

CSCI-351

Data communication and Networks

**Lecture 10: Inter Domain Routing
(It's all about the Money)**

Network Layer, Control Plane

2

Data Plane

Application

Presentation

Session

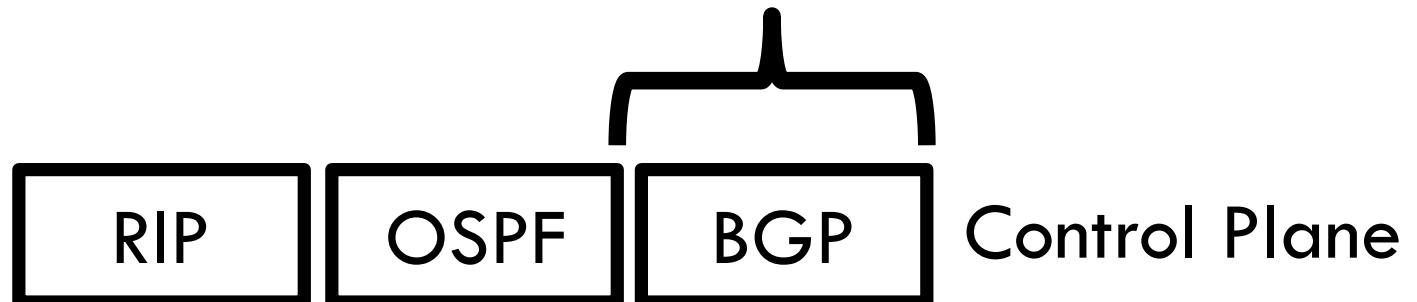
Transport

Network

Data Link

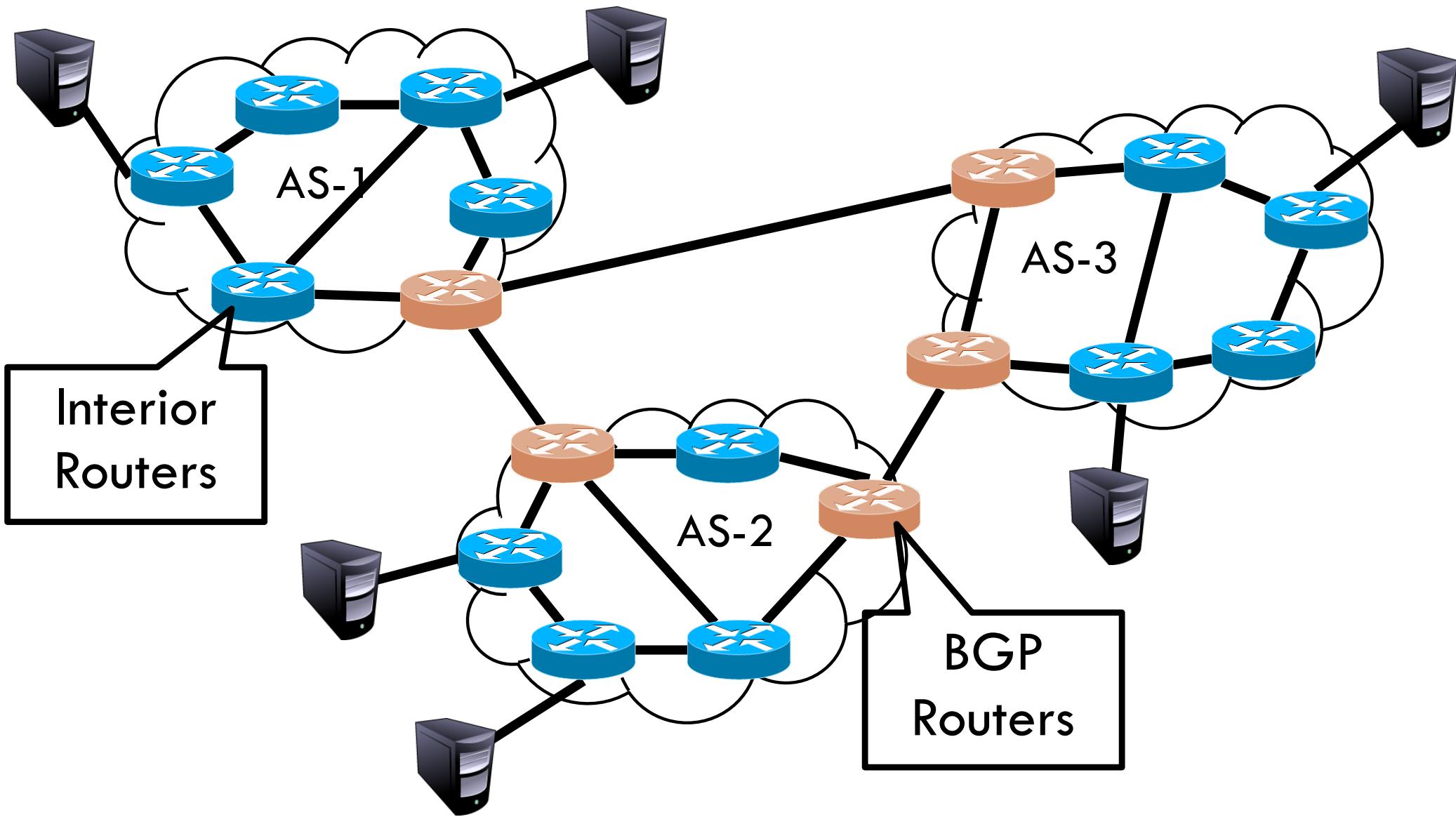
Physical

- Function:
 - Set up routes between networks
- Key challenges:
 - Implementing provider policies
 - Creating stable paths



ASs, Revisited

3



AS Numbers

4

- Each AS identified by an ASN number
 - ▣ 16-bit values (latest protocol supports 32-bit ones)
 - ▣ 64512 – 65535 are reserved
- Currently, there are > 20000 ASNs
 - ▣ AT&T: 5074, 6341, 7018, ...
 - ▣ Sprint: 1239, 1240, 6211, 6242, ...
 - ▣ Northeastern: 156
 - ▣ North America ASs → <ftp://ftp.arin.net/info/asn.txt>

Inter-Domain Routing

5

- Global connectivity is at stake!
 - ▣ Thus, all ASs must use the same protocol
 - ▣ Contrast with intra-domain routing
- What are the requirements?
 - ▣ Scalability
 - ▣ Flexibility in choosing routes
 - Cost
 - Routing around failures
- Question: link state or distance vector?
 - ▣ Trick question: BGP is a path vector protocol

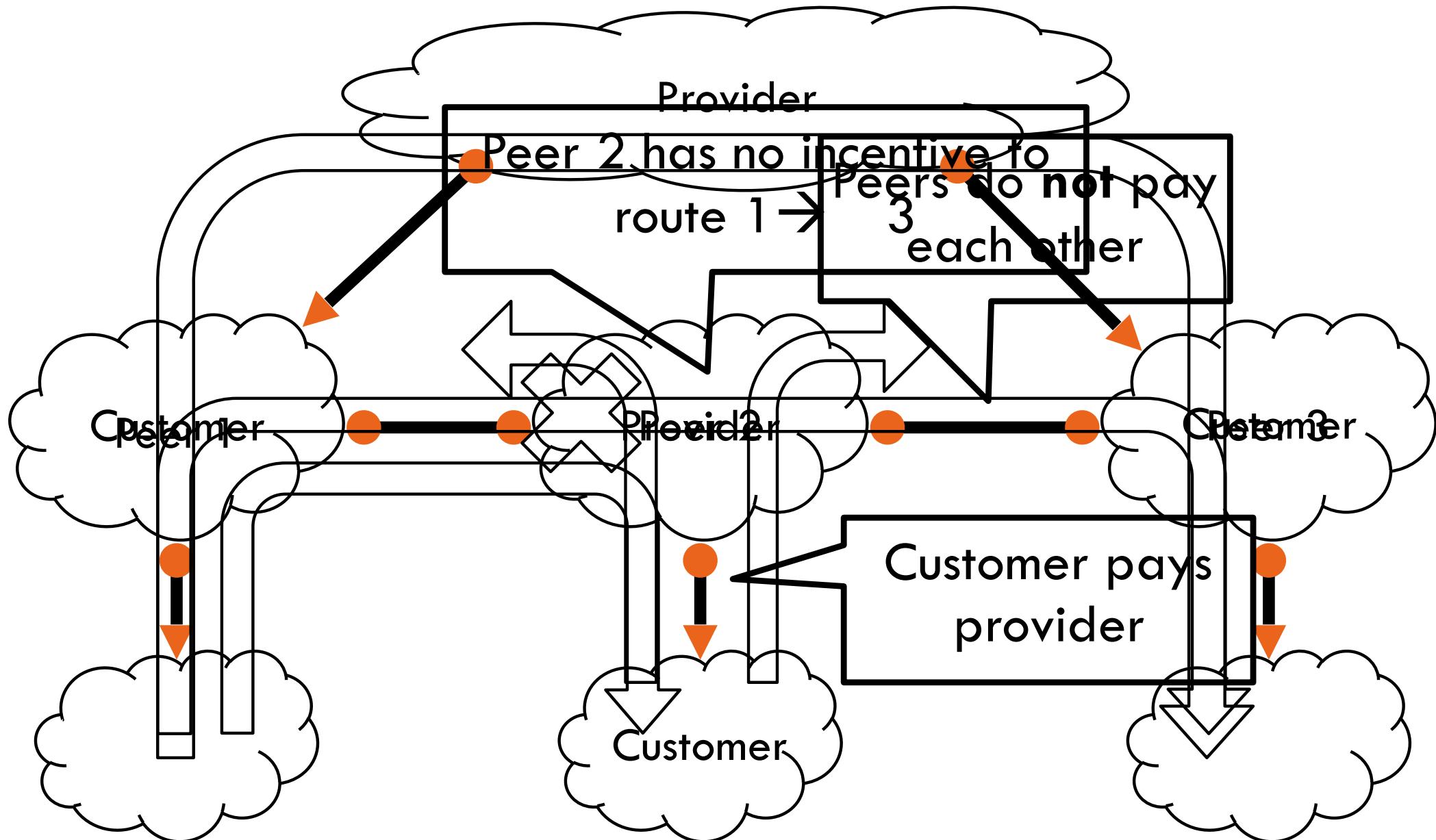
BGP

6

- Border Gateway Protocol
 - De facto inter-domain protocol of the Internet
 - Policy based routing protocol
 - Uses a Bellman-Ford path vector protocol
- Relatively simple protocol, but...
 - Complex, manual configuration
 - Entire world sees advertisements
 - Errors can screw up traffic globally
 - Policies driven by economics
 - How much \$\$\$ does it cost to route along a given path?
 - Not by performance (e.g. shortest paths)

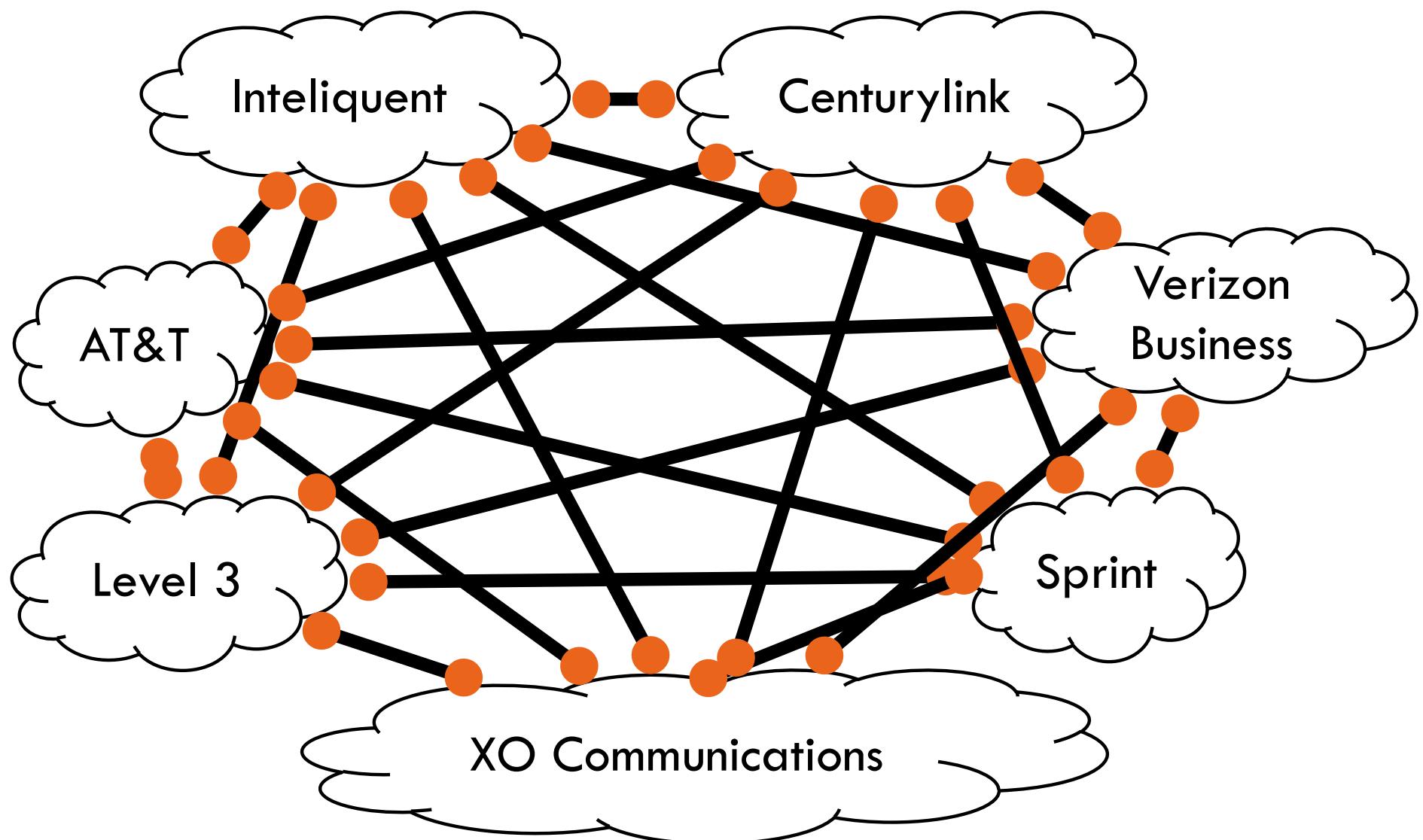
BGP Relationships

7



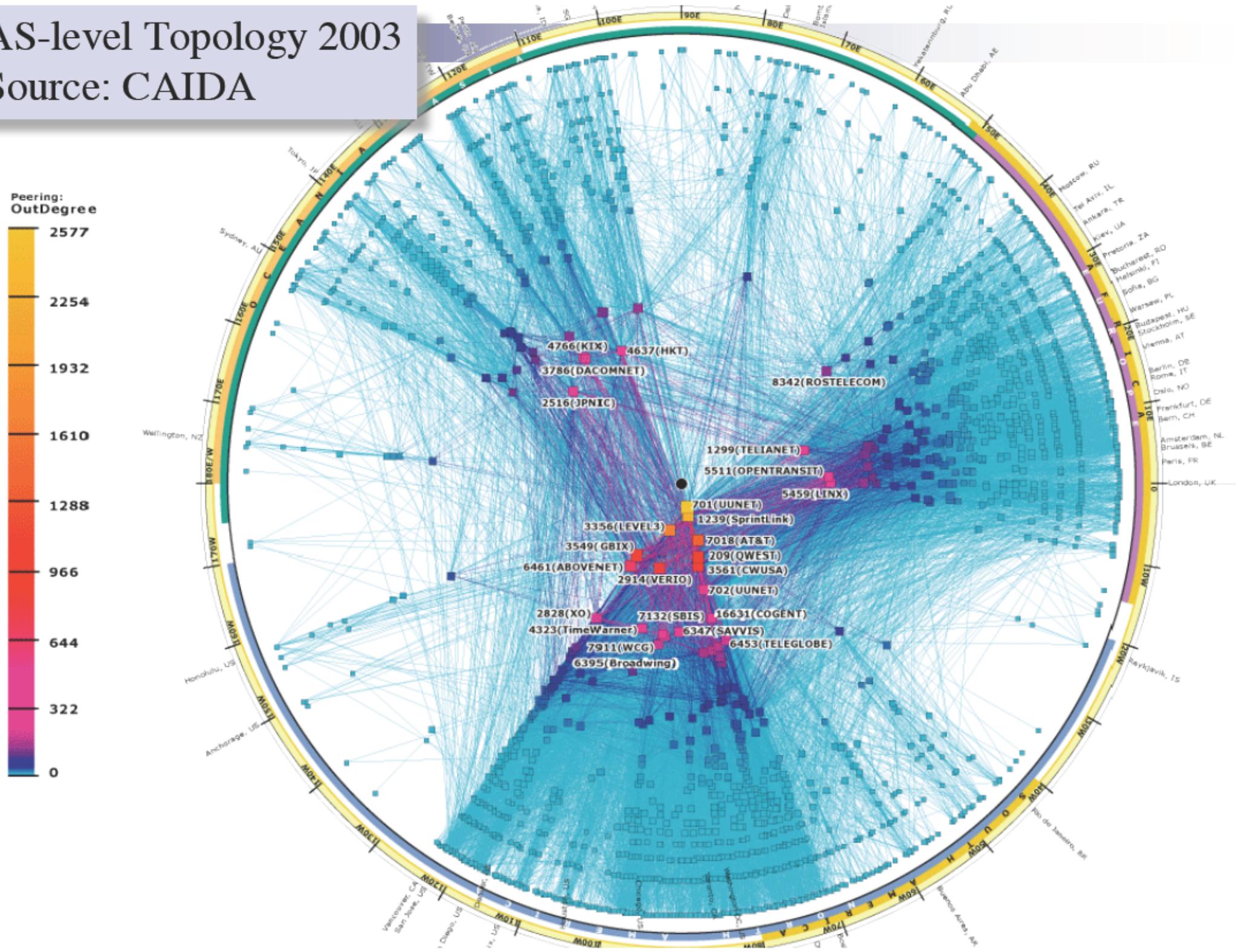
Tier-1 ISP Peering

8



AS-level Topology 2003

Source: CAIDA



Peering Wars

10

Peer

- Reduce upstream costs
- Improve end-to-end performance
- May be the only way to connect to parts of the Internet

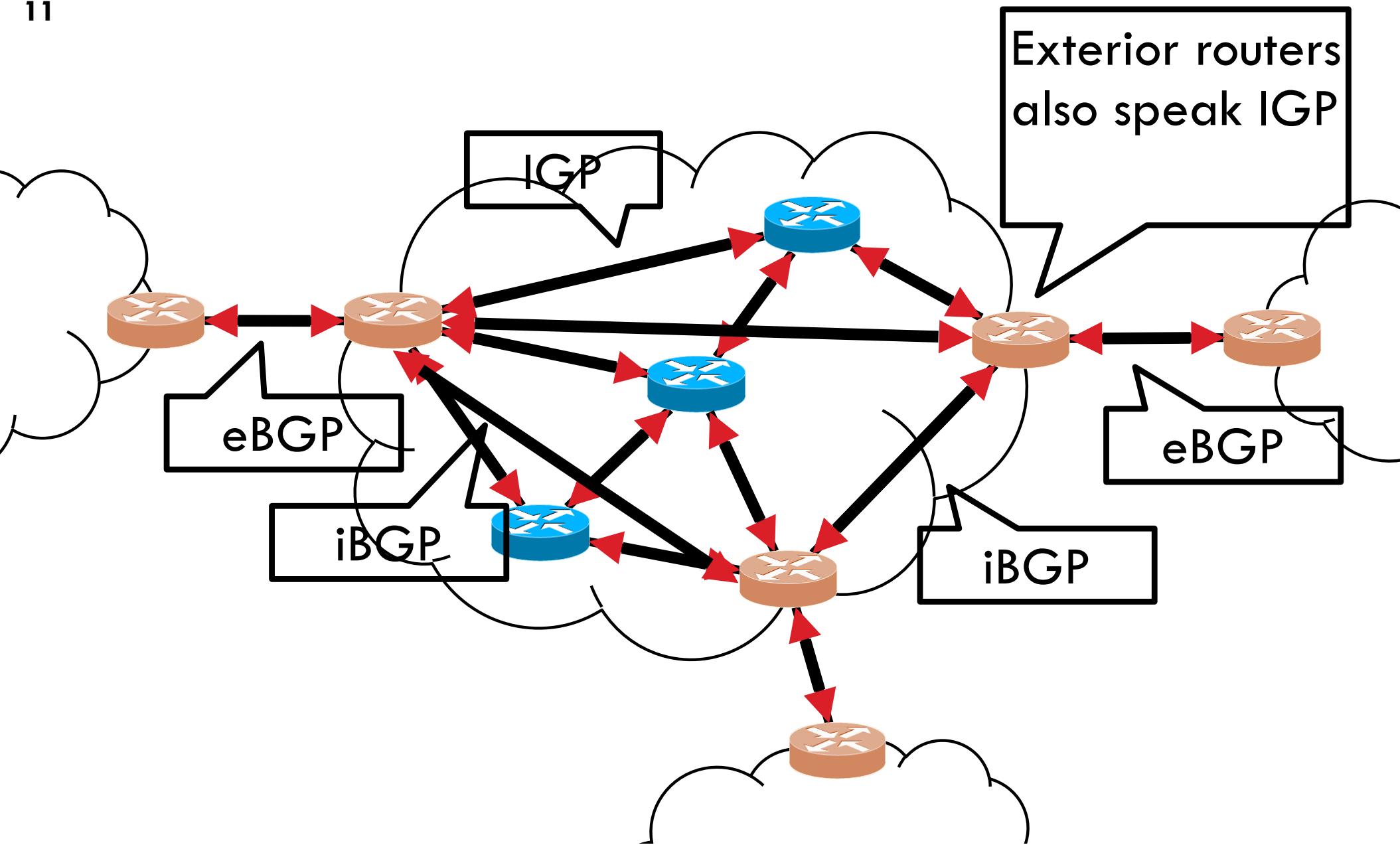
Don't Peer

- You would rather have customers
- Peers are often competitors
- Peering agreements require periodic renegotiation

Peering struggles in the ISP world are extremely contentious, agreements are usually confidential

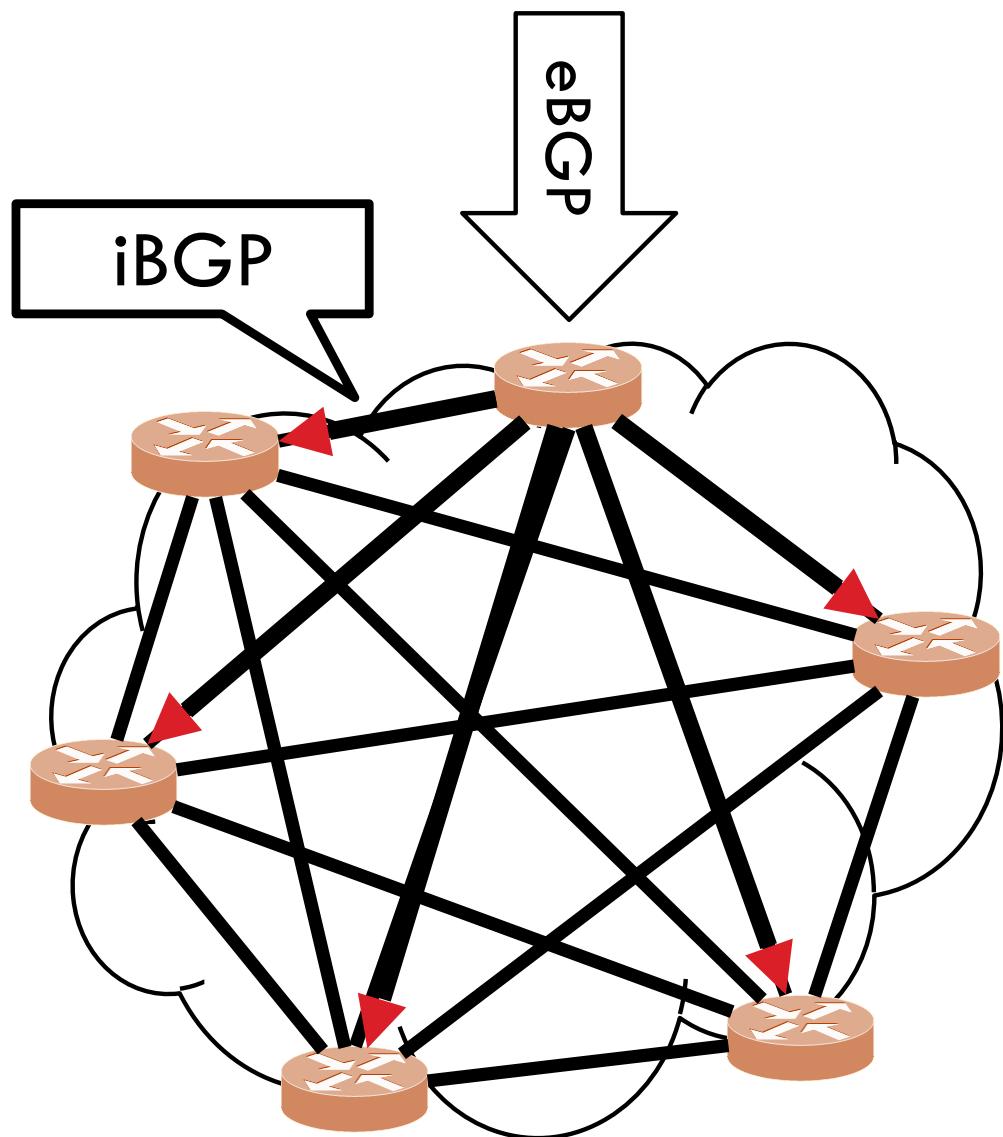
Two Types of BGP Neighbors

11



Full iBGP Meshes

12

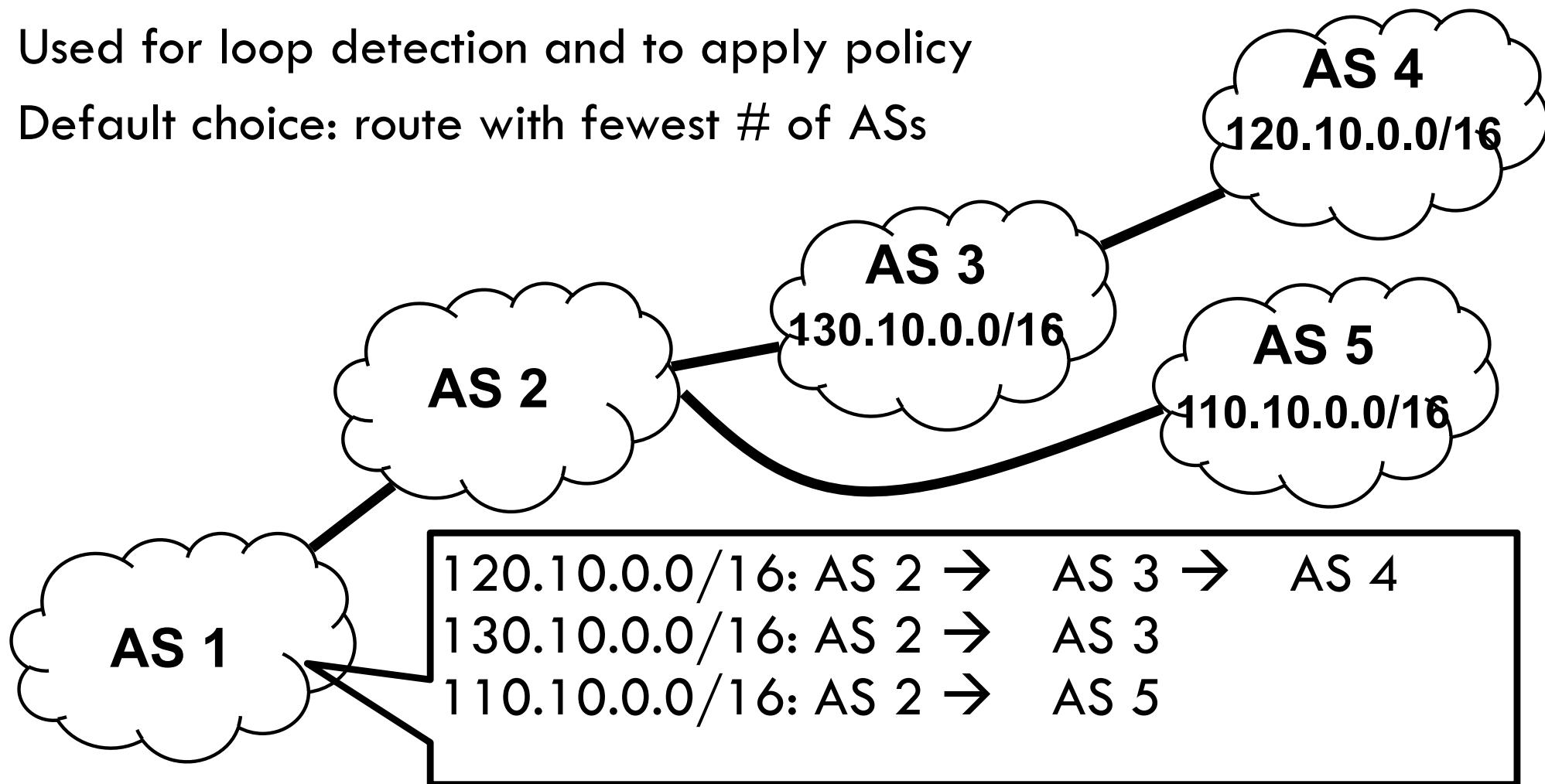


- Question: why do we need iBGP?
 - ▣ OSPF does not include BGP policy info
 - ▣ Prevents routing loops within the AS
- iBGP updates do not trigger announcements

Path Vector Protocol

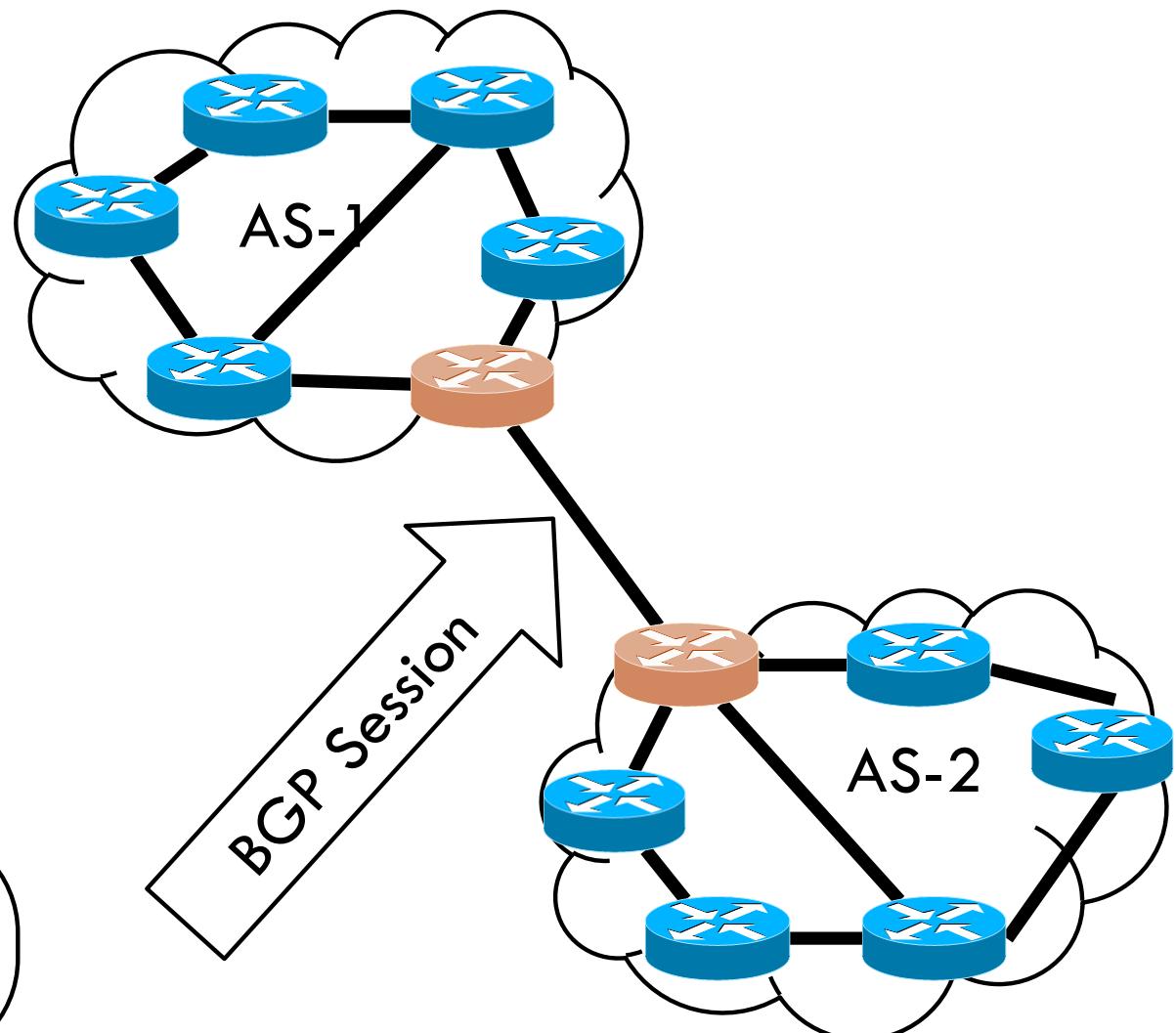
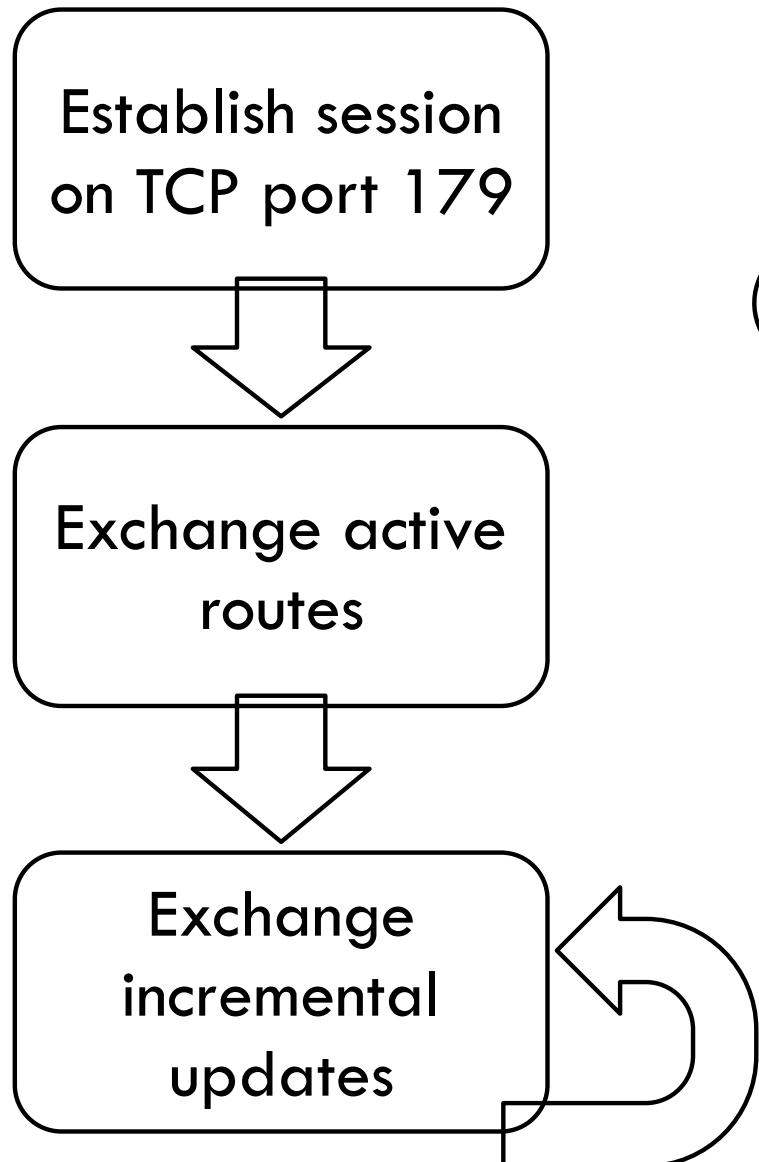
13

- AS-path: sequence of ASes a route traverses
 - Like distance vector, plus additional information
- Used for loop detection and to apply policy
- Default choice: route with fewest # of ASes



BGP Operations (Simplified)

14



Four Types of BGP Messages

15

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

announcement = IP prefix + attributes values

BGP Attributes

16

- Attributes used to select “best” path
 - ▣ LocalPref
 - Local preference policy to choose most preferred route
 - Overrides default fewest AS behavior
 - ▣ Multi-exit Discriminator (MED)
 - Specifies path for external traffic destined for an internal network
 - Chooses peering point for your network
 - ▣ Import Rules
 - What route advertisements do I accept?
 - ▣ Export Rules
 - Which routes do I forward to whom?

Route Selection Summary

18

Highest Local Preference

Enforce relationships

Shortest AS Path

Lowest MED

Lowest IGP Cost to BGP Egress

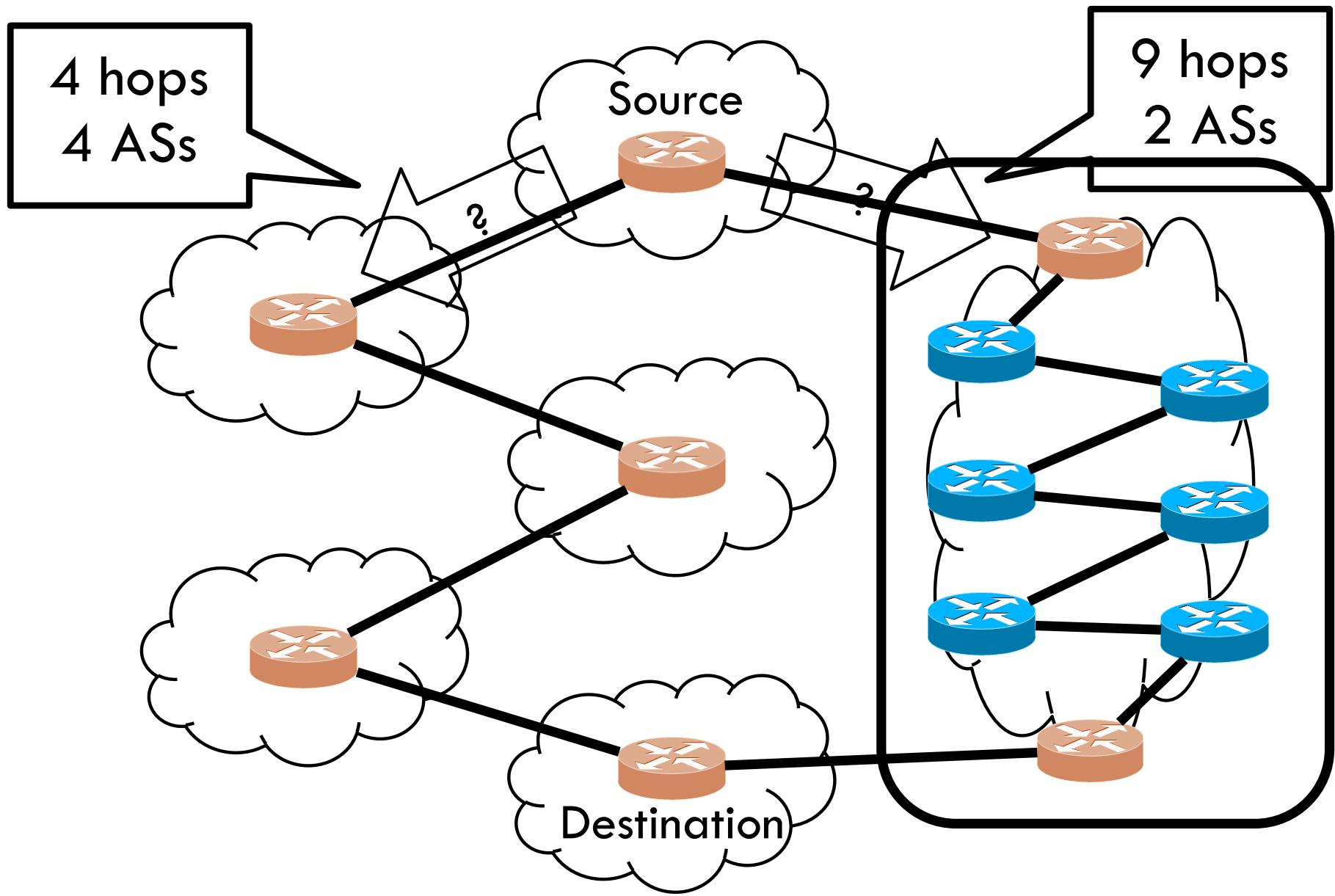
Traffic engineering

Lowest Router ID

**When all else fails,
break ties**

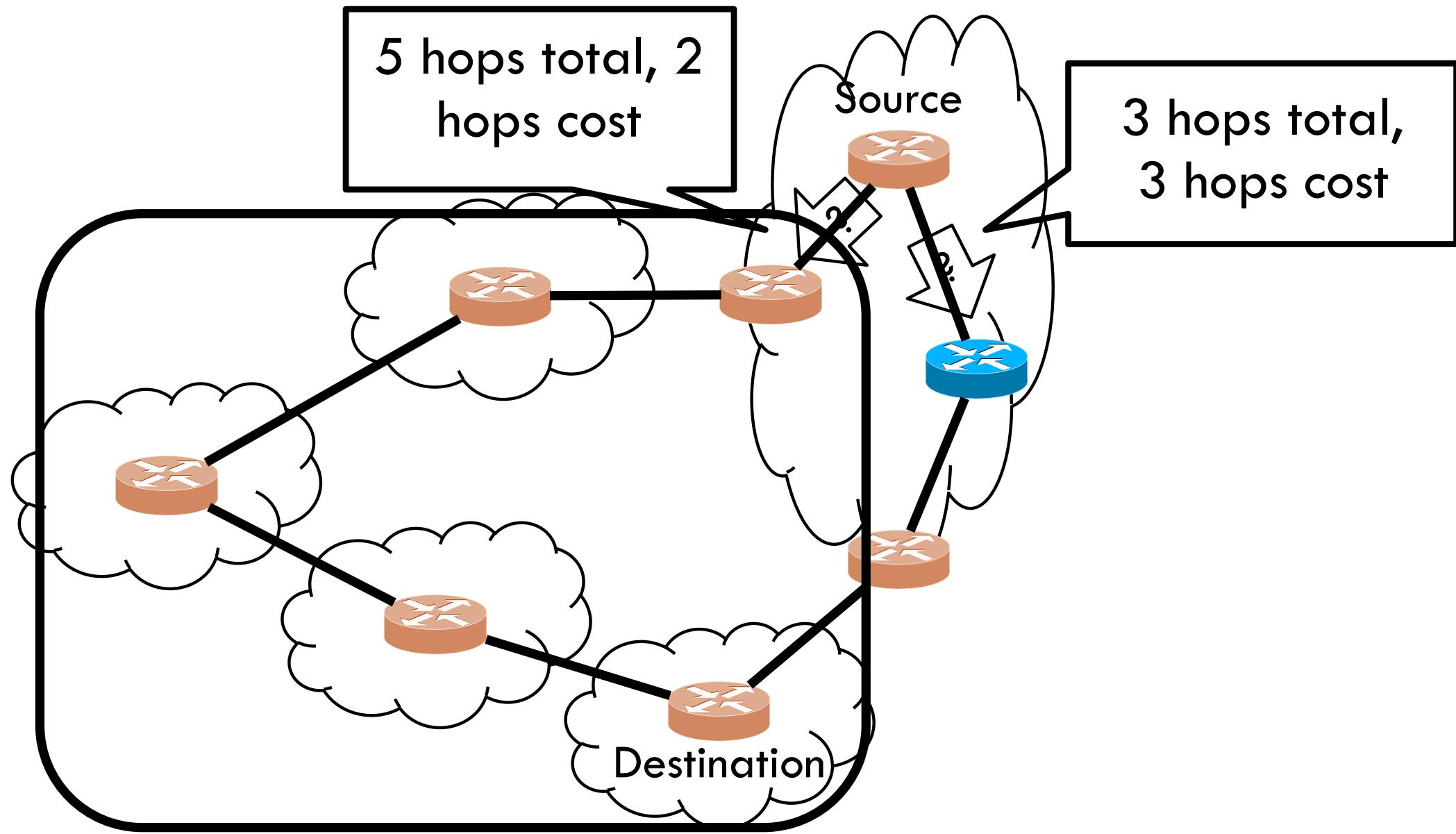
Shortest AS Path != Shortest Path

18



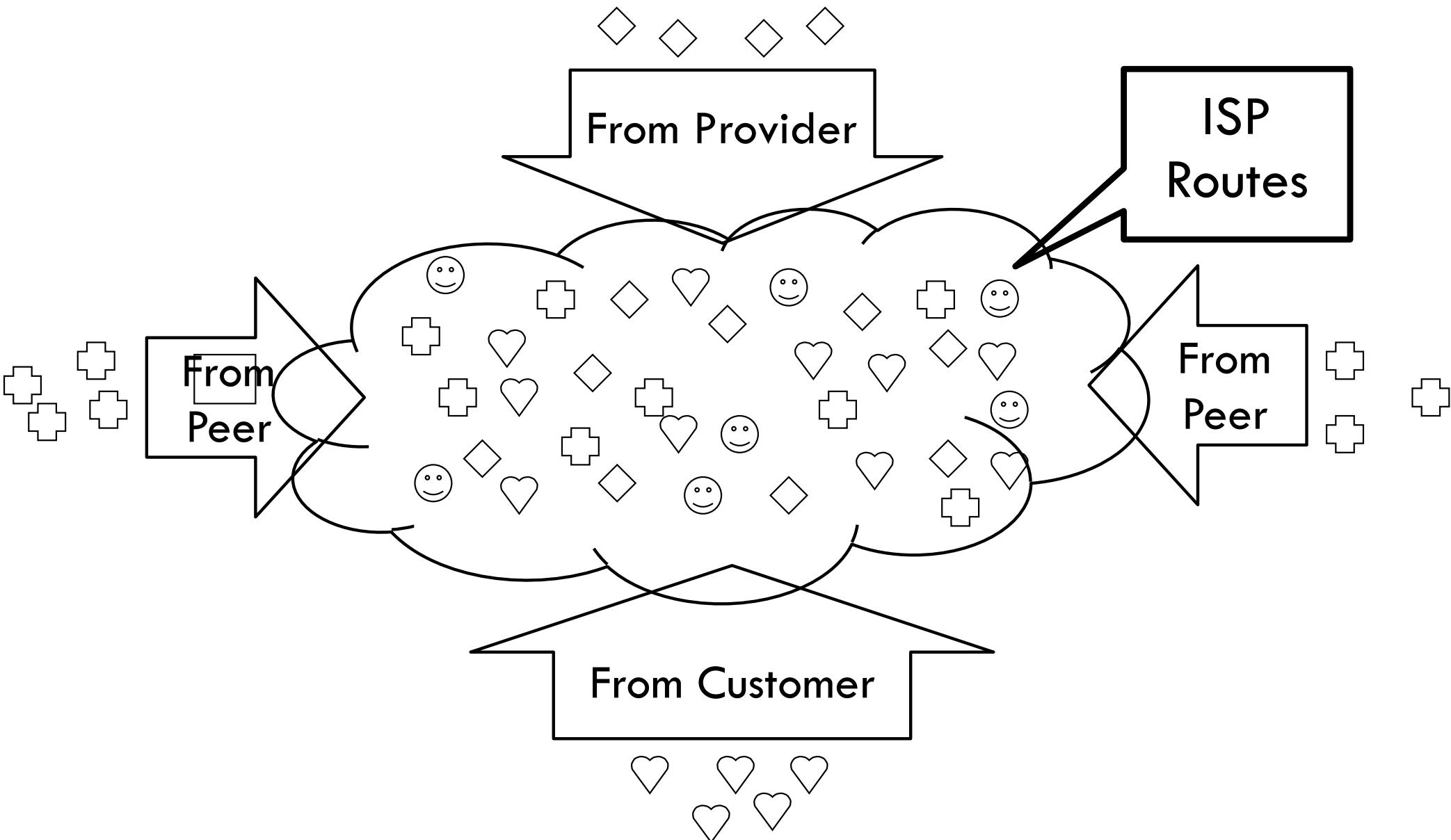
Hot Potato Routing

19



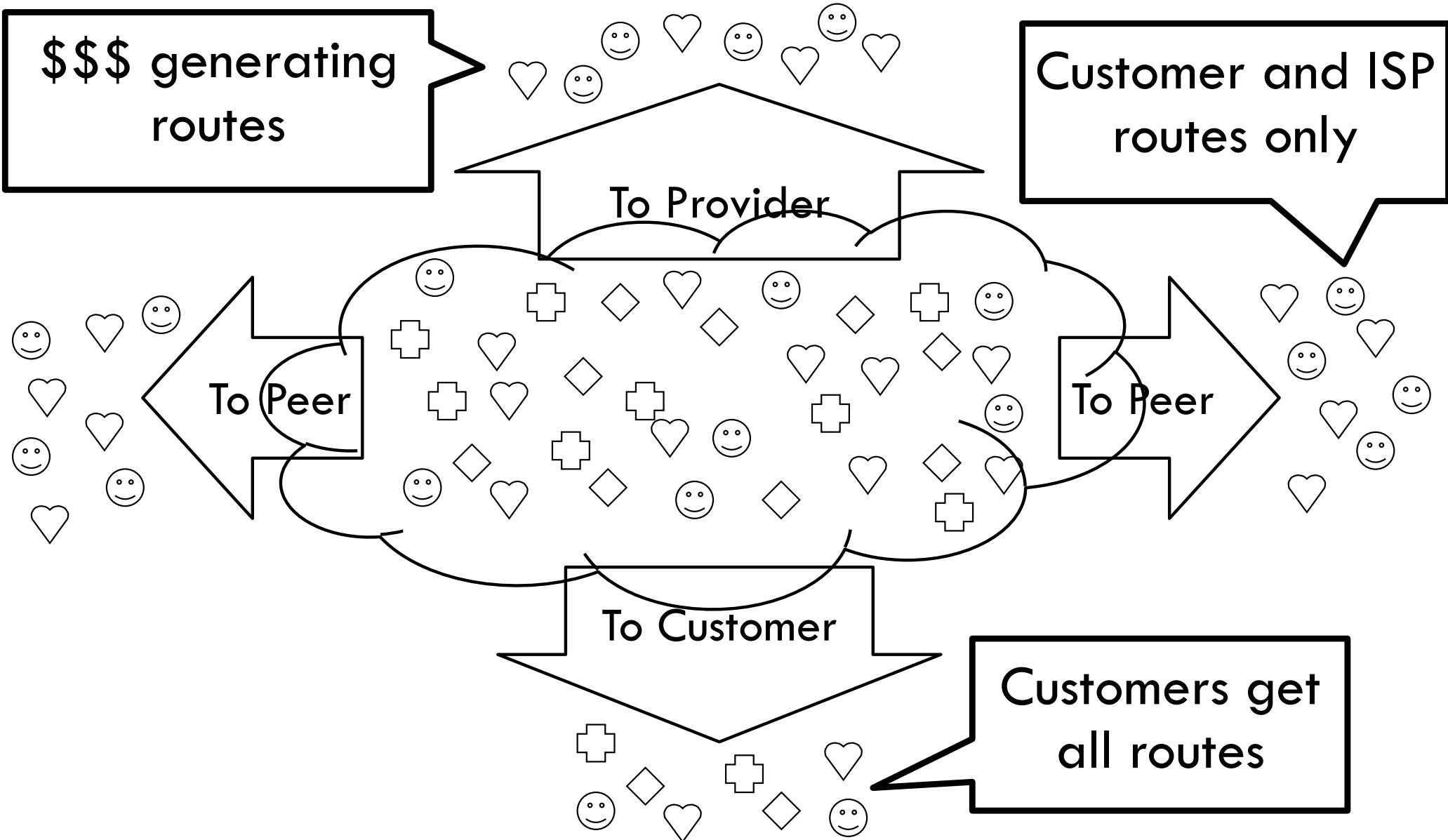
Importing Routes

20



Exporting Routes

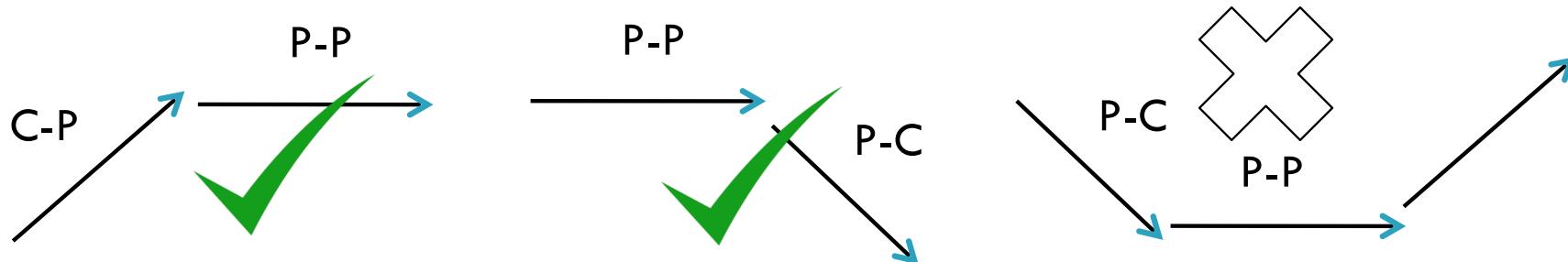
21



Modeling BGP

22

- AS relationships
 - Customer/provider
 - Peer
 - Sibling, IXP
- Gao-Rexford model
 - AS prefers to use customer path, then peer, then provider
 - Follow the money!
 - Valley-free routing
 - Hierarchical view of routing (incorrect but frequently used)



AS Relationships: It's Complicated

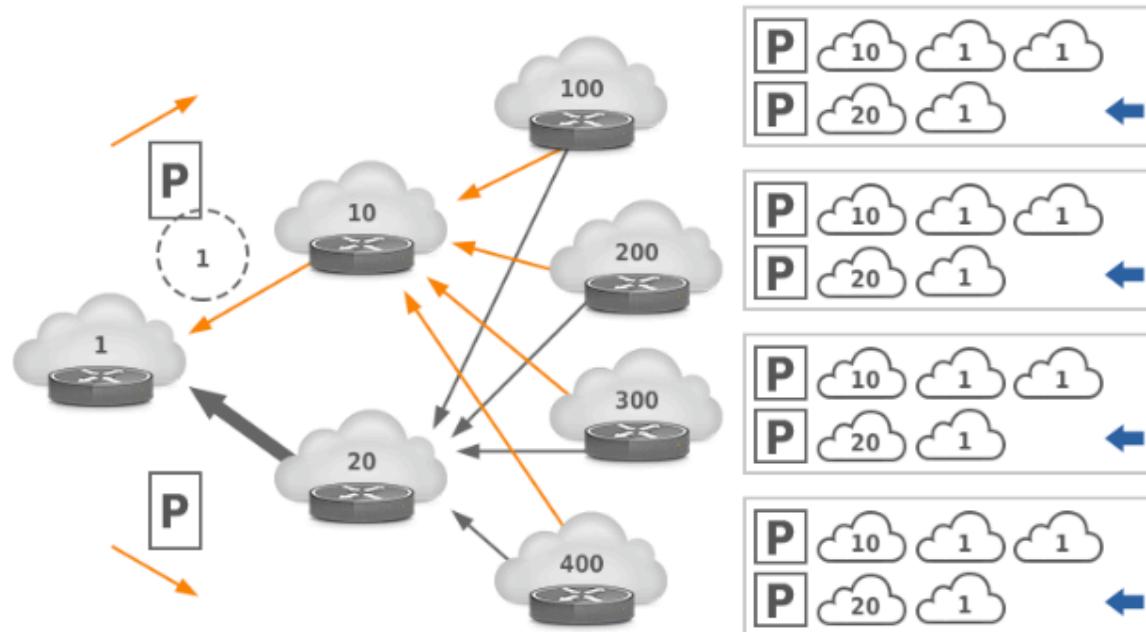
23

- GR Model is strictly hierarchical
 - ▣ Each AS pair has exactly one relationship
 - ▣ Each relationship is the same for all prefixes
- In practice it's much more complicated
 - ▣ Rise of widespread peering
 - ▣ Regional, per-prefix peerings
 - ▣ Tier-1's being shoved out by "hypergiants"
 - ▣ IXPs dominating traffic volume
- Modeling is very hard, very prone to error
 - ▣ Huge potential impact for understanding Internet behavior

Other BGP Attributes

24

- AS_SET
 - ▣ Instead of a single AS appearing at a slot, it's a set of Ases
 - ▣ Why?
- Prepending
 - ▣ Lengthening the route by adding multiple instances of ASN
 - ▣ Why?



Security Challenges?

25

- Any AS can announce ANY IP prefixes

How an Indonesian ISP took down the mighty Google for 30 minutes

26

Internet's web of trust let a company you never heard of block your Gmail.

SEAN GALLAGHER - 11/6/2012, 11:07 AM



Google's services went offline for many users for nearly a half-hour on the evening of November 5, thanks to an erroneous routing message broadcast by [Moratel](#), an Indonesian telecommunications company. The outage might have lasted even longer if it hadn't been spotted by a network engineer at CloudFlare who had a friend in a position to fix the problem.



The root cause of the outage was a configuration change to routers by Moratel, apparently intended to block access to Google's services from within Indonesia. The changes used the Border Gateway Protocol to "advertise" fake routes to Google servers, shunting traffic off to nowhere. But because of a misconfiguration, the BGP advertisements "leaked" through a peering connection in Singapore and spread to the wider Internet through Moratel's connection to the network of Hong Kong-based backbone provider PCCW. Google was interrupted in a similar way in 2008, when Pakistan Telecom moved to [block access to YouTube in Pakistan](#) because of an order from the Pakistani government.

Tom Paseka, a networking engineer at the content distribution network and Web security provider Cloudflare, spotted the source of the outage. "When I figured out the problem," Paseka wrote in [CloudFlare's blog](#) this morning, "I contacted a colleague at Moratel to let him know what was going on. He was able to fix the problem at around 2:50 UTC / 6:50pm PST. Around 3 minutes later, routing returned to normal and Google's services came back online."



By Marie Huillet

APR 24, 2018

MyEtherWallet Warns That A “Couple” Of Its DNS Servers Have Been Hacked

27



Update: [Data from EtherScan](#) shows that over \$150k worth of ETH has been stolen in the DNS hack. Starting from 07:17 this morning, 179 inbound transactions totaling 216.06 ETH were sent to ETH address 0x1d50588C0aa11959A5c28831ce3DC5F1D3120d29. At 10:15, the attacker sent 215 ETH to 0x68ca85dbf8eba69fb70ecdb78e0895f7cd94da83.