

Introduction to Computer Networks

COSC 4377

Lecture 15

Spring 2012

March 19, 2012

Announcements

- HW7 due this week
- HW8 due 3/28
- Exam 2 on 4/23

HW7

- RIP (Routing Information Protocol)
- Components
 - Forwarding
 - Routing
- Port assignments

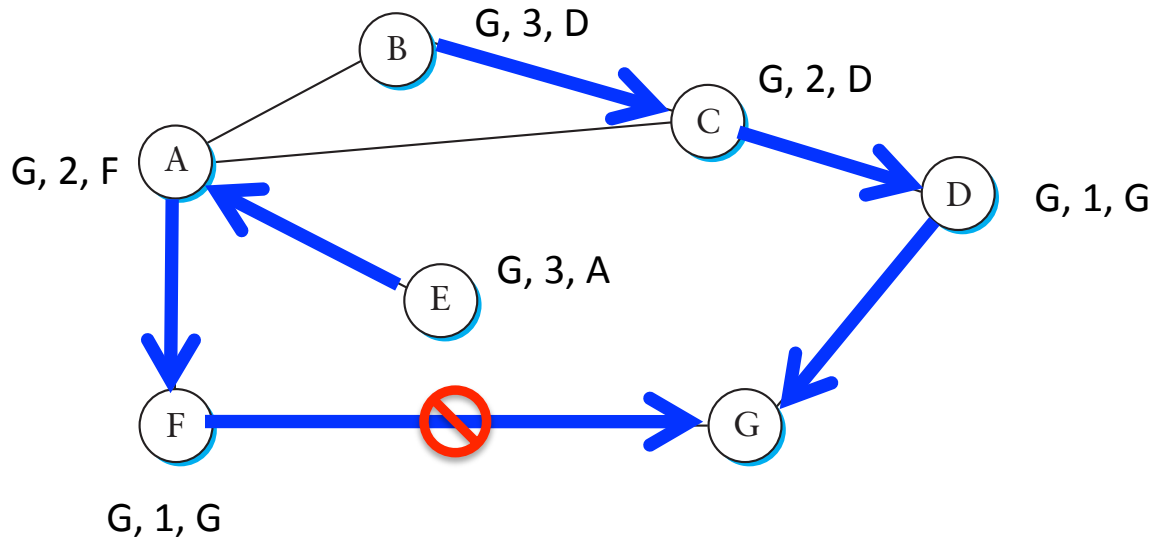
Today's Topics

- Inter-domain Routing

Basic Algorithms

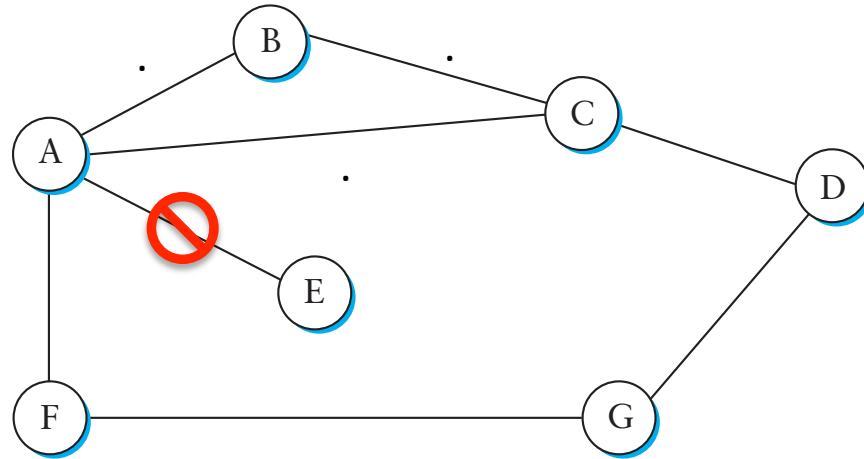
- Two classes of intra-domain routing algorithms
- Distance Vector
 - Requires only local state
 - Harder to debug
 - Can suffer from loops
- Link State
 - Each node has global view of the network
 - Simpler to debug
 - Requires global state

Adapting to Failures



- F-G fails
- F sets distance to G to infinity, propagates
- A sets distance to G to infinity
- A receives periodic update from C with 2-hop path to G
- A sets distance to G to 3 and propagates
- F sets distance to G to 4, through A

Count-to-Infinity



- Link from A to E fails
- A advertises distance of infinity to E
- B and C advertise a distance of 2 to E
- B decides it can reach E in 3 hops through C
- A decides it can reach E in 4 hops through B
- C decides it can reach E in 5 hops through A, ...
- **When does this stop?**

Inter-domain Routing

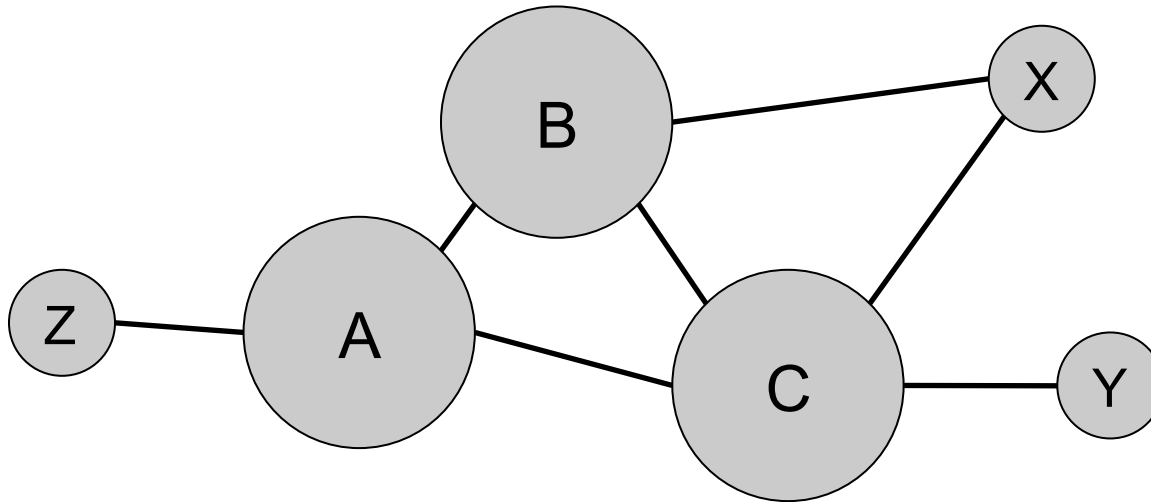
Why Inter vs. Intra

- Trust
- Policy
- Scale
- Performance

Types of ASs

- Local Traffic – source or destination in local AS
- Transit Traffic – passes through an AS
- Stub AS
 - Connects to only a single other AS
- Multihomed AS
 - Connects to multiple ASs
 - Carries no transit traffic
- Transit AS
 - Connects to multiple ASs and carries transit traffic

AS Relationships



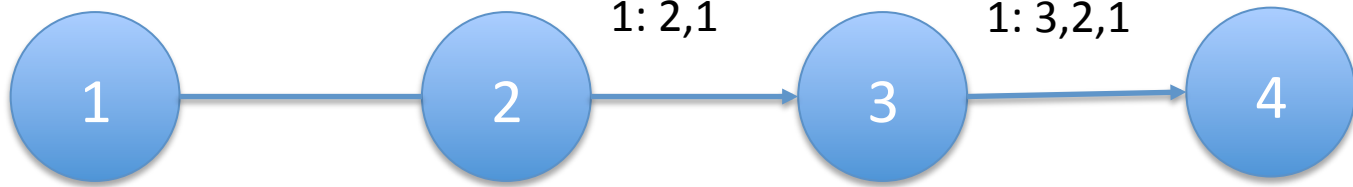
- How to prevent X from forwarding transit between B and C?
- How to avoid transit between CBA ?
 - B: BAZ \rightarrow X
 - B: BAZ \rightarrow C ? (\Rightarrow Y: CBAZ and Y:CAZ)

Autonomous System

- Group of routers / prefixes typically under the control of a single organization
- Example: University of Houston
- Here is one list
 - <http://bgp.potaroo.net/cidr/autnums.html>
- [http://en.wikipedia.org/wiki/Autonomous_system_\(Internet\)](http://en.wikipedia.org/wiki/Autonomous_system_(Internet))

Path Vector Protocol

- Distance vector algorithm with extra information
 - For each route, store the complete path (ASs)
 - No extra computation, just extra storage (and traffic)
- Advantages
 - Can make policy choices based on set of ASs in path
 - Can easily avoid loops



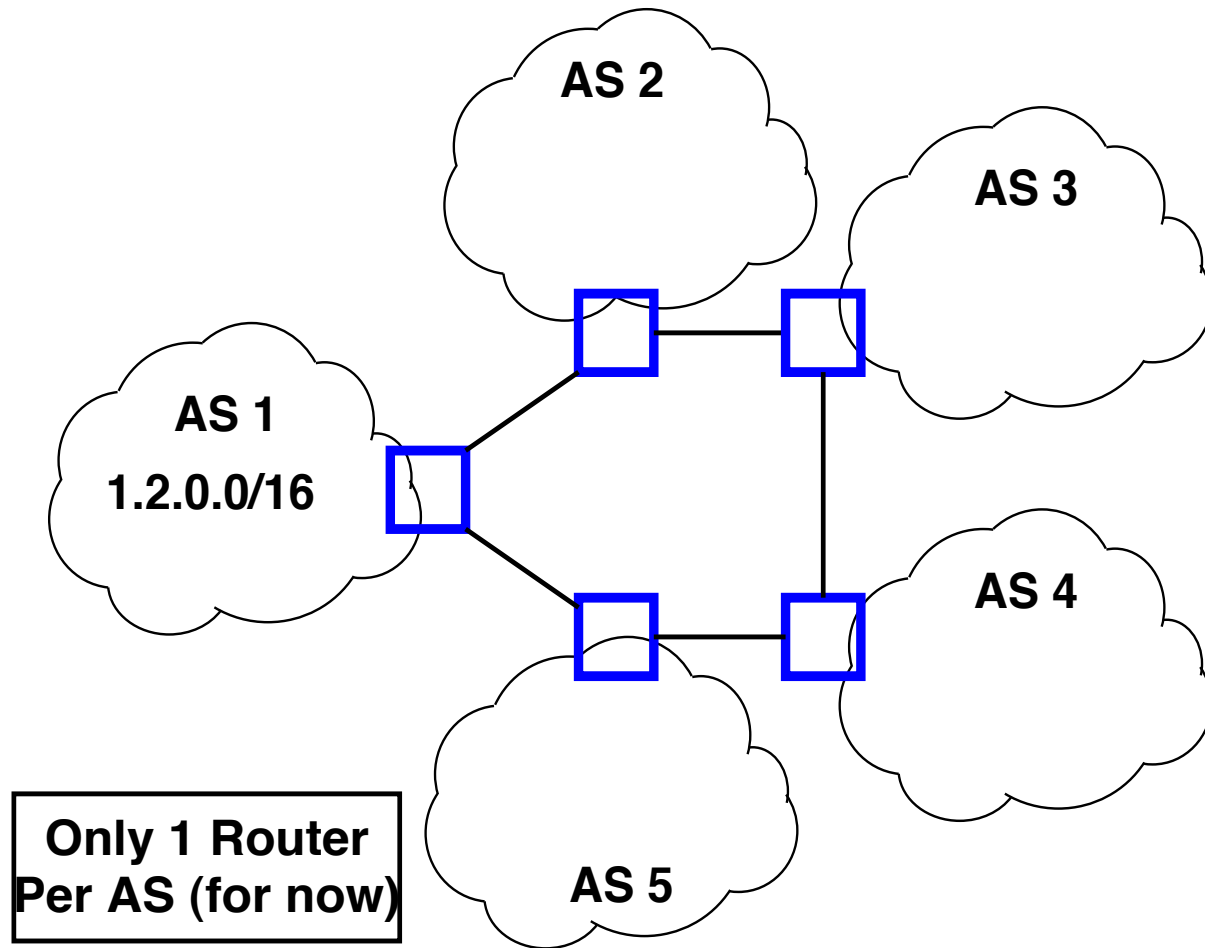
BGP - High Level

- Abstract each AS to a single node
- Destinations are CIDR prefixes
- Exchange prefix *reachability* with all neighbors
 - E.g., “I can reach prefix 128.148.0.0/16 through ASes 44444 3356 14325 11078”
- Select a single path by routing *policy*
- Critical: learn many paths, propagate one
 - Add your ASN to advertised path

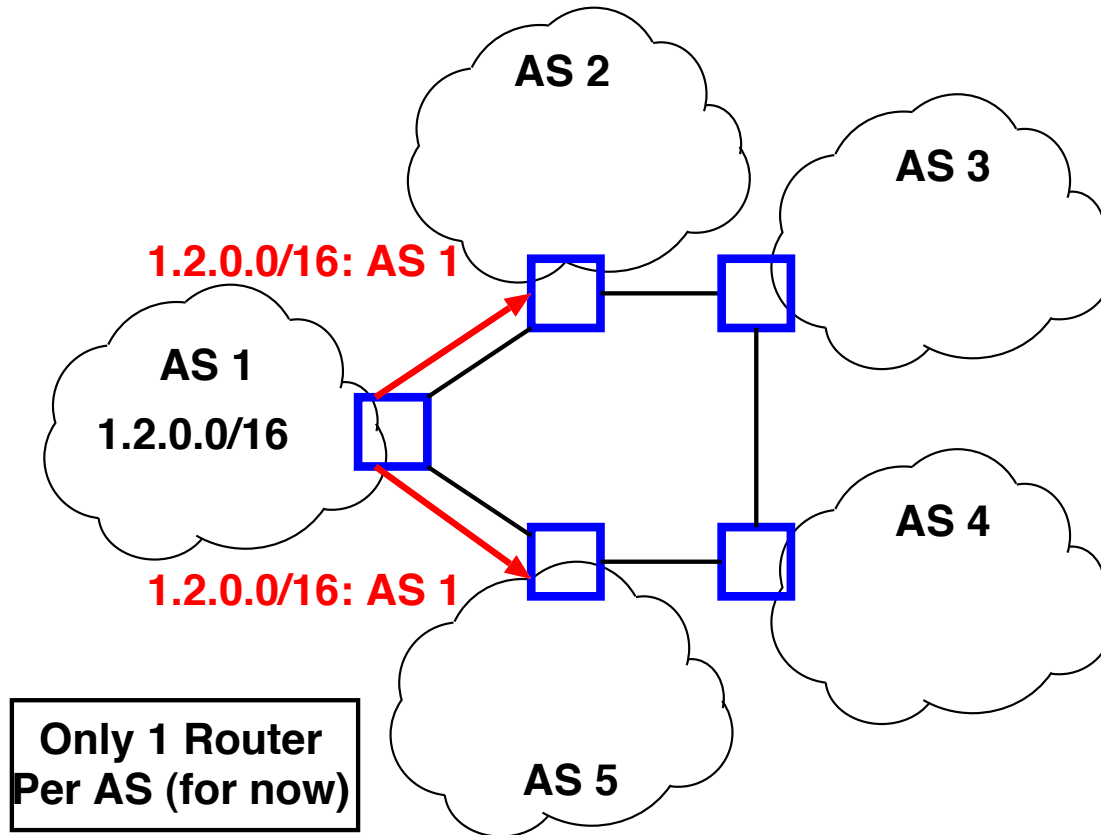
Why study BGP?

- Critical protocol: makes the Internet run
 - Only widely deployed EGP
- Active area of problems!
 - Efficiency
 - Cogent vs. Level3: Internet Partition
 - Spammers use prefix hijacking
 - Pakistan accidentally took down YouTube
 - Egypt disconnected for 5 days

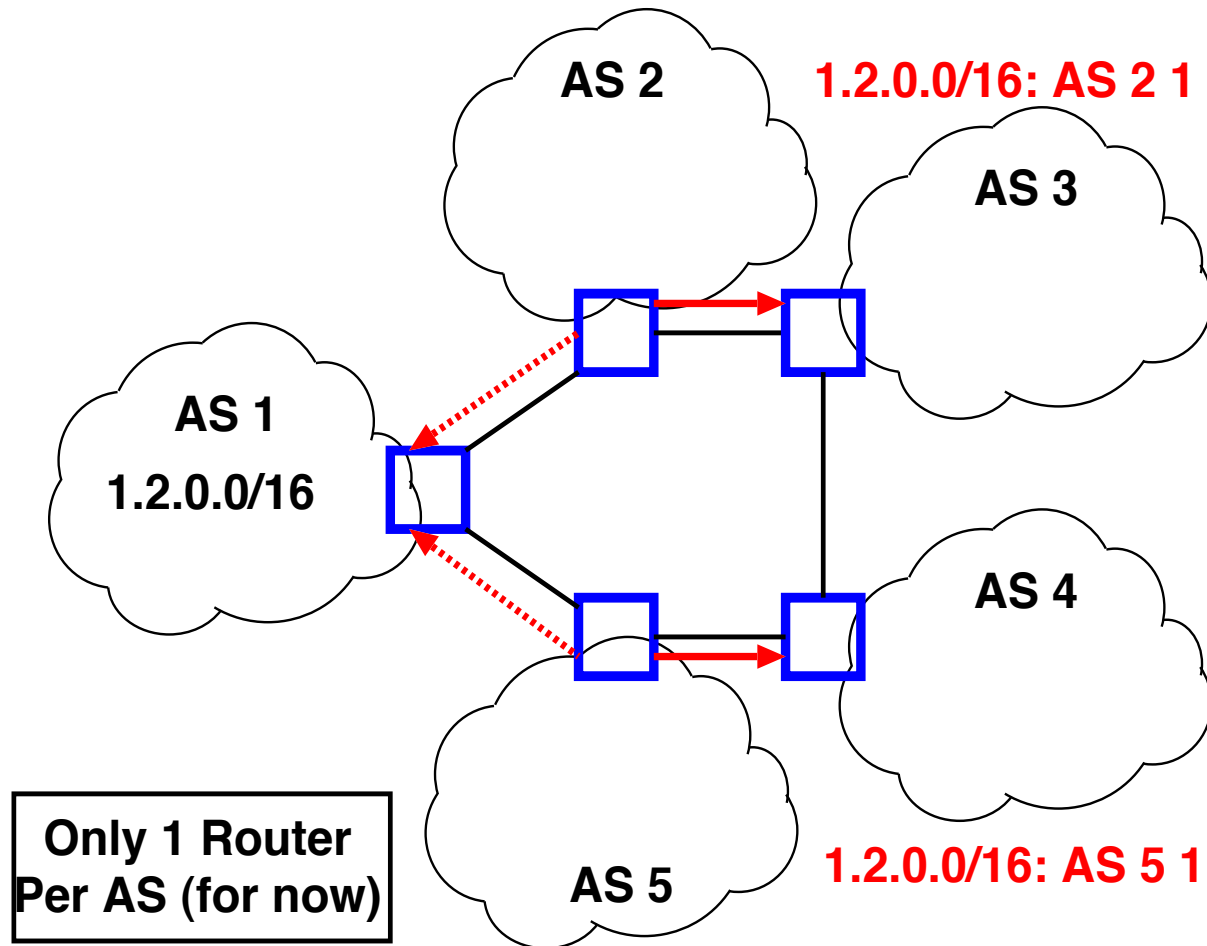
BGP Example



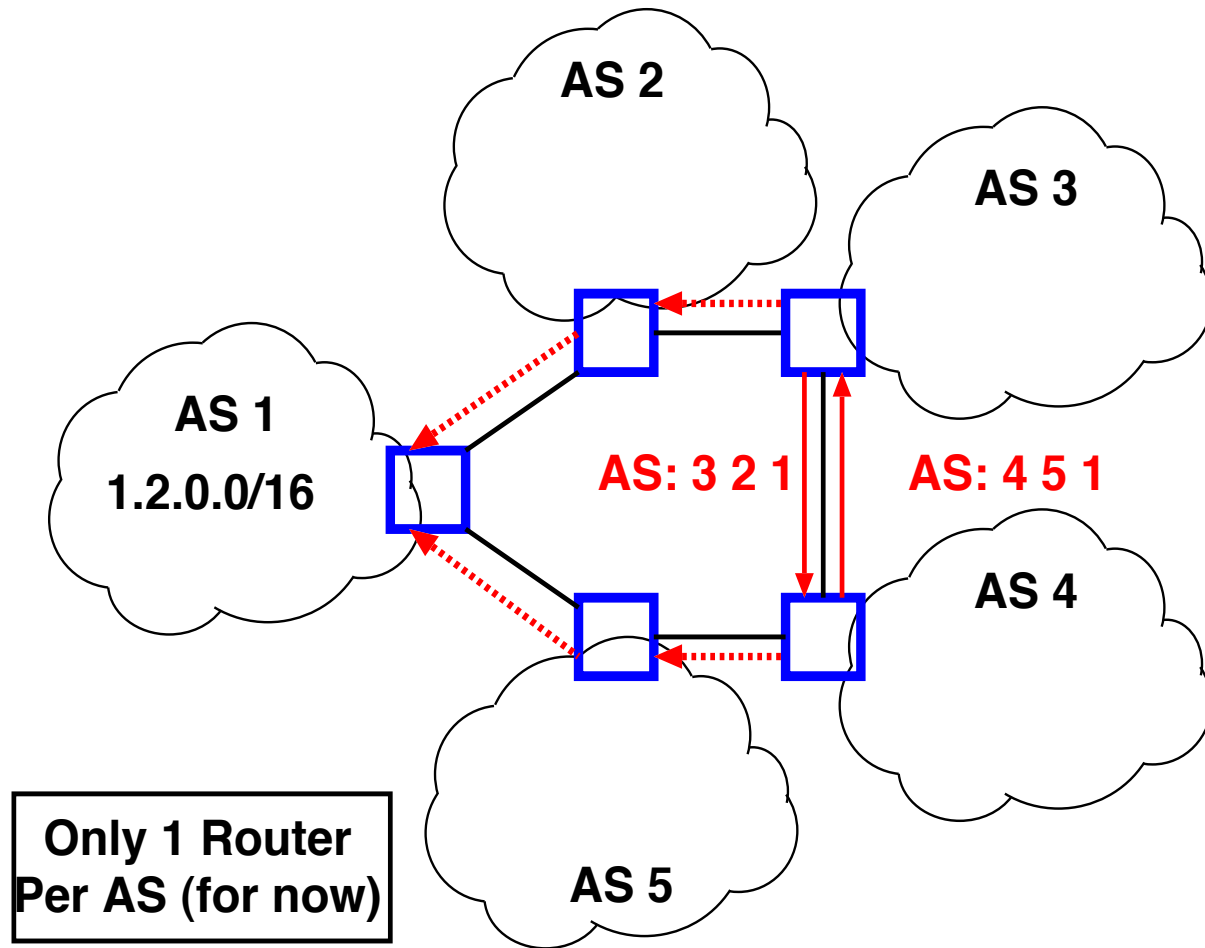
BGP Example



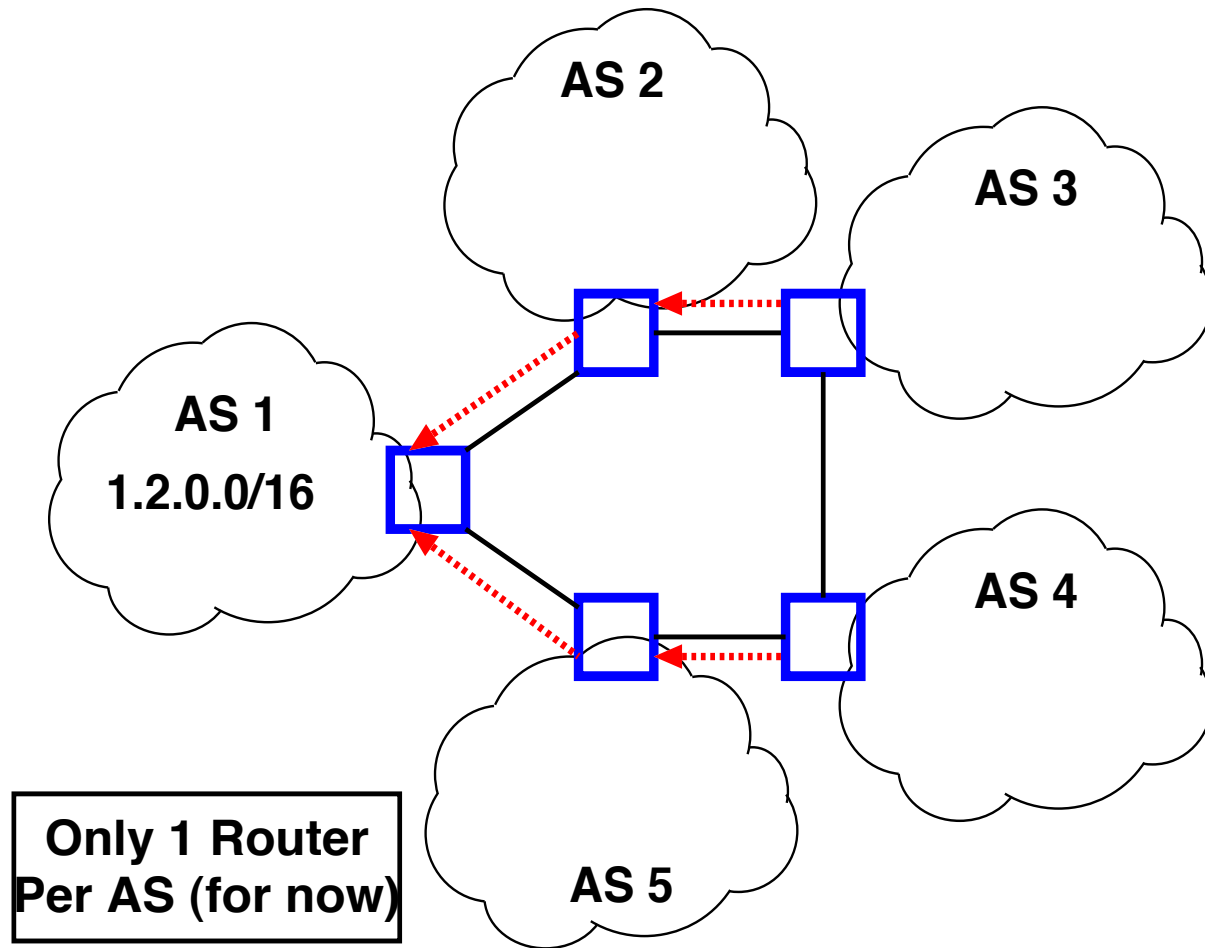
BGP Example



BGP Example



BGP Example



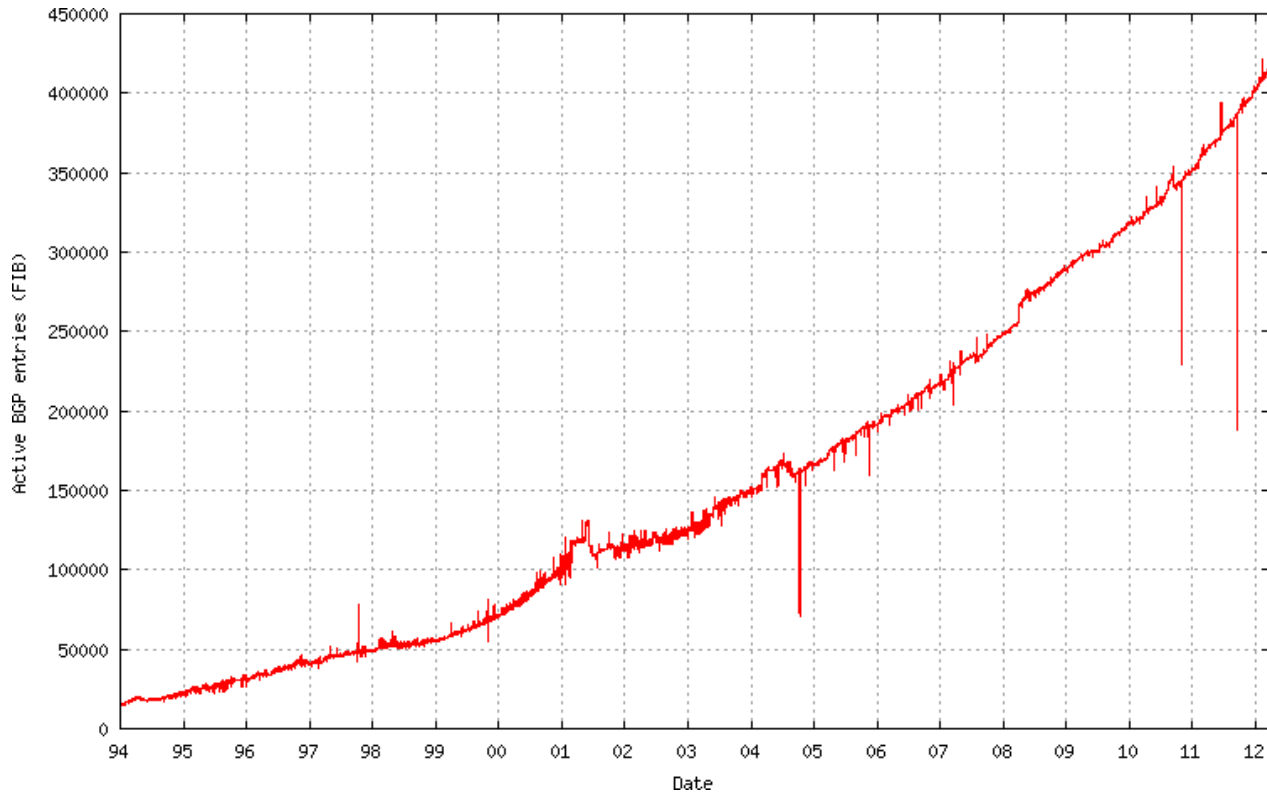
BGP Protocol Details

- Separate roles of *speakers* and *gateways*
 - Speakers talk BGP with other AS
 - Gateways are routers that border other AS
 - Can have more gateways than speakers
 - Speakers know how to reach gateways
- Speakers connect over TCP on port 179
 - Bidirectional exchange over long-lived connection

BGP Implications

- Explicit AS Path == Loop free
 - Except under churn, IGP/EGP mismatch
- Reachability not guaranteed
 - Decentralized combination of policies
- Not all ASs know all paths
- AS abstraction -> loss of efficiency
- Scaling
 - 37K ASs
 - 350K+ prefixes
 - ASs with one prefix: 15664
 - Most prefixes by one AS: 3686 (AS6389, BellSouth)

BGP Table Growth (AS6447)



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

Route Selection

- More specific prefix
- Next-hop reachable?
- Prefer highest weight
 - Computed using some AS-specific local policy
- Prefer highest local-pref
- Prefer locally originated routes
- Prefer routes with shortest AS path length
- Prefer eBGP over iBGP
- Prefer routes with lowest cost to egress point
 - Hot-potato routing
- Tie-breaking rules
 - E.g., oldest route, lowest router-id

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

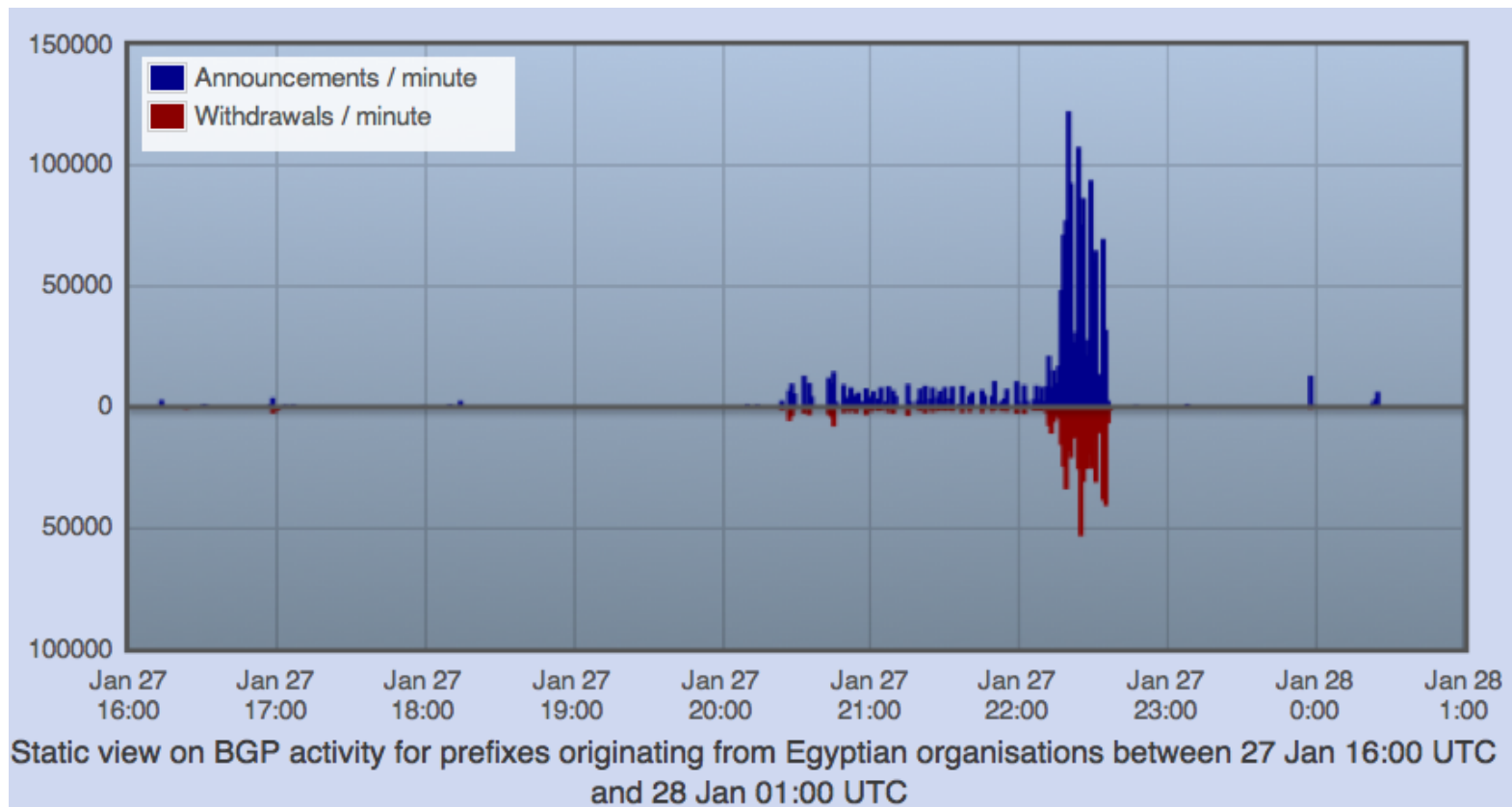
- ASs agree to exchange traffic for free
 - Penalties/Renegotiate if imbalance
- Tier 1 ISPs have no default route: all peer with each other
- You are Tier $i + 1$ if you have a default route to a Tier i
- Typical policies
 - AS only exports customer routes to peer
 - AS exports a peer's routes only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

Peering Drama

- Cogent vs. Level3 were peers
- In 2003, Level3 decided to start charging Cogent
- Cogent said no
- **Internet partition**: Cogent's customers couldn't get to Level3's customers and vice-versa
 - Other ISPs were affected as well
- Took 3 weeks to reach an undisclosed agreement

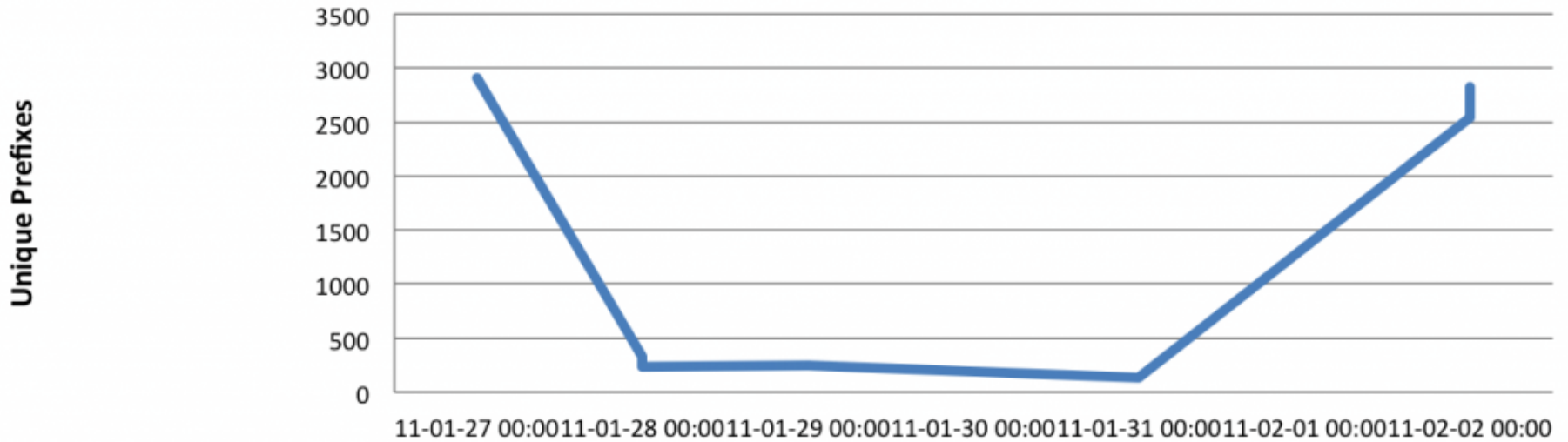
“Shutting off” the Internet

- Starting from Jan 27th, 2011, Egypt was disconnected from the Internet
 - 2769/2903 networks withdrawn from BGP (95%)!



Egypt Incident

Number of Egyptian networks



	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