Class Meeting, Lectures 11 & 12: Wide Area Routing

Kyle Jamieson

COS 461: Computer Networks

Context: Autonomous Systems

- A routing domain is called an Autonomous System (AS)
  - Each AS known by unique 16-bit number
  - AS owns one or handful of address prefixes; allocates addresses under those prefixes
  - AS typically a commercial entity or other organization
  - ASes often competitors (e.g., different ISPs)
- Interior Gateway Protocols (IGPs) (e.g., DV, LS) route within individual ASes
- Exterior Gateway Protocols (EGPs) (e.g., BGP) route among ASes

Today

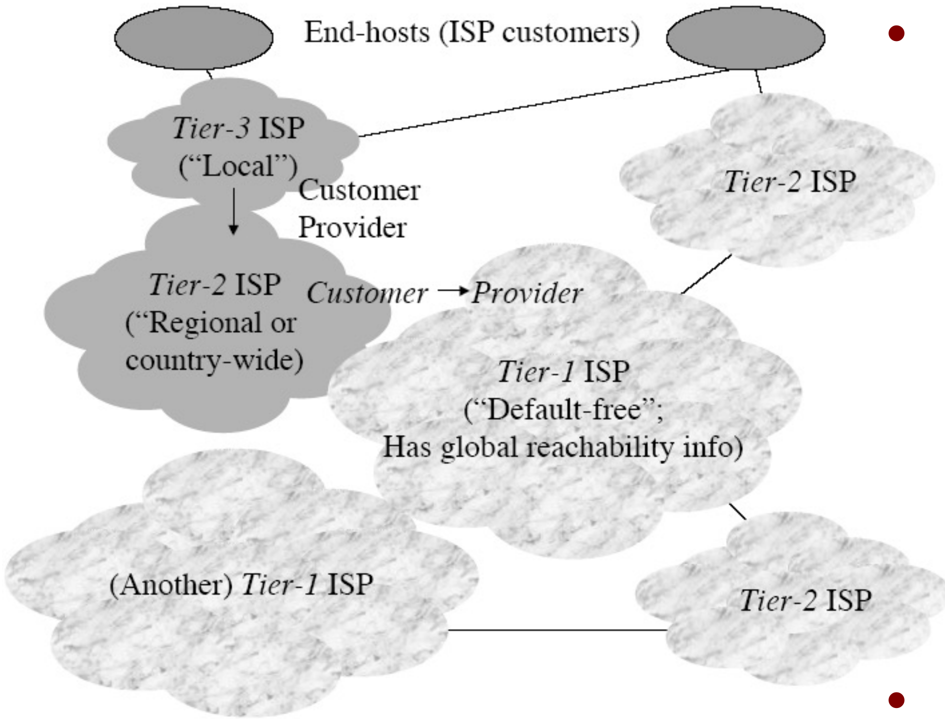
1. Recap of BGP wide area routing

2. Integrating IGP and BGP

3. BGP (in)stability

4. BGP Route monitoring

Global Internet Routing



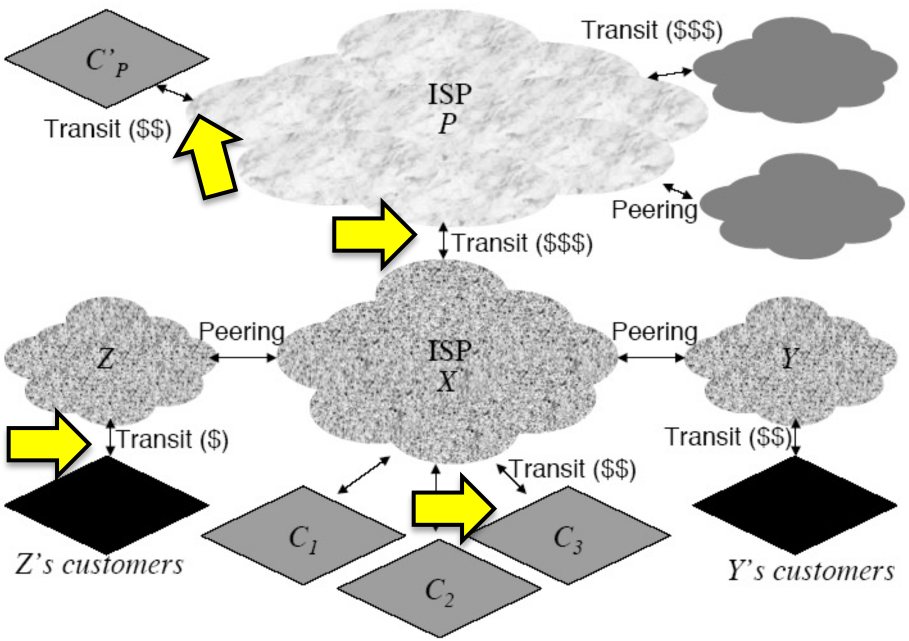
- Tiers of ISPs:
  - Tier 1: geographically global, ISP customers, no default routes
  - Tier 2: regional geographically
  - Tier 3: local geographically, end customers
- Each ISP is an AS
  - AS operator sets policies for how to route to others, how to let others route to them

AS-AS Relationship: Customers and Providers

- Examples:
  - Smaller ASes (corporations, universities) typically purchase connectivity from ISPs
  - Regional ISPs typically purchase connectivity from global ISPs

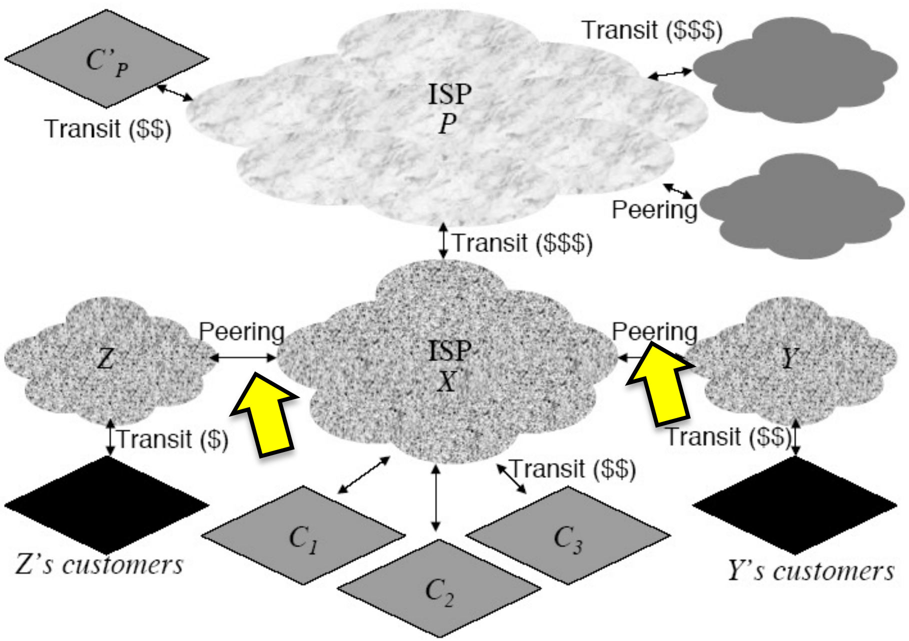
- Each such connection has two roles:
  - Customer: smaller AS paying for connectivity
  - Provider: larger AS being paid for connectivity

AS-AS Relationship: Customer-Provider Transit



- Provider-Customer AS-AS connections are also called **transit**
- Provider allows customer to route to (nearly) all destinations in its routing tables
- Transit nearly always involves payment from customer to provider

AS-AS Relationship: Peering



- Peering: two ASes (usually ISPs) mutually allow one another to route to some of the destinations in their routing tables

- By contract, but usually no money changes hands, so long as traffic ratio is narrower than, e.g., 4:1

Financial Motives: Peering and Transit

- Peering relationship often between competing ISPs
- Incentives to peer:
  - Typically, two ISPs notice their own direct customers originate a lot of traffic for the other
  - Each can avoid paying transit costs to others for this traffic; shunt it directly to one another
  - Often better performance (shorter latency, lower loss rate) as avoid transit via another provider
  - Easier than stealing one another's customers
- Tier 1s must typically peer with one another to build complete, global routing tables

The Meaning of Advertising Routes

- AS **A** advertises a route for destination **D** to AS **B**: effectively an offer to forward all traffic from AS **B** to **D**
- Forwarding traffic costs bandwidth
- AS' incentive to control which routes they advertise:
  - no one wants to forward packets without being compensated to do so
  - e.g., when peering, only let neighboring AS send to specific own customer destinations enumerated peering contract

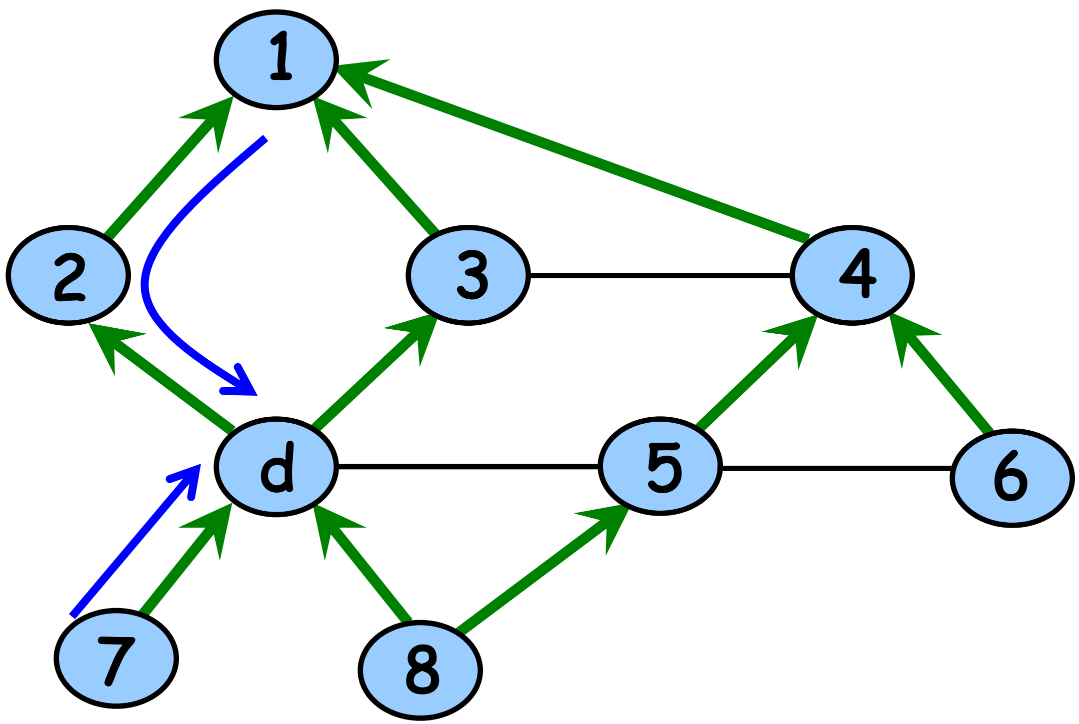
Example -- AS Relationships: Valid and Invalid Paths

Path 1 2 d
Path 7 d
Path 5 8 d
Path 6 4 3 d
Path 8 5 d
Path 6 5 d
Path 1 4 3 d

→ Customer-Provider

Arrow goes from customer to provider

—— Peer-Peer



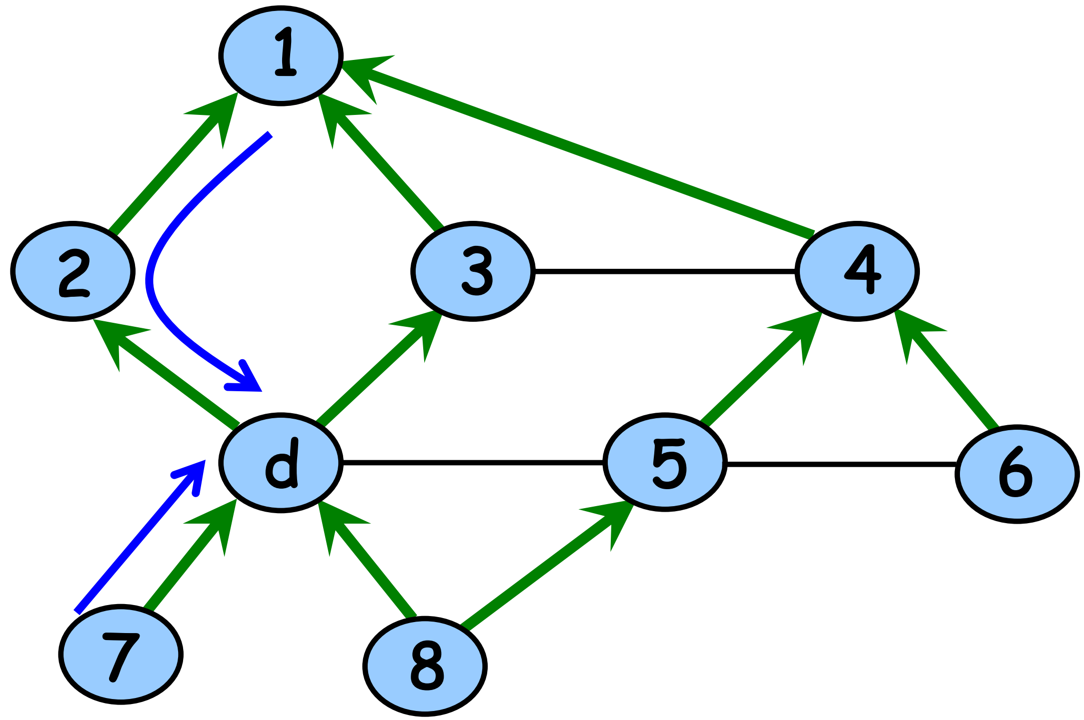
Example -- AS Relationships: Valid and Invalid Paths

Path 1 2 d	Valid
Path 7 d	Valid
Path 5 8 d	Invalid
Path 6 4 3 d	Valid
Path 8 5 d	Valid
Path 6 5 d	Invalid
Path 1 4 3 d	Invalid

→ Customer-Provider

Arrow goes from customer to provider

—— Peer-Peer



Using Route Attributes

- Recall: BGP route advertisement is simply:
  - IP Prefix: [Attribute 0] [Attribute 1] […]

- Administrators enforce policy routing using attributes:
  - filter and rank routes based on attributes
  - modify "next hop" IP address attribute
  - tag a route with attribute to influence ranking and filtering of route at other routers

Synthesis: Multiple Attributes into Policy Routing

- How do attributes interact? Priority order:

Priority	Rule	Details
1	LOCAL PREF	**Highest LOCAL PREF (e.g., prefer transit customer routes over peer and provider routes)**
2	ASPATH	**Shortest ASPATH length**
3	MED	**Lowest MED**
4	eBGP > iBGP	**Prefer routes learned over eBGP vs. over iBGP**
5	IGP path	**"Nearest" egress router**
6	Router ID	**Smallest router IP address**

Today

1. Recap of BGP wide area routing

2. Integrating IGP and BGP

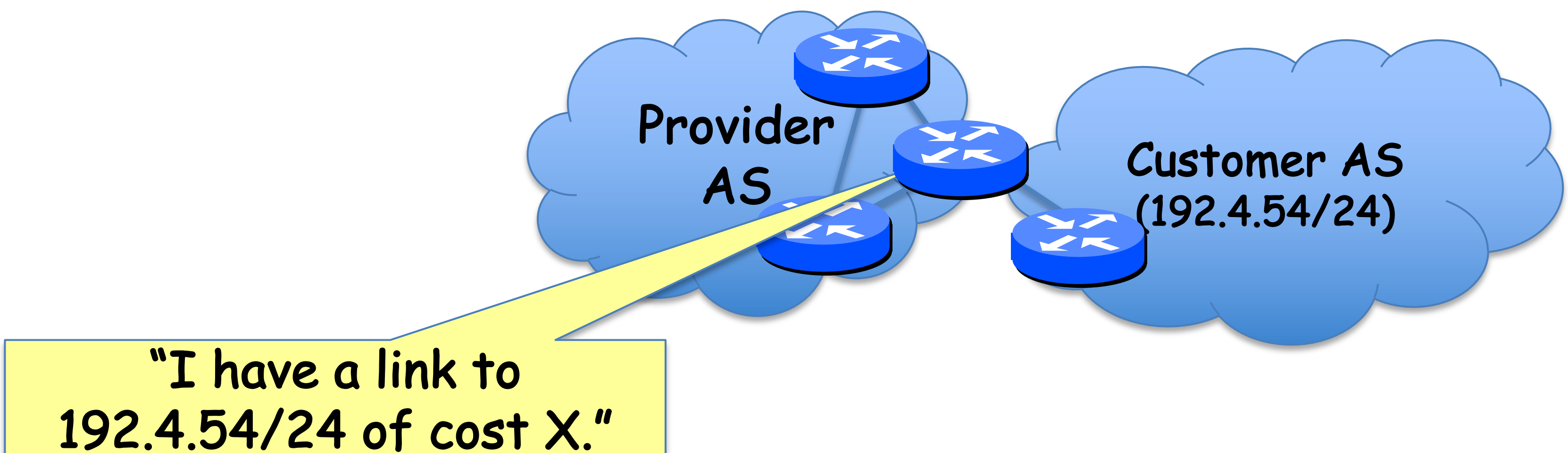
3. BGP (in)stability

4. BGP Route monitoring

Integrating Interdomain and Intradomain Routing (1/2)

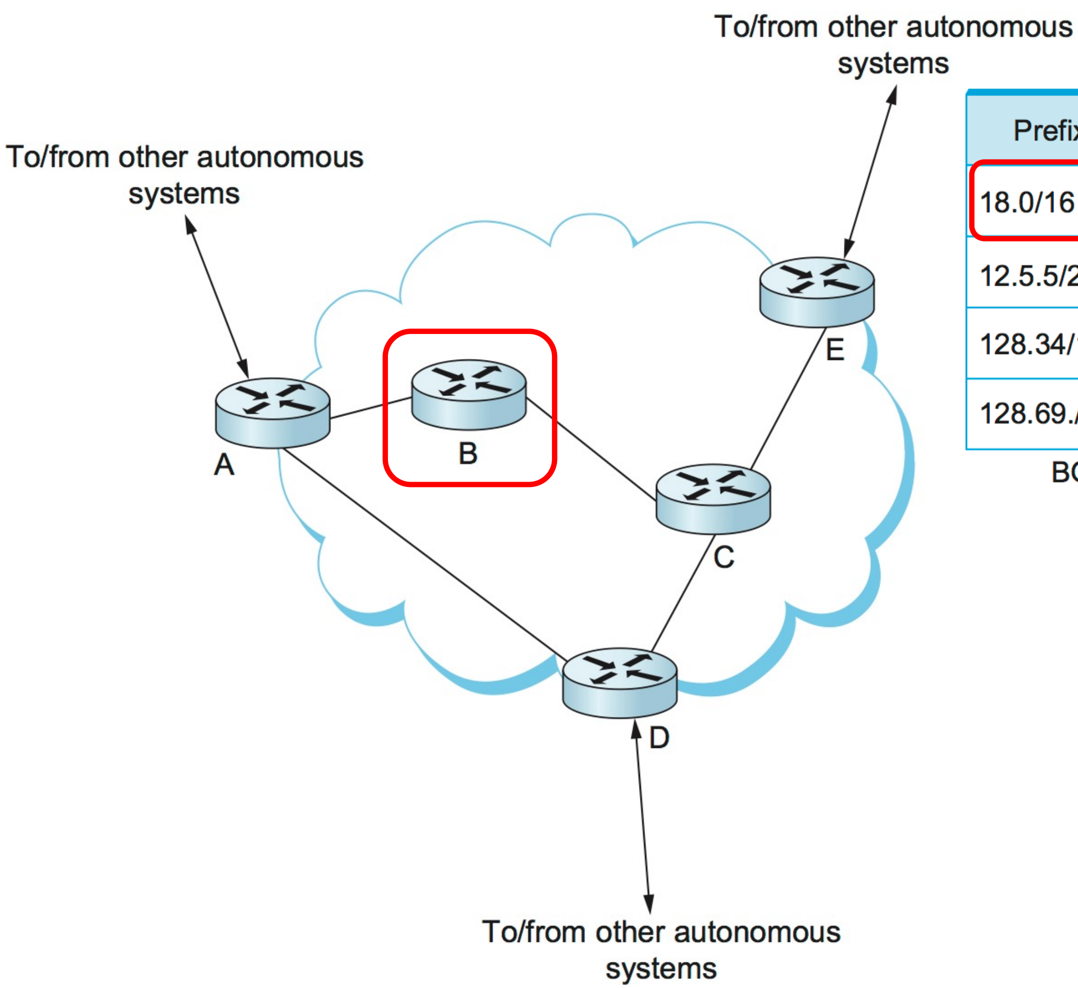
- Option #1: Stub AS that only connects to other autonomous systems at a single point
  - Inject *default route* to border router into the intradomain routing protocol

- Ooption #2: Inject BGP routes into IGP



Integrating Interdomain and Intradomain Routing: Backbone Networks (2/2)

Option #3: Interior BGP (iBGP): redistribute routes learned by the ***BGP speakers*** at the AS edges to all the other routers



Prefix	BGP Next Hop
18.0/16	E
12.5.5/24	A
128.34/16	D
128.69./16	A

BGP table for the AS

Router	IGP Path
A	A
C	C
D	C
E	C

IGP table for router B

Prefix	IGP Path
18.0/16	C
12.5.5/24	A
128.34/16	C
128.69./16	A

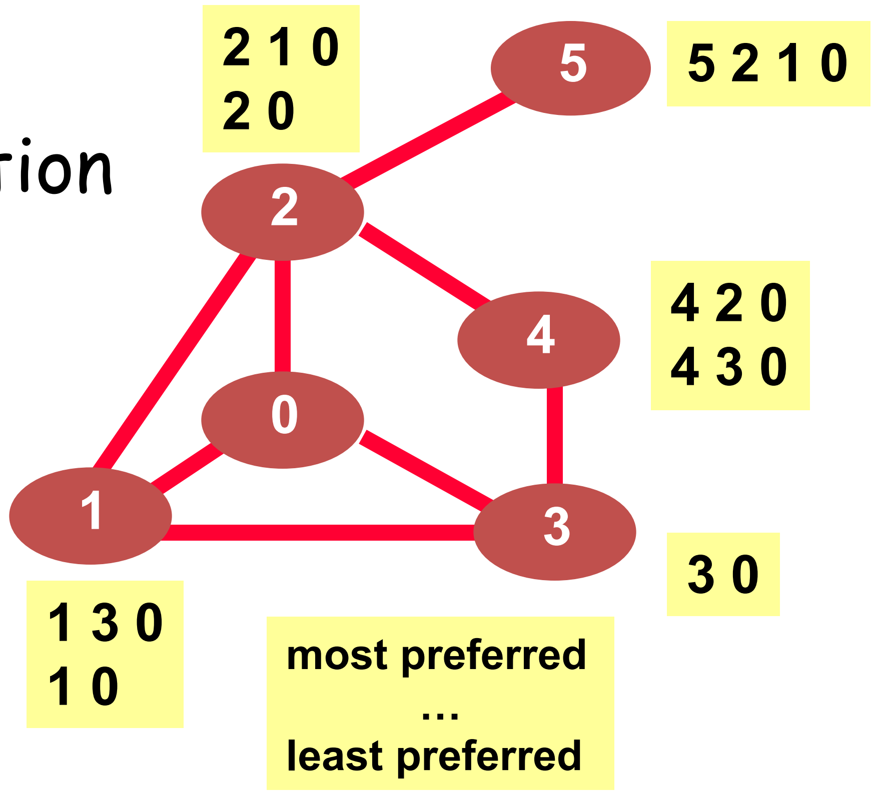
Combined table for router B

Today

1. Recap of BGP wide area routing
2. Integrating IGP and BGP
3. BGP (in)stability
4. BGP Route monitoring

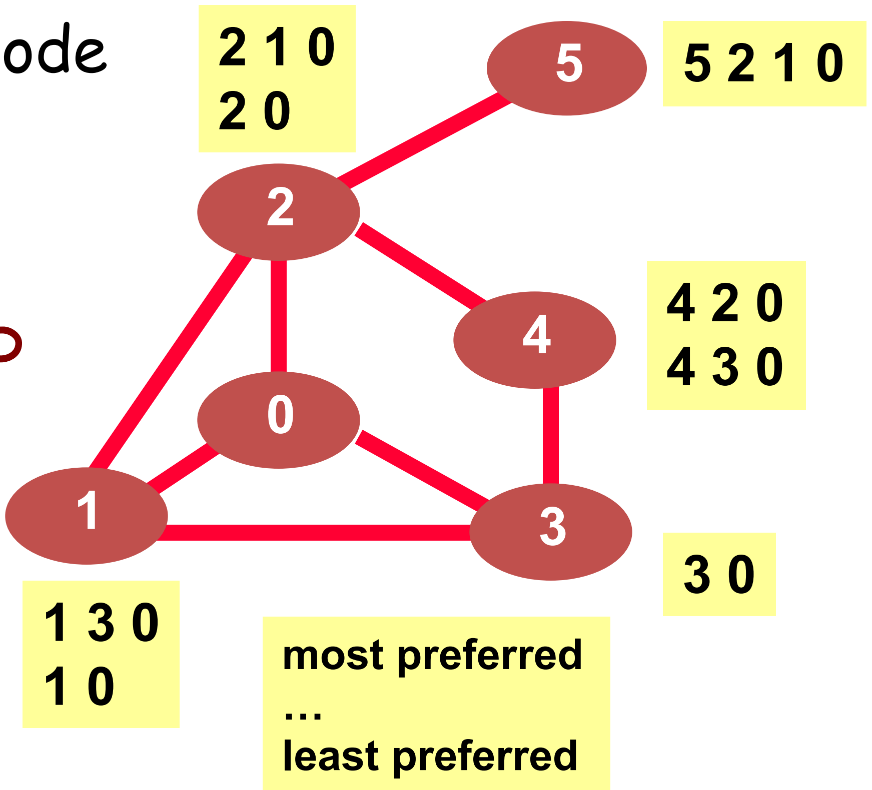
Stable Paths Problem (SPP) Instance

- Node
  - BGP-speaking router
  - E.g.: node 0 = destination
- Edge
  - BGP adjacency
- Permitted paths
  - **Set** of routes to 0 at each node
  - **Ranking** of the paths



SPP Solution

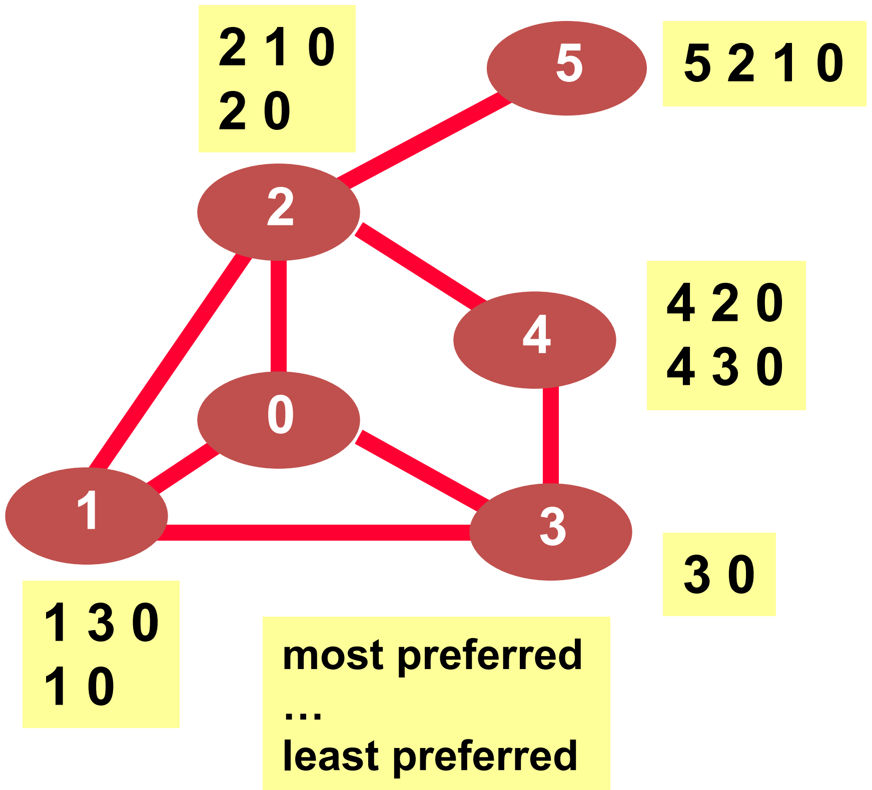
- Solution is:
  - Choice of path to 0, per node
    - Can be the "null" path
- If node u has path {u,w}P
  - {u,w} is edge in graph
- then, node w assigned P
- Each node is assigned
  - Highest ranked path consistent with its neighbors



Stable Paths Problem (SPP) Instance

1. Does 5 have a path to 0?

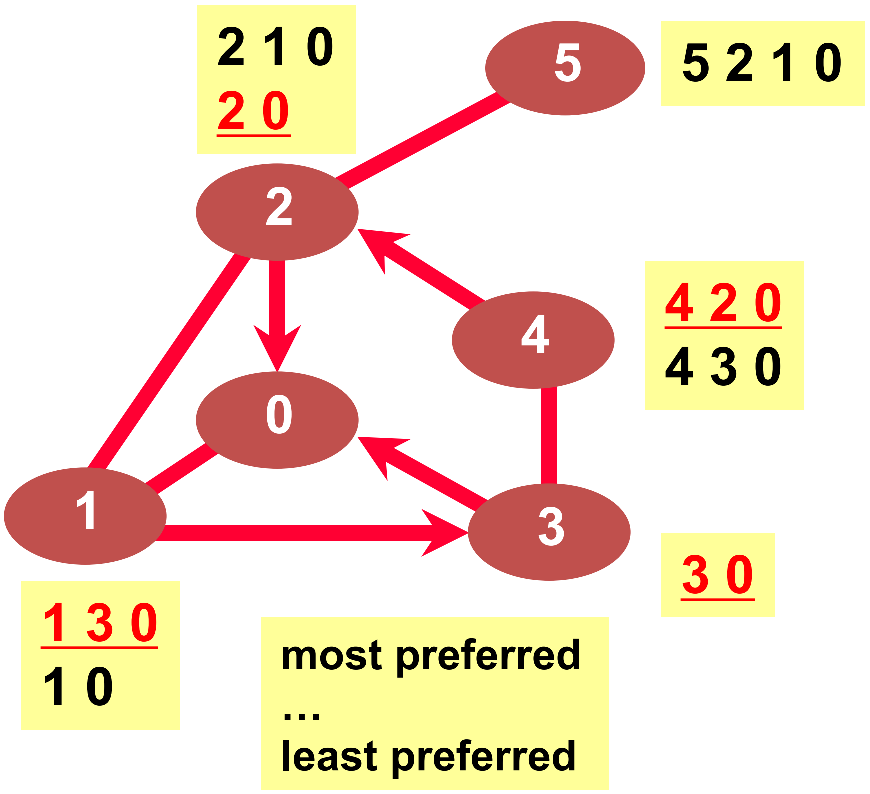
2. Will 1 use the direct path to 0?



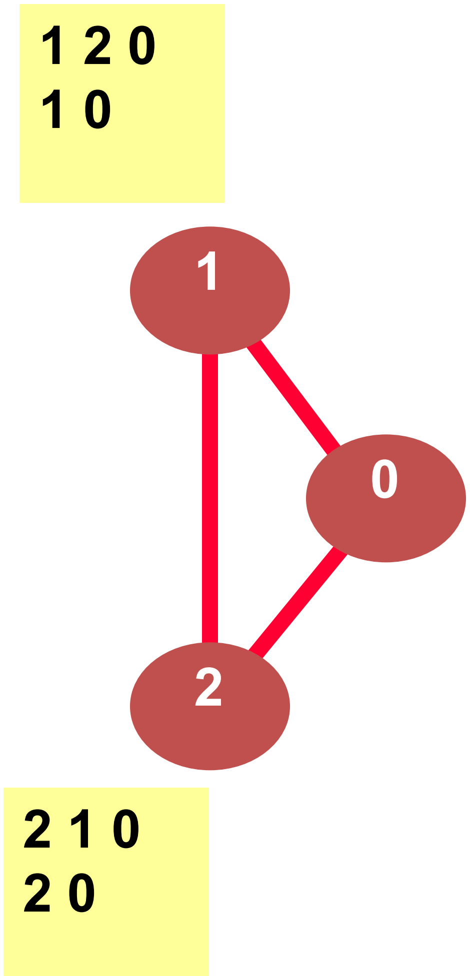
Stable Paths Problem (SPP) Instance

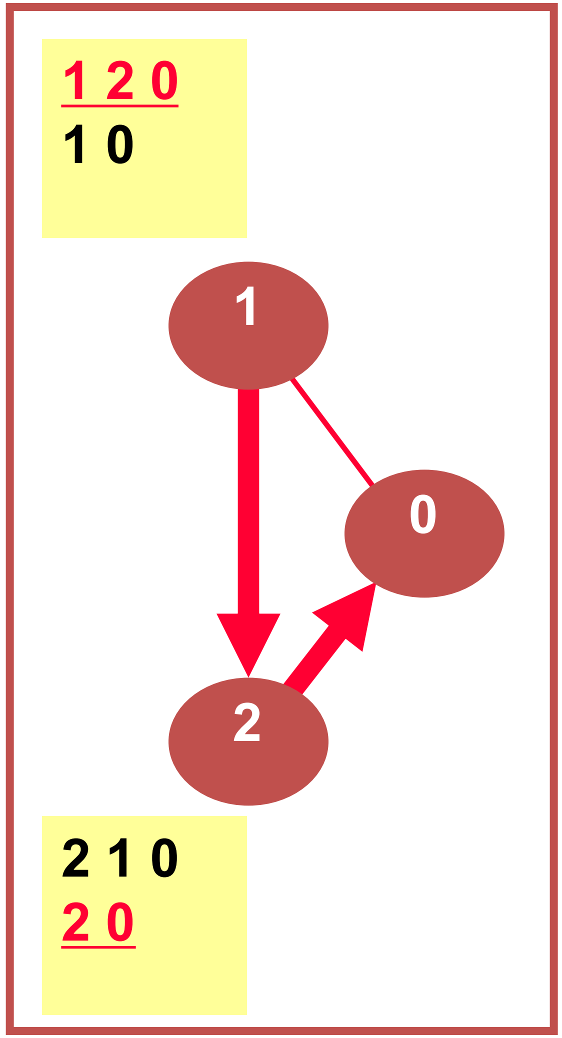
1. Does 5 have a path to 0?

2. Will 1 use the direct path to 0?

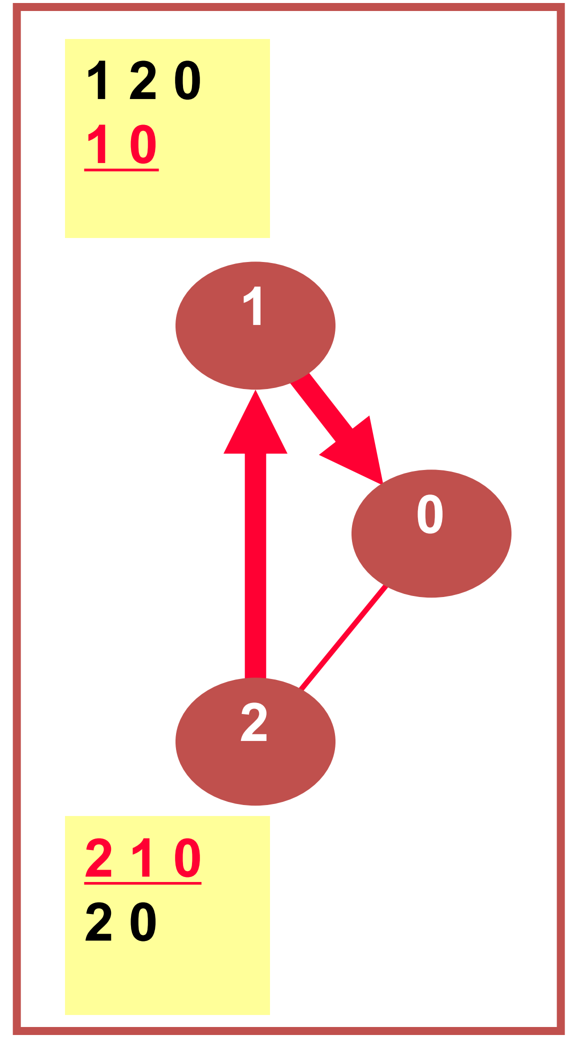


SPP May Have Multiple Solutions





First solution



Second solution

Avoiding BGP Instability

- Detecting conflicting policies
  - Computationally expensive
  - Requires too much cooperation
- Detecting oscillations
  - Observing the repetitive BGP routing messages
- Restricted routing policies and topologies
  - Policies based on business relationships

Today

1. Recap of BGP wide area routing

2. Integrating IGP and BGP

3. BGP (in)stability

4. <u>BGP Route monitoring</u>

Motivation for BGP Monitoring

- Visibility into external destinations
  - What neighboring ASes are telling your AS
  - How you are reaching external destinations
- Detecting routing anomalies
  - Increases in number of destination prefixes
  - Lost reachability or instability of some destinations
- Input to traffic-engineering tools
  - Knowing the current routes in the network

BGP Monitoring: A Wish List

- Ideally: know what the router knows
  - All externally-learned routes
  - Before applying policy and selecting best route
- How to achieve this
  - Special monitoring session on routers that tells everything they have learned
  - Packet monitoring on all links with BGP sessions
- If you can't do that, you could always do…
  - Periodic dumps of routing tables, or
  - BGP session to learn best route from router

BGP Table ("show ip bgp" at RouteViews)

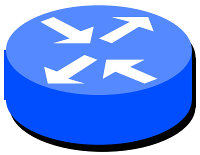
```
 Network             Next Hop        Metric  LocPrf  Weight    Path
*  3.0.0.0           205.215.45.50                            0    4006 701 80 i
*                    167.142.3.6                              0    5056 701 80 i
*                    157.22.9.7                               0    715 1 701 80 i
*                    195.219.96.239                           0    8297 6453 701 80 i
*                    195.211.29.254                           0    5409 6667 6427 3356 701 80 i
*>                   12.127.0.249                             0    7018 701 80 i
*                    213.200.87.254                  929      0    3257 701 80 i
* 9.184.112.0/20     205.215.45.50                            0    4006 6461 3786 i
*                    195.66.225.254                           0    5459 6461 3786 i
*>                   203.62.248.4                             0    1221 3786 i
*                    167.142.3.6                              0    5056 6461 6461 3786 i
*                    195.219.96.239                           0    8297 6461 3786 i
*                    195.211.29.254                           0    5409 6461 3786 i
```

AS 80 is General Electric, AS 701 is UUNET, AS 7018 is AT&T

AS 3786 is DACOM (Korea), AS 1221 is Telstra

Using Routers to Monitor BGP

Talk to operational routers using SNMP or telnet at command line

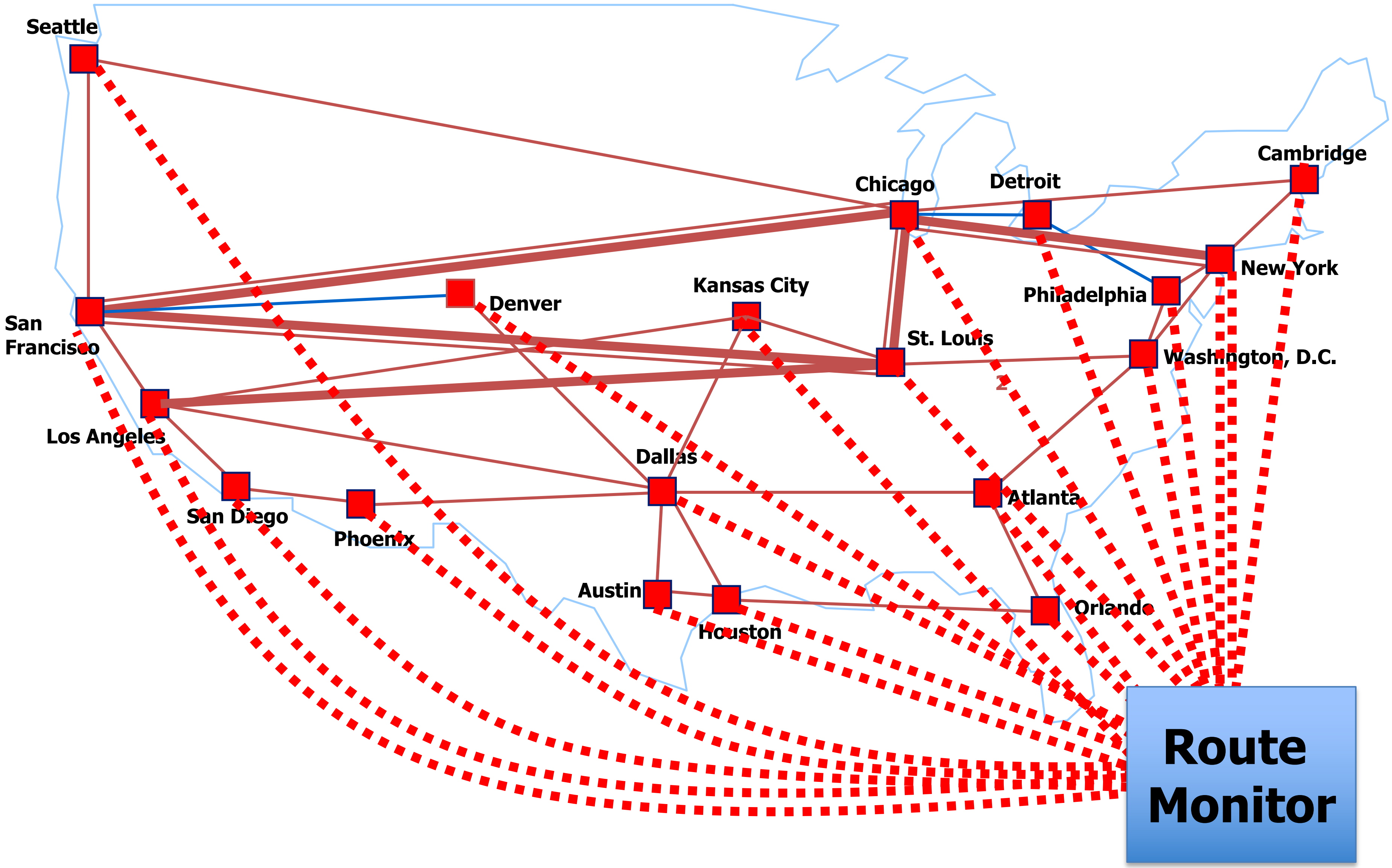


- (-) BGP table dumps are expensive
- (+) Table dumps show all alternate routes
- (-) Update dynamics lost
- (-) Restricted to interfaces provided by vendors

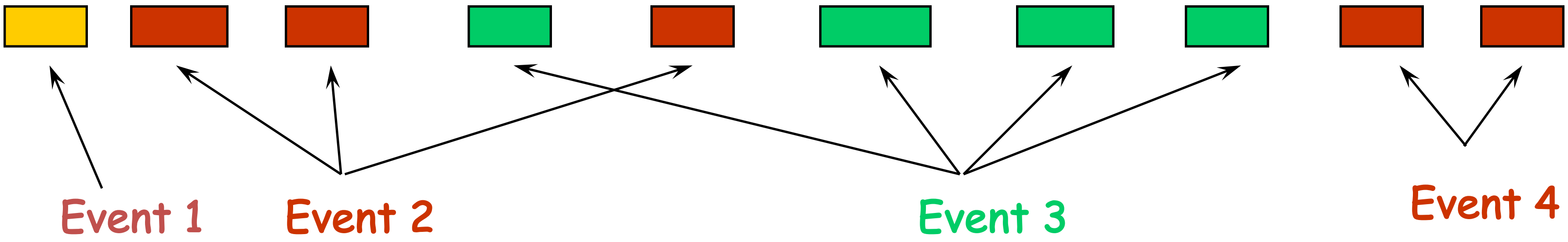
Establish a "passive" BGP session from a workstation running BGP software

- (+) BGP table dumps do not burden operational routers
- (-) Receives only best route from BGP neighbor
- (+) Update dynamics captured
- (+) Not restricted to interfaces provided by vendors

Collect BGP Data From Many Routers



BGP Events



- Group BGP updates that "belong together"
  - Same IP prefix, originating AS, or AS_PATH
- Updates that are "close" together in time
  - Maximum spacing between packets (e.g. 30 sec)
  - E.g.: events 2 and 4 are separated in time

Summary- Today

- Inter-domain routing chiefly concerned with policy, not optimality
- Behavior and configuration of BGP is complex and not fully understood
- Measurement is crucial to BGP network operations