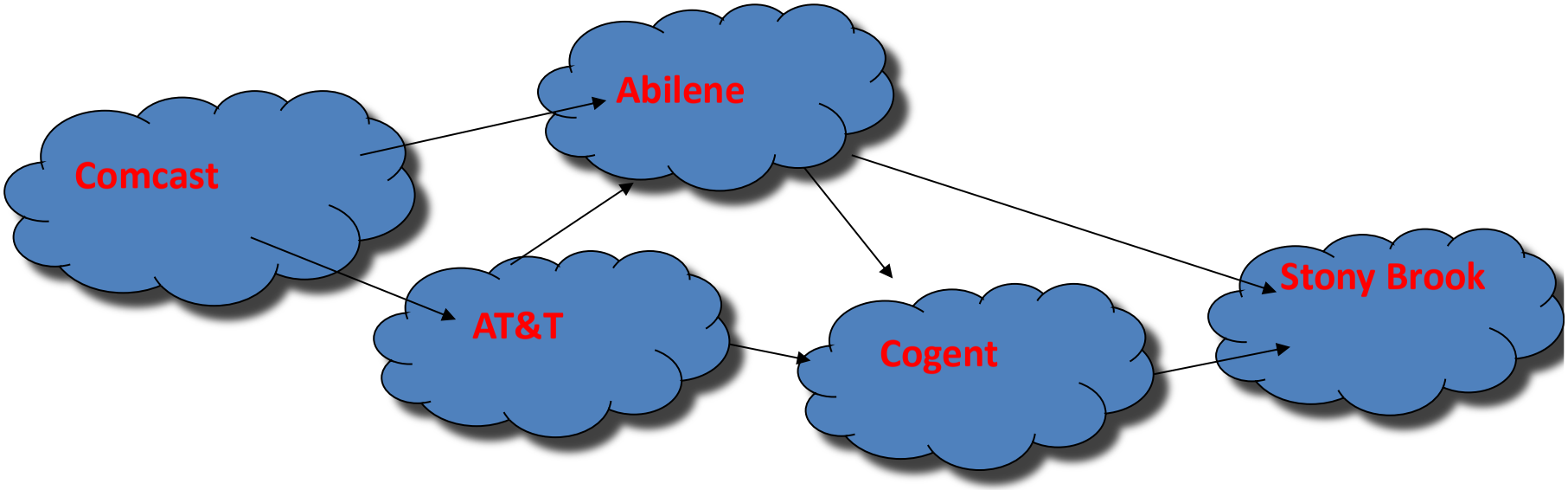
Inter-Domain Routing: Routing between ASes

More about ASes

- Independent networks with independent self interests.



Autonomous System (AS)

- Each AS identified by an ASN number
  - 16-bit values (latest protocol supports 32-bit ones)
- Currently, there are ~ 40000 ASNs
  - AT&T: 5074, 6341, 7018, …
  - Sprint: 1239, 1240, 6211, 6242, …
  - Stony Brook U: 5719
  - Google 15169, 36561 (formerly YT), + others
  - Facebook 32934
  - North America ASs → ftp://ftp.arin.net/info/asn.txt

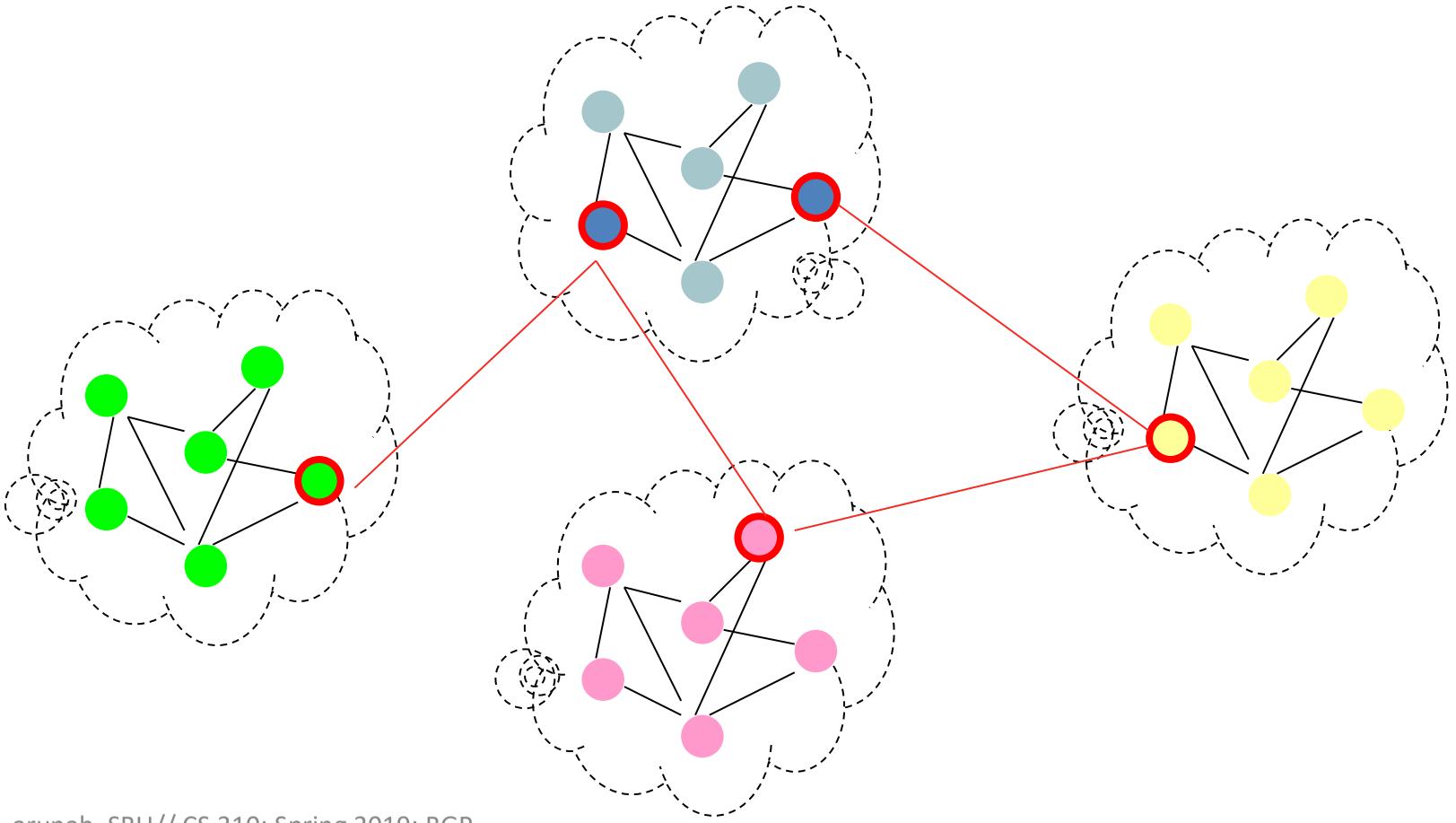
Inter-domain routing: Between ASes

- eBGP: External BGP routing
  - Runs BGP routing protocol
  - Send route advertisements to border routers of adjacent ASes
  - Border routers are in the same subnet as their adjacent AS border routers
- iBGP: Routes from an internal router in one AS to another AS
  - Uses routes created by eBGP routers to get a different AS
  - iBGP advertises external routes to the internal routers
  - Uses an intra-AS protocol to get the border (eBGP) router, and then BGP routes to get to the destination AS

IGP vs iBGP vs eBGP

- IGP (also intra-AS protocols): Routes inside an AS to internal destinations
- iBGP: Routes inside an AS to external destinations
- eBGP: Routes outside an AS

Intra-domain and Inter-domain routing



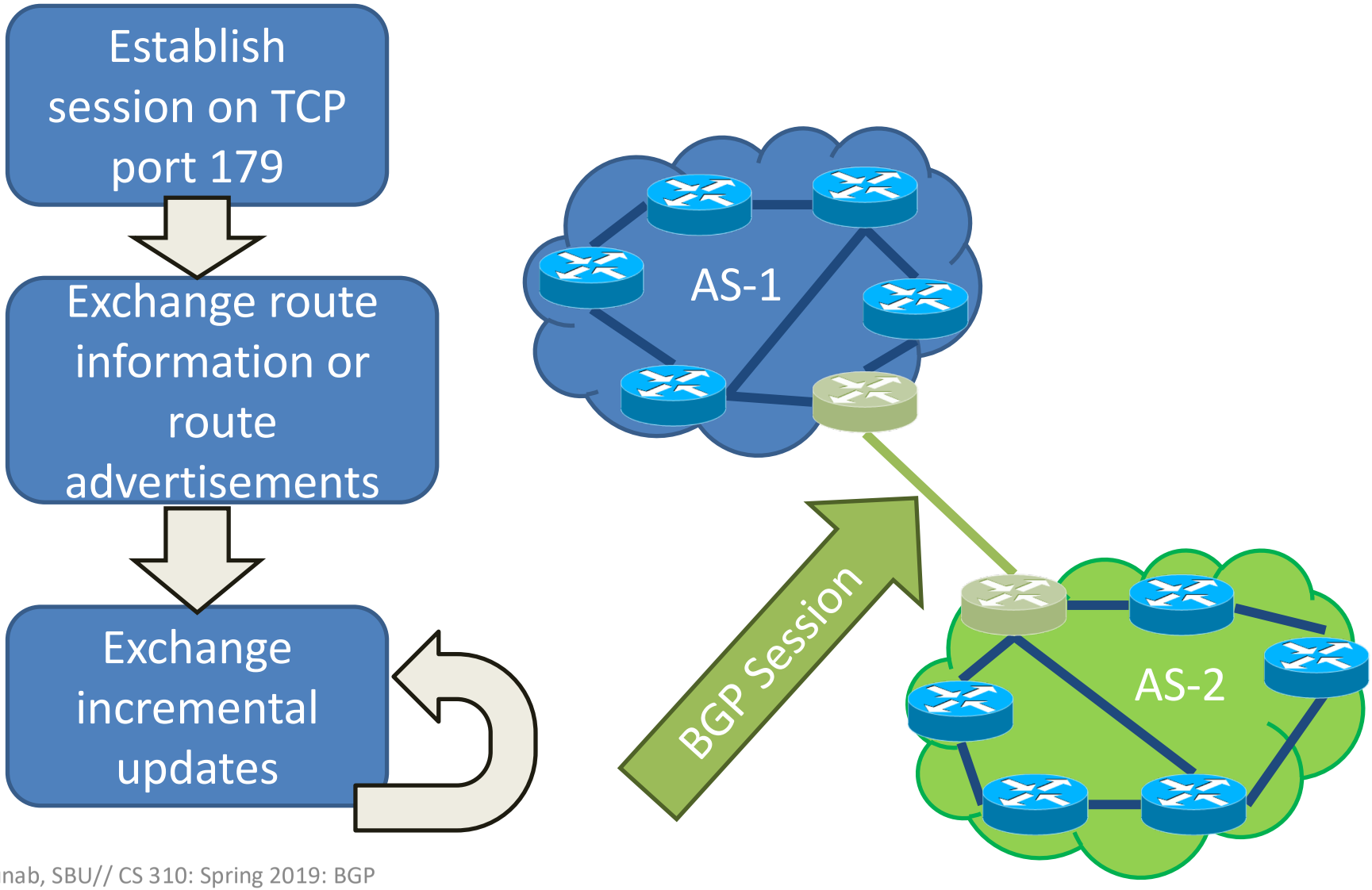
Inter-Domain Routing

- Global connectivity is at stake!
  - Thus, all ASs must use the same protocol
  - Contrast with intra-domain routing
- BGP is a path vector protocol
  - Similar to distance vector but when the entire path is included in the advertisement

Inter-domain routing (continued)

- Goal:
  - Provide reachability for different ASes
  - Comply to polices of different ASes
- BGP: path vector based on ASes

BGP Operations

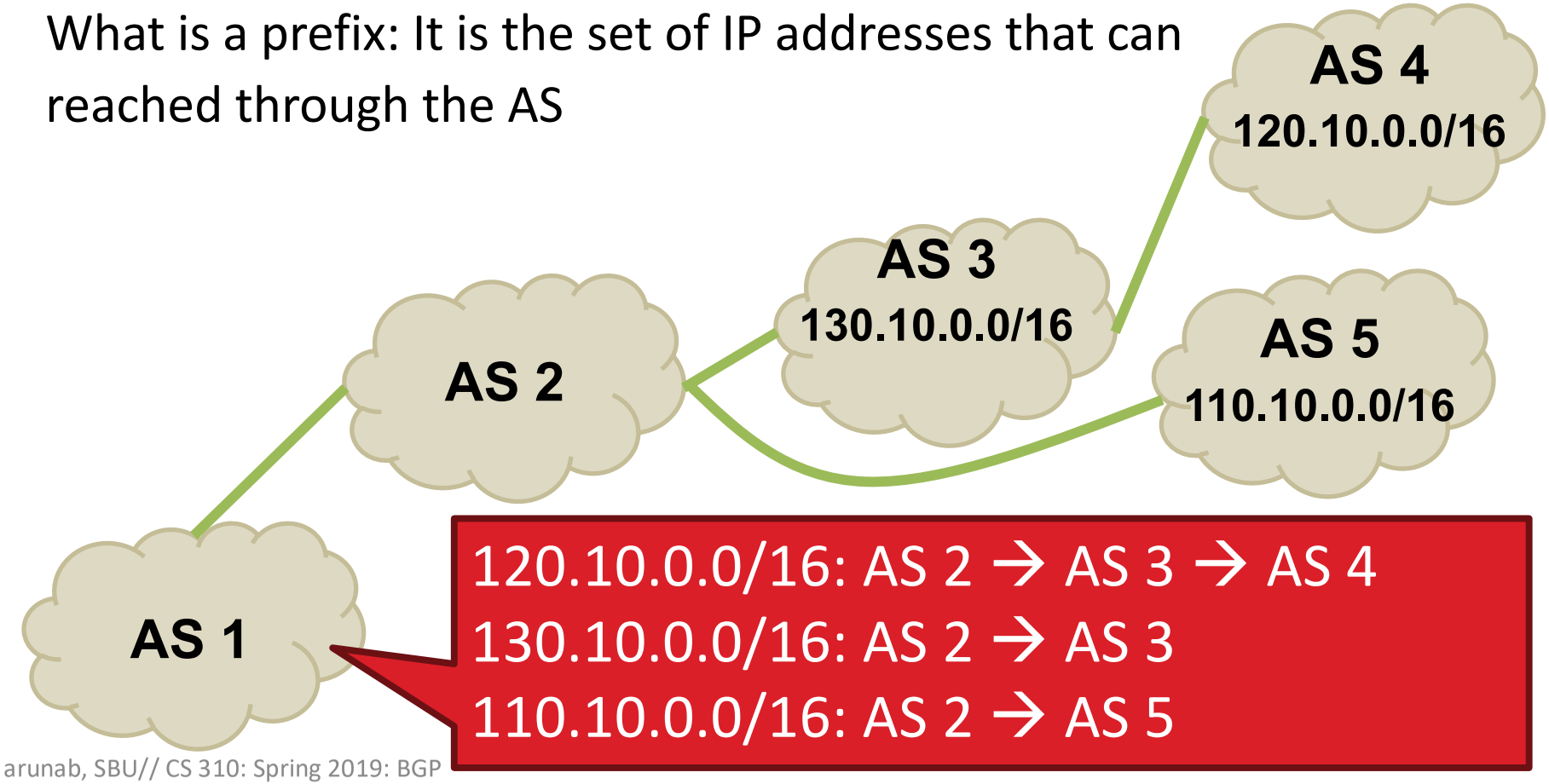


Route advertisements

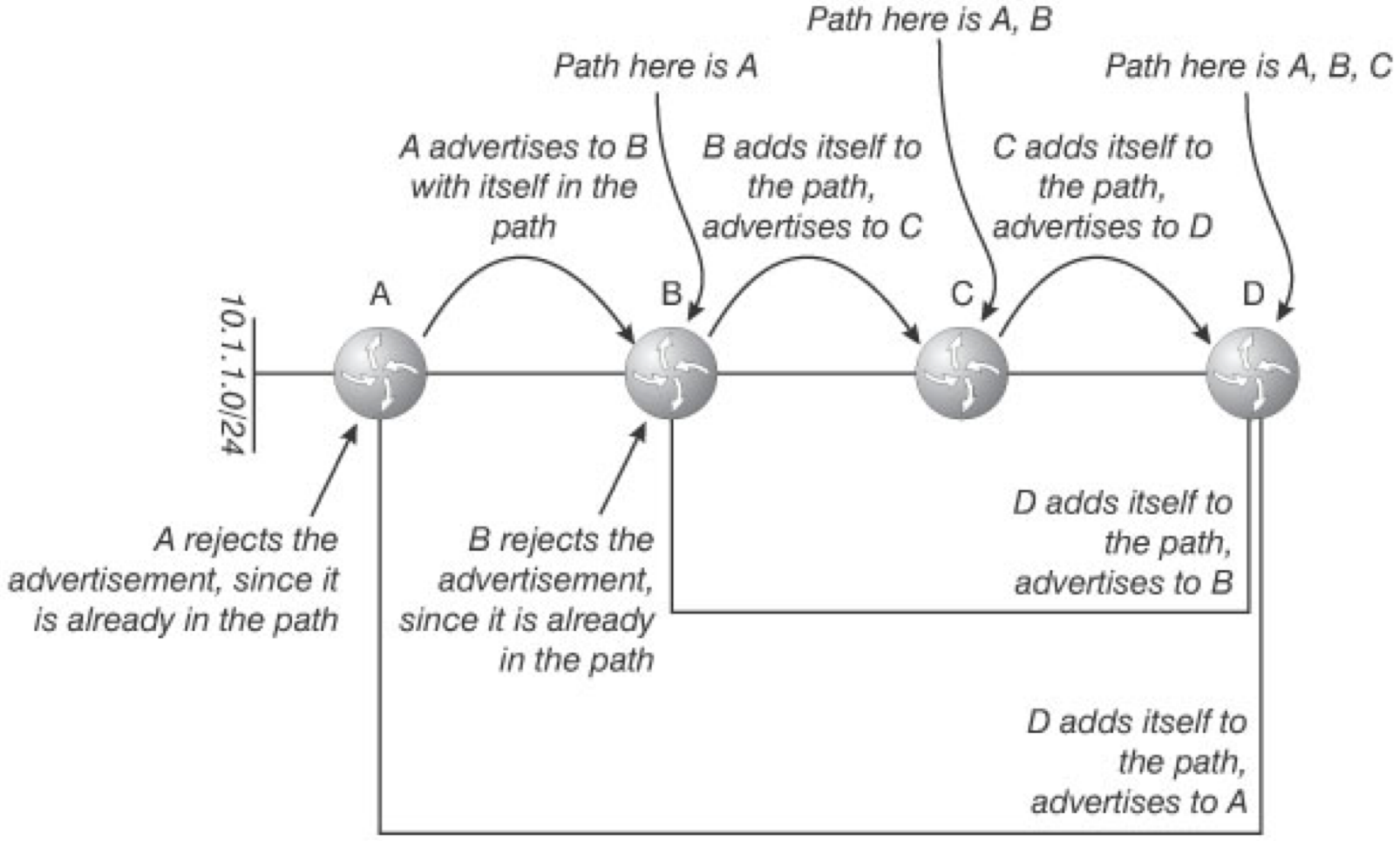
- AS-path: sequence of ASs a route traverses
- Origin AS: The AS that originates the advertisement

for a prefix

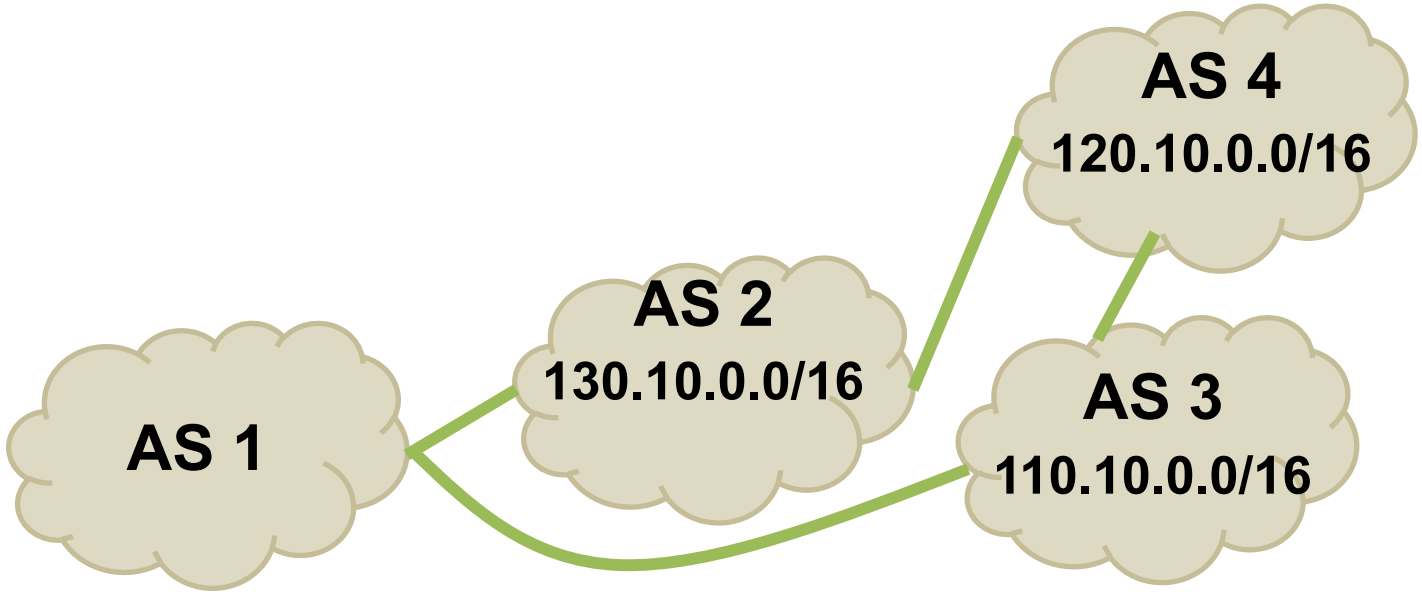
What is a prefix: It is the set of IP addresses that can reached through the AS



Why is the the entire path to destination included?



After route advertisement: Route selection



Route 1 **Dest:120.10.0.0/16, Next:130.10.0.0/16, AS2-AS4**
Route 2 **Dest:120.10.0.0/16, Next:110.10.0.0/16, AS3-AS4**
Which route to choose?

Route Selection no longer shortest path

- Prefer a route with **highest local preference** value
- Among routes with equal local preference
  - Pick the path that has smaller hops
- For multiple paths with different gateway routers in the same AS
  - Use **hot potato** routing
- Arbitrary Tie breaker

Route selection based on attributes

Local Pref

- Used to prefer customer > peer > provider
- high values are better

ASPATH

- Prefer paths with lowest # of ASes

MED

- Tell others to choose one exit point over another
- low values are better

IGP path cost

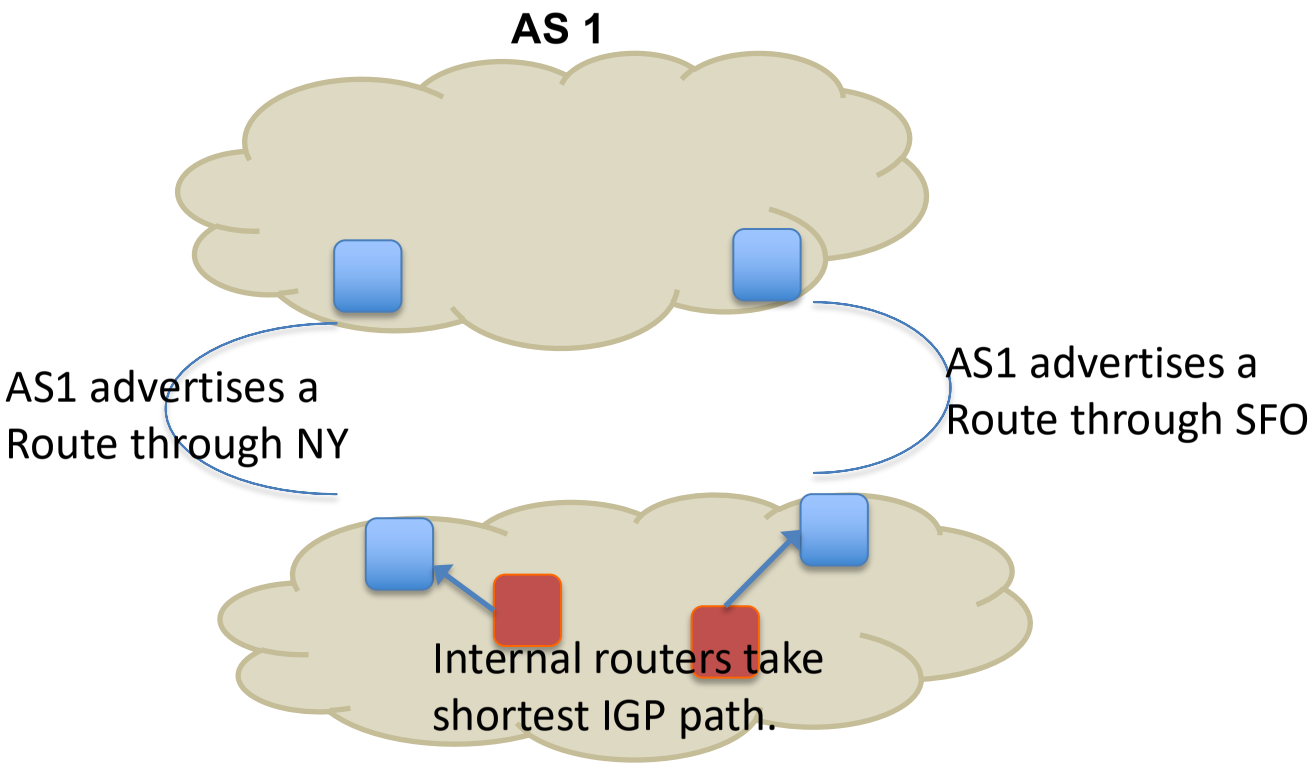
- Lower values are better
- leads to “hot potato” routing

Router ID

Local preference value

- Network operator in the local AS assigns to a particular route (based on policy)
- Operator sets a local preference value

What is Hot Potato Routing?



Hot Potato: An AS tries to get packets out of its network as soon as possible.

Beyond path selection, BGP routes are determined by policy

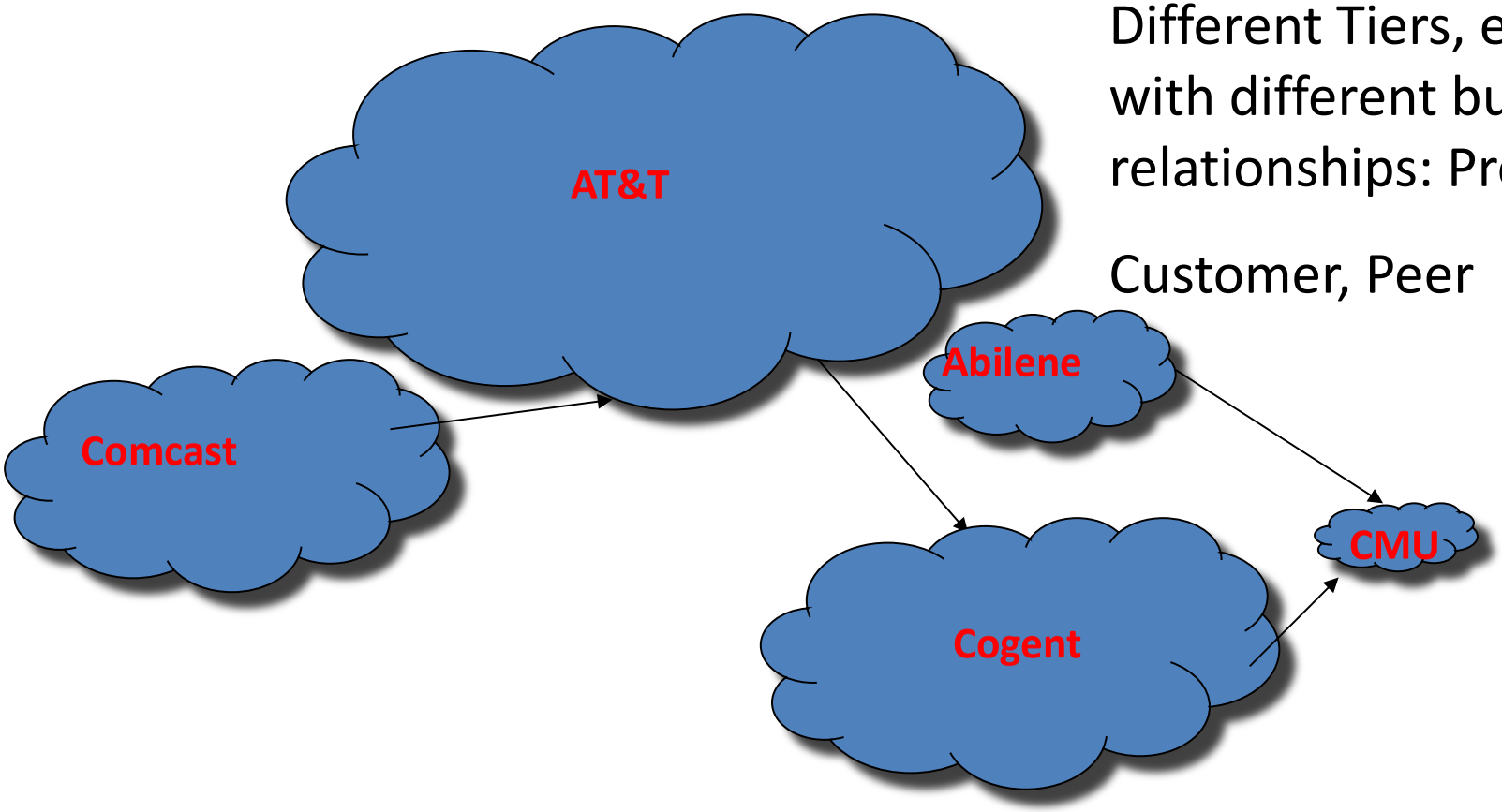
- We saw how an AS chooses a route, but a policy determines how routes are advertised
- What is a policy?
  - Each AS gets a set of route advertisements, but filters them when sending to out based on AS policy
  - These policies are determined by business relationships.
- Policies are the key difference between BGP and other routing protocols

The Internet: Zooming In 2x

All ASes are not equal

Different Tiers, each with different business relationships: Provider,

Customer, Peer



AS business relationships: customer, provider, and peer

- Customer AS pays the provider AS to send their packets through
  - This is usually called paying for Transit services
  - A Provider AS gives transit to its customers
  - The customer pays the provider no matter the direction of the data flows.
- Who is the customer and who is the provider?
  - Usually, the one who can “live without” the other.
- What if both need each other? Peering.
  - When two ASes peer with each other, they can exchange traffic for free
  - Peering arrangements are typically confidential.

More on AS relationships

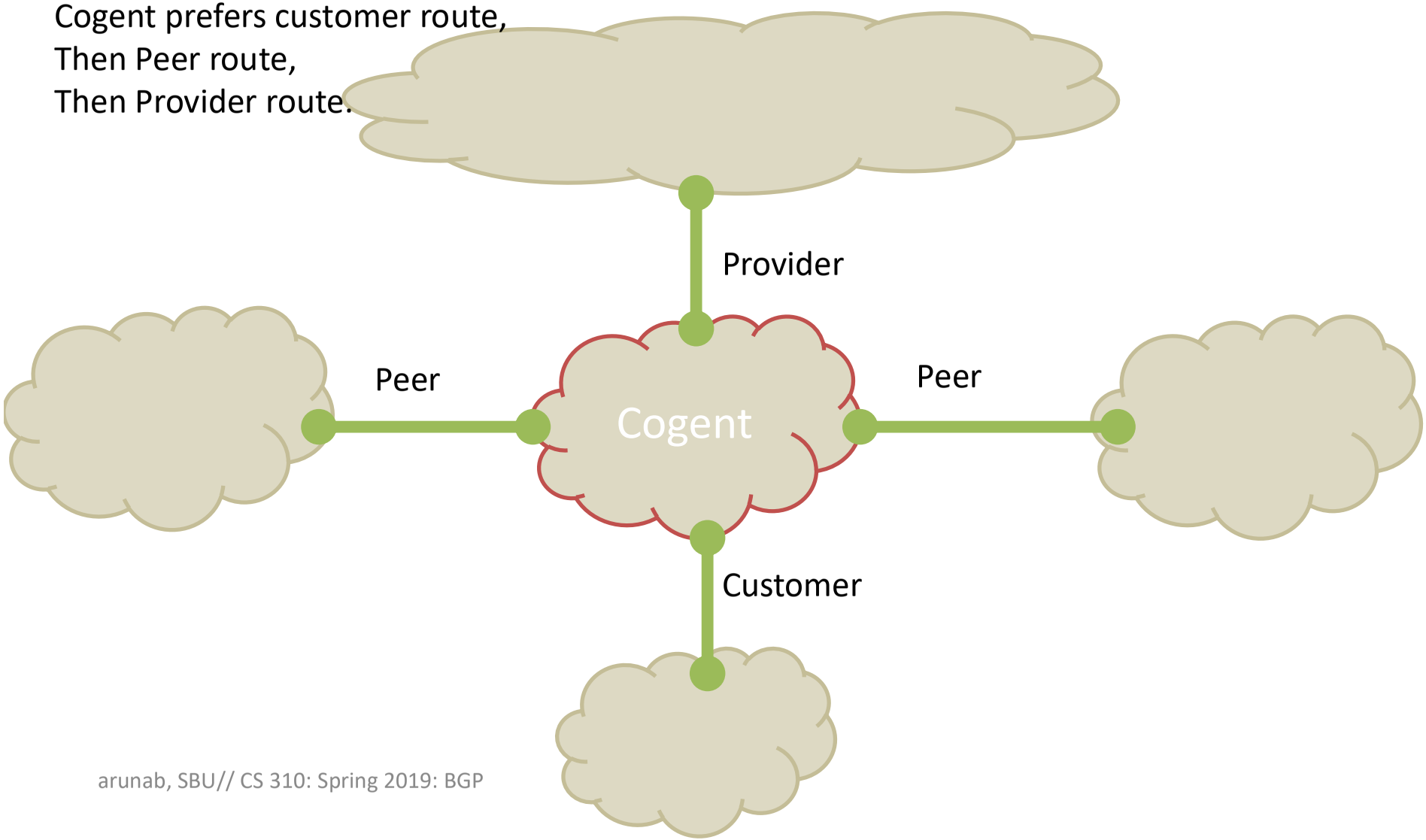
- **Transit/Provider AS**: "You pay me to carry your traffic" (provider-customer)
- **Peering AS:** "For free, I carry your packets to my customers only." (peer-peer)
- **Customer AS:** "I pay you to carry my traffic (customer-provider)

Tiers

- ASes are also categorized into different tiers
  - Technical definition of tier-1 ISP: They have no providers, only peers.
  - All Tier 1 ASes peer with each other by definition
- Other Tier arrangements are for convenience
  - A Tier 3 ISP can be a customer of a Tier 1 ISP
  - Typically determined based on size
- ASes have pairwise relationships with each other
  - A AS can be a provider, a customer, and a peer to different ASes

Routing Policies (Customer > Peer > Provider)

Cogent prefers customer route,
Then Peer route,
Then Provider route.



Filter and export decisions

- Given that an AS learns a route from its neighbor, to whom should I advertise this decision to?
  - Routes learnt from a customer gets advertised to everybody
  - Routes learnt from a provider only will be advertised to its customers
    - Otherwise, the AS will provide transit services to two of its providers and pay for both
  - Routes learnt from a peer only will be advertised to its customers

BGP in one page

- Border Gateway 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)
- BPG security is still a million dollar question

Modeling BGP

- Difficult to determine because AS relationships are not known
  - Gao-Rexford model helps in reverse engineering AS relationships
- What is BGP optimizing?
  - The Stable Path Problem (SPP)
  - Modeling BGP as a SPP also helps explain the reason for BGP oscillations.

Some IP layer protocols for management

ICMP: internet control message protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Two useful applications using ICMP

Ping

- Used to find the RTT to a server
- Source sends a ICMP echo message
- Ping inserts the send time
- If the host is alive it sends a echo-reply message---you can use this to estimate RTT
- Ping then sends another echo message and on and on

Traceroute

- use to find the route to a destination

Traceroute and ICMP

- source sends series of UDP segments to destination
  - first set has TTL =1
  - second set has TTL=2, etc.
  - unlikely port number
- when datagram in *n*th set arrives to nth router:
  - router discards datagram and sends source ICMP message
  - ICMP message include name of router & IP address

- when ICMP message arrives, source records RTTs

stopping criteria:

- UDP segment eventually arrives at destination host
- destination returns ICMP "port unreachable" message
- source stops