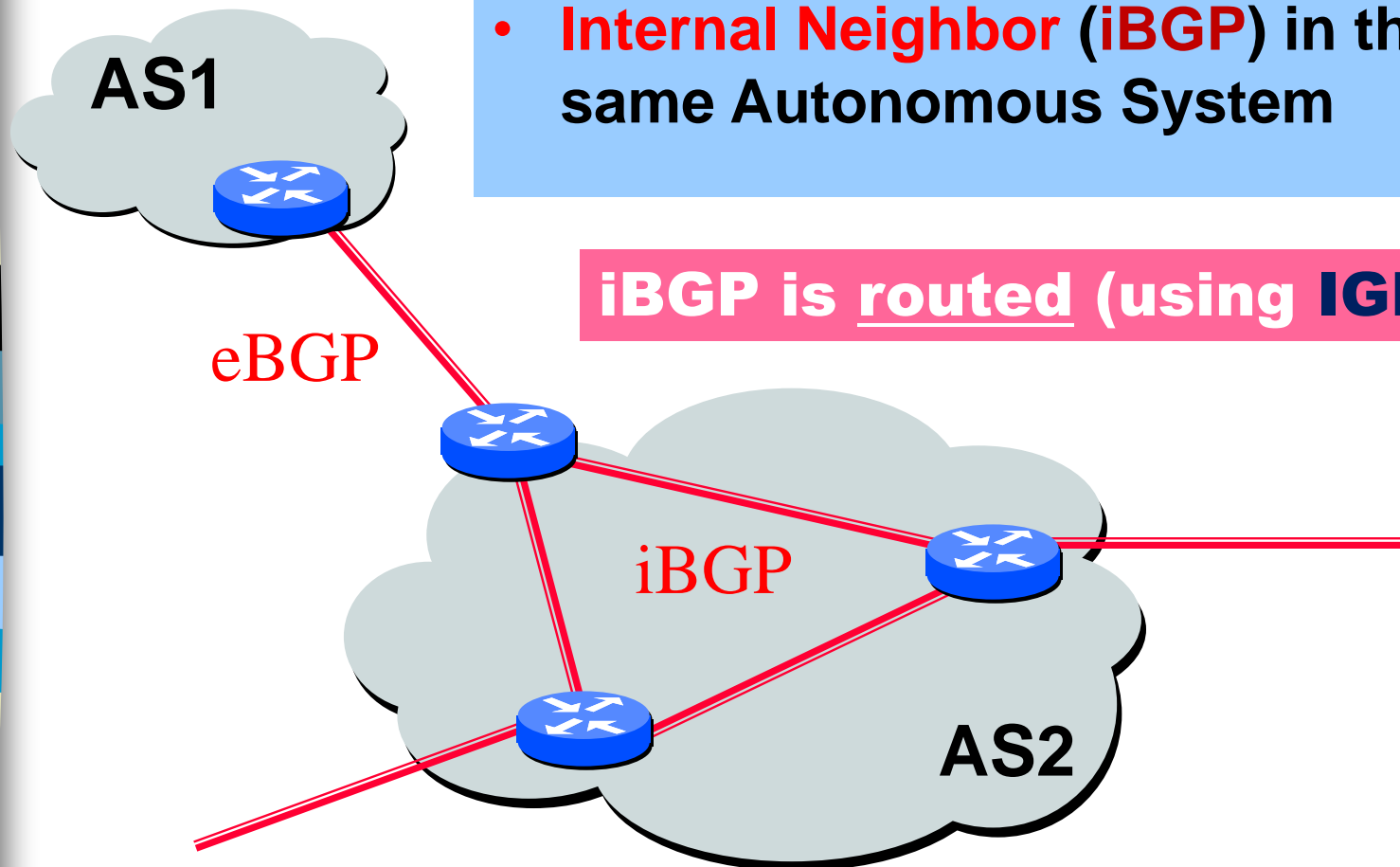
## BGP (Border Gateway Protocol) routing

# BGP Operations (Simplified)



# Two Types of BGP Neighbor Relationships

- **External Neighbor (eBGP)** in a different Autonomous Systems
- **Internal Neighbor (iBGP)** in the same Autonomous System



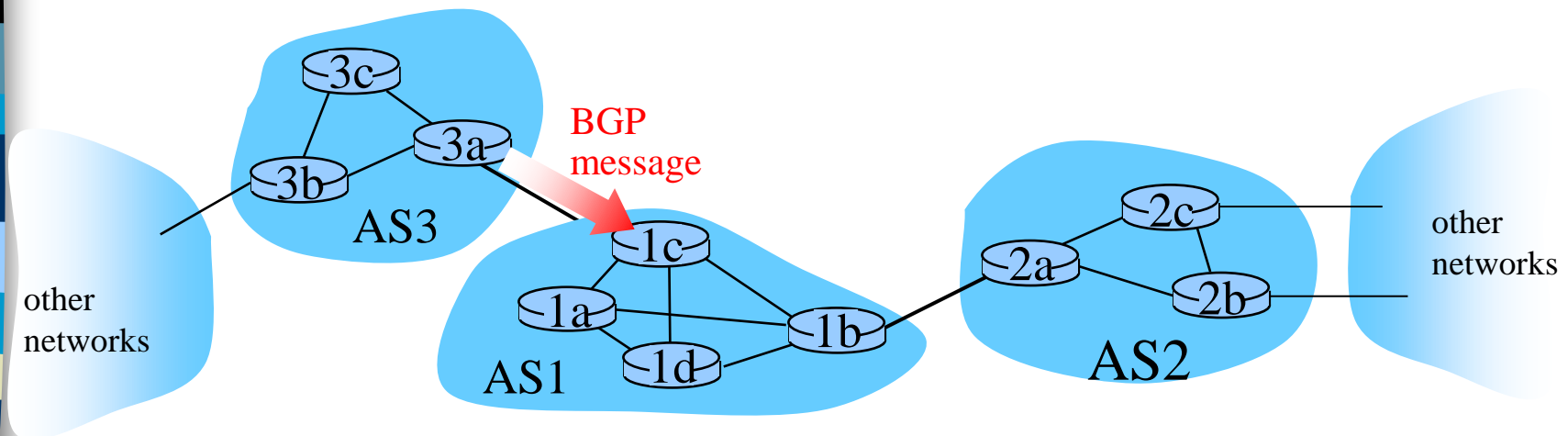


# Internet inter-AS routing: BGP

- **BGP (Border Gateway Protocol):** *the* de facto inter-domain routing protocol
  - “glue that holds the Internet together”
- 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*
- Allows subnet to advertise its existence to rest of Internet: *“I am here”*

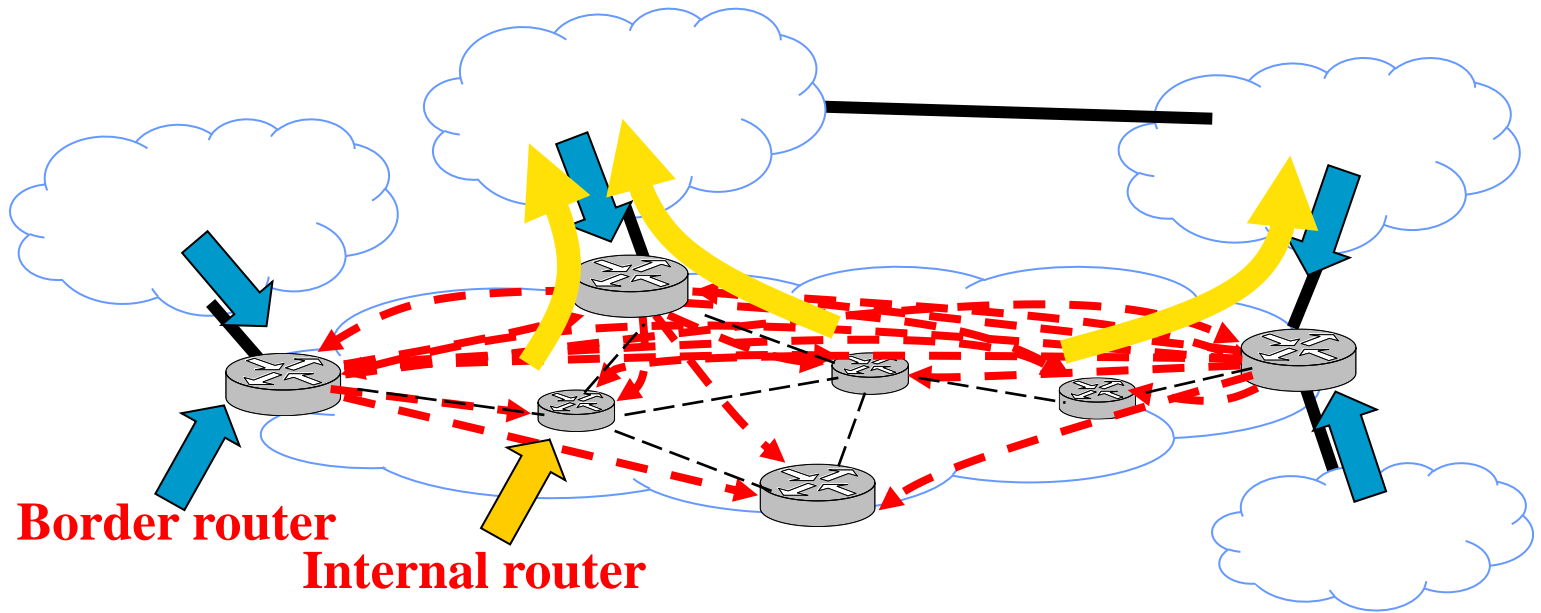
# BGP basics

- ❖ **BGP session:** two BGP routers (“peers”) exchange BGP messages
  - Advertising *paths* to different destination network prefixes (“path vector” protocol)
  - Exchanged over semi-permanent TCP connections
- When AS3 advertises a prefix to AS1:
  - AS3 *promises* it will forward datagrams towards that prefix
  - AS3 can aggregate prefixes in its advertisement





# BGP Session



**BGP “Managed Session” is used to**

- 1) Distribute externally learned routes internally
- 2) Distribute routes to external destinations

# BGP messages

- BGP messages **exchanged between peers** over TCP connection
- 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

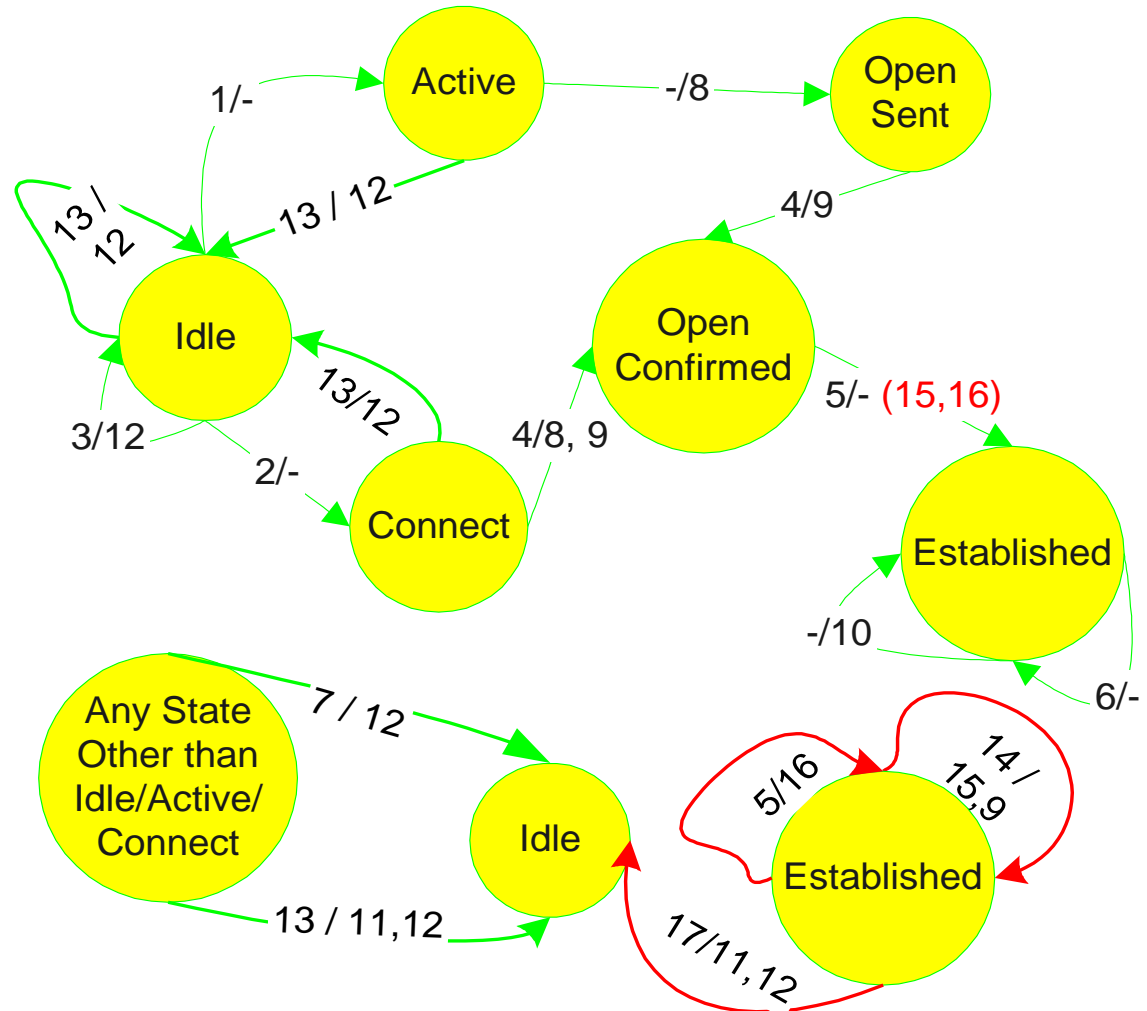
**announcement**  
=  
**prefix + attributes values**

# BGP FSM

## Event/Action List

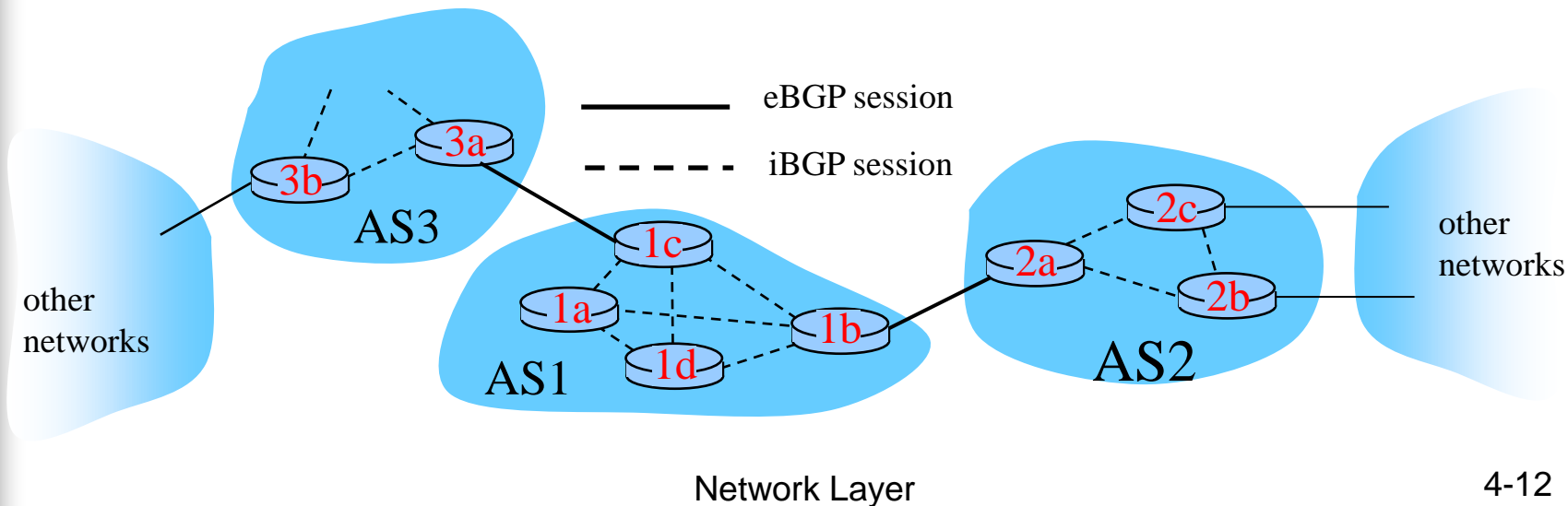
- 1 Request for Connection
- 2 Connection Succeeded
- 3 Connection Failed
- 4 "Open" Received
- 5 "KeepAlive" Received
- 6 "Update" Received
- 7 "Notification" Received
- 8 "Open" Sent
- 9 "KeepAlive" Sent
- 10 "Update" Sent
- 11 "Notification" Sent
- 12 Close connection
- 13 Any Error, including:  
TCP operation error,  
Packets format error,  
Packets type error.

- 14 KeepAlive Timer Expire
- 15 KeepAlive Timer Reset
- 16 HoldOn Timer Reset
- 17 HoldOn Timer Expire

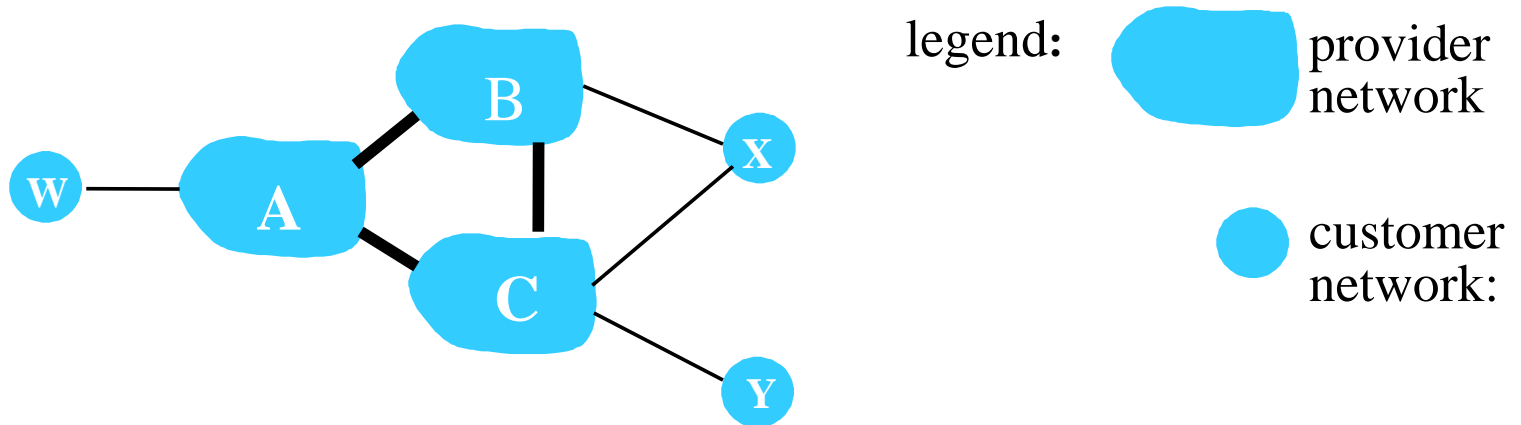


# BGP basics: distributing path information

- Using eBGP session between **3a** and **1c**, **AS3** sends prefix reachability info to **AS1**
  - **1c** can then use iBGP to 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

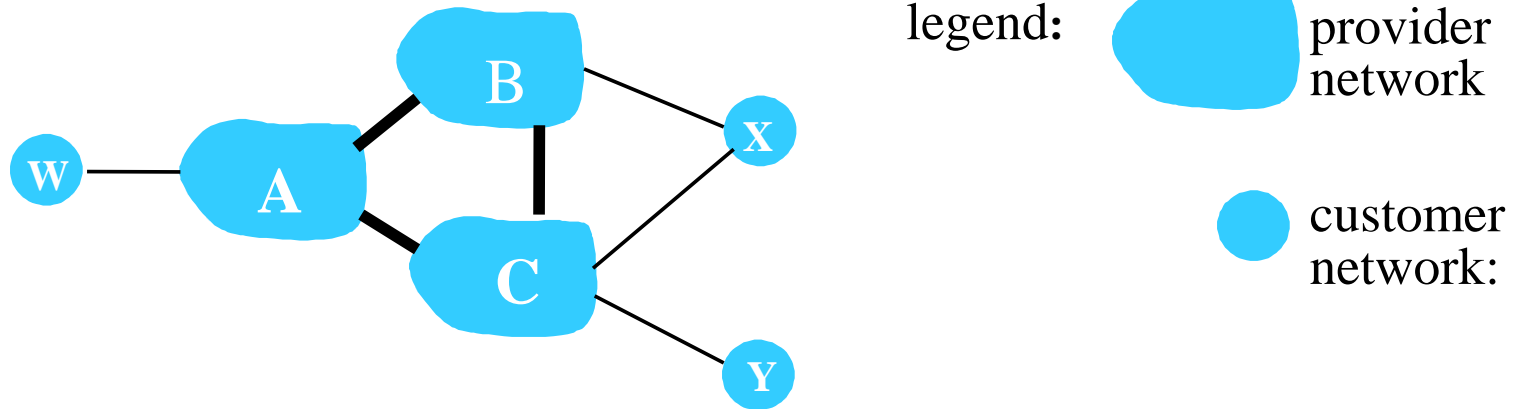


# BGP routing policy



- ❖ 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)



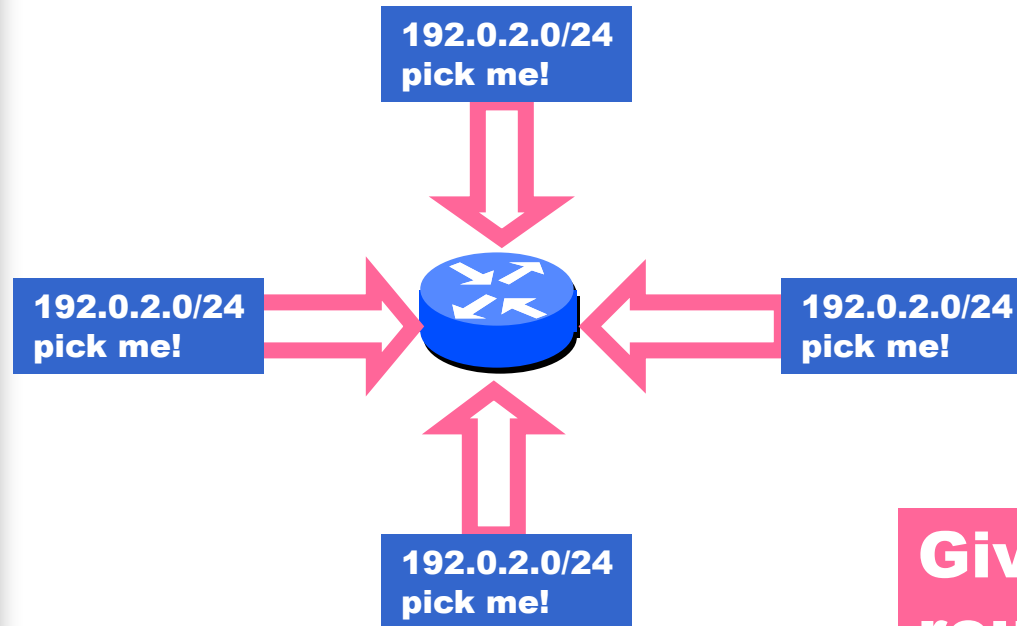
- ❖ A advertises path AW to B
- ❖ B advertises path BAW to X
- ❖ Should B advertise path BAW to C?



# 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)
- Gateway router receiving route advertisement uses **import policy** to accept/decline
  - e.g., never route through AS x
  - *policy-based* routing

# Attributes are Used to Select Best Routes

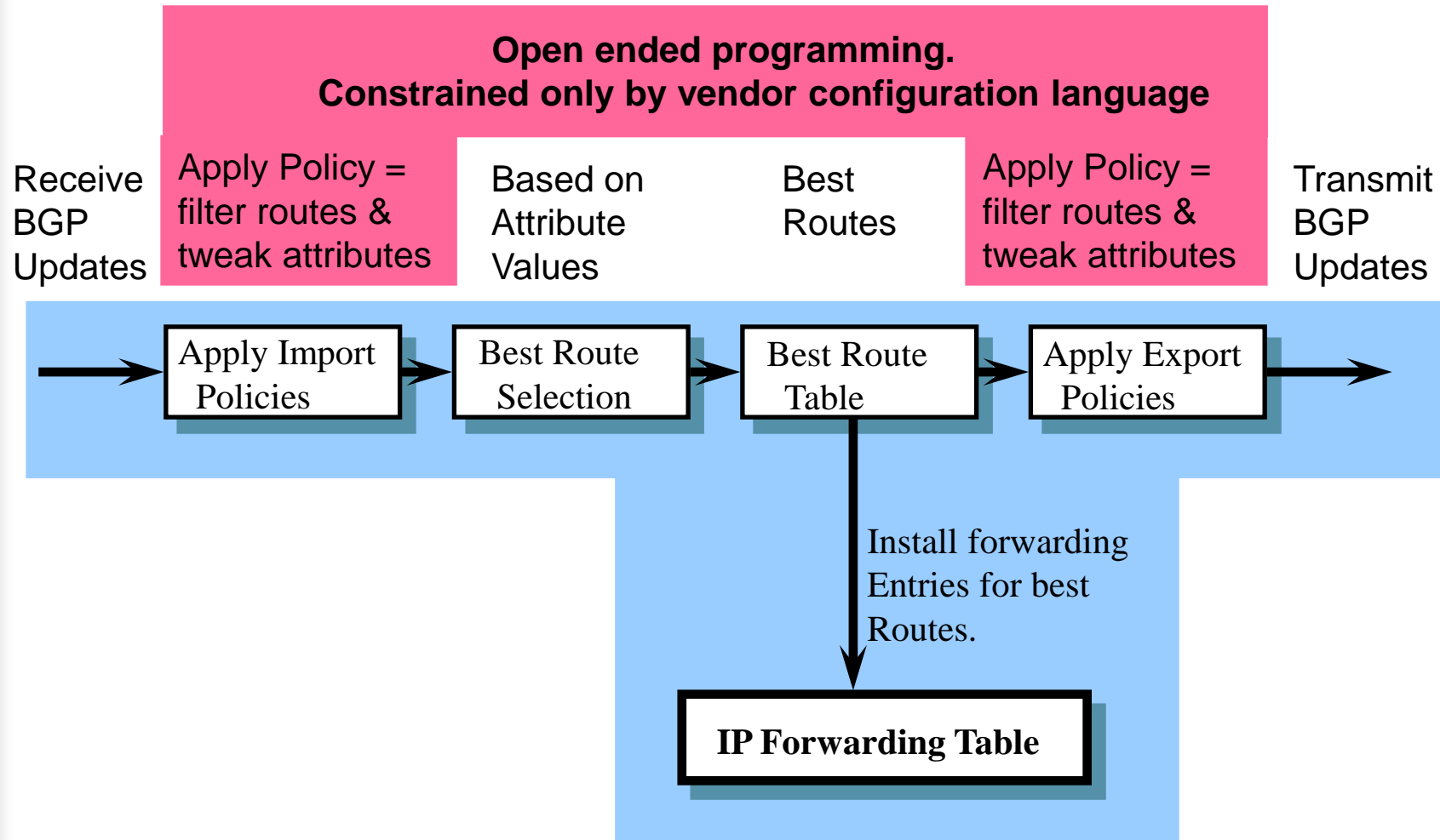


**Given multiple routes to the same prefix, a BGP speaker must pick at most one best route**

**(Note: it could reject them all!)**

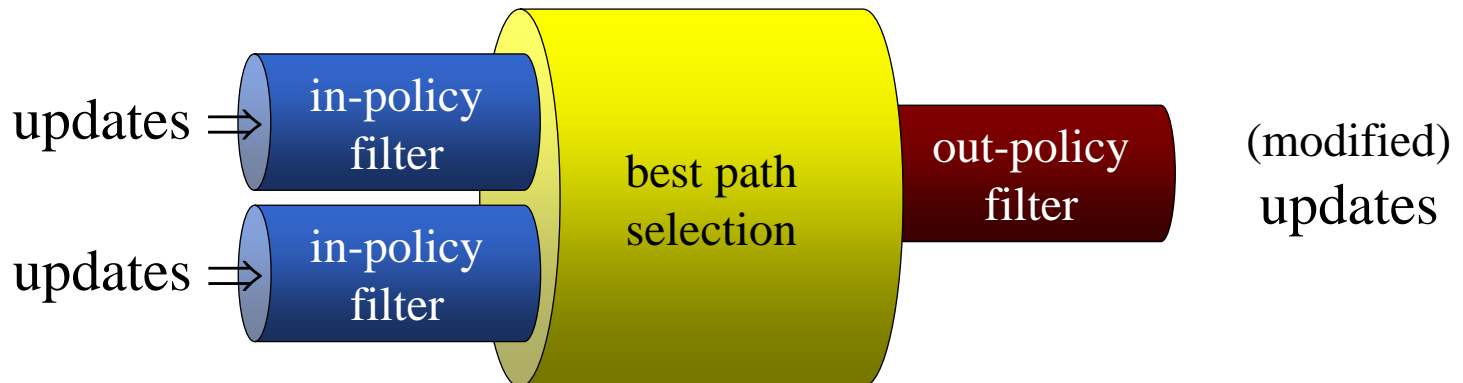


# BGP Route Processing



# BGP Design

- Distance vector protocol
- Implements policies
  - to enable politics and traffic engineering
- Decision process (in each router):



# Route Selection Summary



**Highest Local Preference**

**Enforce relationships**

**Shortest ASPATH**

**Lowest MED**

**i-BGP < e-BGP**

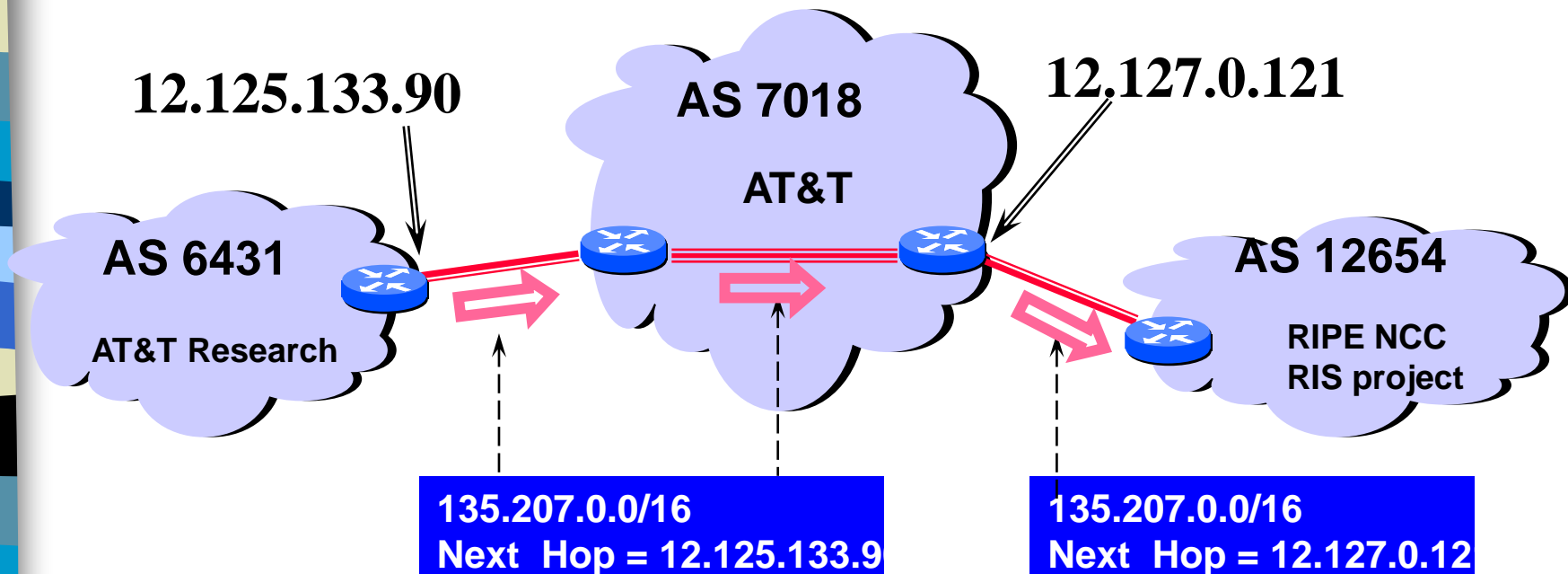
**Lowest IGP cost**

**Traffic engineering**

**Lowest router ID**

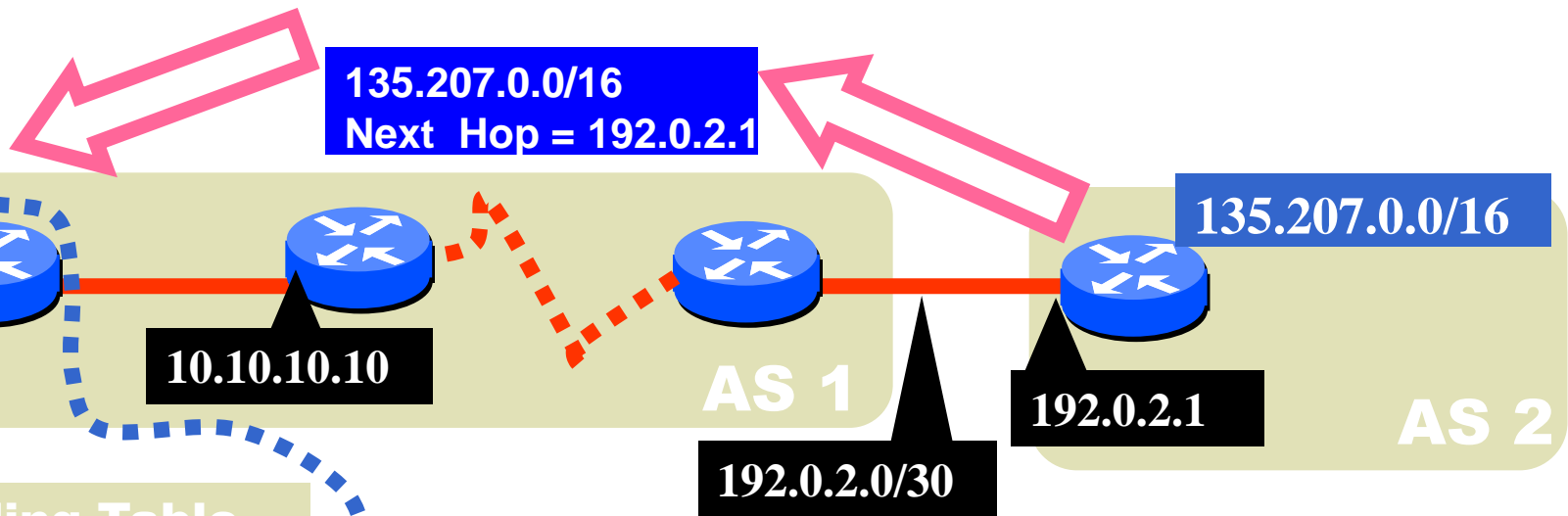
**Throw up hands and  
break ties**

# 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.

# Join EGP with IGP For Connectivity



**Forwarding Table**

destination	next hop
192.0.2.0/30	10.10.10.10

+

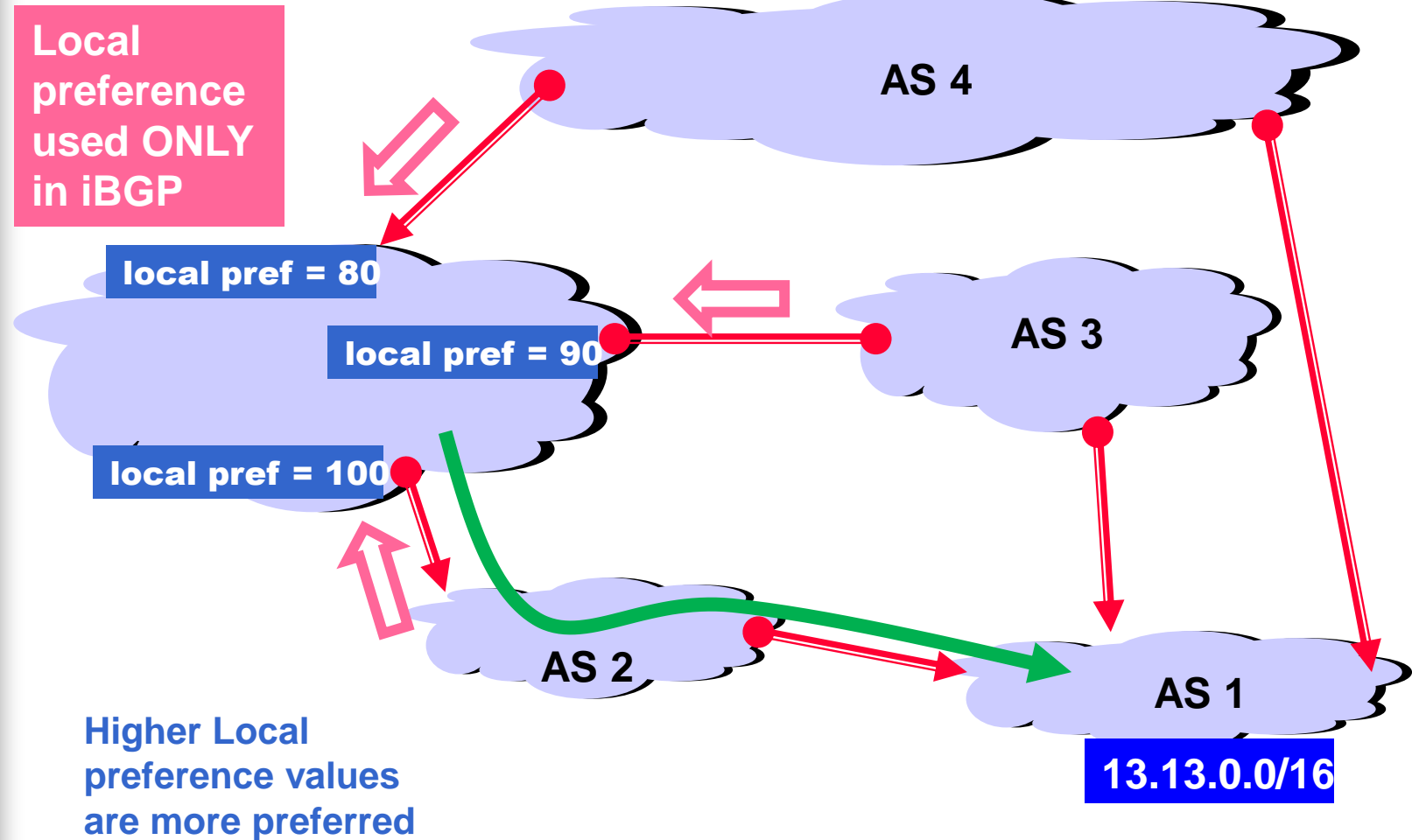
**EGP**

destination	next hop
135.207.0.0/16	192.0.2.1

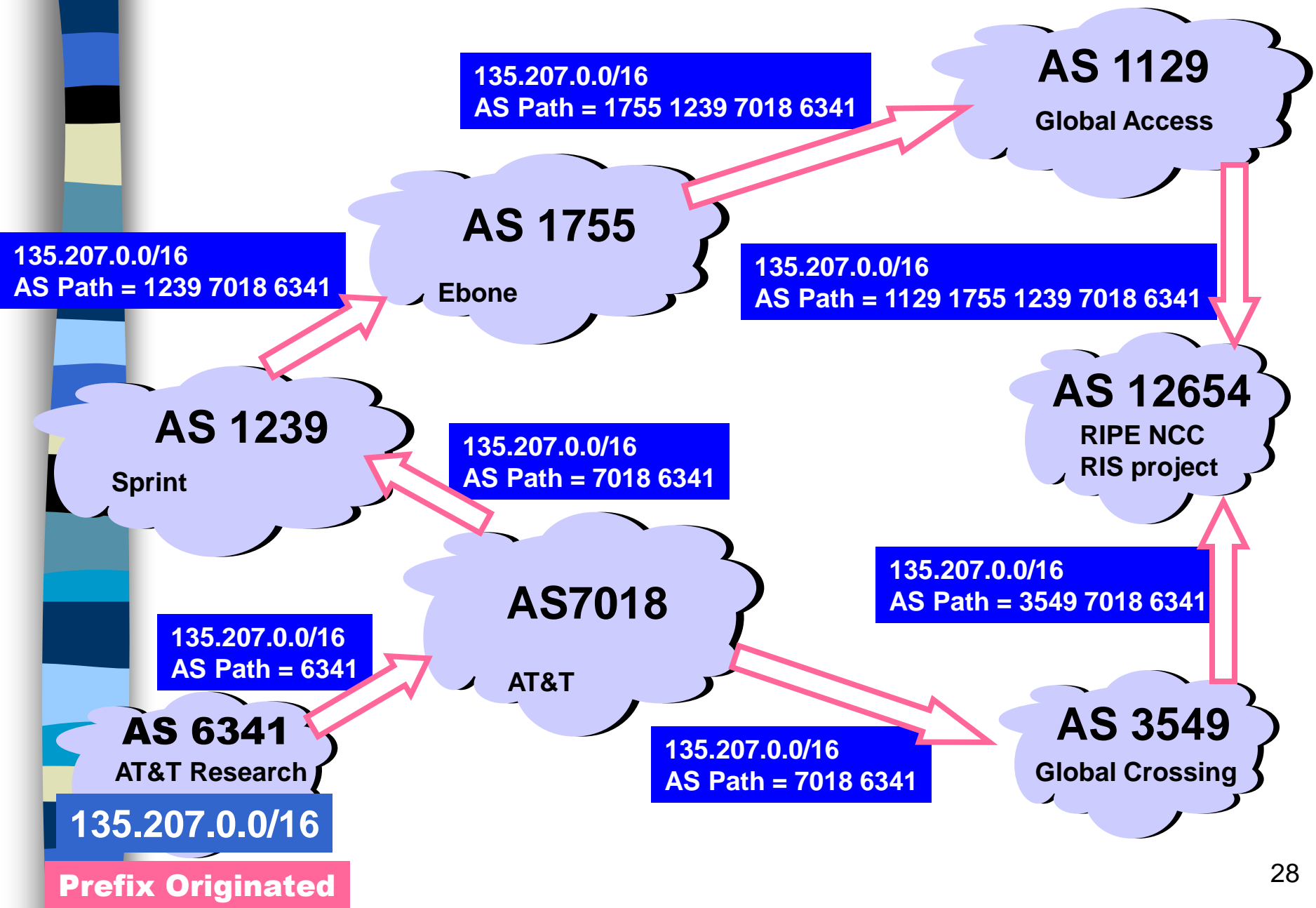
**Forwarding Table**

destination	next hop
135.207.0.0/16	10.10.10.10

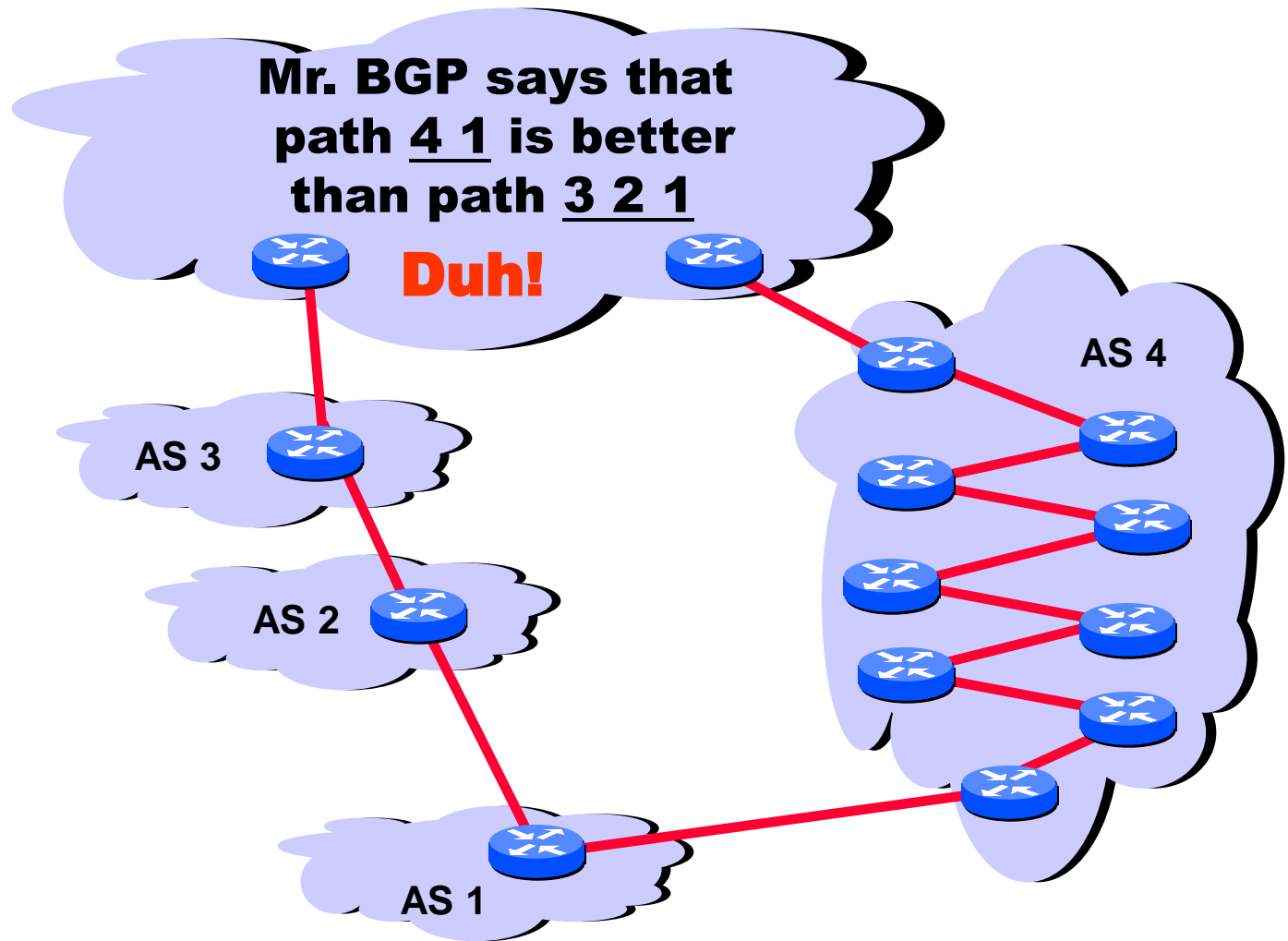
# LOCAL PREFERENCE



# ASPATH Attribute

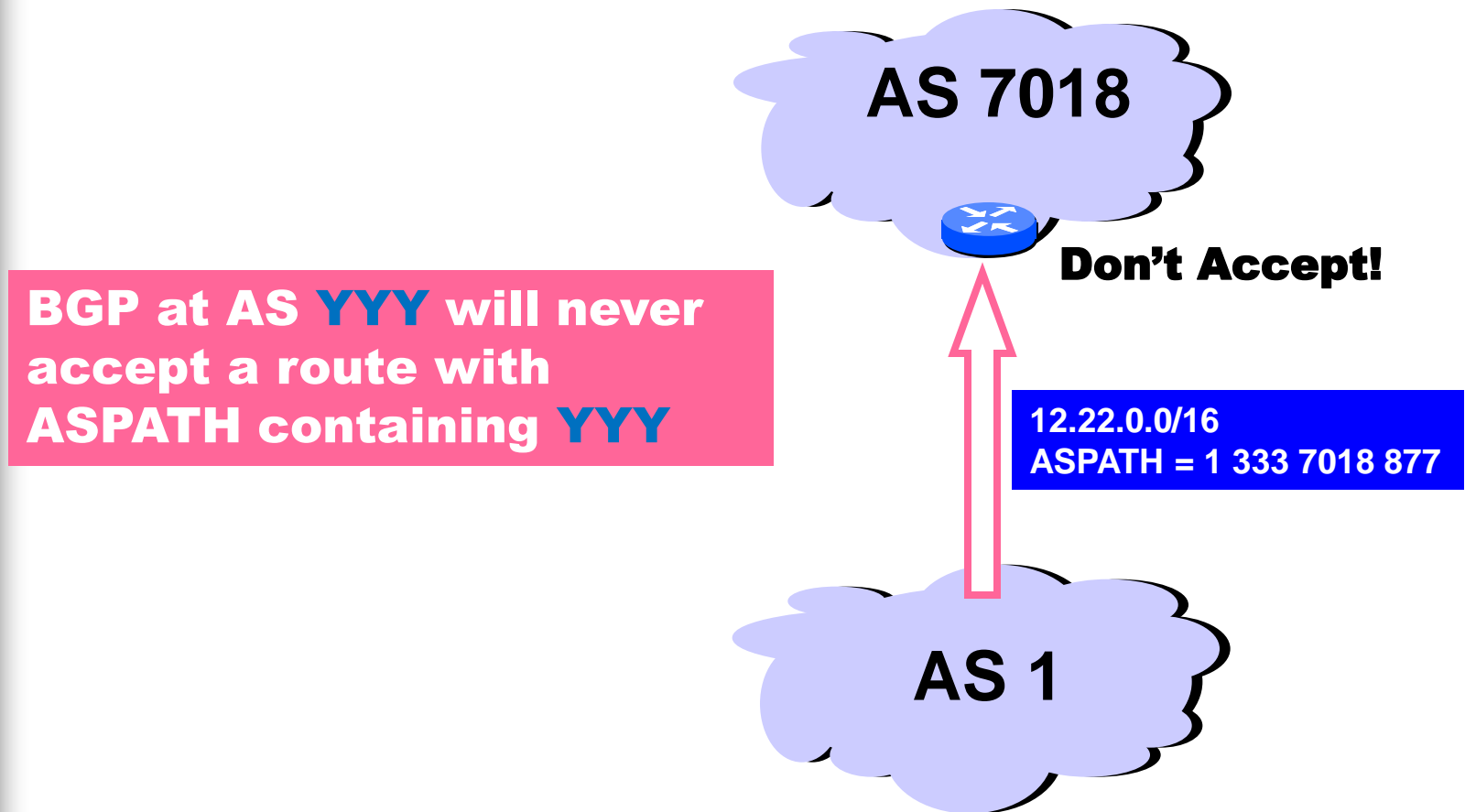


# Shorter Doesn't Always Mean Shorter

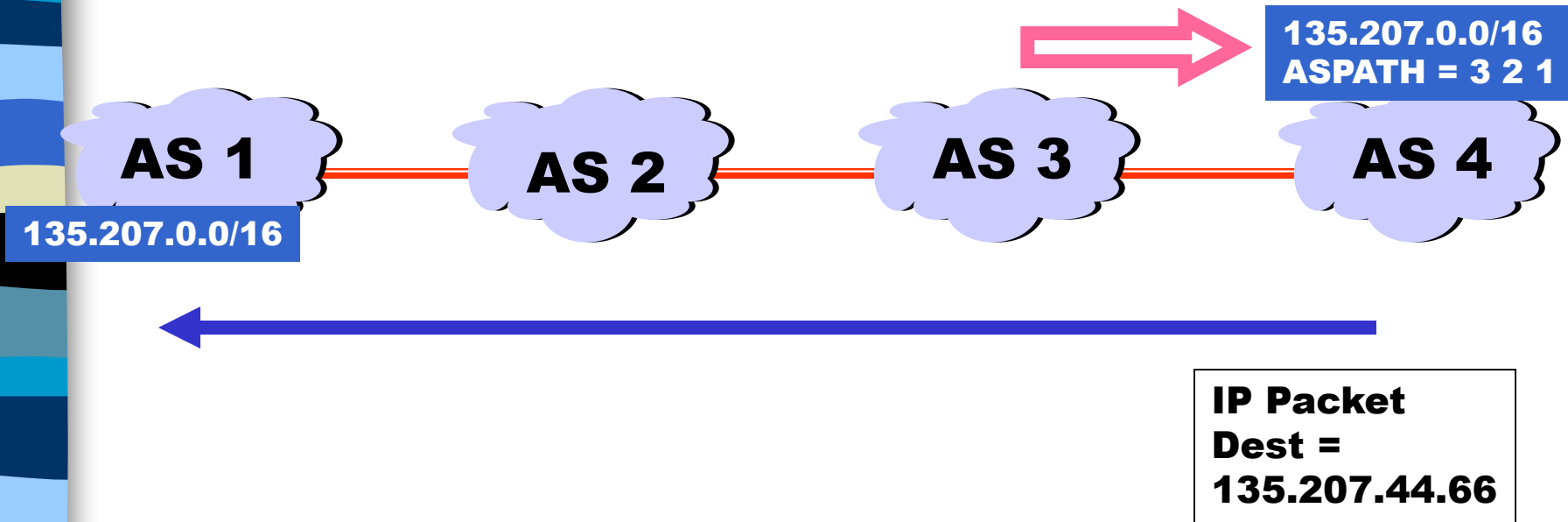




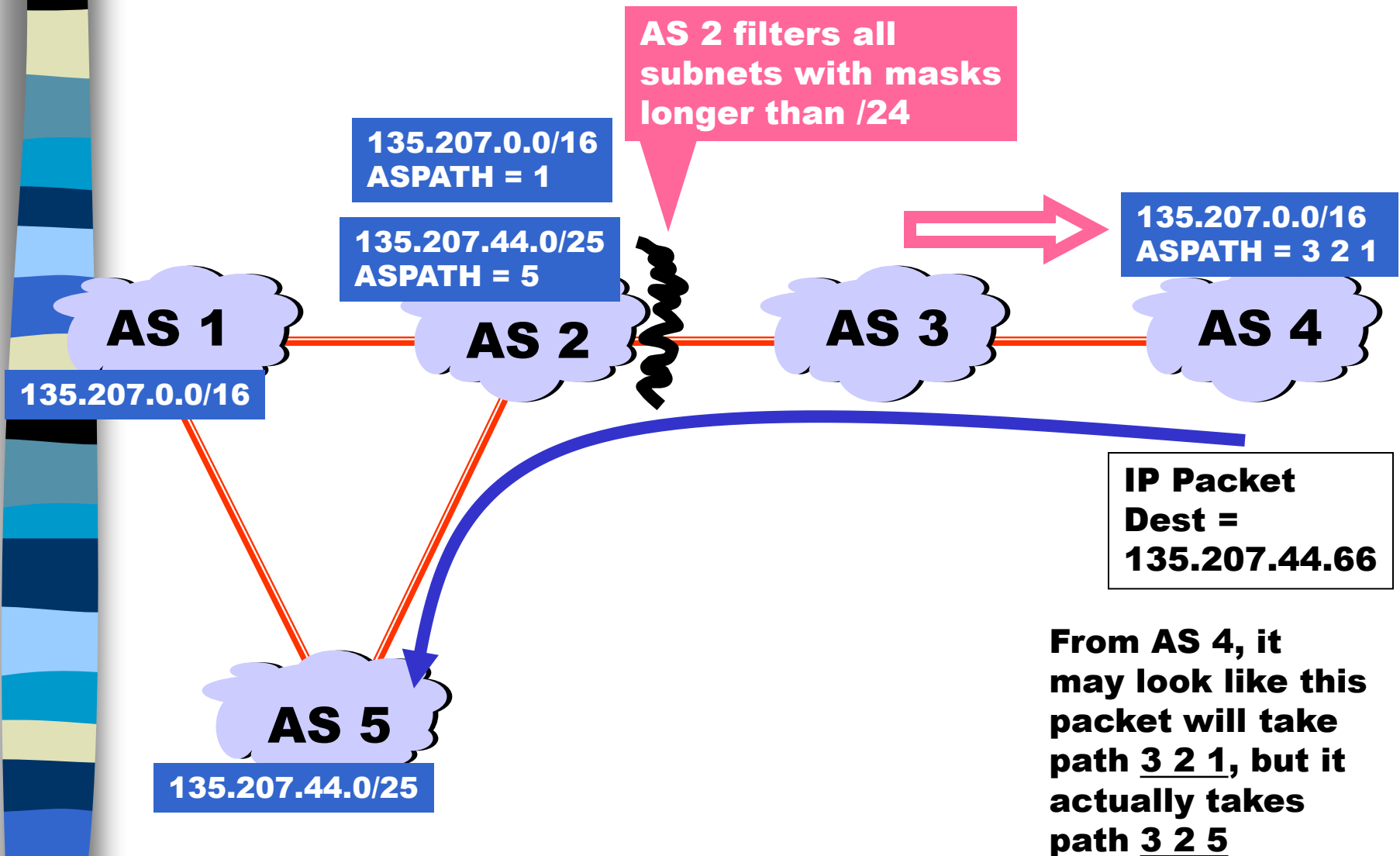
# Interdomain Loop Prevention



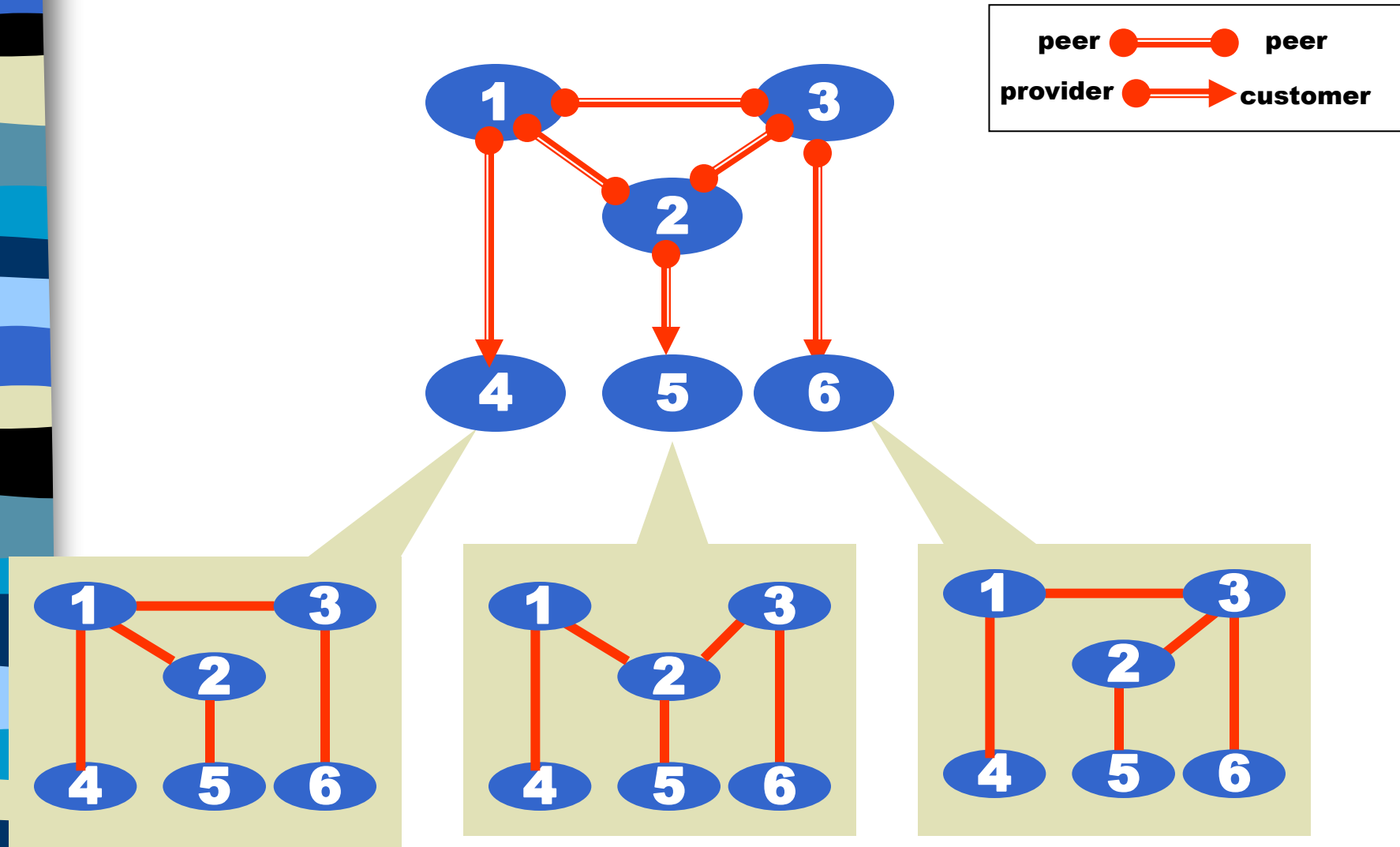
# Traffic Often Follows ASPATH



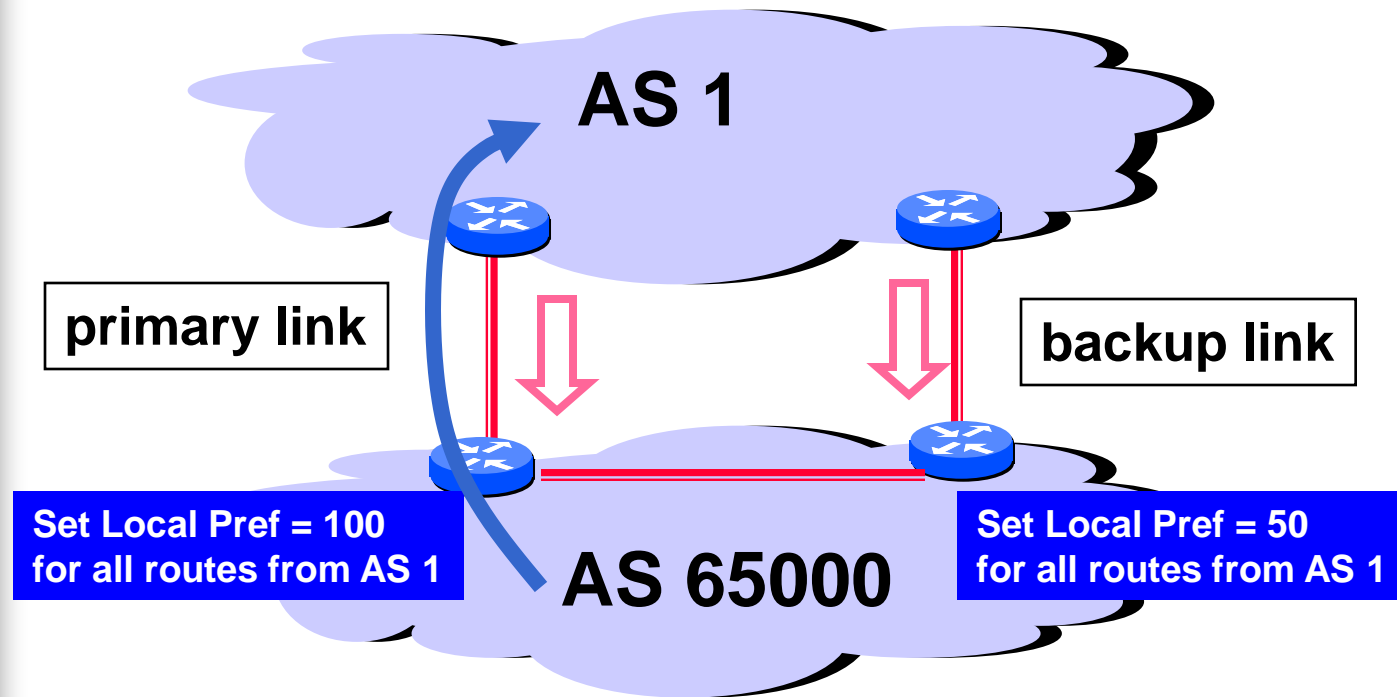
# ... But It Might Not



# AS Graphs Depend on Point of View



# Implementing Backup Links with Local Preference (Outbound Traffic)



Forces outbound traffic to take primary link, unless link is down

**We'll talk about inbound traffic soon ...**



# Why different Intra- and Inter-AS routing ?

## Policy

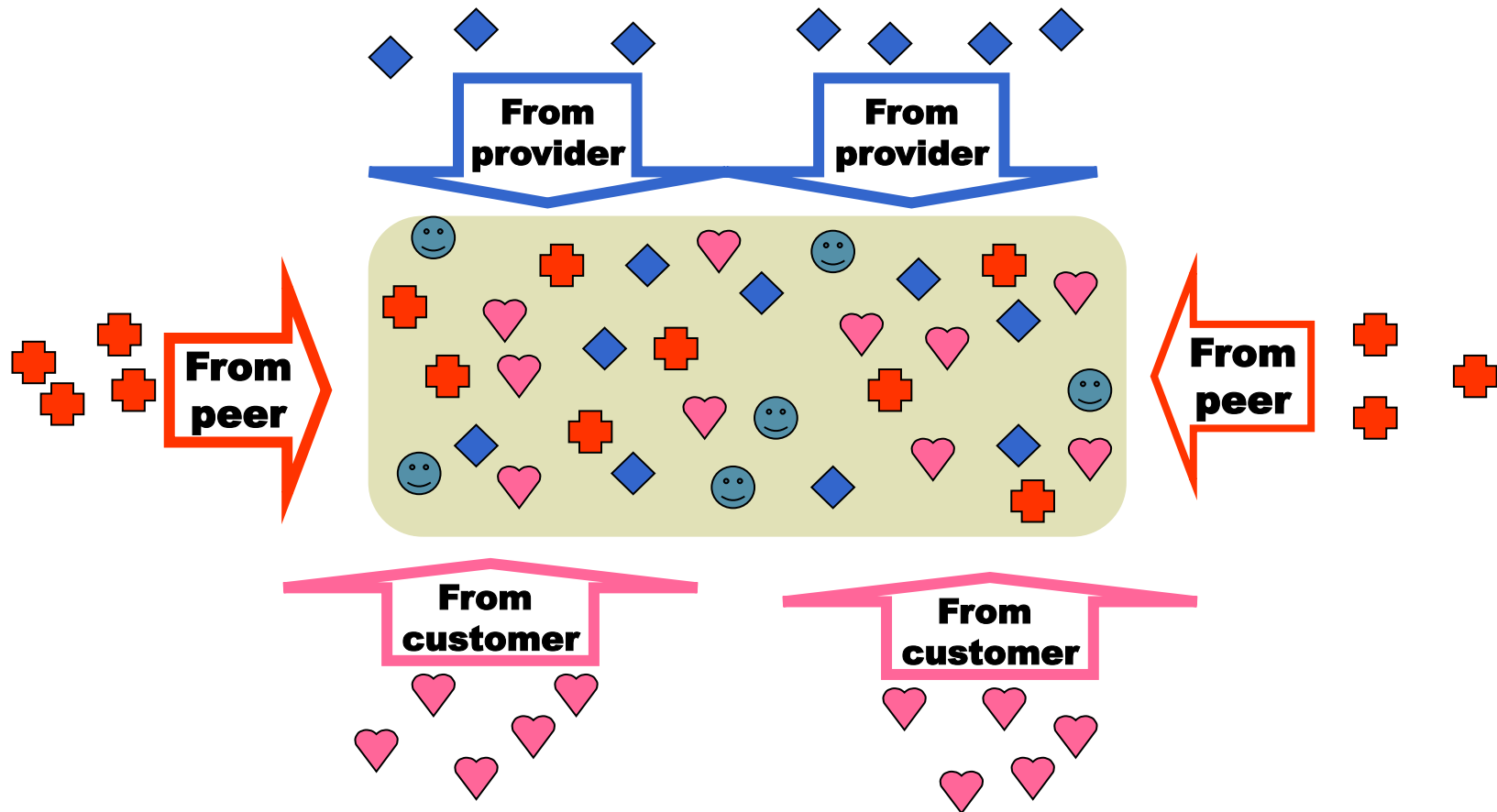
- Inter-AS: admin wants control over how its traffic routed, who routes through its net.
- Intra-AS: single admin, so no policy decisions needed

## Performance

- Intra-AS: can focus on performance
- Inter-AS: policy may dominate over performance

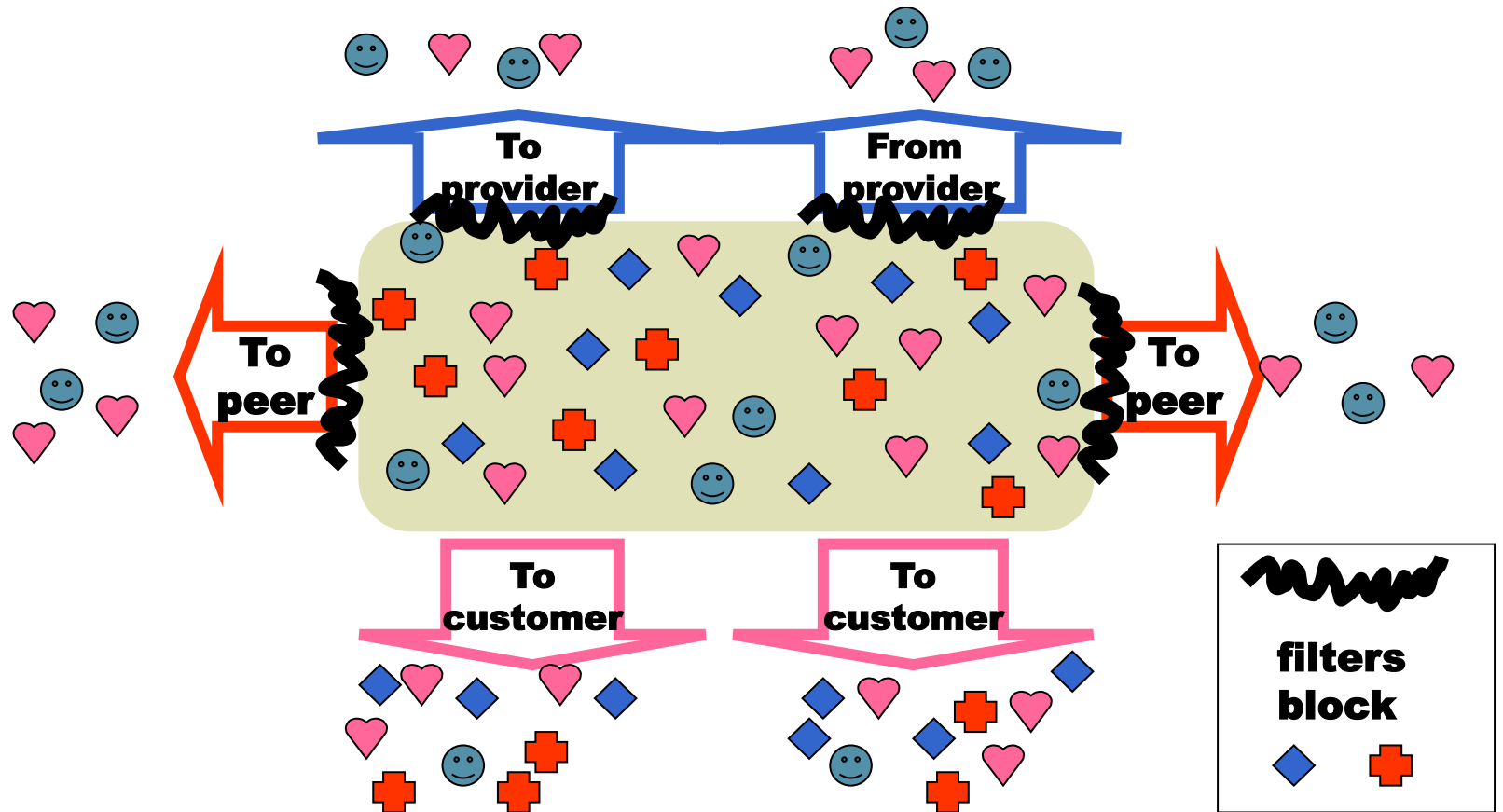
# Import Routes

◆ provider route    + peer route    ♥ customer route    ☺ ISP route



# Export Routes

◆ provider route    + peer route    ♥ customer route    ☺ ISP route







# Basic BGP commands

## ■ Basic Commands

- router bgp <as-number>
- neighbor <ip address> remote-as <as-number>
- Examples
  - router bgp 400
  - neighbor 120.5.1.1 remote-as 300
  - neighbor 5.5.5.3 remote-as 3
  - network 2.2.2.0 mask 255.255.255.0

## ■ show commands

- show ip bgp
- show ip bgp neighbors

# Protocoles et Interconnexions



## Course Overview and Introduction

Dario Vieira

Department of Computer Science

EFREI