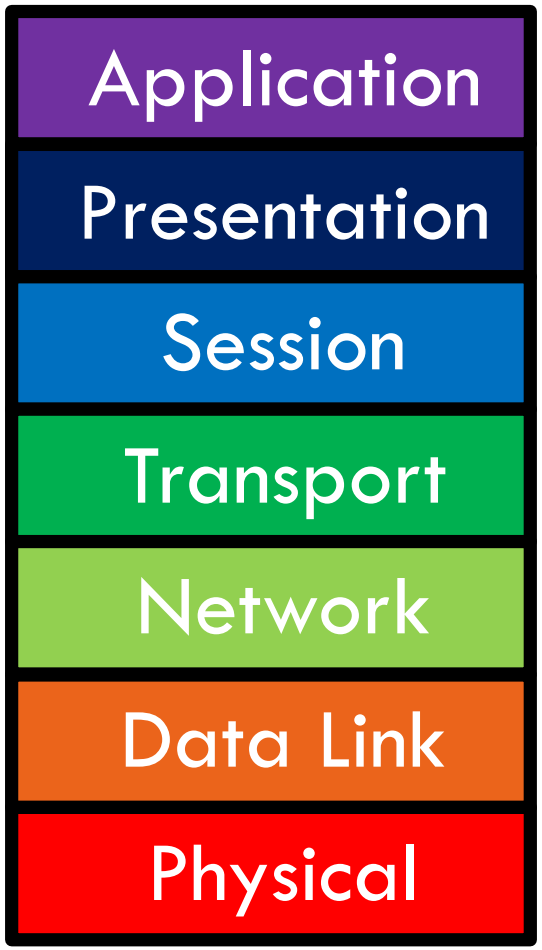
CSCI-351
Data communication and Networks

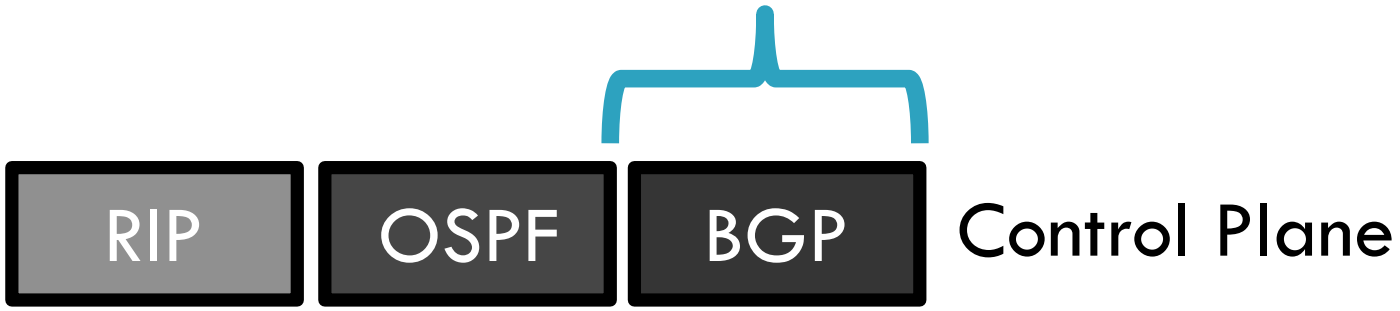
Lecture 10: Inter Domain Routing (It's all about the Money)

The slide is built with the help of Prof. Alan Mislove, Christo Wilson, and David Choffnes's class

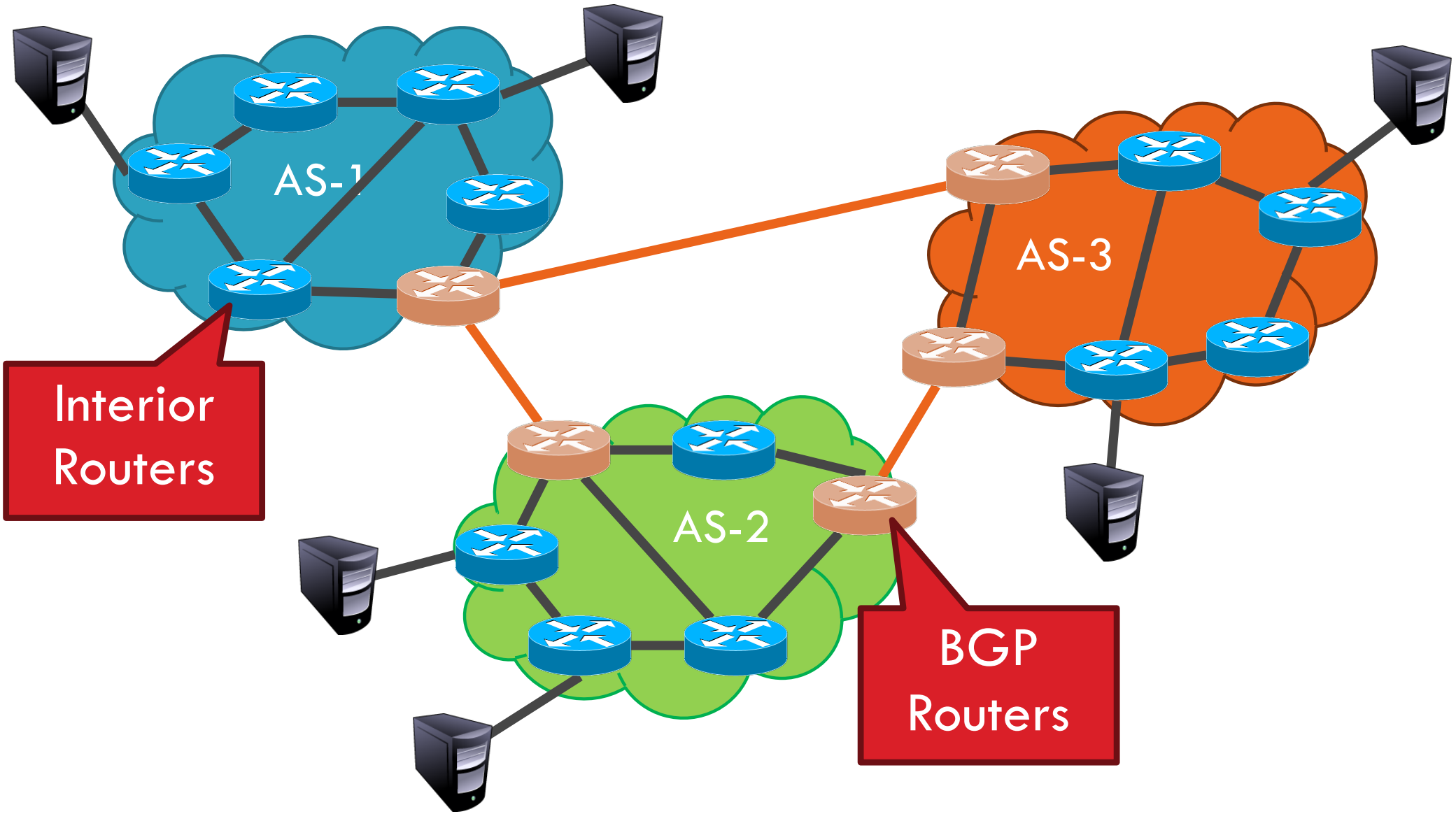
Network Layer, Control Plane

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





ASs, Revisited



AS Numbers

- Each AS identified by an ASN number
  - 16-bit values (latest protocol supports 32-bit ones)
  - 64512 – 65535 are reserved
- Currently, there are > 60000 ASNs
  - AT&T: 5074, 6341, 7018, …
  - Sprint: 1239, 1240, 6211, 6242, …
  - RIT: 4385,
  - North America ASs → ftp://ftp.arin.net/info/asn.txt
  - bgp.he.net is a good tool

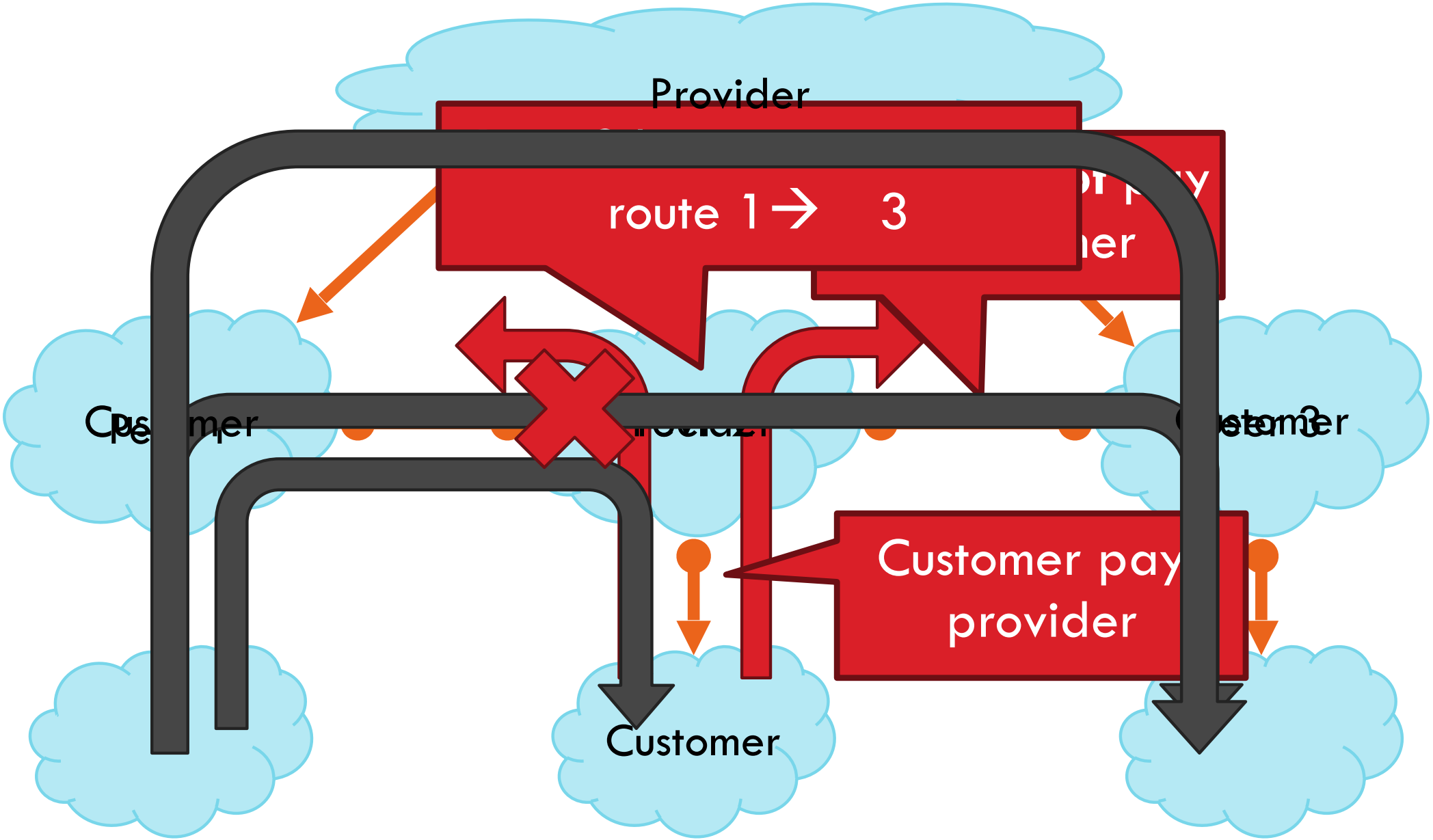
Inter-Domain Routing

- 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

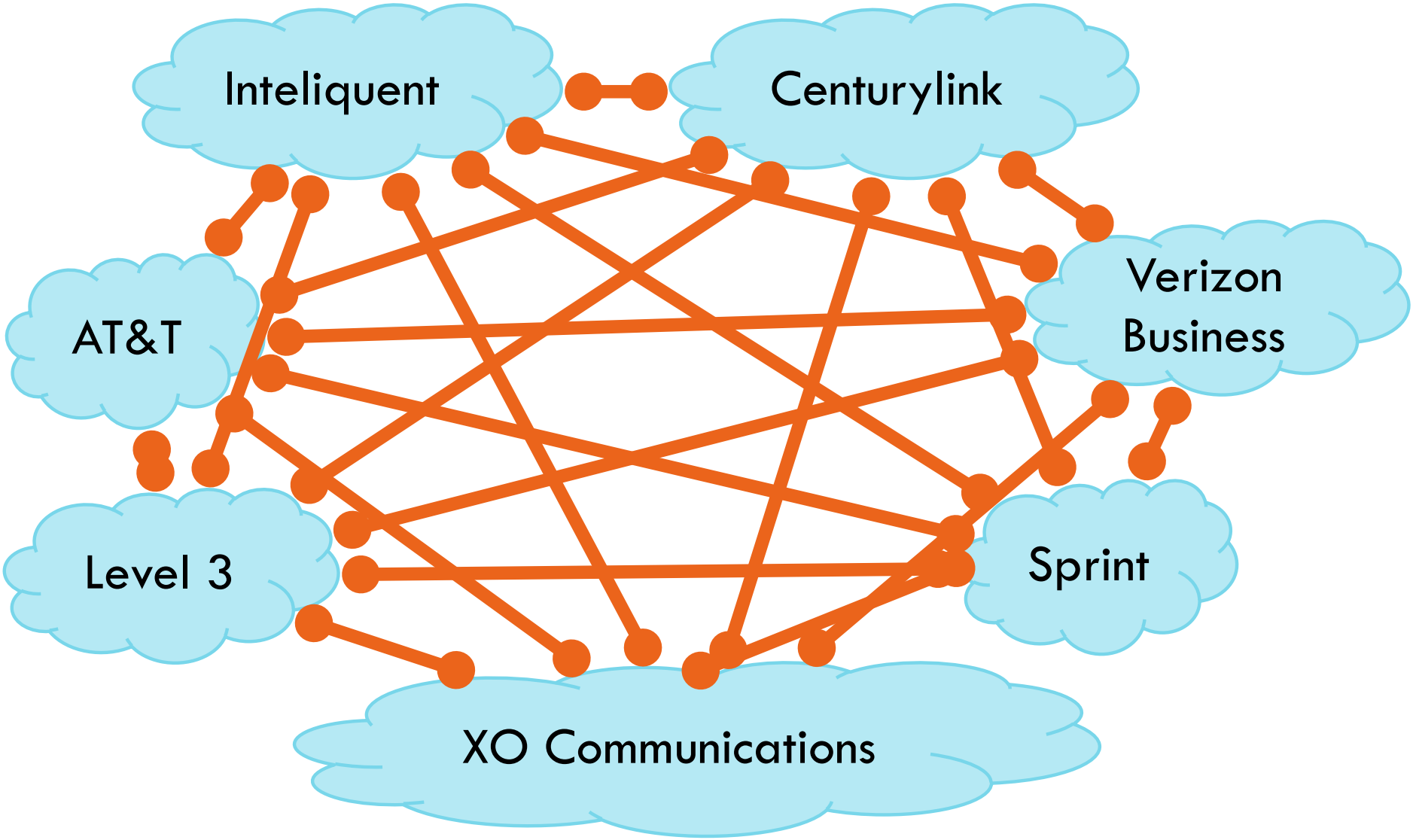
BGP

- 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)

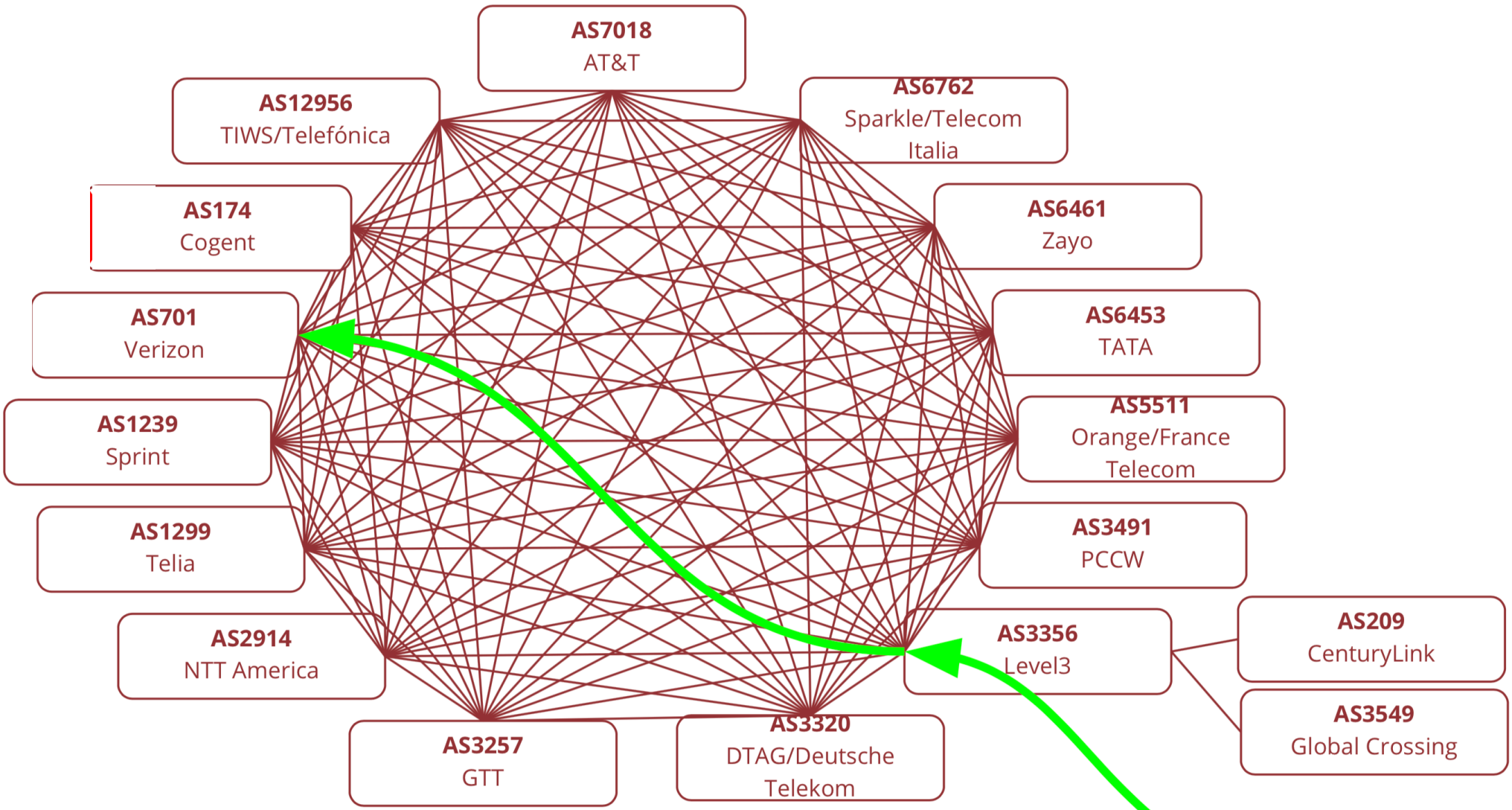
BGP Relationships



Tier-1 ISP Peering



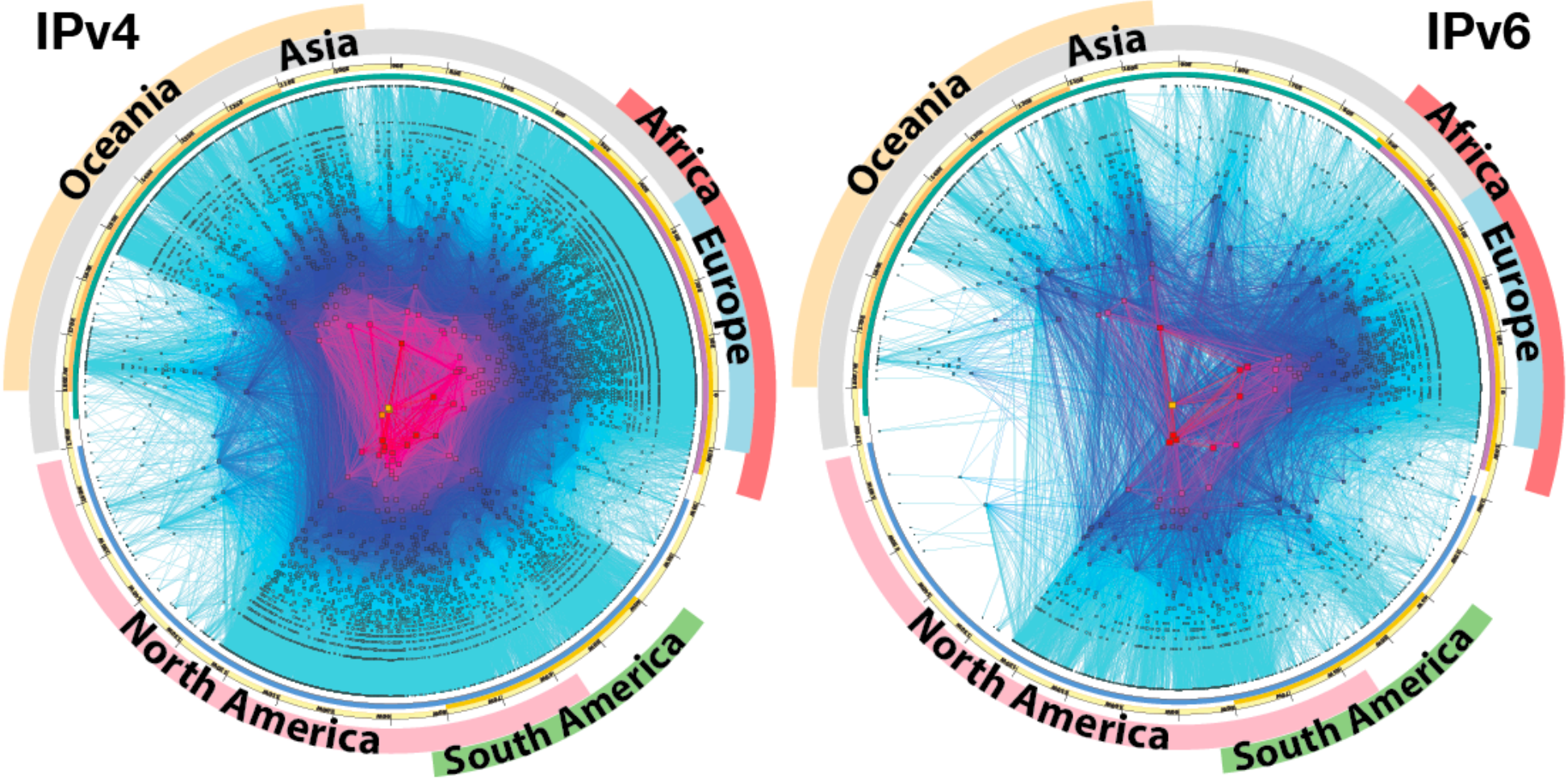
Tier-1 ISP Peering



Slide is from Martin Levy's Talk (Cloudflare) at APNIC48

CAIDA's IPv4 vs IPv6 AS Core AS-level Internet Graph

Archipelago July 2015

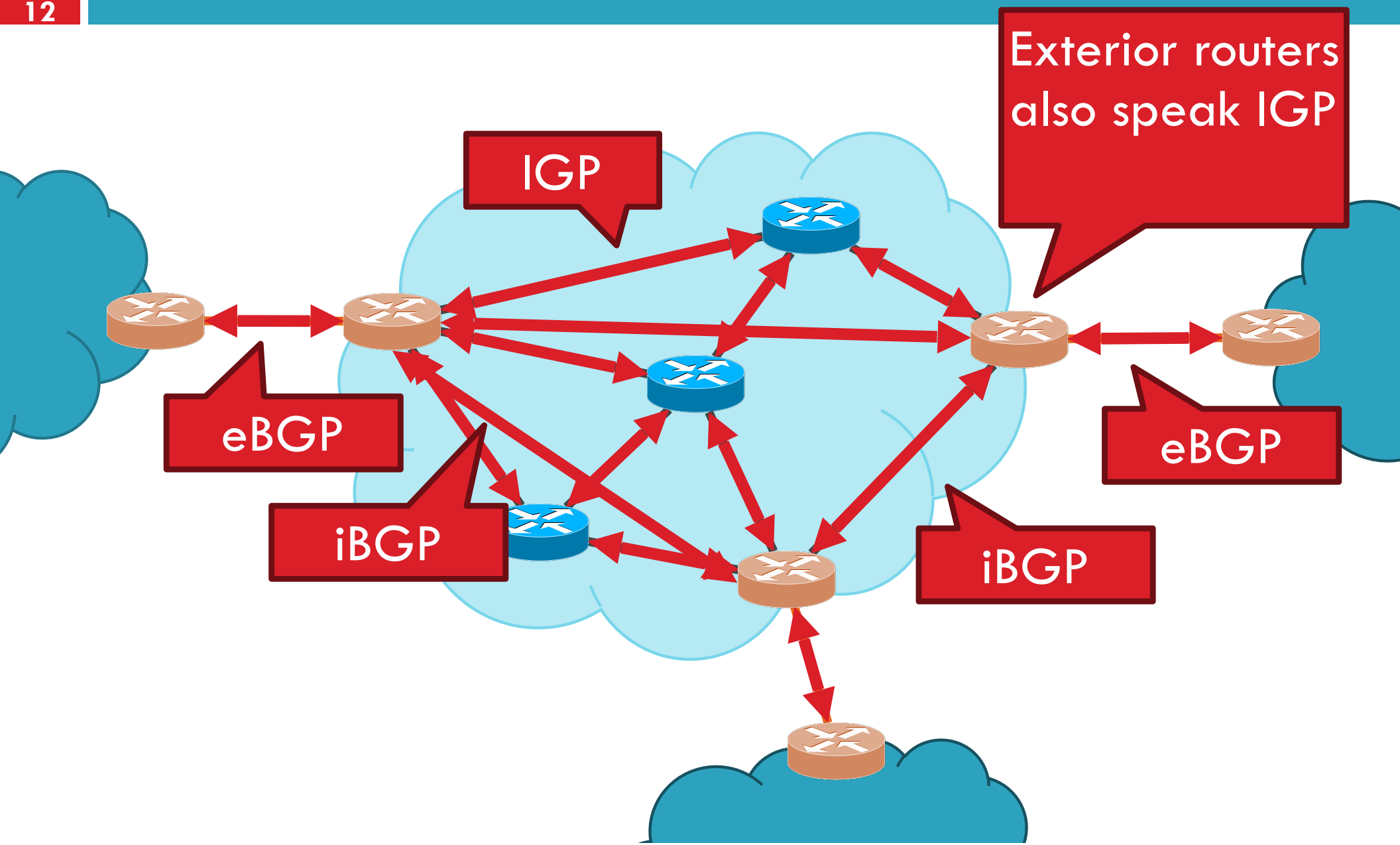


Peering Wars

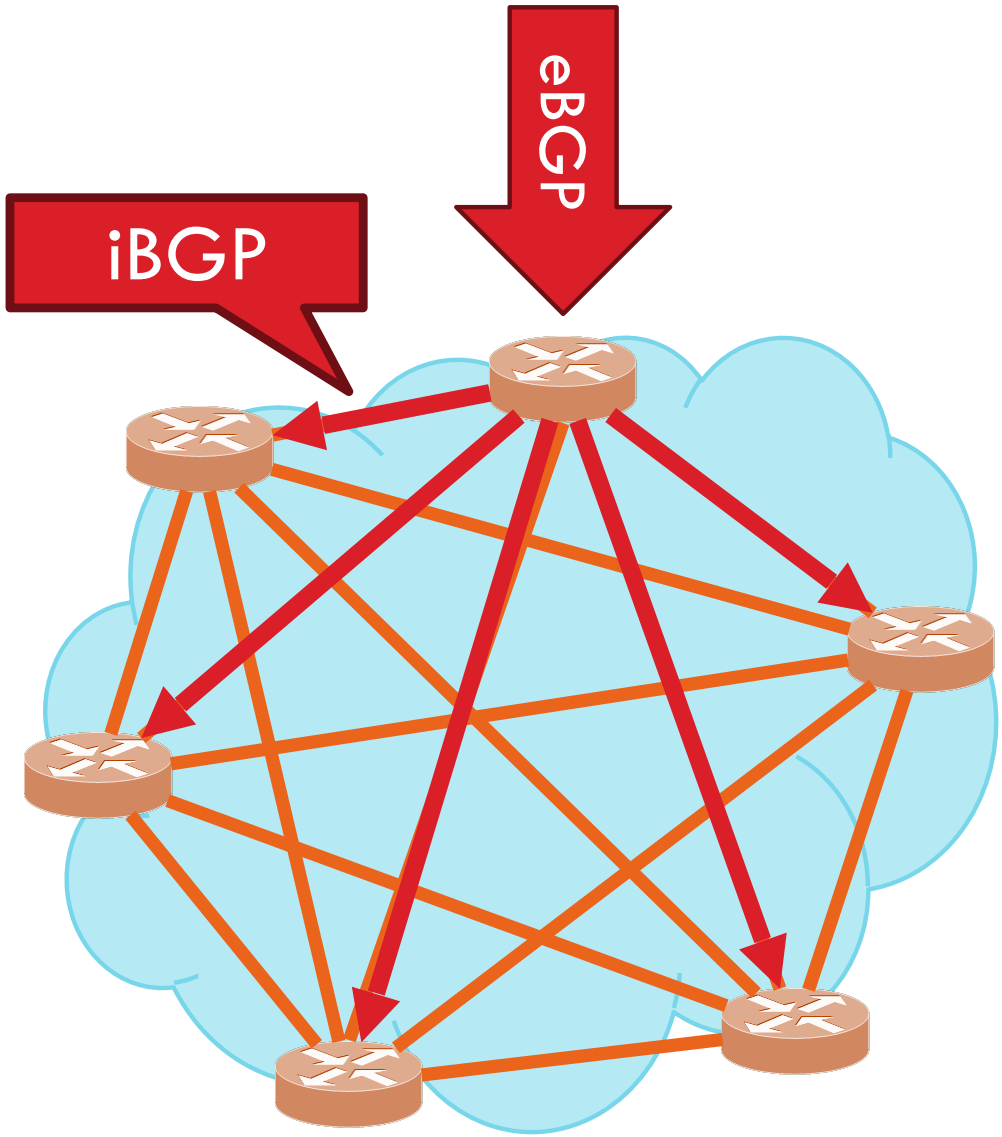
Peer	Don't Peer
Reduce upstream costs	You would rather have customers
Improve end-to-end performance	Peers are often competitors
May be the only way to connect to parts of the Internet	Peering agreements require periodic renegotiation

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

Two Types of BGP Neighbors



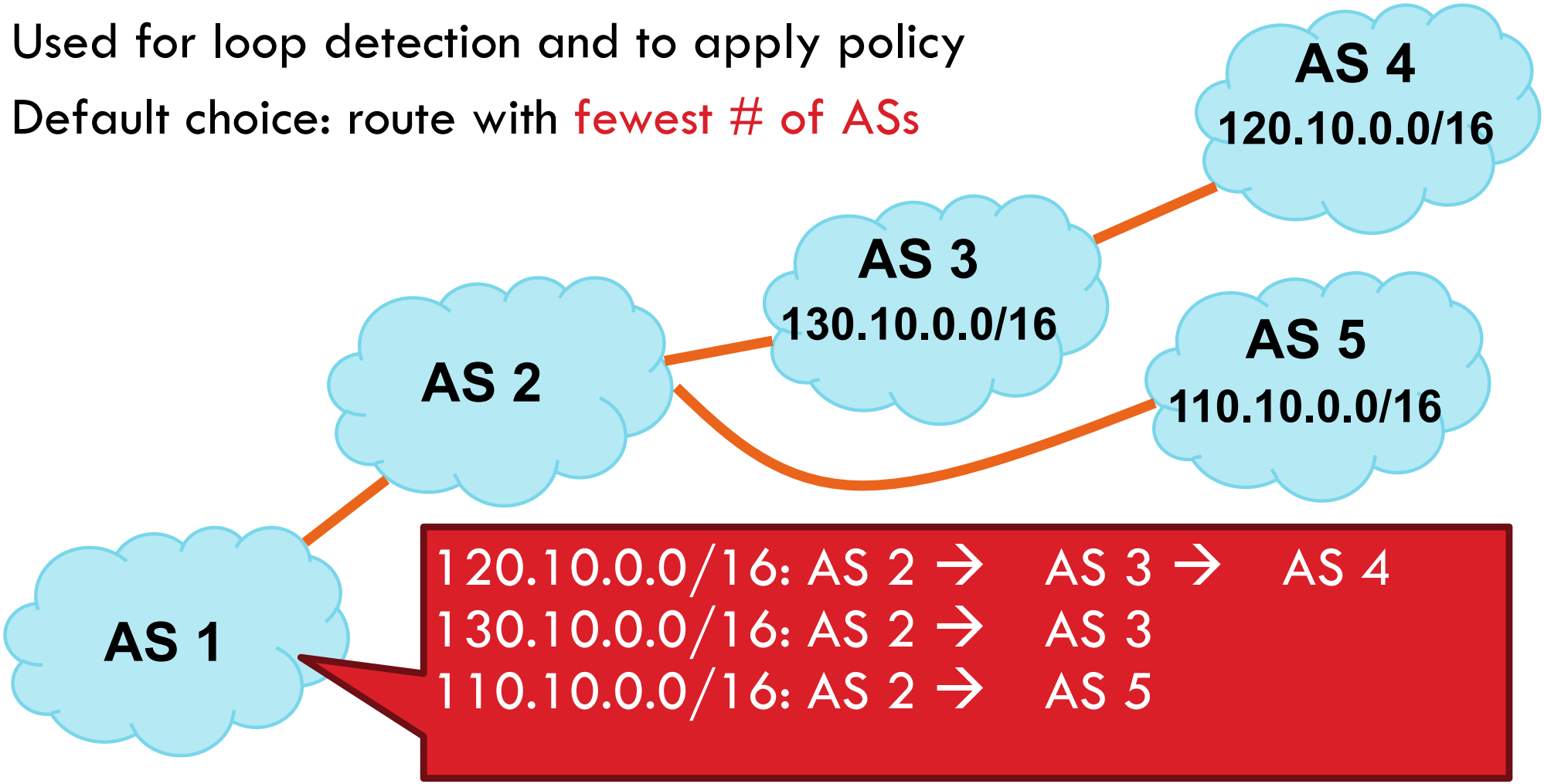
Full iBGP Meshes



- 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

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

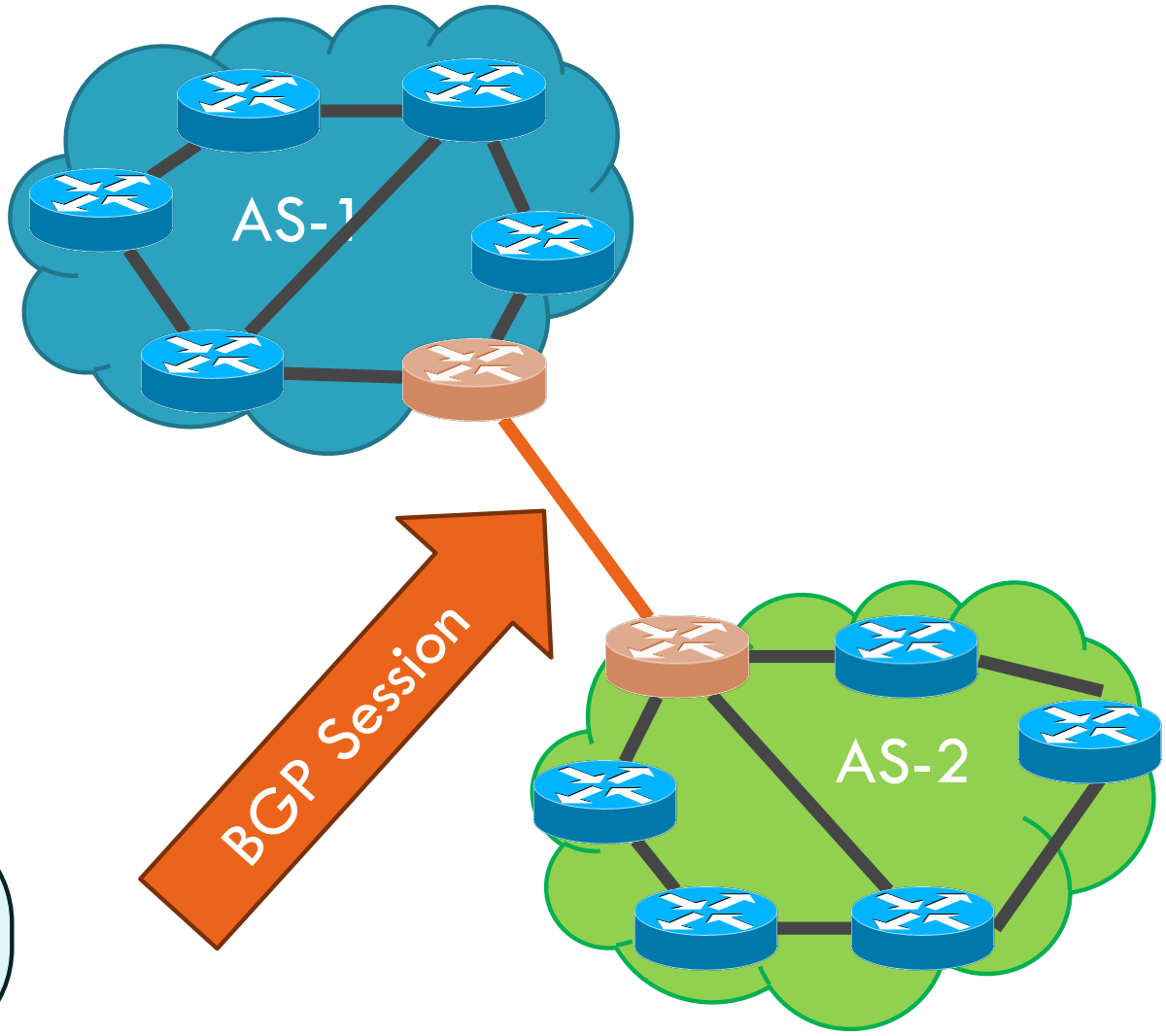


BGP Operations (Simplified)

Establish session on TCP port 179

Exchange active routes

Exchange incremental updates



Four Types of BGP Messages

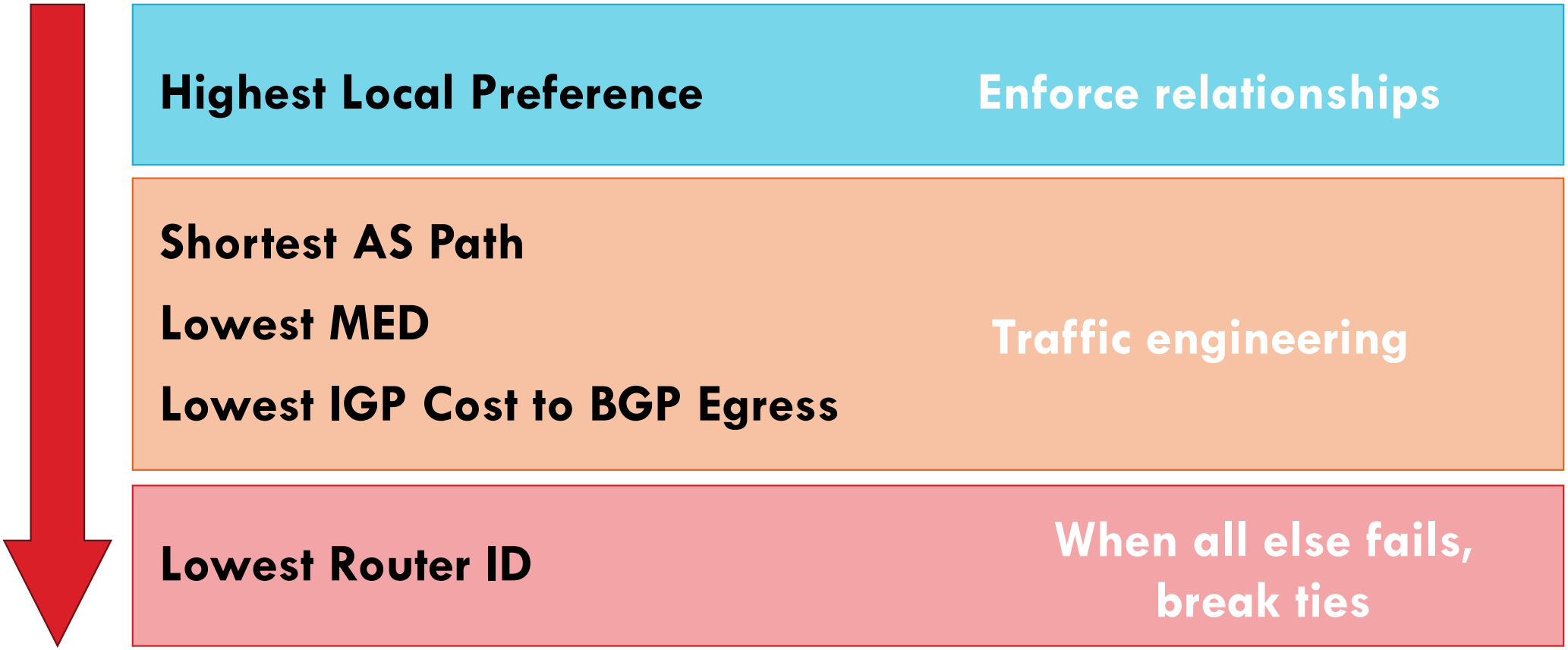
- Open: Establish a peering session.
- Keep Alive: Handshake at regular intervals.
- Notification: Shuts down a peering session.
- Update: Announce new routes or withdraw previously announced routes.

announcement = IP prefix + attributes values

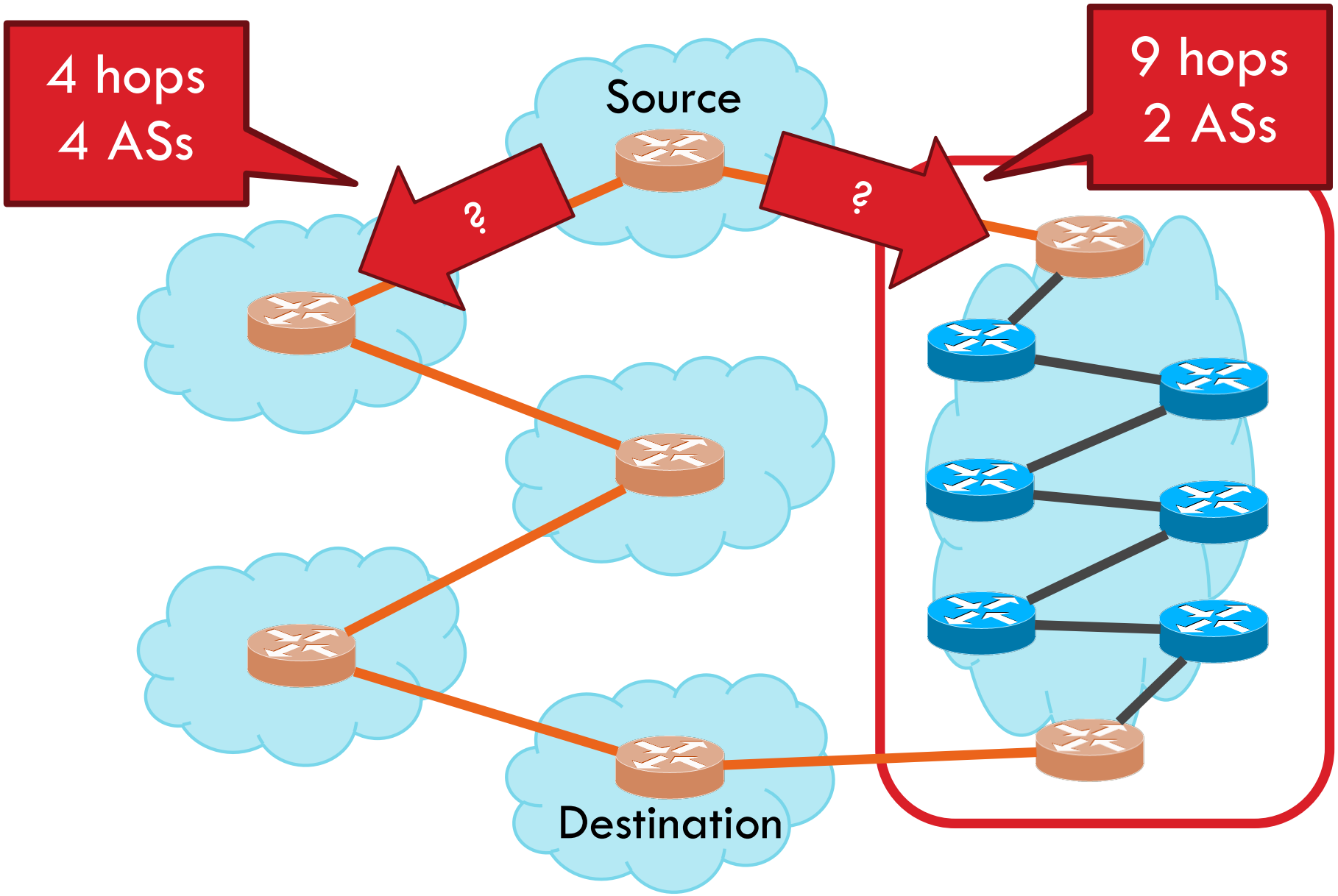
BGP Attributes

- 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?

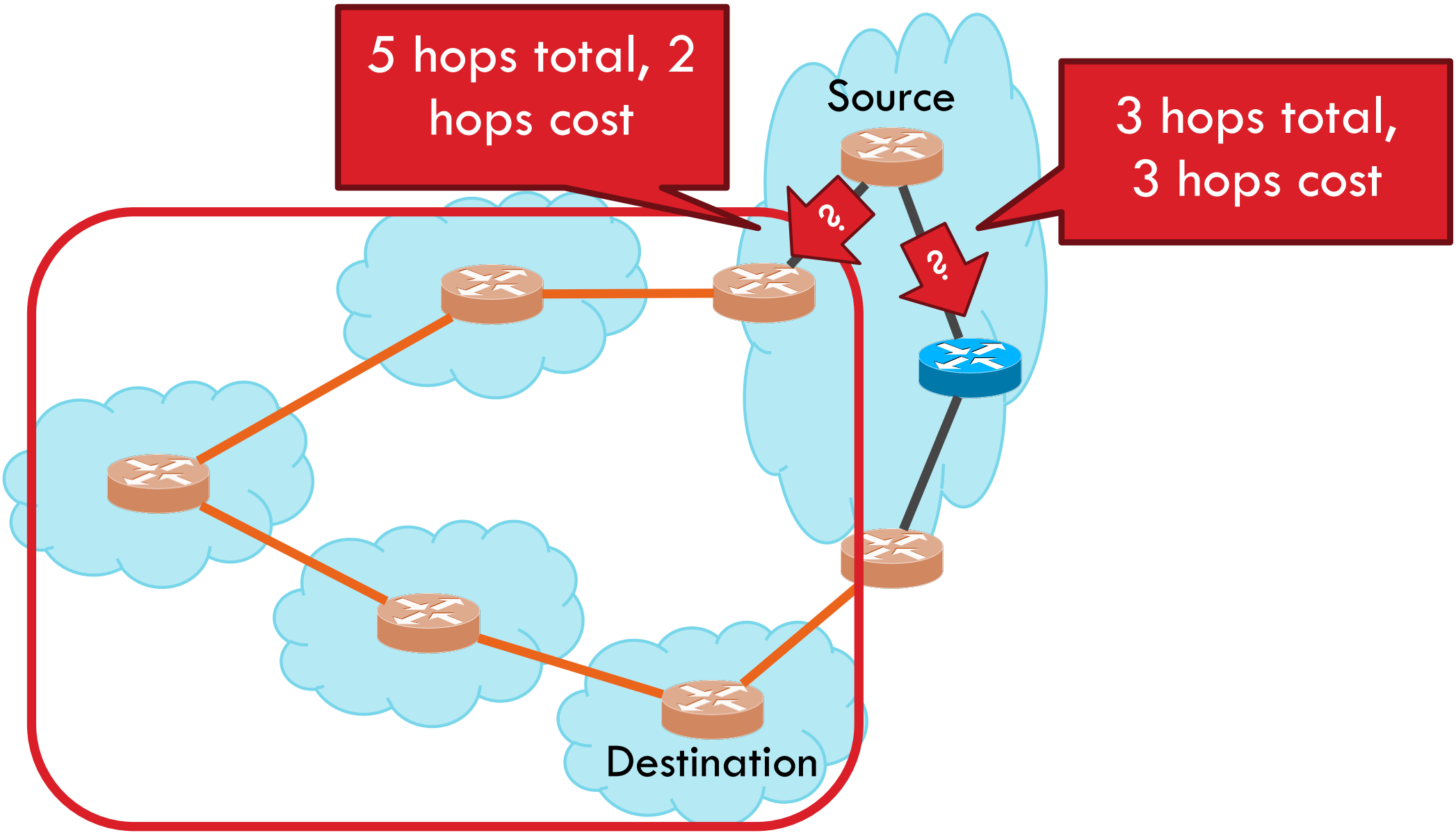
Route Selection Summary



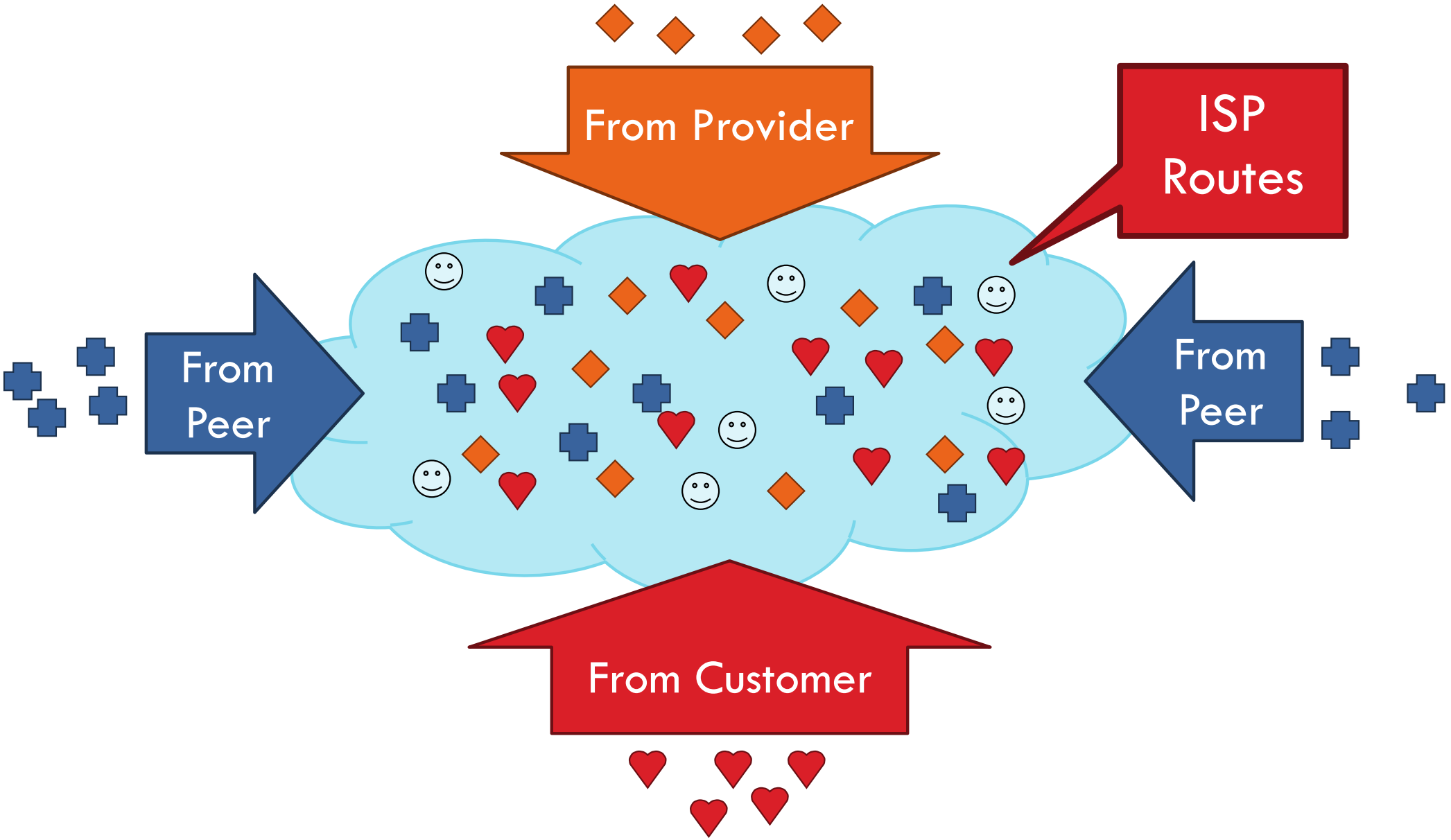
Shortest AS Path != Shortest Path



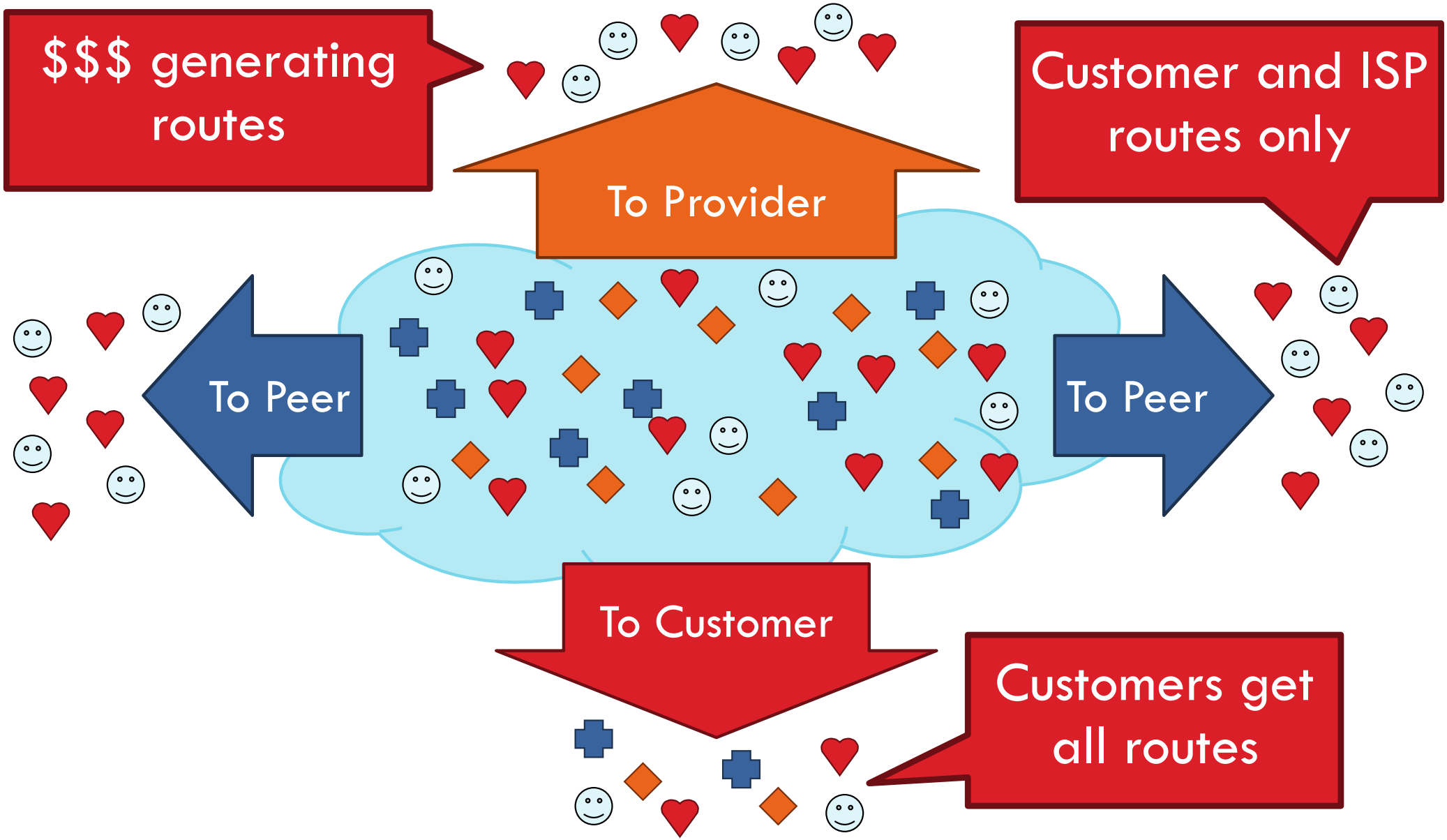
Hot Potato Routing



Importing Routes

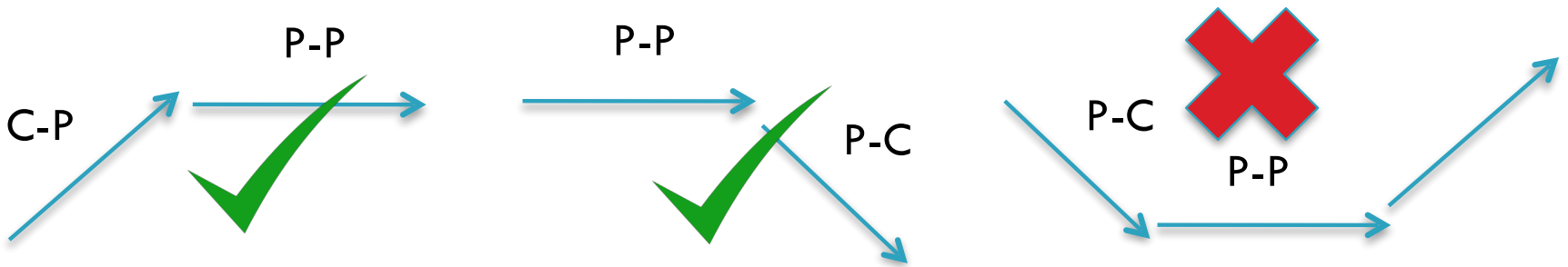


Exporting Routes



Modeling BGP

- 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)

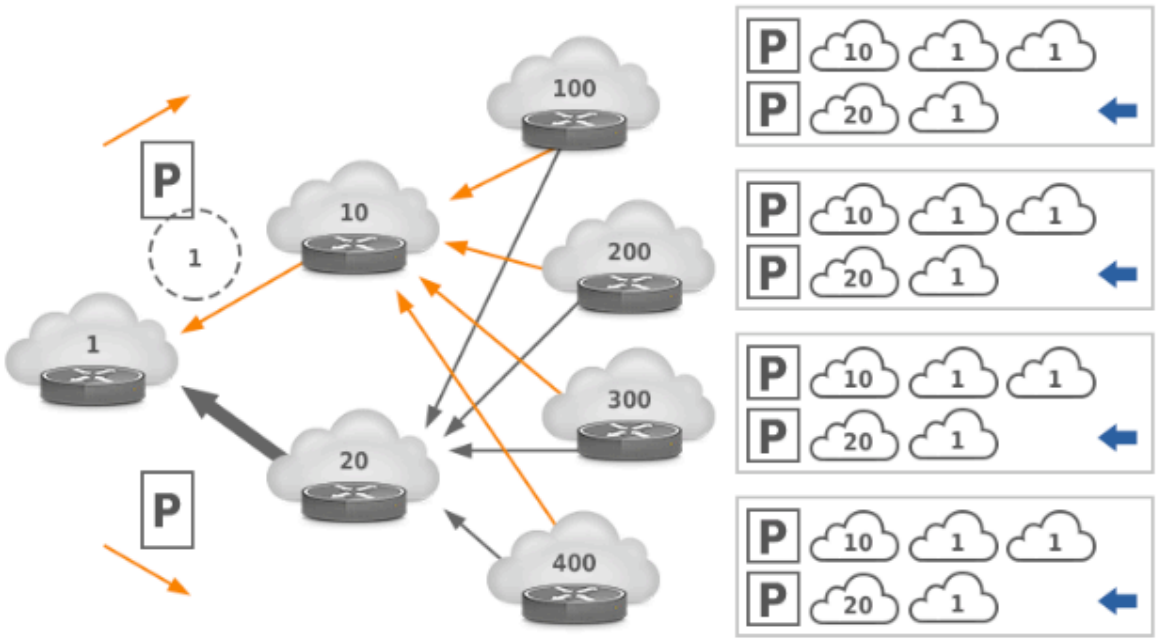


AS Relationships: It's Complicated

- 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

- 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?



An example of BGP updates

Advertised IP Prefix

01/01/19 15:03:04 | A | 218.189.6.2 | 9304 | 217.171.93.0/24 | 9304 6453 701 22351 25395 43256 | IGP

The last hop of BGP updates:
AS9304 HGC Global Communications Limited

The owner ASN:
AS43256 Global Broadband Solution Inc (In Belgium)

Security Challenges?

- Any AS can announce ANY IP prefixes
  - What?

BGP Route leak

THE POWER OF FALSE ADVERTISING —

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

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."

BGP Subprefix Hijacking

By Marie Huillet

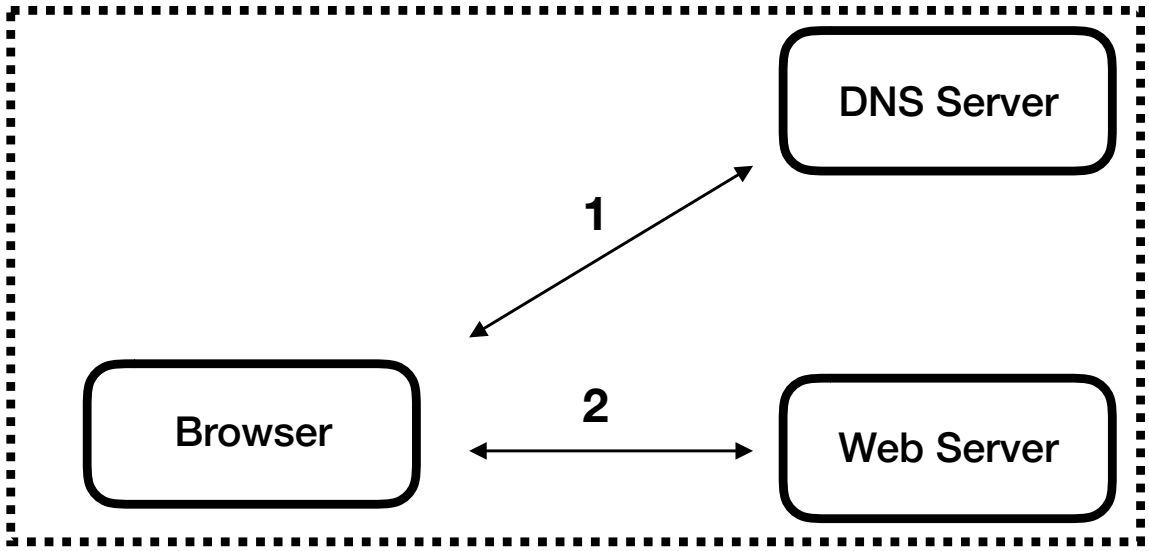
APR 24, 2018

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

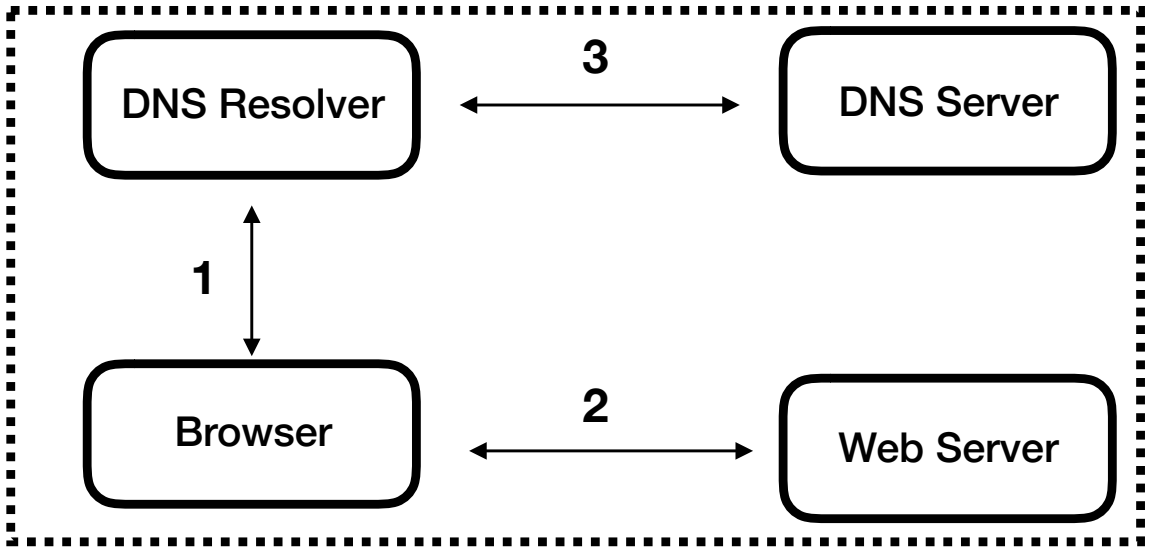
Update: Data from EtherScan shows that over $150k worth of ETH has been stolen in the DNS hack. Starting from 07:17 this morning, 179 inbound transactions totaling 216.06 ETH were sent to ETH address 0x1d50588C0aa11959A5c28831ce3DC5F1D3120d29. At 10:15, the attacker sent 215 ETH to 0x68ca85dbf8eba69fb70ecdb78e0895f7cd94da83.

Sneak peek of security-related lectures

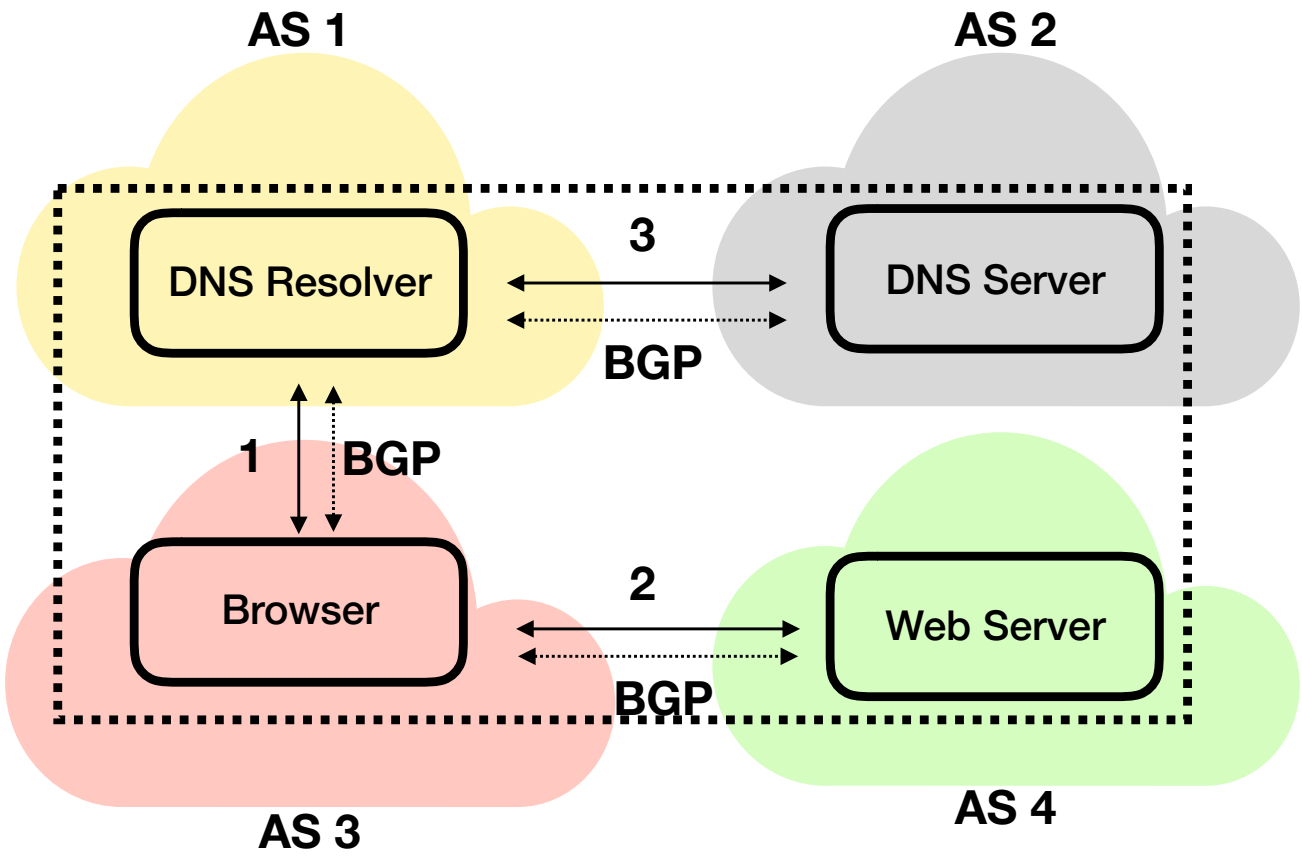
When you request a webpage (1)



When you request a webpage (2)



Underlying structure

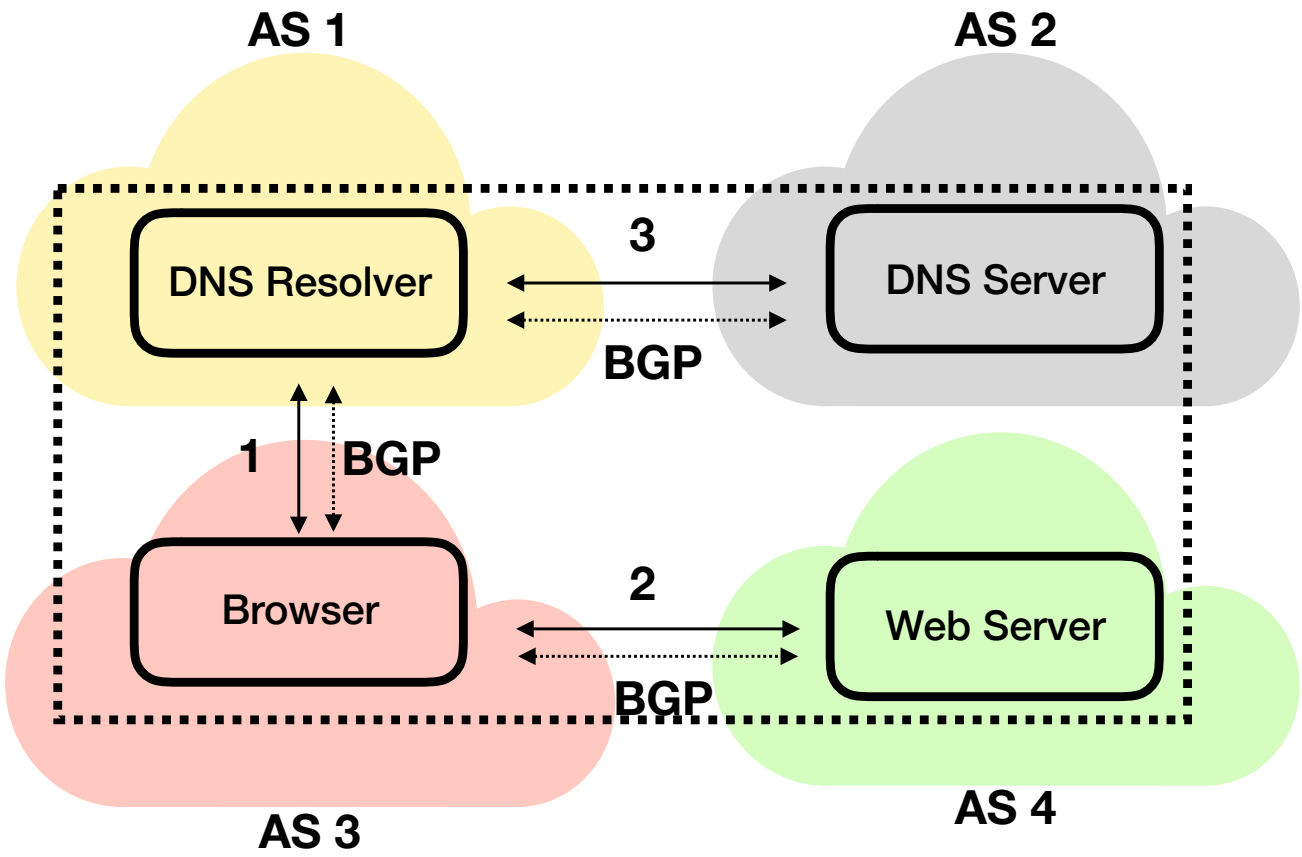


By Marie Huillet APR 24, 2018

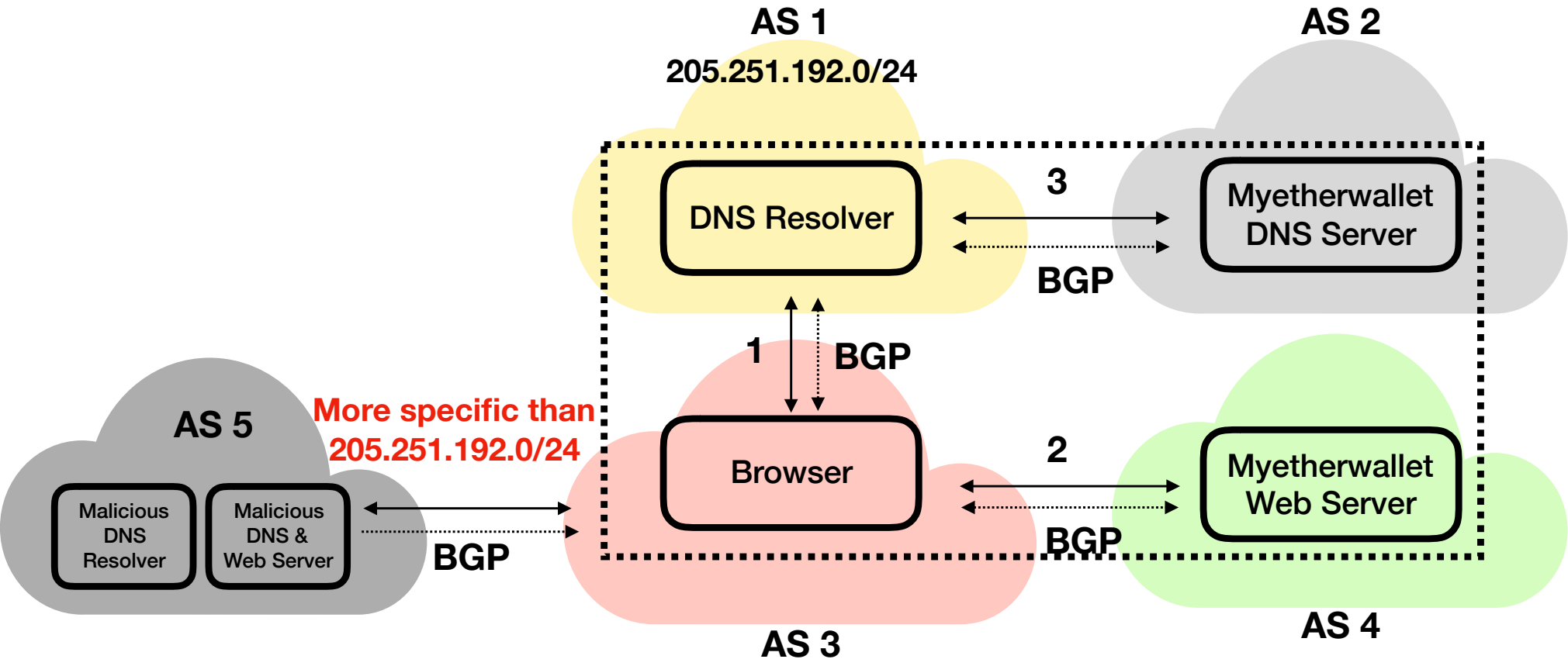
MyEtherWallet Warns That A “Couple” Of Its DNS Servers Have Been Hacked

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.

Underlying structure



Underlying structure



Another Routeleak

HOME > NEWS > OUTAGES

Verizon BGP route leak causes Cloudflare customer outages, AWS issues

Another week, another BGP issue

June 24, 2019 By: Sebastian Moss

How Verizon and a BGP Optimizer Knocked Large Parts of the Internet Offline Today

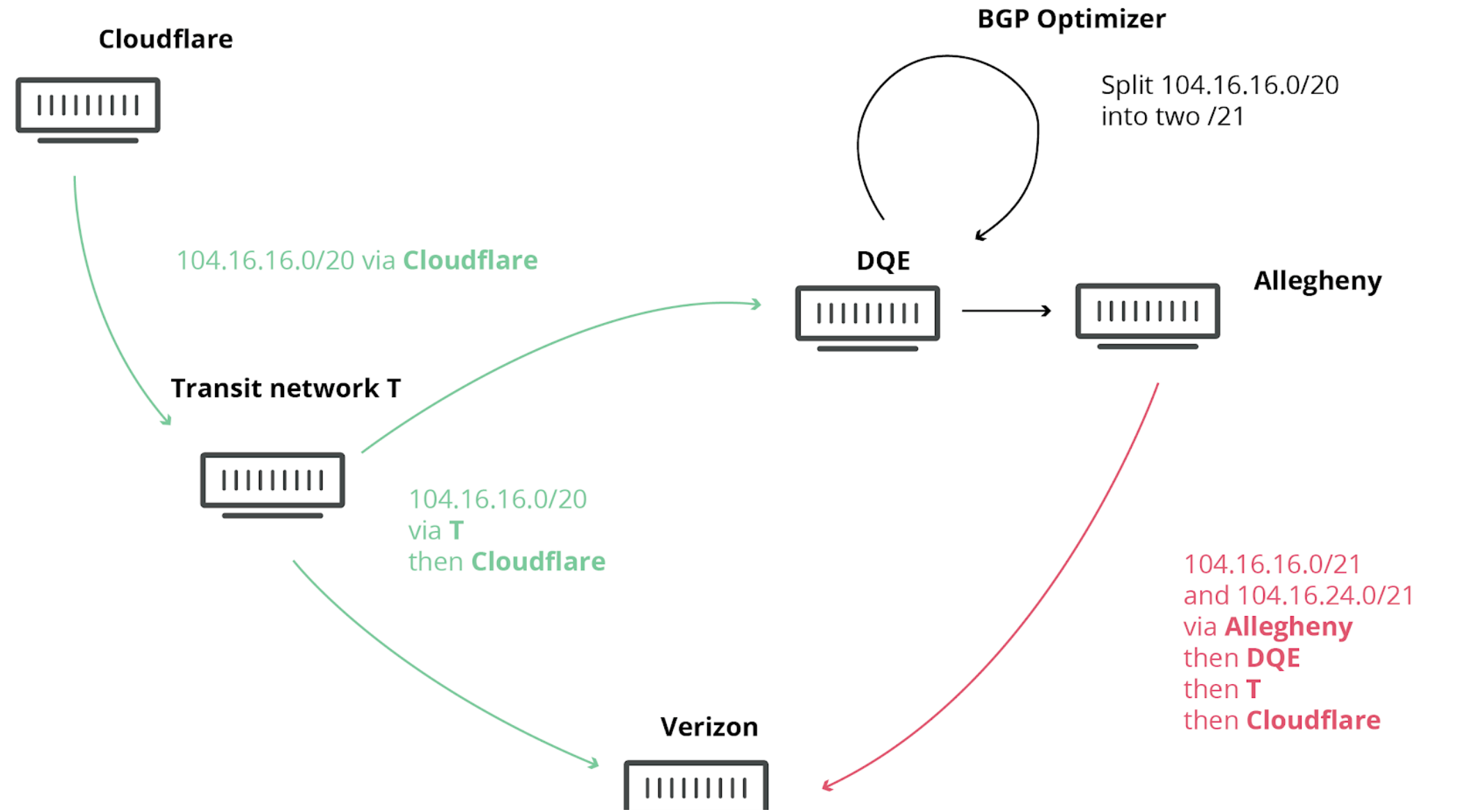
Share Like 4.8K Tweet

Tom Strickx

June 24, 2019 7:58PM

Massive route leak impacts major parts of the Internet, including Cloudflare

How it works (from 10,000 ft view)



More details in https://blog.cloudflare.com/how-verizon-and-a-bgp-optimizer-knocked-large-parts-of-the-internet-offline-today/

How can we prevent from these attacks?

- Here RPKI (Resource PKI) comes to save you
  - Cryptographic signature to sign “who” is authorized to announce “which” IP prefixes
  - (may) talk about this near the end of the semester

Again

- Why this is happening?
- Security matters
  - after it performs very well? (my thoughts)

TINDER'S LACK OF ENCRYPTION LETS STRANGERS SPY ON YOUR SWIPES

MAI SCHOTZ

IN 2018, YOU'D be forgiven for assuming that any sensitive app encrypts its connection from your phone to the cloud, so

Again

- Why this is happening?
- Security matters
  - after it performs very well?
    (my thoughts)
- Privacy matters
- Network Security course CSCI-759 (Topics in System)