Introduction to Computer Networks

COSC 4377

Lecture 16

Spring 2012

March 21, 2012

Announcements

- HW8 and HW9 are out
- HW deadlines

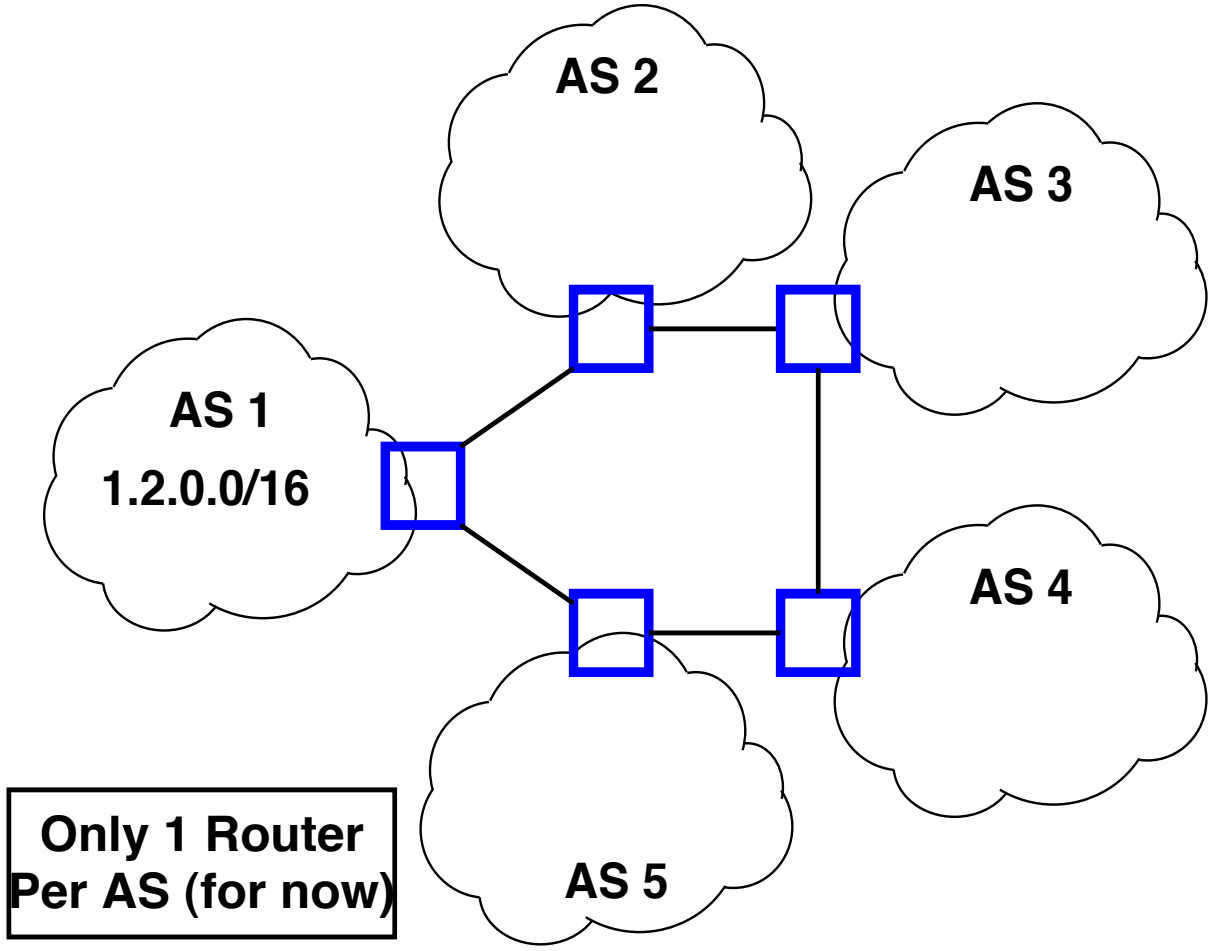
HW8

- Distance Vector Routing
- Count-to-infinity
- Split-horizon

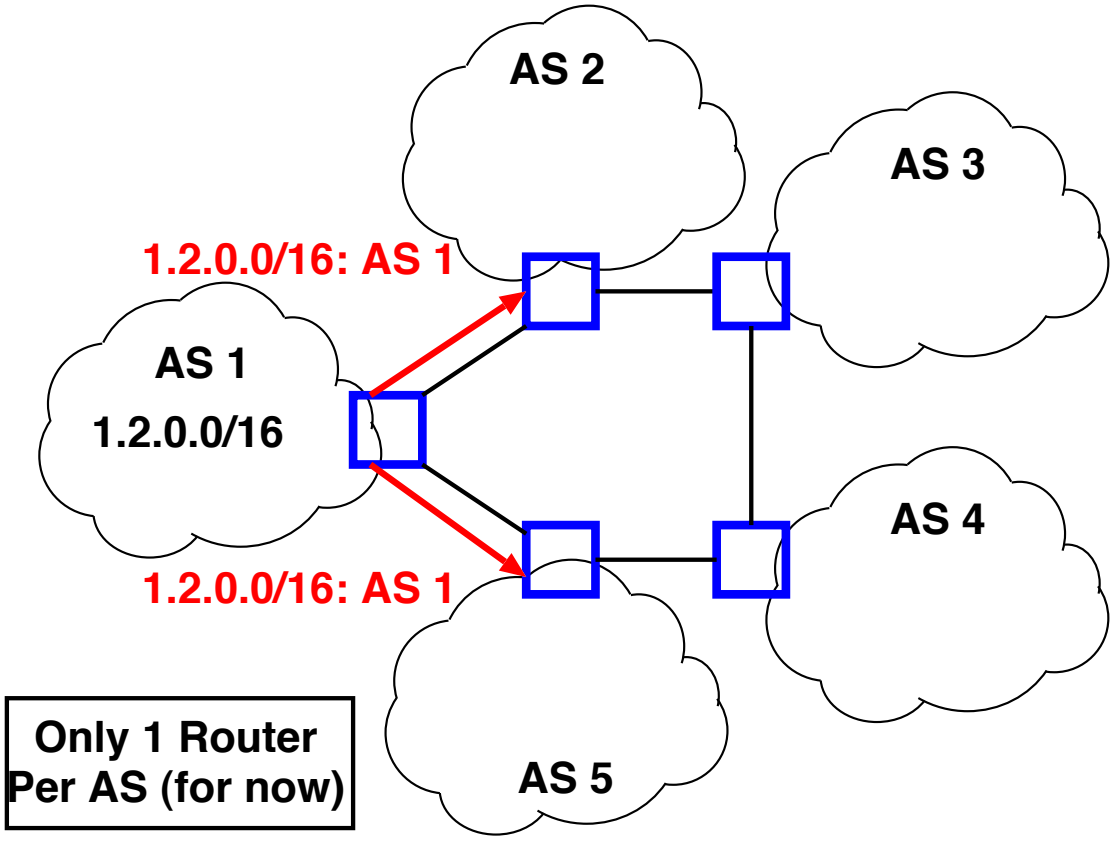
Today’s Topics

- BGP

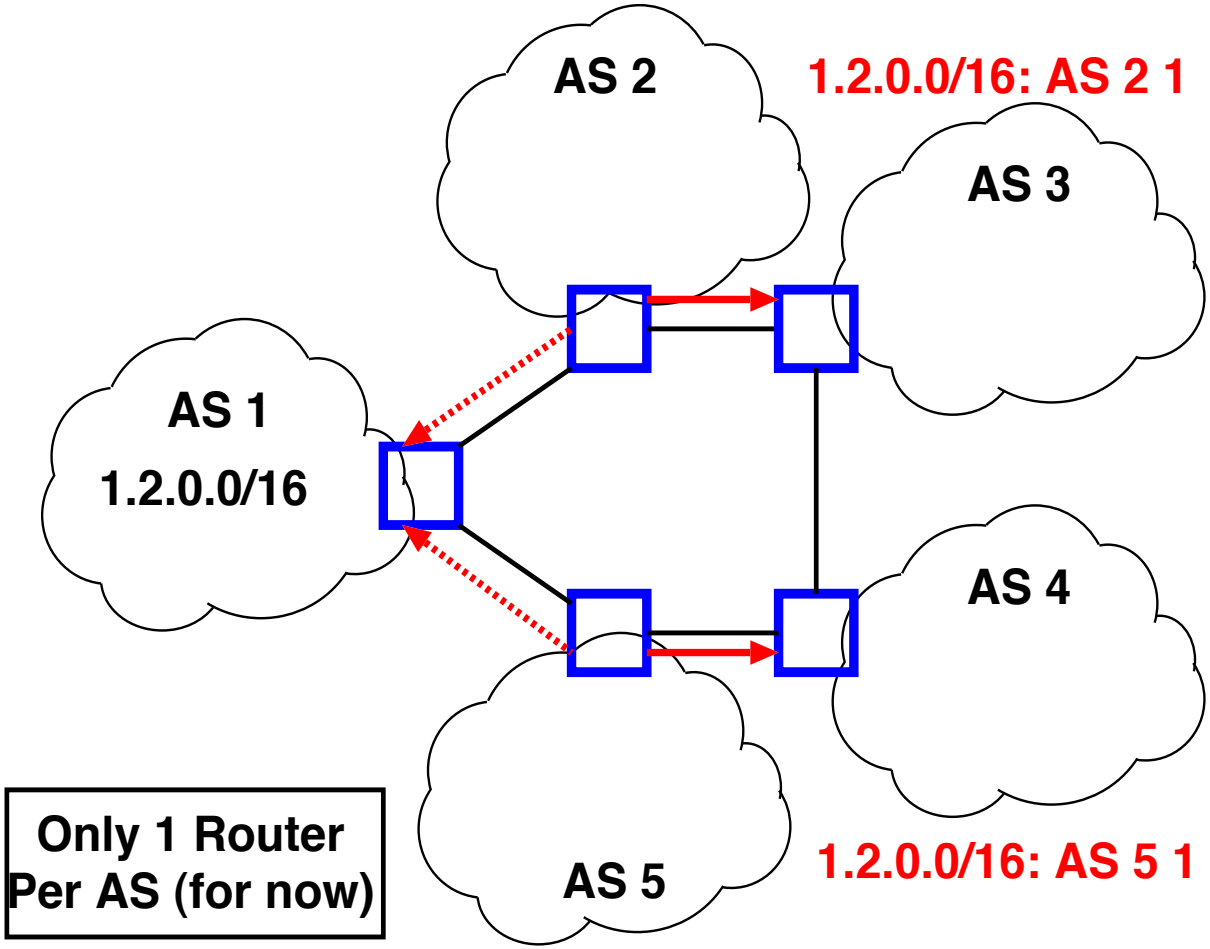
BGP Example



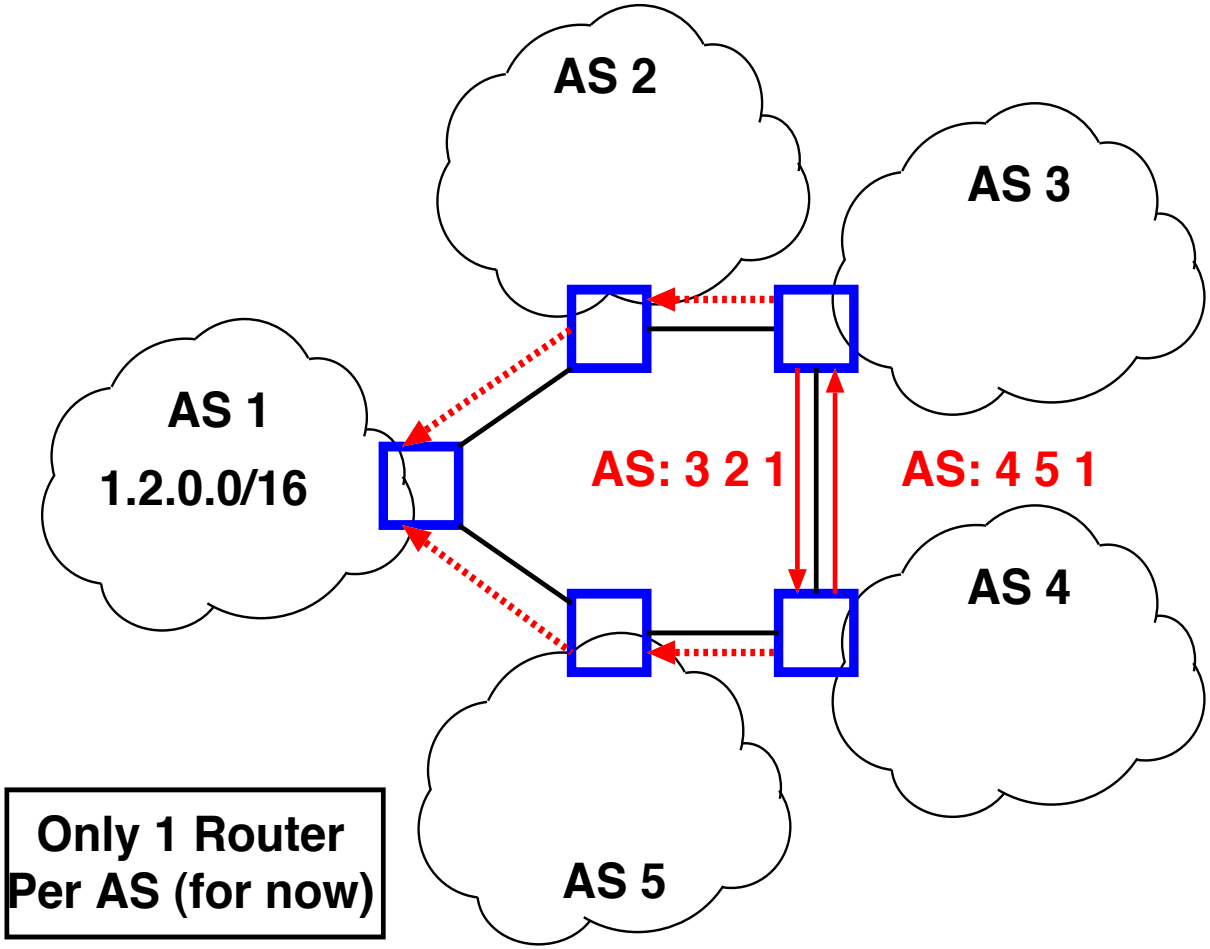
BGP Example


AS 2
AS 3
1.2.0.0/16: AS 1
AS 1
1.2.0.0/16
AS 4
1.2.0.0/16: AS 1
AS 5
Only 1 Router
Per AS (for now)

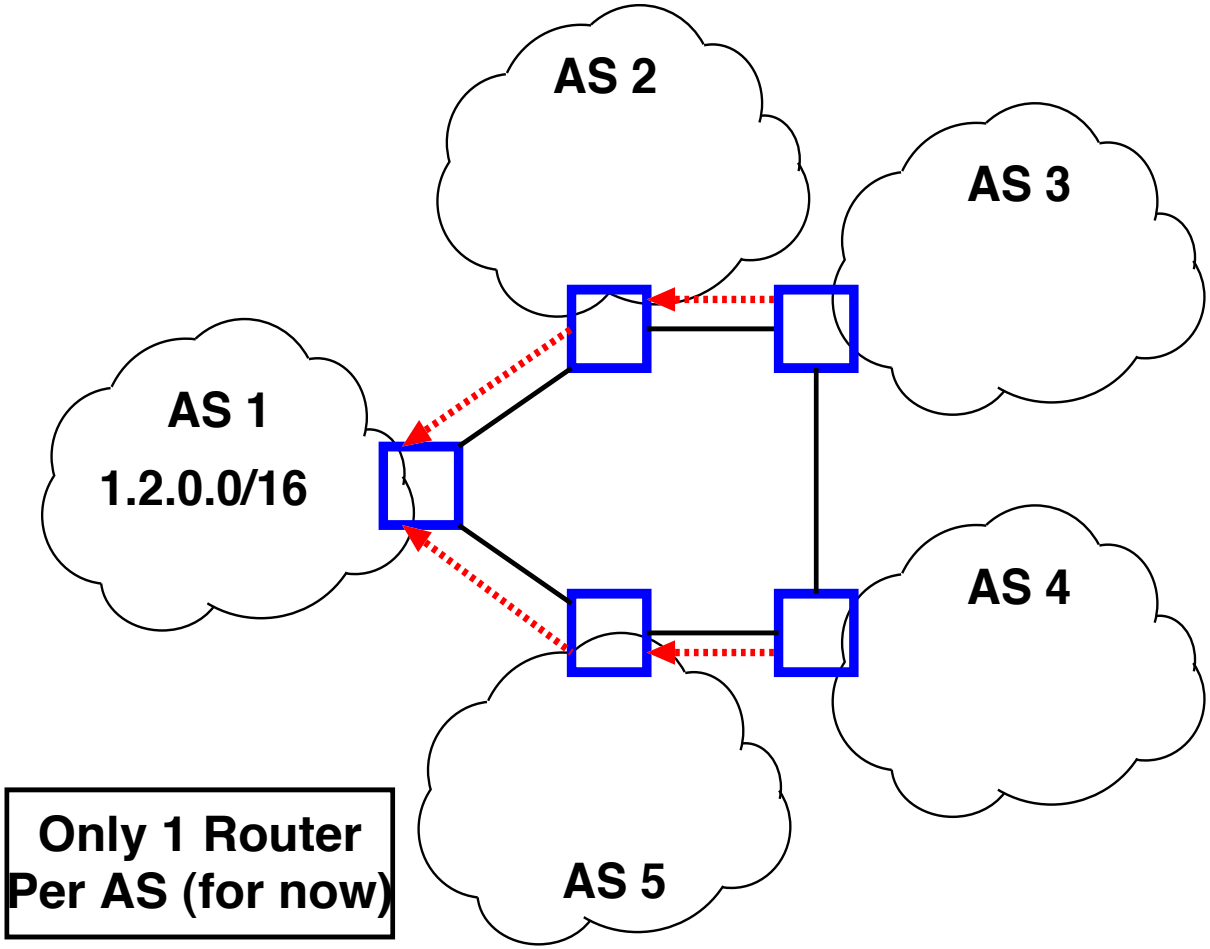
BGP Example


AS 2
1.2.0.0/16: AS 2 1
AS 3
AS 1
1.2.0.0/16
AS 4
AS 5
1.2.0.0/16: AS 5 1
Only 1 Router
Per AS (for now)

BGP Example


AS 2
AS 3
AS 1
1.2.0.0/16
AS: 3 2 1
AS: 4 5 1
AS 4
AS 5
Only 1 Router
Per AS (for now)

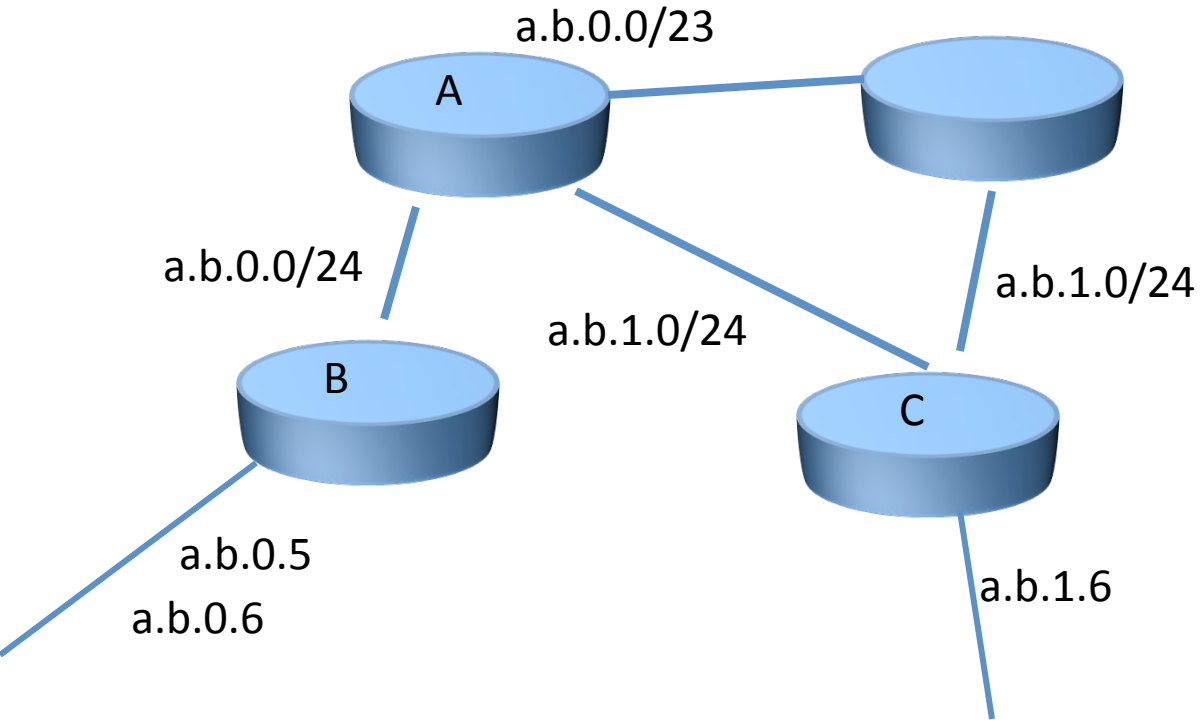
BGP Example



Forwarding with CIDR

- Longest Prefix Match

Prefix	Nexthop
a.b.0.0/23	A
a.b.1.0/24	C



Where to forward these packets?
- dst: a.b.0.5
- dst: a.b.1.6

BGP and Policy

- BGP provides capability for enforcing various policies
- Policies are not part of BGP: they are provided to BGP as configuration information
- BGP enforces policies by choosing paths from multiple alternatives and controlling advertisement to other AS's

BGP Path Selection

- Policies determined by path selection
- Information based on path attributes
- Attributes + external (policy) information

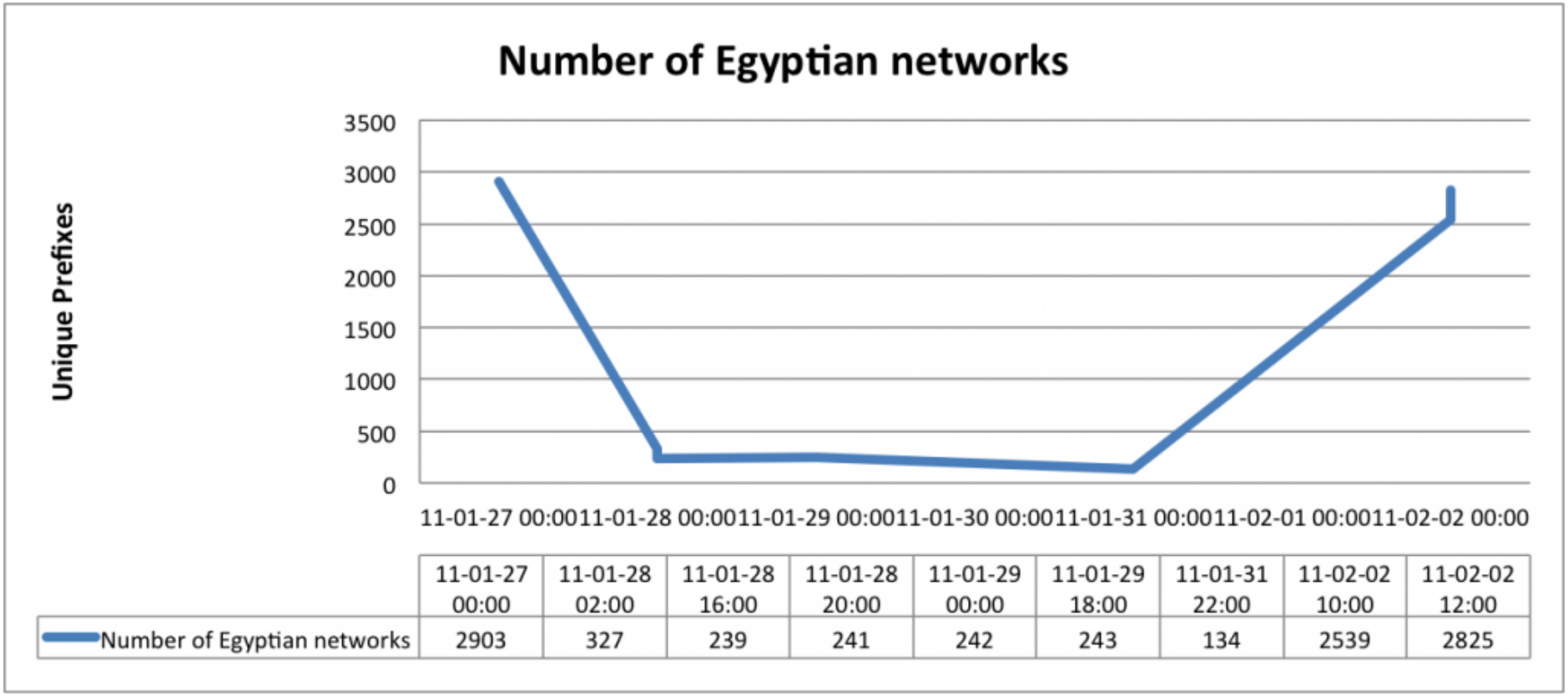
Customer/Provider AS relationships

- Customer pays for connectivity
  - E.g. University of Houston contracts with AboveNet and TW Telecom
  - Customer is stub, provider is a transit
- Many customers are multi-homed
  - E.g., AboveNet connects to Level3, Cogent,…
- Typical policies:
  - Provider tells all neighbors how to reach customer
  - Provider prefers routes from customers ($$)
  - Customer does not provide transit service

Peer Relationships

- ASs agree to exchange traffic for free
  - Penalties/Renegotiate if imbalance
- Tier 1 ISPs have no default route: all peer with each other
- You are Tier *i* + 1 if you have a default route to a Tier *i*
- Typical policies
  - AS only exports customer routes to peer
  - AS exports a peer's routes only to its customers
  - Goal: avoid being transit when no gain

Egypt Incident



	11-01-27 00:00	11-01-28 02:00	11-01-28 16:00	11-01-28 20:00	11-01-29 00:00	11-01-29 18:00	11-01-31 22:00	11-02-02 10:00	11-02-02 12:00
Number of Egyptian networks	2903	327	239	241	242	243	134	2539	2825

Source: BGPMon (http://bgpmon.net/blog/?p=480)

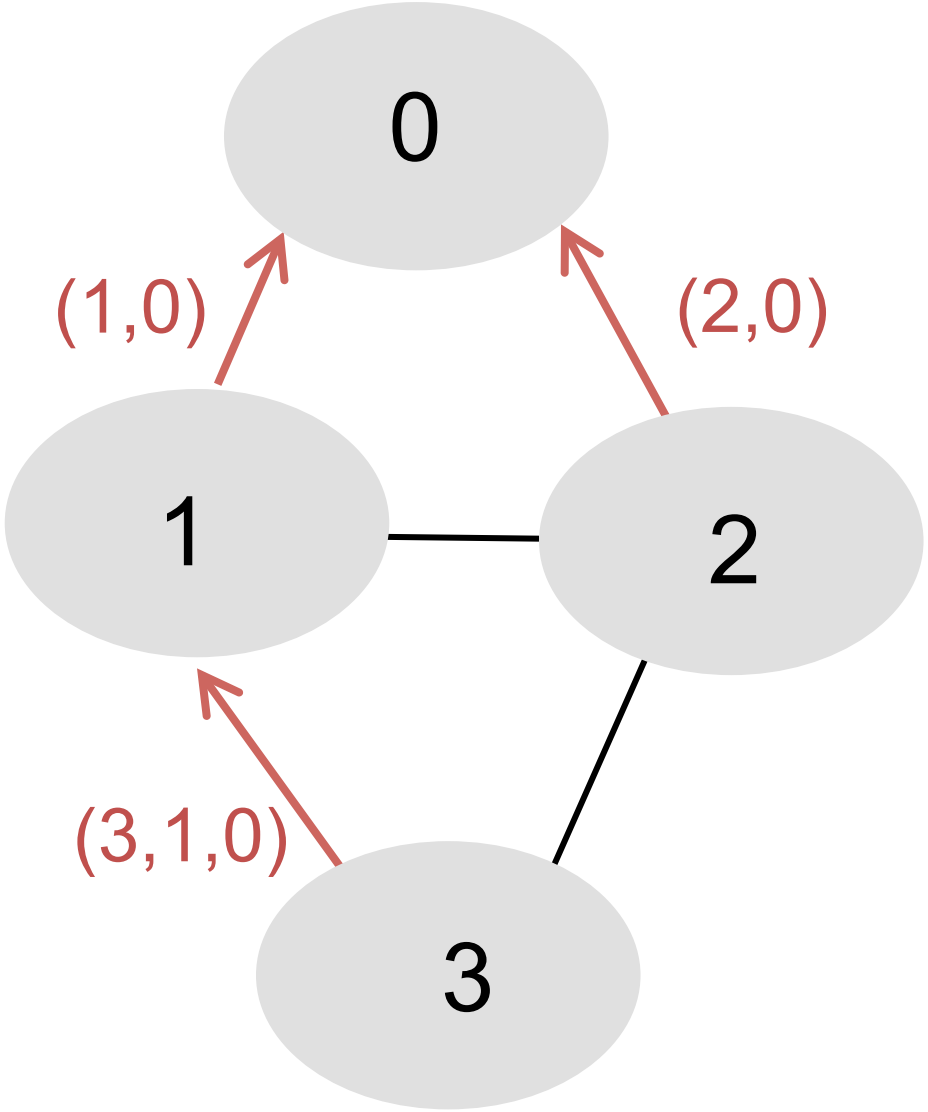
Some BGP Challenges

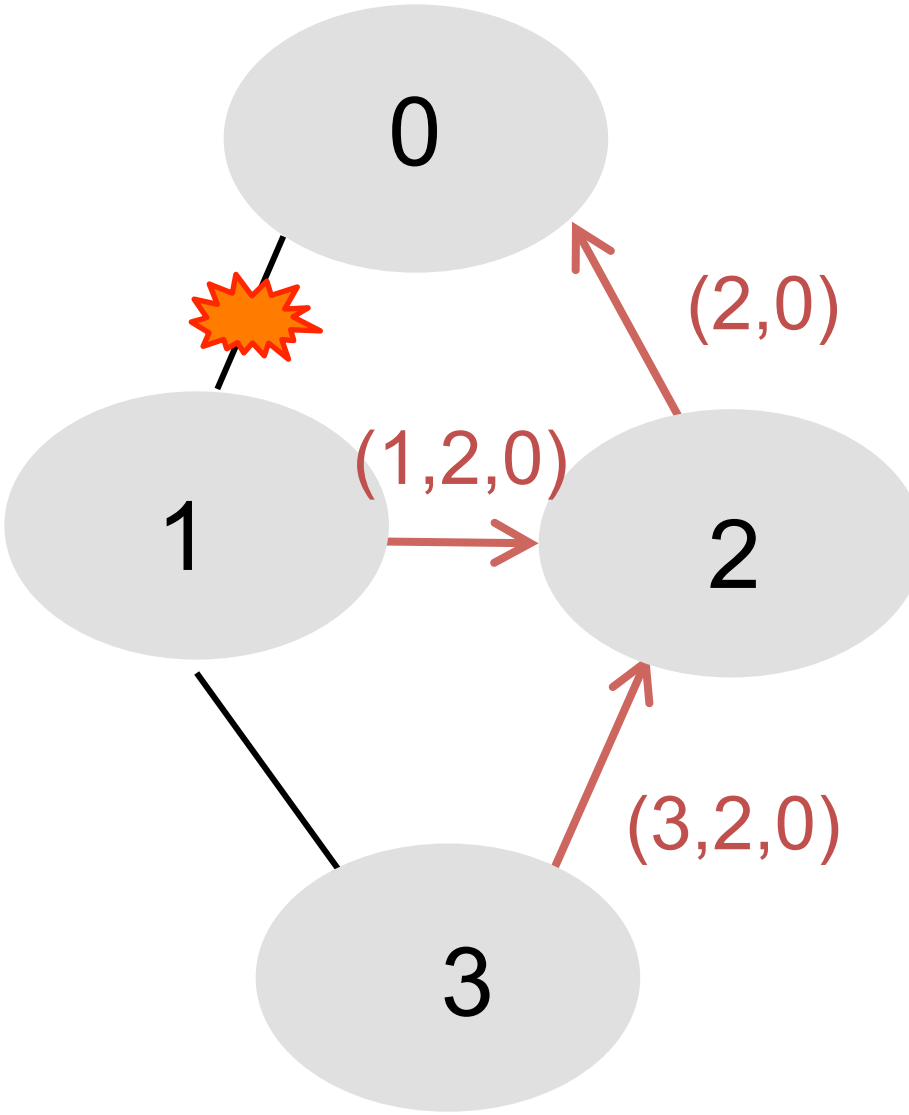
- Convergence
- Traffic engineering
  - How to assure certain routes are selected
- Scaling (route reflectors)
- Security

Convergence

- Given a change, how long until the network re-stabilizes?
  - Depends on change: sometimes never
  - Open research problem: “tweak and pray”
  - Distributed setting is challenging
- Some reasons for change
  - Topology changes
  - BGP session failures
  - Changes in policy
  - Conflicts between policies can cause oscillation

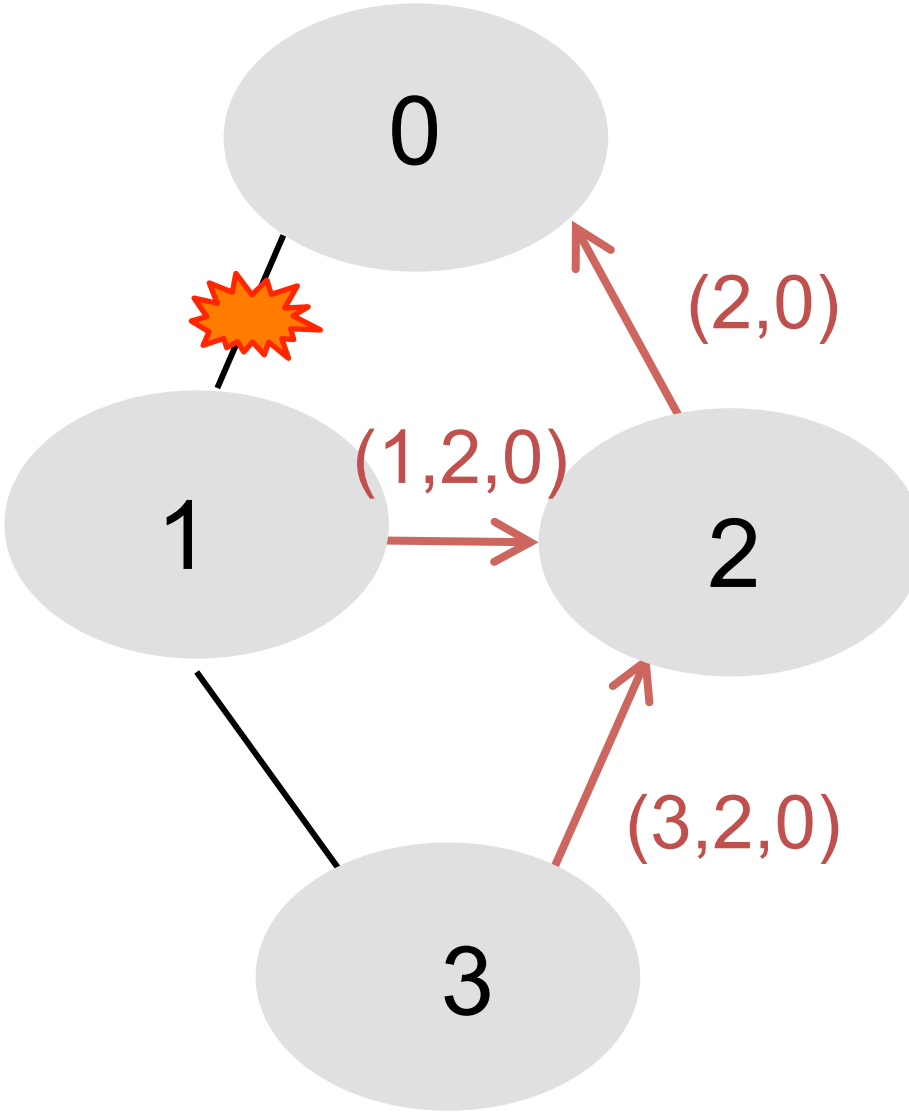
Routing Change: Before and After





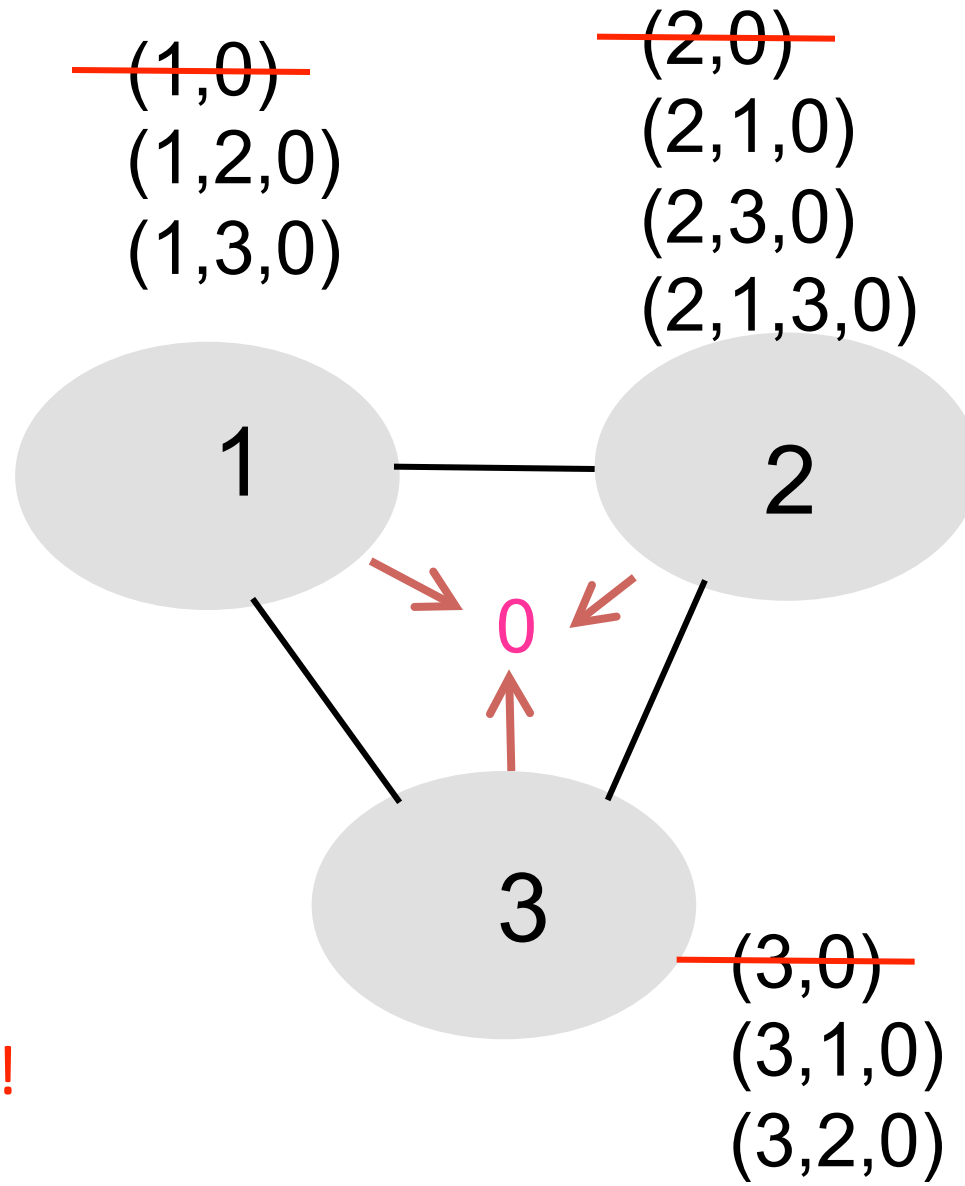
Routing Change: Path Exploration

- AS 1
  - Delete the route (1,0)
  - Switch to next route (1,2,0)
  - Send route (1,2,0) to AS 3
- AS 3
  - Sees (1,2,0) replace (1,0)
  - Compares to route (2,0)
  - Switches to using AS 2



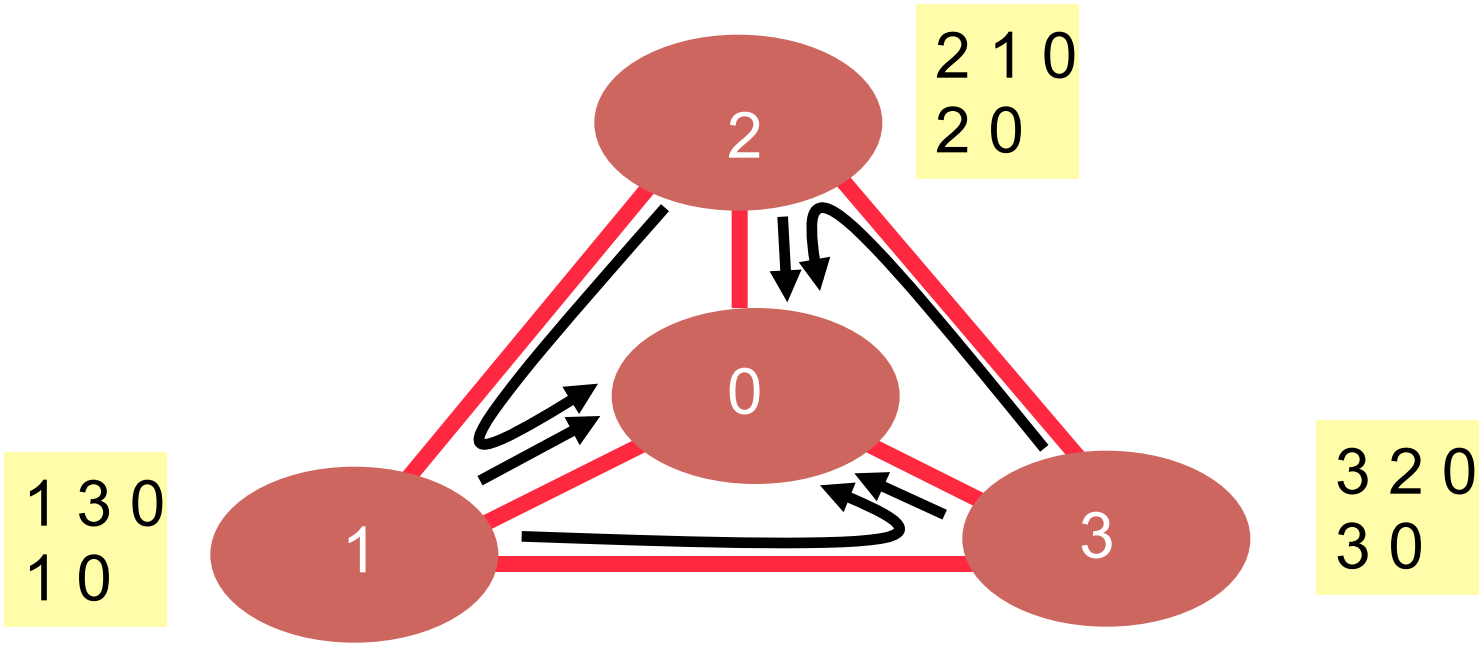
Routing Change: Path Exploration

- Initial situation
  - Destination 0 is alive
  - All ASes use direct path
- When destination dies
  - All ASes lose direct path
  - All switch to longer paths
  - Eventually withdrawn
- E.g., AS 2
  - (2,0) → (2,1,0)
  - (2,1,0) → (2,3,0)
  - (2,3,0) → (2,1,3,0)
  - (2,1,3,0) → null
- Convergence may be slow!



Unstable Configurations

- Due to policy conflicts



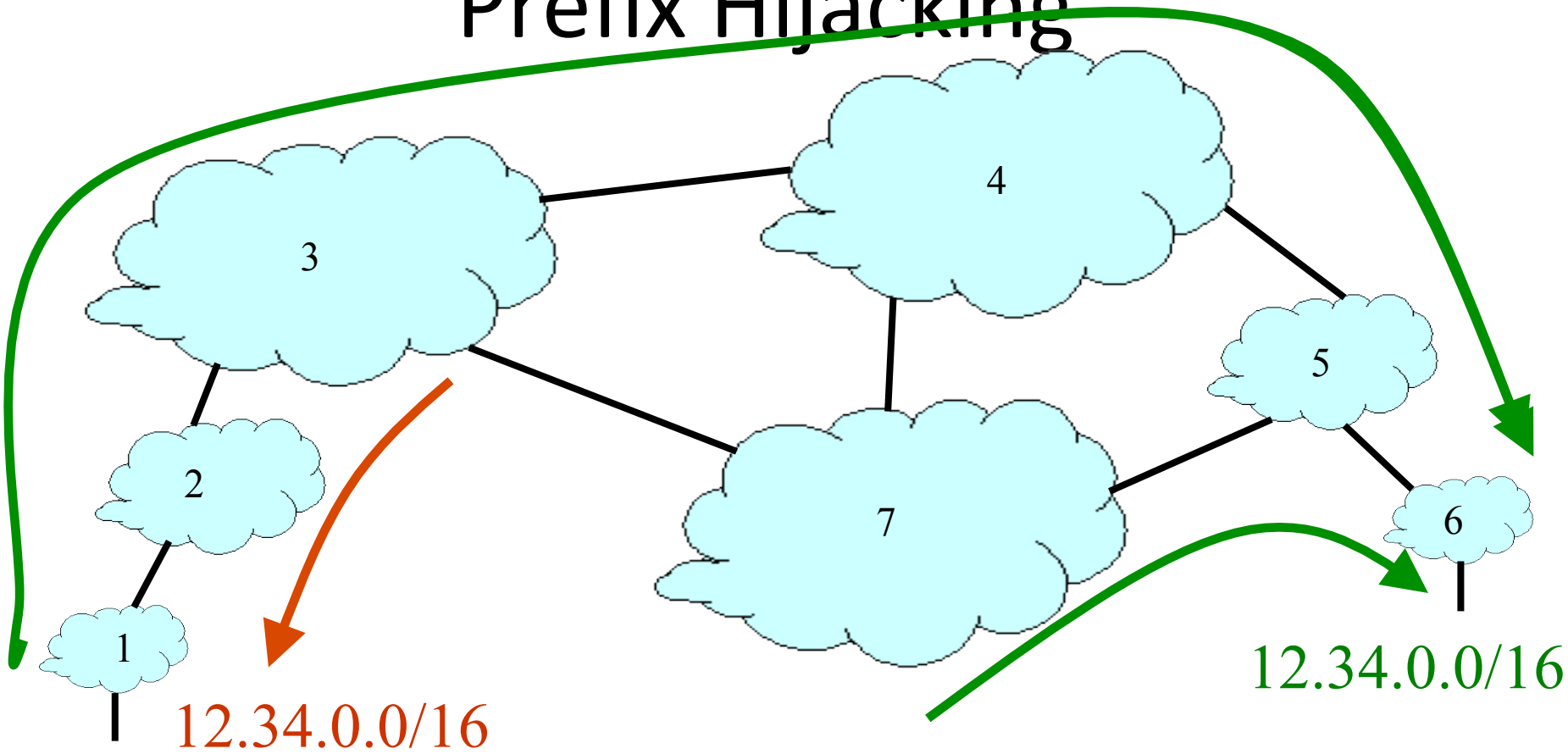
BGP Security Goals

- Confidential message exchange between neighbors
- Validity of routing information
  - Origin, Path, Policy
- Correspondence to the data path

Origin: 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

Prefix Hijacking

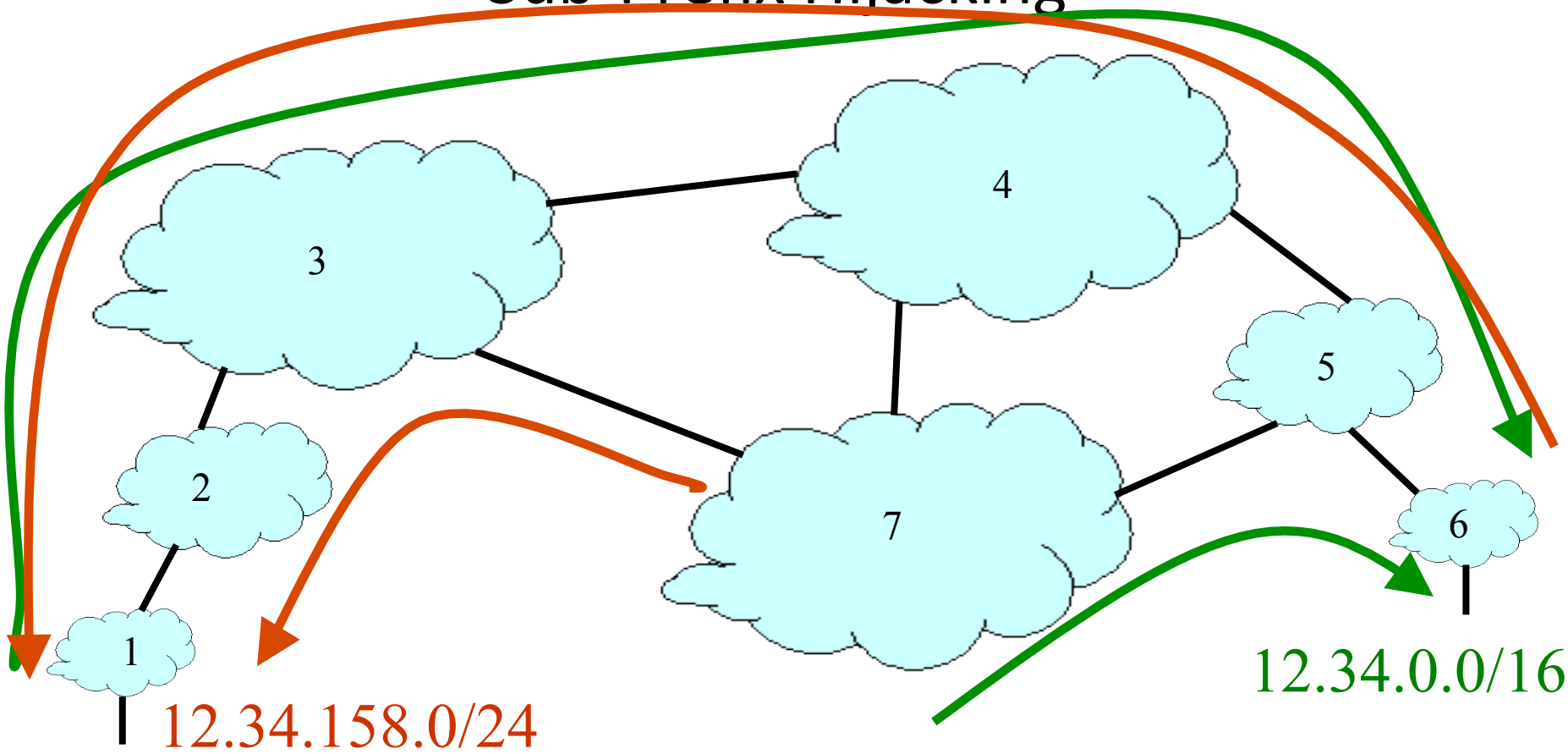


- Consequences for the affected ASes
  - Blackhole: data traffic is discarded
  - Snooping: data traffic is inspected, and then redirected
  - Impersonation: data traffic is sent to bogus destinations

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

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

Pakistan Youtube incident

- Youtube's has prefix 208.65.152.0/22
- Pakistan's government order Youtube blocked
- Pakistan Telecom (AS 17557) announces 208.65.153.0/24 in the wrong direction (outwards!)
- Longest prefix match caused worldwide outage
- http://www.youtube.com/watch?v=IzLPKuAOe50

Many other incidents

- Spammers steal unused IP space to hide
  - Announce very short prefixes (e.g., /8). Why?
  - For a short amount of time
- China incident, April 8th 2010
  - China Telecom's AS23724 generally announces 40 prefixes
  - On April 8th, announced ~37,000 prefixes
  - About 10% leaked outside of China
  - Suddenly, going to www.dell.com might have you routing through AS23724!

Attacks on BGP Paths

- Remove an AS from the path
  - E.g., 701 3715 88 -> 701 88
- Why?
  - Attract sources that would normally avoid AS 3715
  - Make AS 88 look like it is closer to the core
  - Can fool loop detection!
- May be hard to tell whether this is a lie
  - 88 could indeed connect directly to 701!

Attacks on BGP Paths

- Adding ASes to the path
  - E.g., 701 88 -> 701 3715 88
- Why?
  - Trigger loop detection in AS 3715
    - This would block unwanted traffic from AS 3715!
  - Make your AS look more connected
- Who can tell this is a lie?
  - AS 3715 could, if it could see the route
  - AS 88 could, but would it really care?

Attacks on BGP Paths

- Adding ASes at the end of the path
  - E.g., 701 88 into 701 88 3
- Why?
  - Evade detection for a bogus route (if added AS is legitimate owner of a prefix)
- Hard to tell that the path is bogus!