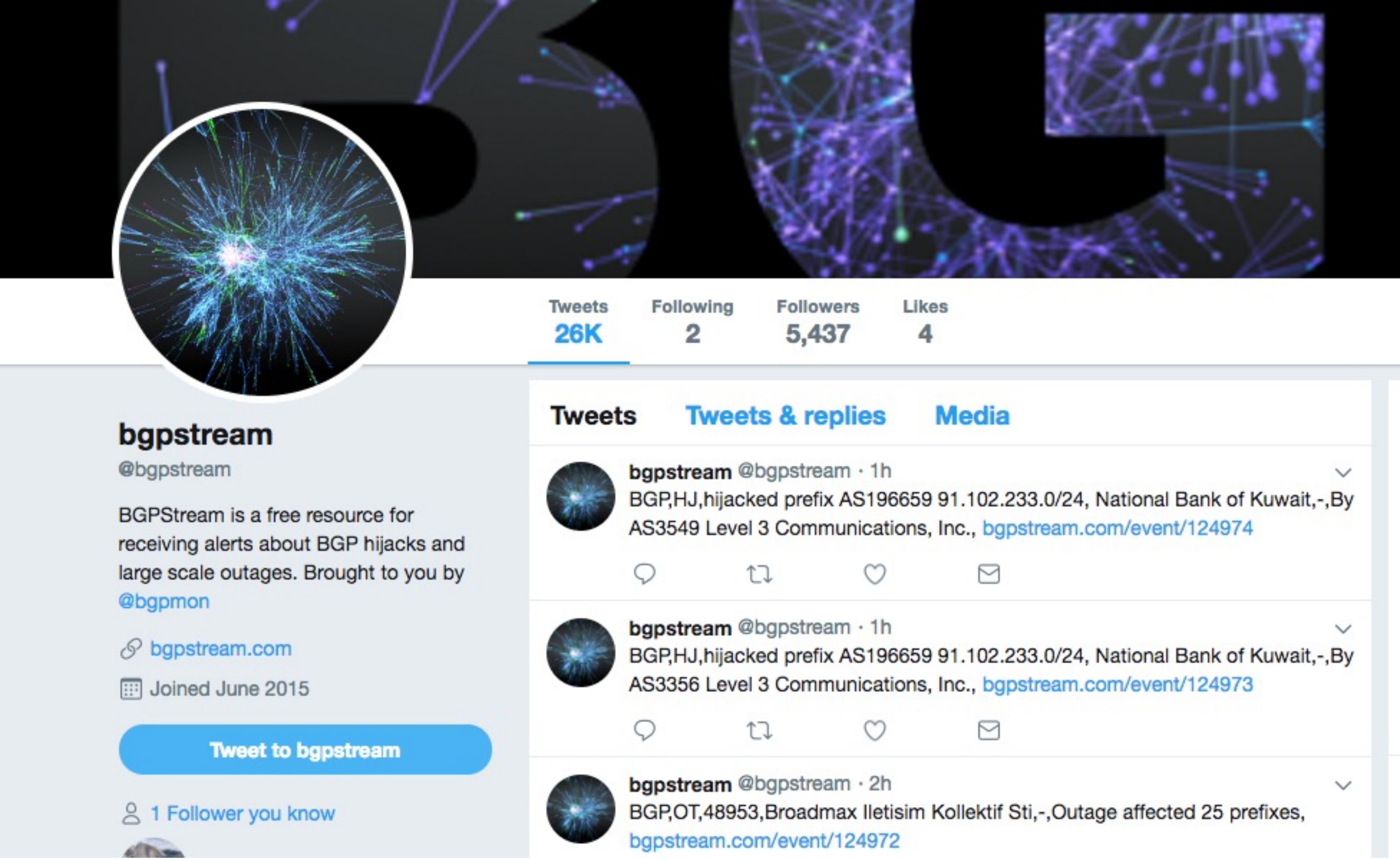
Routing and Security



bgpstream
@bgpstream

BGPStream is a free resource for receiving alerts about BGP hijacks and large scale outages. Brought to you by @bgpmon

bgpstream.com
Joined June 2015

bgpstream @bgpstream · 1h
BGP,HJ,hijacked prefix AS196659 91.102.233.0/24, National Bank of Kuwait,-,By AS3549 Level 3 Communications, Inc., bgpstream.com/event/124974

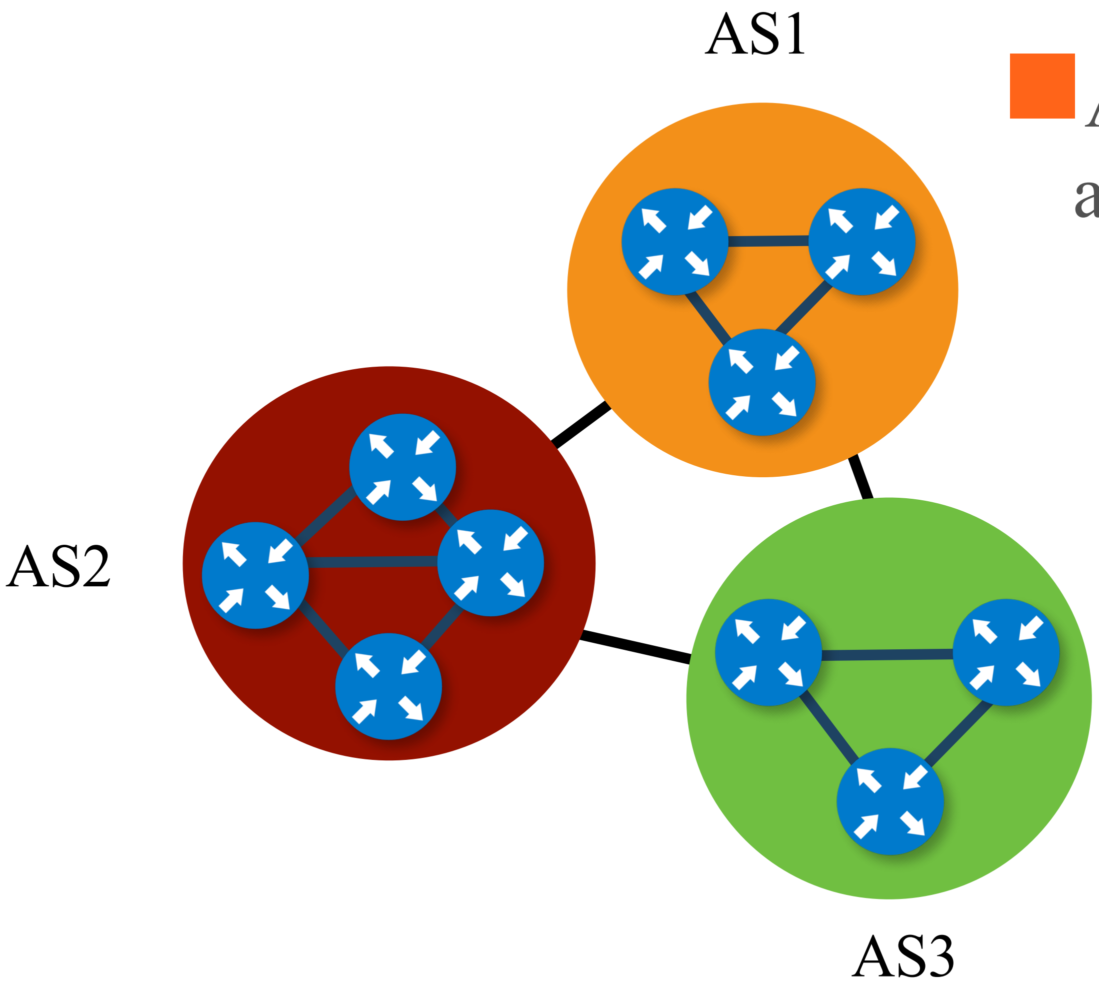
bgpstream @bgpstream · 1h
BGP,HJ,hijacked prefix AS196659 91.102.233.0/24, National Bank of Kuwait,-,By AS3356 Level 3 Communications, Inc., bgpstream.com/event/124973

bgpstream @bgpstream · 2h
BGP,OT,48953,Broadmax Iletisim Kollektif Sti,-,Outage affected 25 prefixes, bgpstream.com/event/124972

The Internet

- Who runs the Internet?
  - No one (in particular), not ICANN, nor the EU
- Any help to keep it working?
  - No central coordination
  - Many individuals and organisations

Internet relations



- A network controlled by a single entity
  - Same interior and exterior routing policy
  - Can also be a group of networks

The Scale of the “Internet”

- 20464 Autonomous Systems
- 167138 IP Address Prefixes announced

- Every single AS must maintain the routing table such that it knows how to route the traffic toward any one of the 167138 prefixes to the right destination.

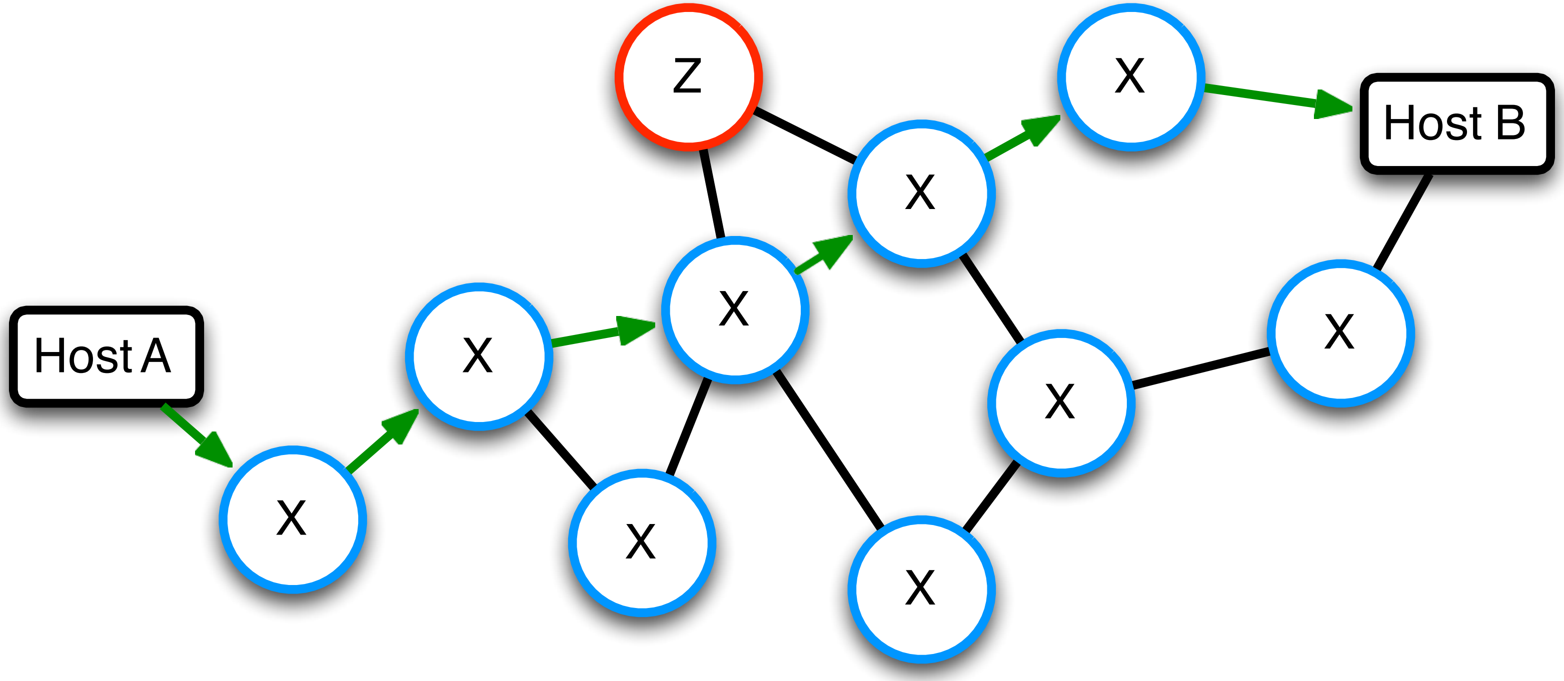
Who distributes AS numbers?

- AS numbers are distributed by Regional Internet Registries

- In our region:

CNNIC
中国互联网络信息中心
CHINA INTERNET NETWORK INFORMATION CENTER

Routing ...

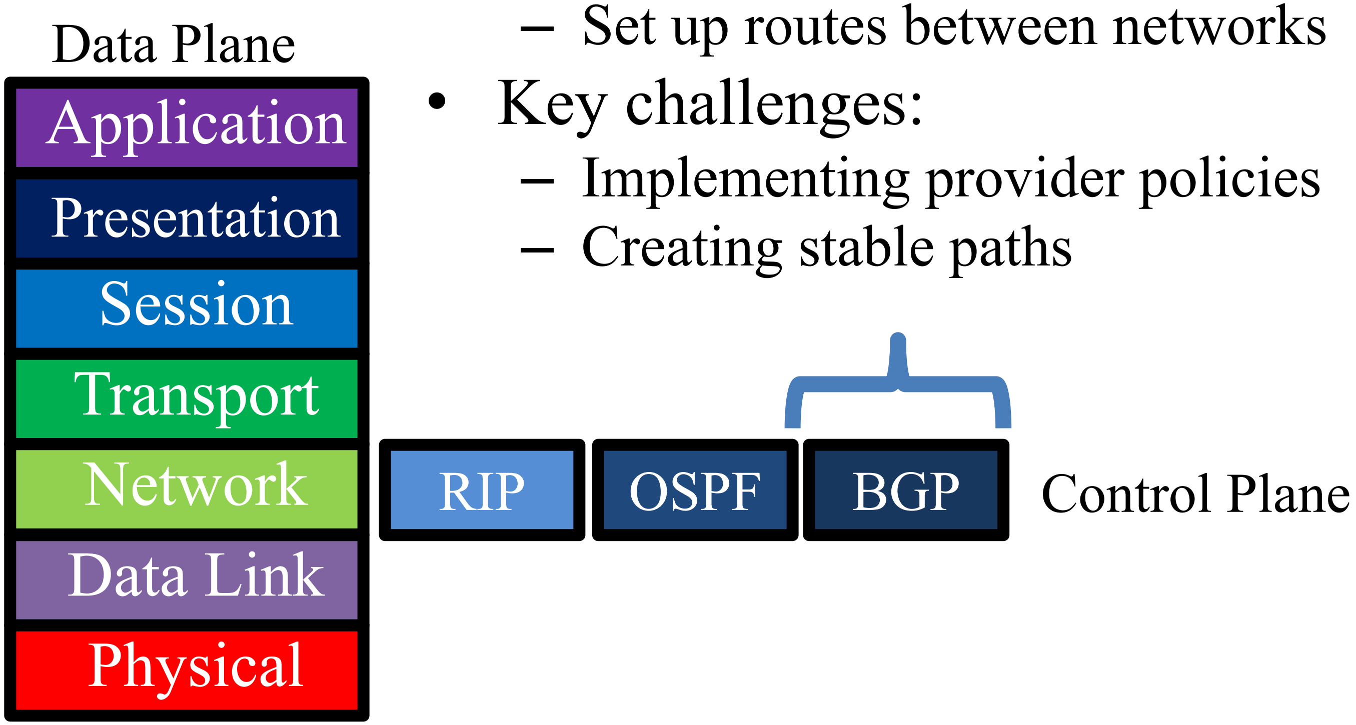


RoutingProtocols

- **Routers speak to each other.**
- **They exchange topology information and cost information.**
- **Each router calculates the shortest path to each destination.**
- **Routers forward packets along locally shortest path.**
- **Attacker can lie to other routers.**

Network Layer, Control Plane

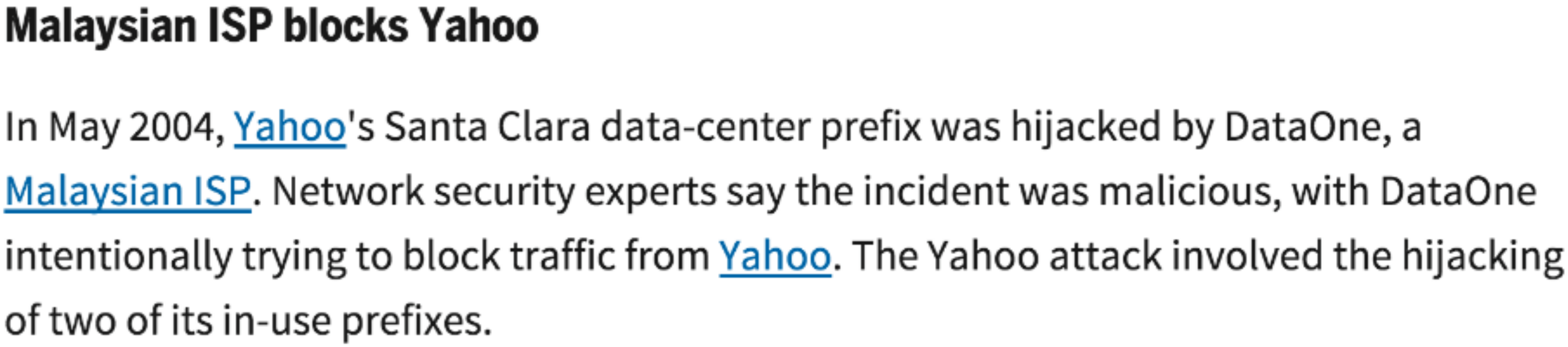
- Function:
  - Set up routes between networks
- Key challenges:
  - Implementing provider policies
  - Creating stable paths



IBGP - InternalGateway Protocol

VS

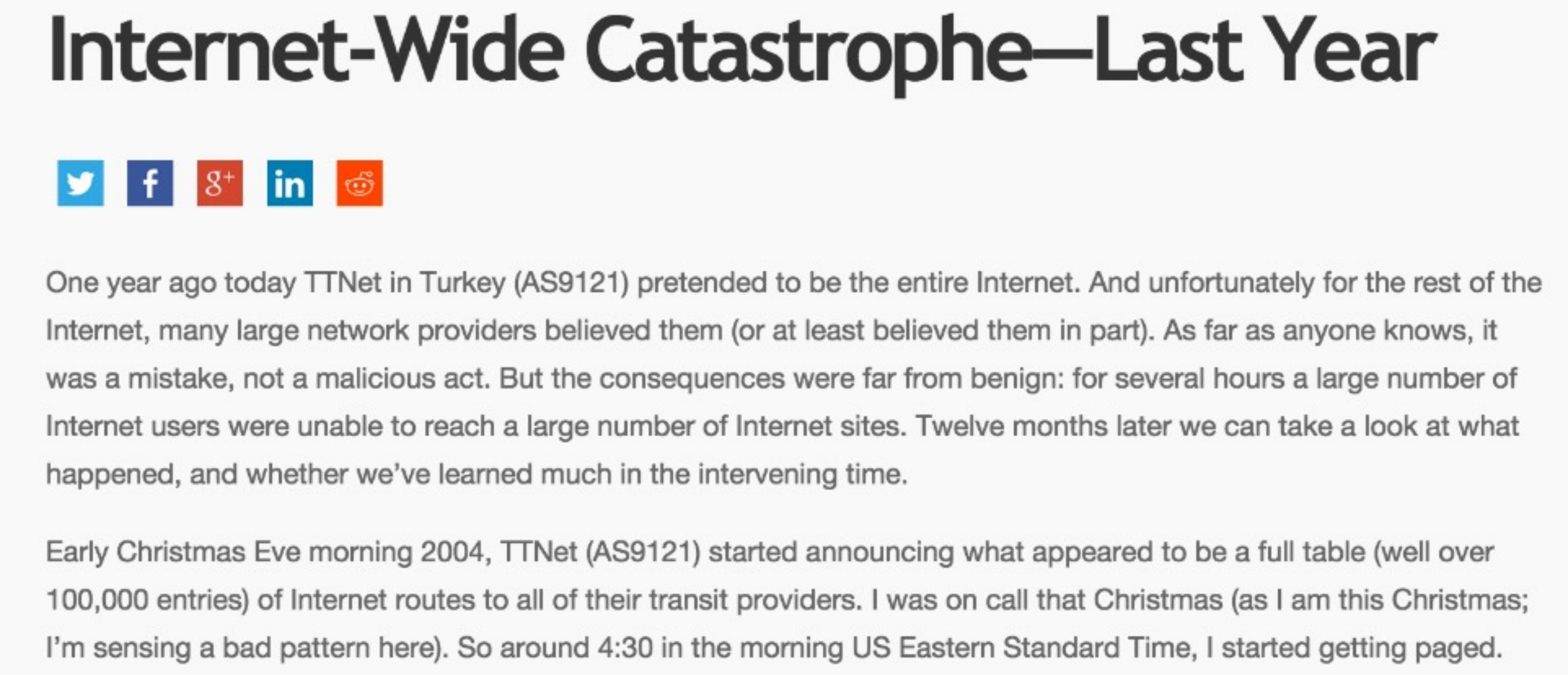
EBGP - ExternalGateway Protocol



Malaysian ISP blocks Yahoo

In May 2004, Yahoo's Santa Clara data-center prefix was hijacked by DataOne, a Malaysian ISP. Network security experts say the incident was malicious, with DataOne intentionally trying to block traffic from Yahoo. The Yahoo attack involved the hijacking of two of its in-use prefixes.

http://www.networkworld.com/article/2272520/lan-wan/six-worst-internet-routing-attacks.html



Internet-Wide Catastrophe—Last Year

One year ago today TTNet in Turkey (AS9121) pretended to be the entire Internet. And unfortunately for the rest of the Internet, many large network providers believed them (or at least believed them in part). As far as anyone knows, it was a mistake, not a malicious act. But the consequences were far from benign: for several hours a large number of Internet users were unable to reach a large number of Internet sites. Twelve months later we can take a look at what happened, and whether we've learned much in the intervening time.

Early Christmas Eve morning 2004, TTNet (AS9121) started announcing what appeared to be a full table (well over 100,000 entries) of Internet routes to all of their transit providers. I was on call that Christmas (as I am this Christmas; I'm sensing a bad pattern here). So around 4:30 in the morning US Eastern Standard Time, I started getting paged.

BGP

Border Gateway Protocol

Routing Protocol for exchanging information between networks

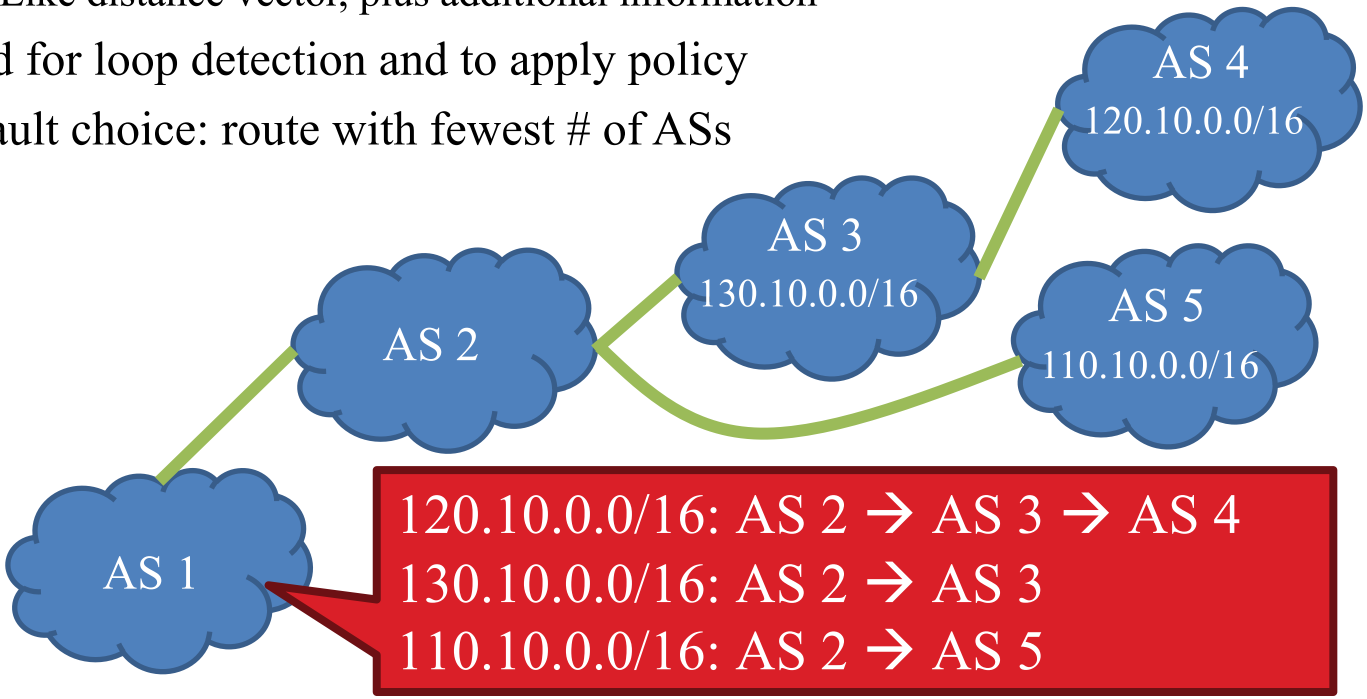
RFC4271

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
  - Policies driven by economics
    - How much $$$ does it cost to route along a given path?
    - Not by performance (e.g. shortest paths)
  - Entire world sees advertisements
    - Errors can screw up traffic globally
  - No authentication of 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



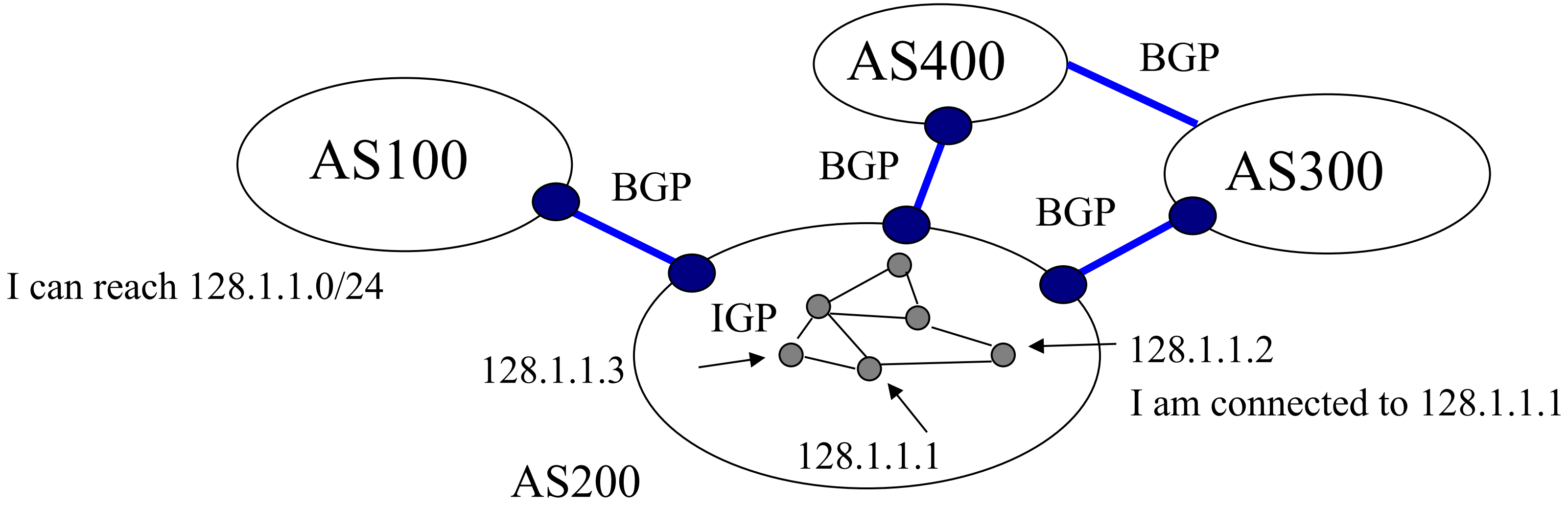
120.10.0.0/16: AS 2 → AS 3 → AS 4

130.10.0.0/16: AS 2 → AS 3

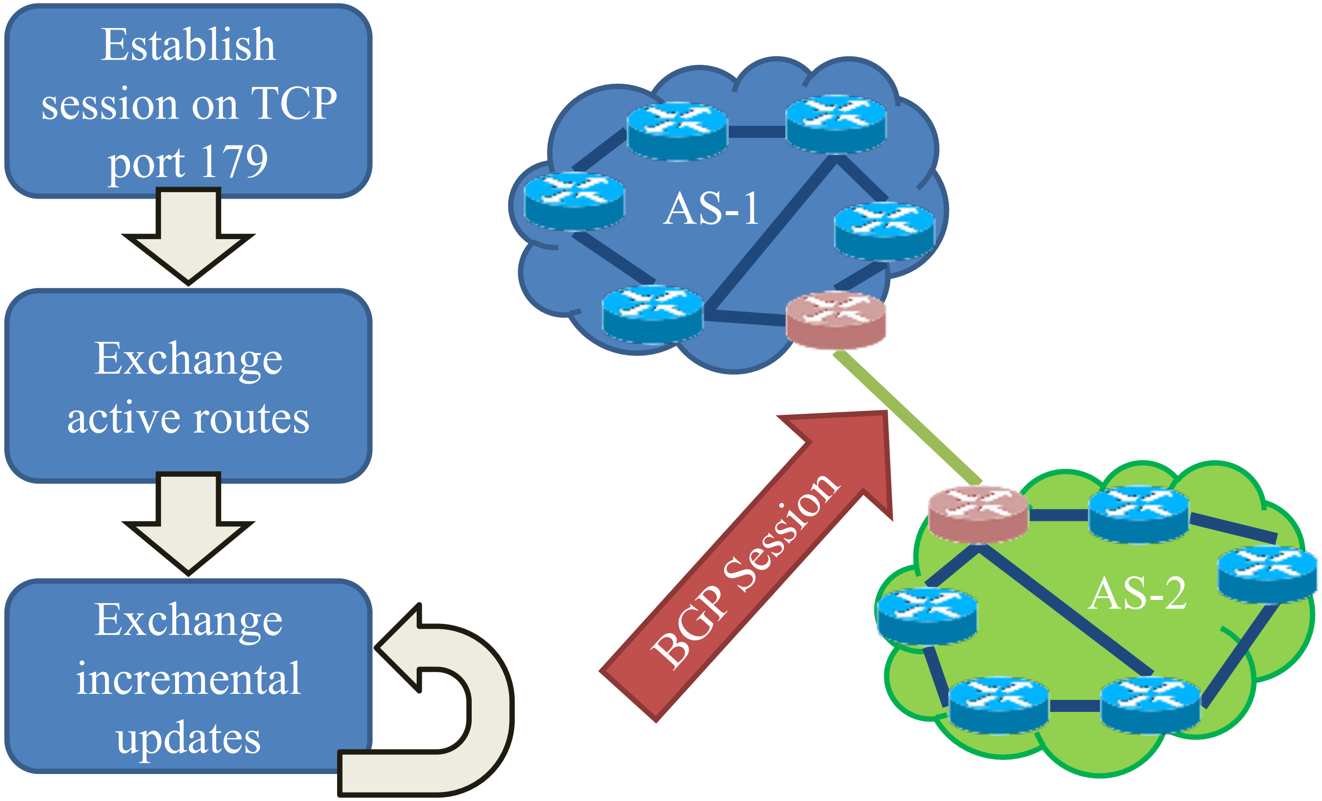
110.10.0.0/16: AS 2 → AS 5

Intra-/Inter-domain Routing

- AS 200 is assigned to use 128.1.1.0/24
- Uses IBGP to reach local destinations, 128.1.1.1, 128.1.1.2, 128.1.1.3, etc
- Outside world only needs to know how to reach the aggregate 128.1.1.0/24



BGP Operations (Simplified)



Establish session on TCP port 179

Exchange active routes

Exchange incremental updates

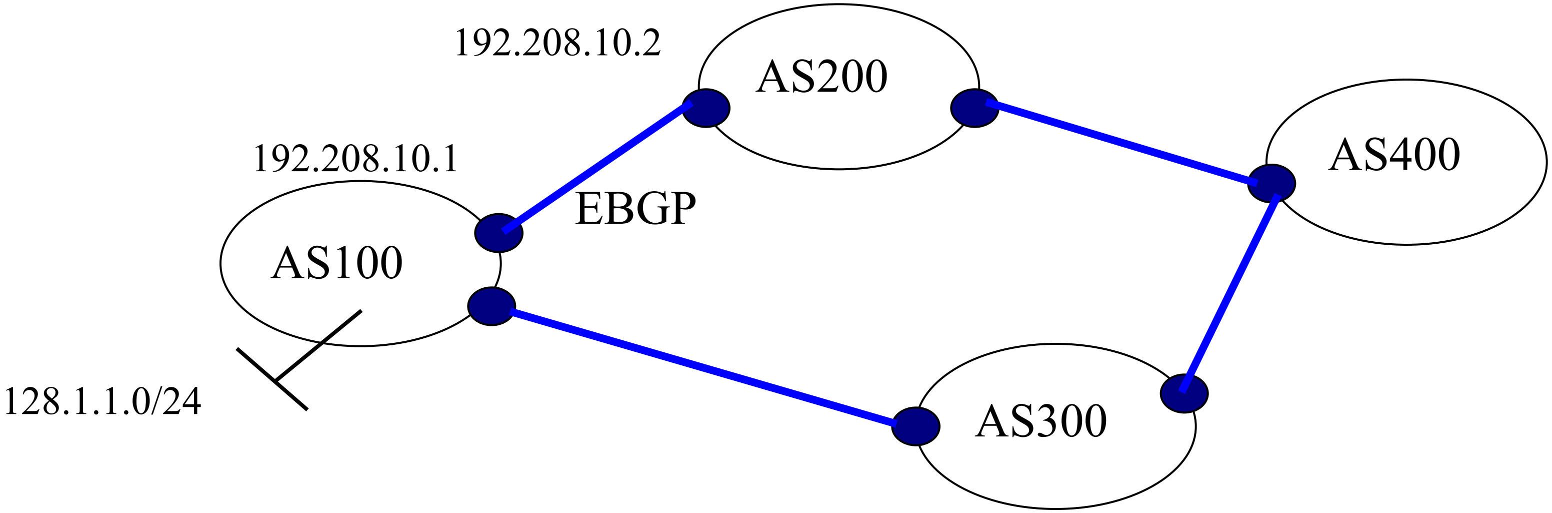
Four Types of BGP Messages

- 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

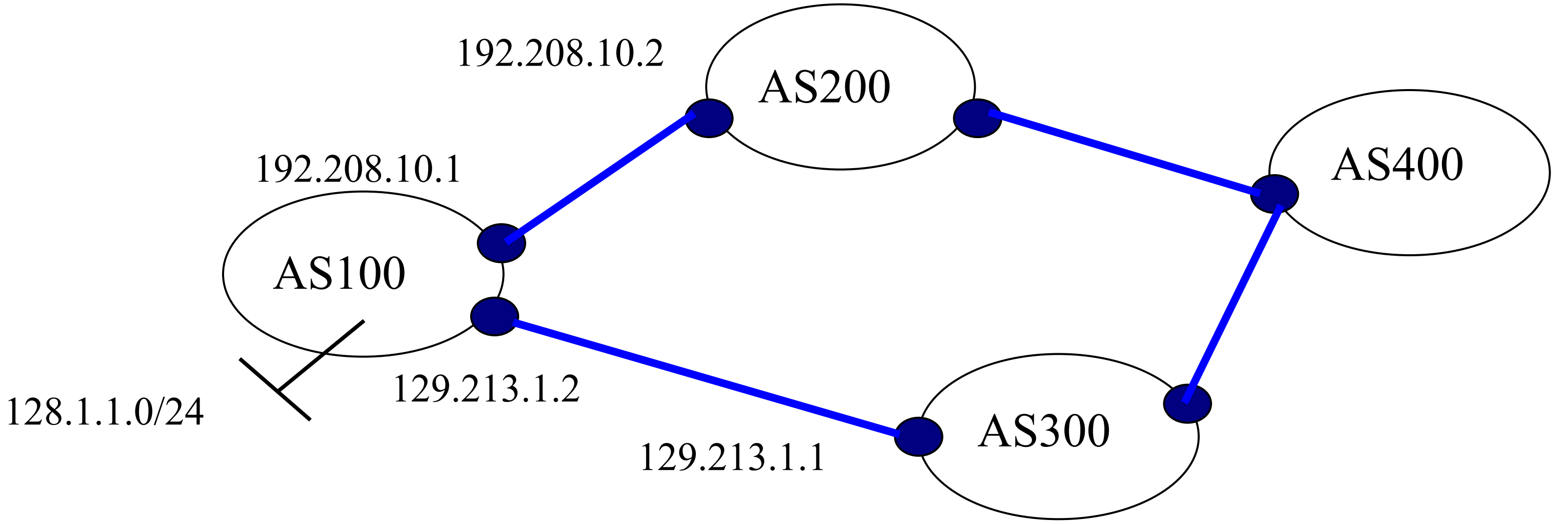
Principles of operation

- Two BGP routers establish a peering session over TCP
  - Periodic KEEPALIVEs to keep up session
  - NOTIFICATION to do a graceful close



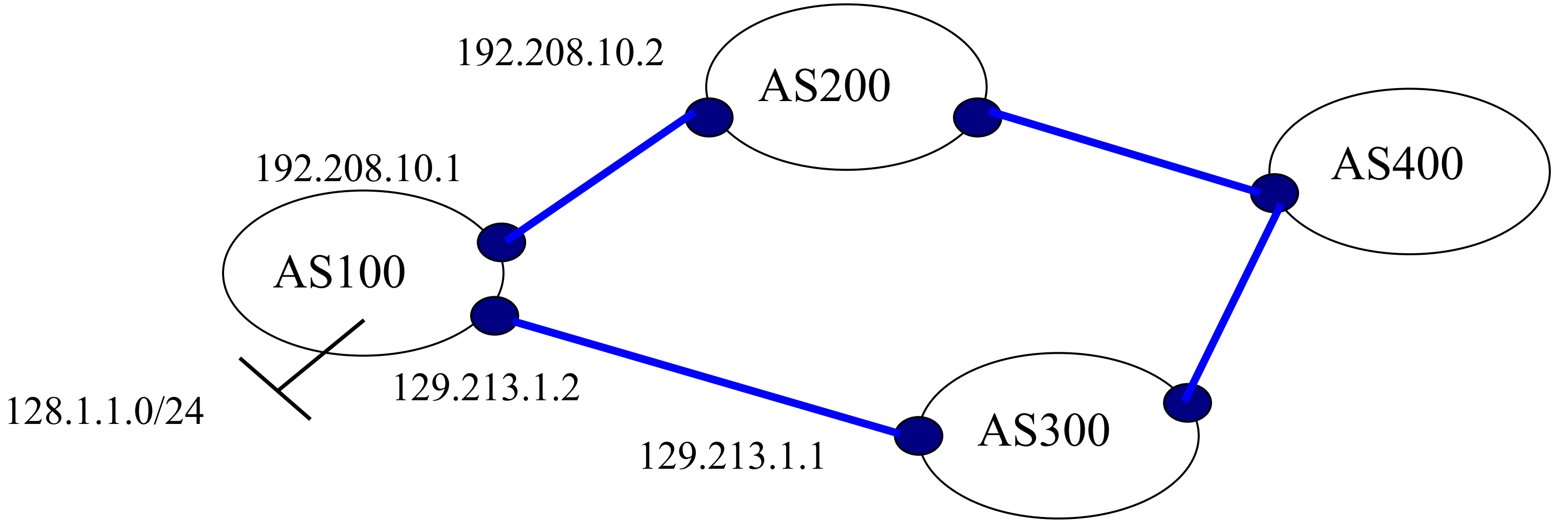
Principles of operation

- Exchange routes
  - AS100 announces 128.1.1.0/24 prefix to AS200 and AS300, etc
- Incremental updates



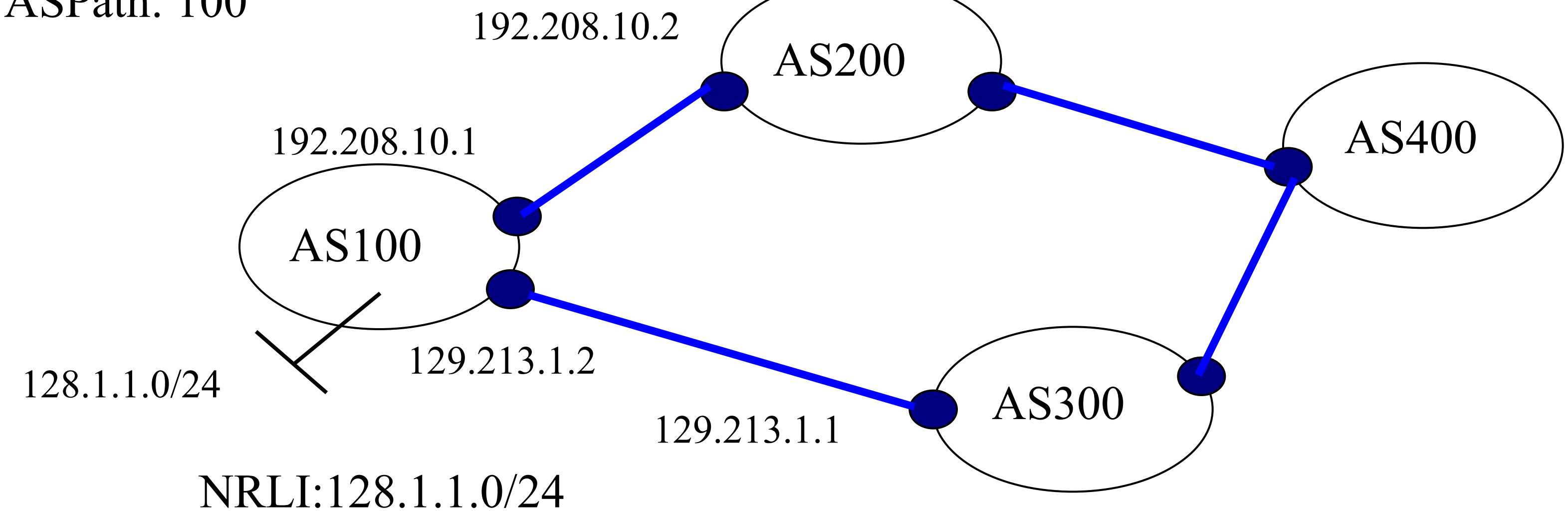
BGP UPDATE message

- Announced prefixes (NLRI:Network Layer Reachability Information)
- Path attributes associated with NLRI
- Withdrawn prefixes



UPDATE message example

NLRI: 128.1.1.0/24
Nexthop: 192.208.10.1
ASPath: 100

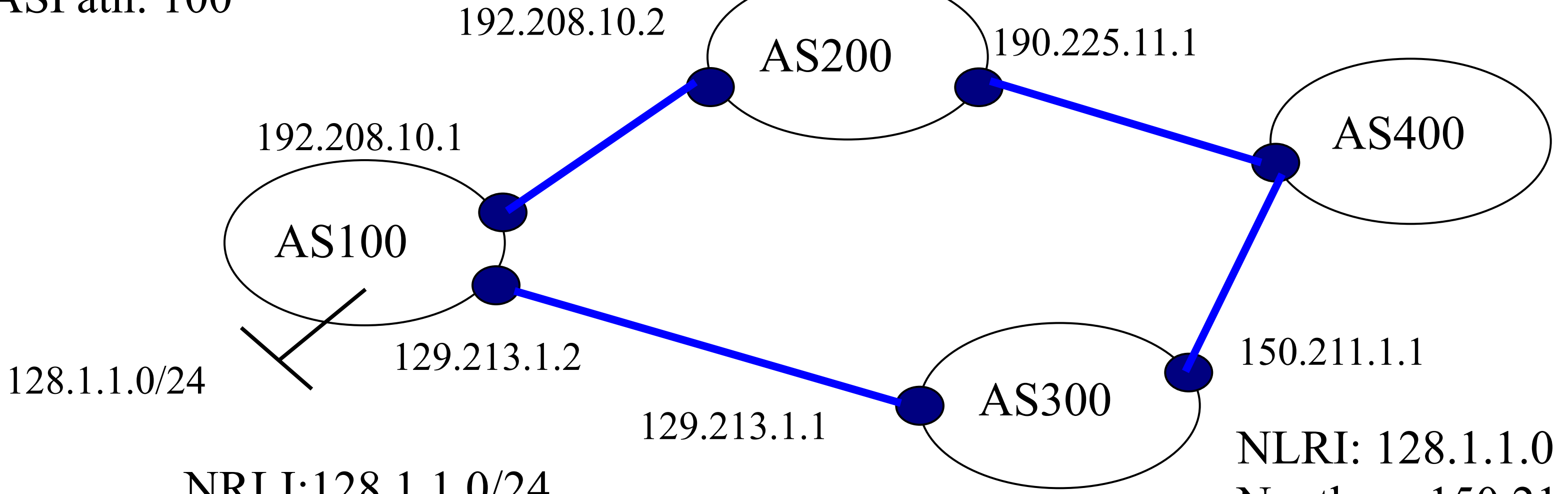


NRLI:128.1.1.0/24
Nexthop: 129.213.1.2
ASPath: 100

Route propagation

NLRI: 128.1.1.0/24
Nexthop: 192.208.10.1
ASPath: 100

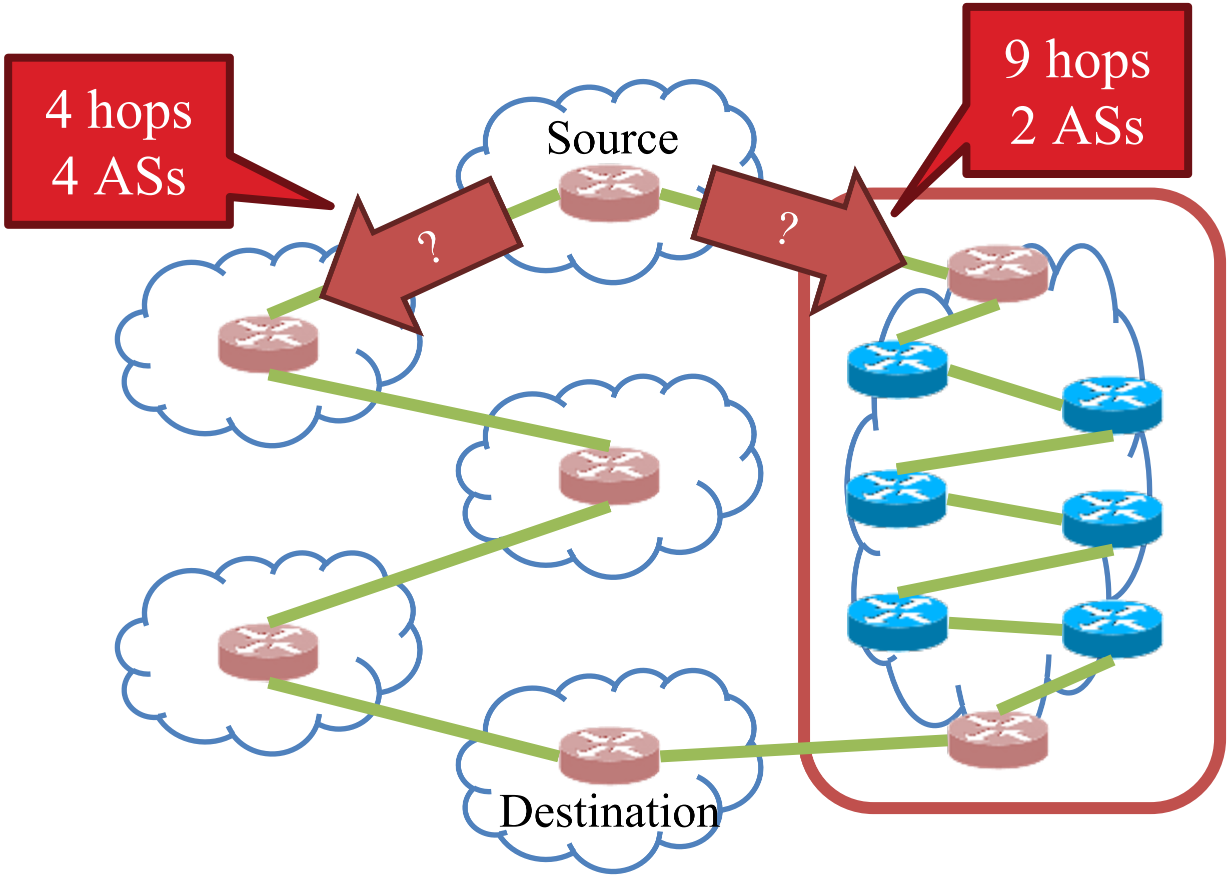
NLRI: 128.1.1.0/24
Nexthop: 190.225.11.1
ASPath: 200 100



NRLI:128.1.1.0/24
Nexthop: 129.213.1.2
ASPath: 100

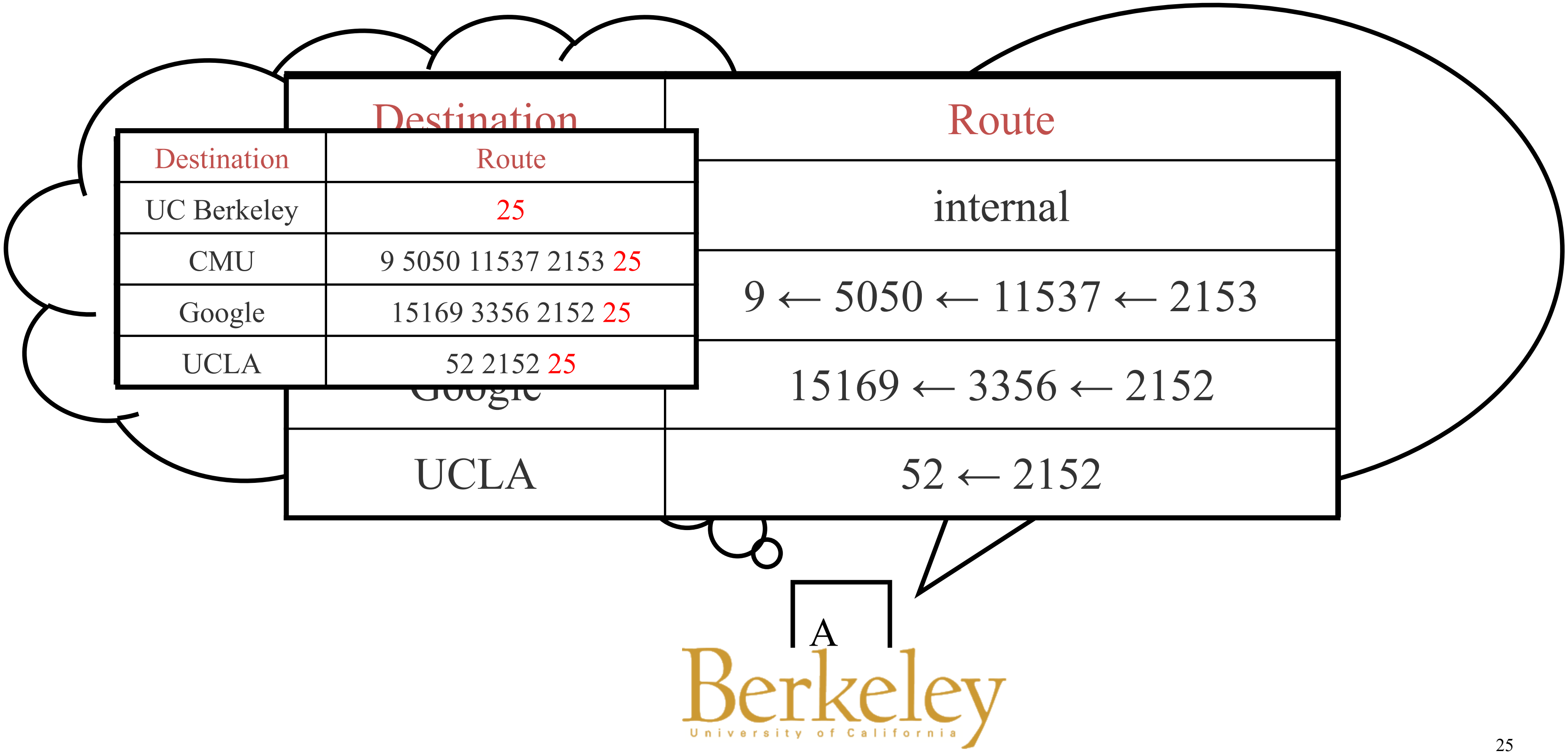
NLRI: 128.1.1.0/24
Nexthop: 150.212.1.1
ASPath: 300 100

Shortest AS Path != Shortest Path



Sample BGP routing table

- Routing table at UC Berkeley:



Destination	Route
UC Berkeley	25
CMU	9 5050 11537 2153 25
Google	15169 3356 2152 25
UCLA	52 2152 25

Destination	Route
	internal
	9 ← 5050 ← 11537 ← 2153
Google	15169 ← 3356 ← 2152
UCLA	52 ← 2152

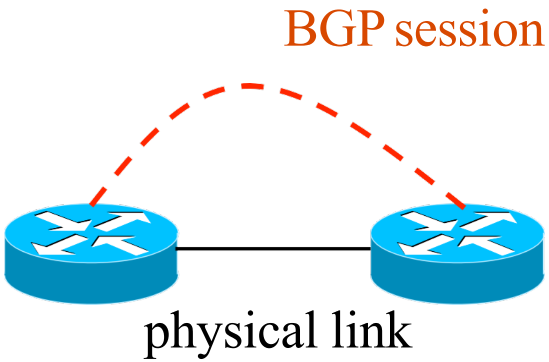
BGP Session Security

TCP Connection Underlying BGP Session

- BGP session runs over TCP
  - TCP connection between neighboring routers
  - BGP messages sent over TCP connection
  - Makes BGP vulnerable to attacks on TCP
- Main kinds of attacks
  - Against confidentiality: eavesdropping
  - Against integrity: tampering
  - Against performance: denial-of-service
- Main defenses
  - Message authentication or encryption
  - Limiting access to physical path between routers
  - Defensive filtering to block unexpected packets

Attacks Against Confidentiality

- Eavesdropping
  - Monitoring the messages on the BGP session
  - … by tapping the link(s) between the neighbors
- Reveals sensitive information
  - Inference of business relationships
  - Analysis of network stability
- Reasons why it may be hard
  - Challenging to tap the link
    - Often, eBGP session traverses just one link and may be hard to get access to tap it
    - Encryption may obscure message contents
      - BGP neighbors may run BGP over IPSec



Attacking Message Integrity

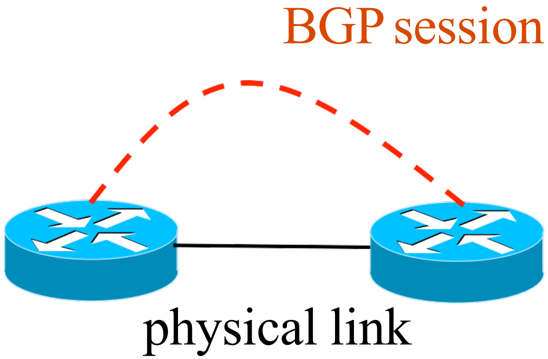
- Tampering
  - Man-in-the-middle tampers with the messages
  - Insert, delete, modify, or replay messages
- Leads to incorrect BGP behavior
  - Delete: neighbor doesn’t learn the new route
  - Insert/modify: neighbor learns bogus route
- Reasons why it may be hard
  - Getting in-between the two routers is hard
  - Use of authentication (signatures) or encryption
  - Spoofing TCP packets the right way is hard

Attacking Message Integrity

- Tampering
  - Man-in-the-middle tampers with the messages
  - Insert, delete, modify, or replay messages
- Leads to incorrect BGP behavior
  - Delete: neighbor doesn’t learn the new route
  - Insert/modify: neighbor learns bogus route
- Reasons why it may be hard
  - Getting in-between the two routers is hard
  - Use of authentication (signatures) or encryption
  - Spoofing TCP packets the right way is hard
    - Getting past source-address packet filters
    - Generating the right TCP sequence number

Denial-of-Service Attacks, Part1

- Overload the link between the routers
  - To cause packet loss and delay
  - … disrupting the performance of the BGP session
- Relatively easy to do
  - Can send traffic between end hosts
  - As long as the packets traverse the link
  - (which you can figure out from traceroute)
- Easy to defend
  - Give higher priority to BGP packets
  - E.g., by putting packets in separate queue



Denial-of-Service Attacks, Part2

- Third party sends bogus TCP packets
  - FIN/RST to close the session
  - SYN flooding to overload the router
- Leads to disruptions in BGP
  - Session reset, causing transient routing changes
  - Route-flapping
- Reasons why it may be hard
  - Spoofing TCP packets the right way is hard
    - Difficult to send FIN/RST with the right TCP header
  - Packet filters may block the SYN flooding
    - Filter packets to BGP port from unexpected source
    - … or destined to router from unexpected source

Exploiting the IP TTL Field as a Defense

- BGP speakers are usually one hop apart
  - To thwart an attacker, can check that the packets carrying the BGP message have not traveled far
- IP Time-to-Live (TTL) field
  - Decremented once per hop
  - Avoids packets staying in network forever
- Generalized TTL Security Mechanism (RFC 3682)
  - Send BGP packets with initial TTL of 255
  - Receiving BGP speaker checks that TTL is 254
  - … and flags and/or discards the packet others
- Hard for third-party to inject packets remotely

Validity of the routing information: Origin authentication

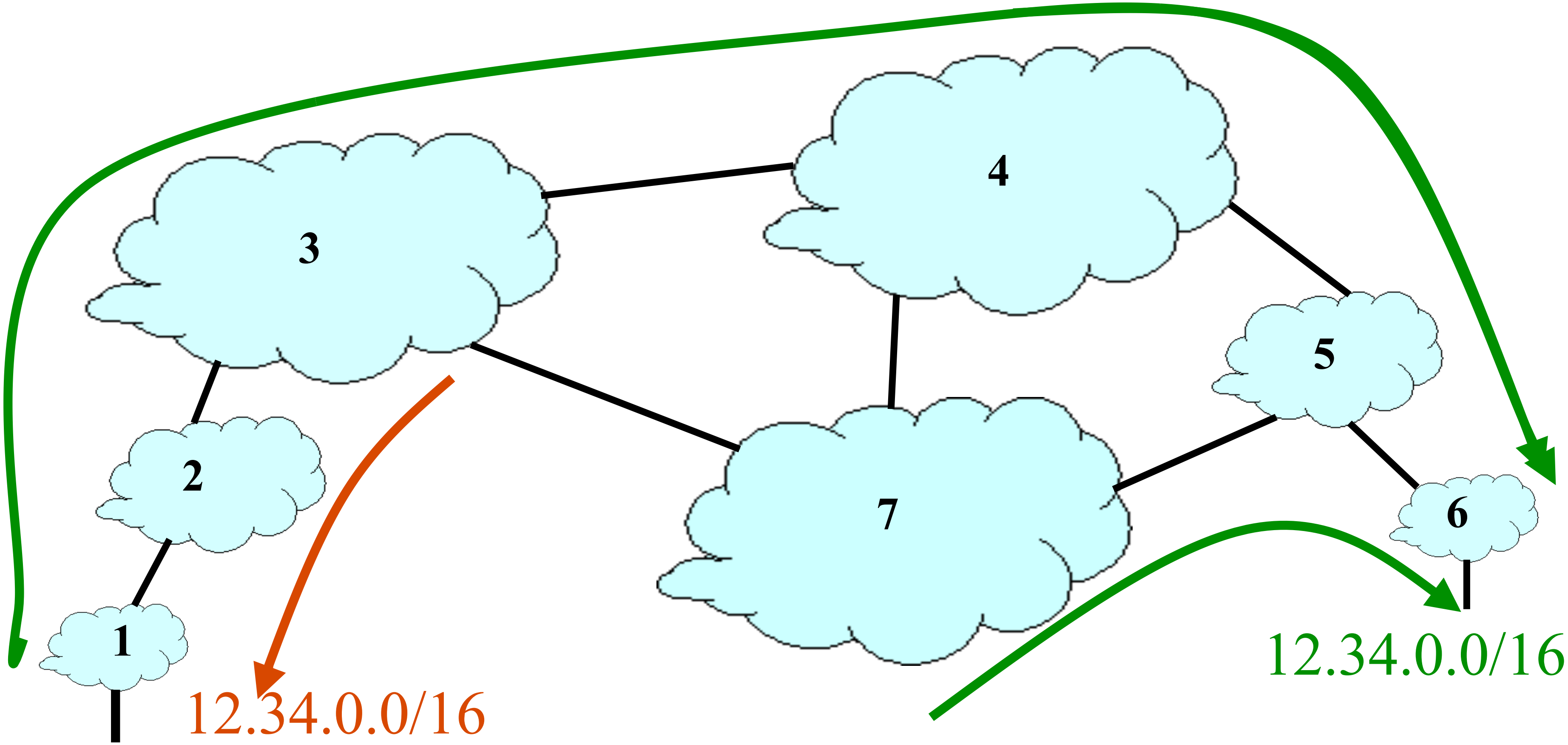
IP Address Ownership and Hijacking

- IP address block assignment
  - Regional Internet Registries (ARIN, RIPE, APNIC)
  - Internet Service Providers
- Proper origination of a prefix into BGP
  - By the AS who owns the prefix
  - … or, by its upstream provider(s) in its behalf
- However, what's to stop someone else?
  - Prefix hijacking: another AS originates the prefix
  - BGP does not verify that the AS is authorized
  - Registries of prefix ownership are inaccurate

How to Hijack a Prefix

- The hijacking AS has
  - Router with eBGP session(s)
  - Configured to originate the prefix
- Getting access to the router
  - Network operator makes configuration mistake
  - Disgruntled operator launches an attack
  - Outsider breaks in to the router and reconfigures
- Getting other ASes to believe bogus route
  - Neighbor ASes not filtering the routes
  - … e.g., by allowing only expected prefixes
  - But, specifying filters on peering links is hard

Prefix ijacking

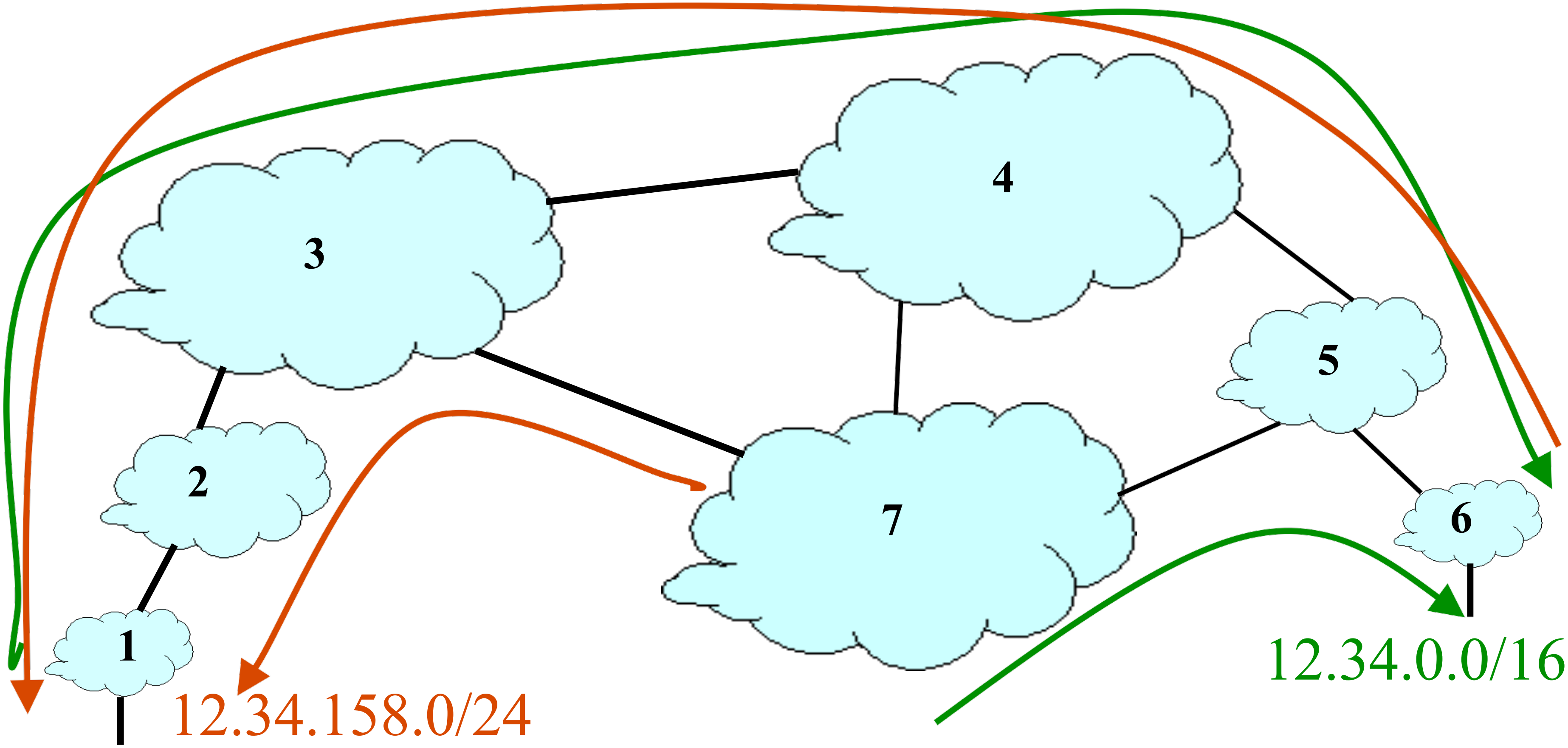


- Consequences for the affected ASes
  - Blackhole: data traffic is discarded
  - Snooping: data traffic is inspected, and then redirected
  - Impersonation: data traffic is sent to bogus destinations

Hijacking is Hard to Debug

- Real origin AS doesn’t see the problem
    - Picks its own route
    - Might not even learn the bogus route
- May not cause loss of connectivity
    - E.g., if the bogus AS snoops and redirects
    - … may only cause performance degradation
- Or, loss of connectivity is isolated
    - E.g., only for sources in parts of the Internet
- Diagnosing prefix hijacking
    - Analyzing updates from many vantage points
    - Launching traceroute from many vantage points

Sub-Prefix Hijacking



- Originating a more-specific prefix
  - Every AS picks the bogus route for that prefix
  - Traffic follows the longest matching prefix

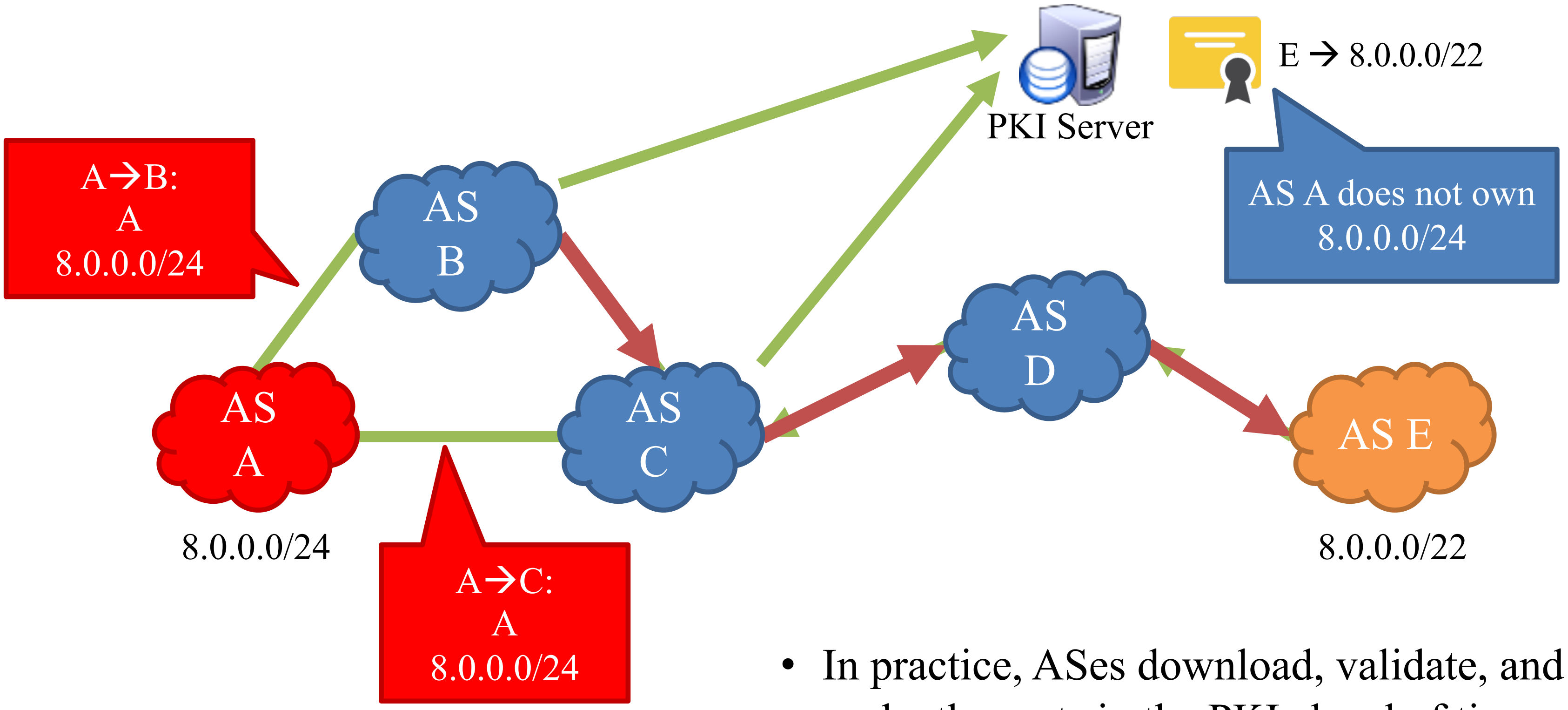
BGP Security Today

- Applying best common practices (BCPs)
  - Securing the session (authentication, encryption)
  - Filtering routes by prefix and AS path
  - Packet filters to block unexpected control traffic
- This is not good enough
  - Depends on vigilant application of BCPs
    - … and not making configuration mistakes!
  - Doesn’t address fundamental problems
    - Can’t tell who owns the IP address block
    - Can’t tell if the AS path is bogus or invalid
    - Can’t be sure the data packets follow the chosen route

Secure BGP

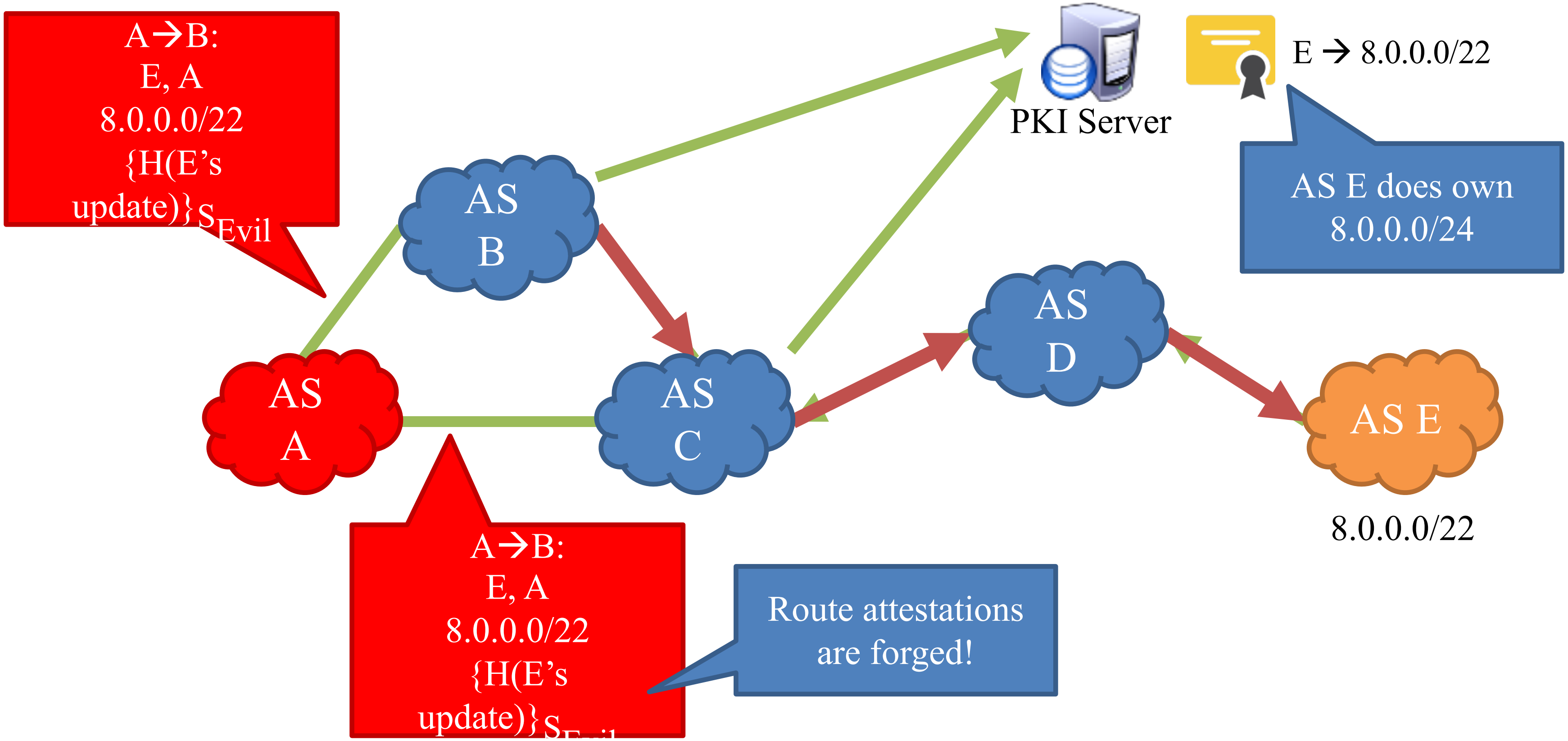
1. Use PKI to authenticate BGP
   - Dual hierarchies of certificates bind prefix ownership to ASes and routers to ASes
   - Certificate hierarchy distributed and validated out-of-band
   - Routers only accept updates that are covered by valid certificates
2. Route attestations using “onion” signatures
   - Each BGP update is signed by the announcer
   - These signatures accumulate as the update propagates
   - Any AS receiving the announcement can verify the signature added by each AS back to the source

S-BGP vs. Subprefix Hijack Example



- In practice, ASes download, validate, and cache the certs in the PKI ahead of time

S-BGP vs. Short Path Hijack



A→B:
E, A
8.0.0.0/22
{H(E’s
update)}S_{Evil}

AS E does own
8.0.0.0/24

A→B:
E, A
8.0.0.0/22
{H(E’s
update)}S_{Evil}

Route attestations
are forged!

(The Lack of) S-BGP Deployment

- S-BGP was proposed at least a decade ago, and implementations were available soon afterwards
- But, it was never deployed. Why?
  - Trust rooted in ICANN, a US organization
    - Other countries are wary of centralizing power in the US
  - Verification of signed attestations is costly in terms of CPU
    - Routers are expensive and resource constrained
    - Entire chain of attestations must be cryptographically validated for each received update
    - In contrast, PKI validation can be done out of band and applied using simple filters