Routing: BGP
Routing Policies

Prof. Anja Feldmann, Ph.D.

Balakrishnan Chandrasekaran, Ph.D.

Inter-AS routing: BGP

The *de facto* standard: ***Border Gateway Protocol (BGP)***

BGP provides each ***AS*** a means to:

- Obtain subnet reachability information from neighboring ASs
- Propagate reachability information to all routers in the AS
- Determine *"good"* routes to subnets based on reachability information and routing policy.

Allows a subnet to advertise its existence to rest of the Internet: *"I am here"*

Issues:

- Which routing algorithm?
- How are routes advertised?
- How to implement routing policies?

BGP: A path-vector protocol

Distance vector algorithm with extra information

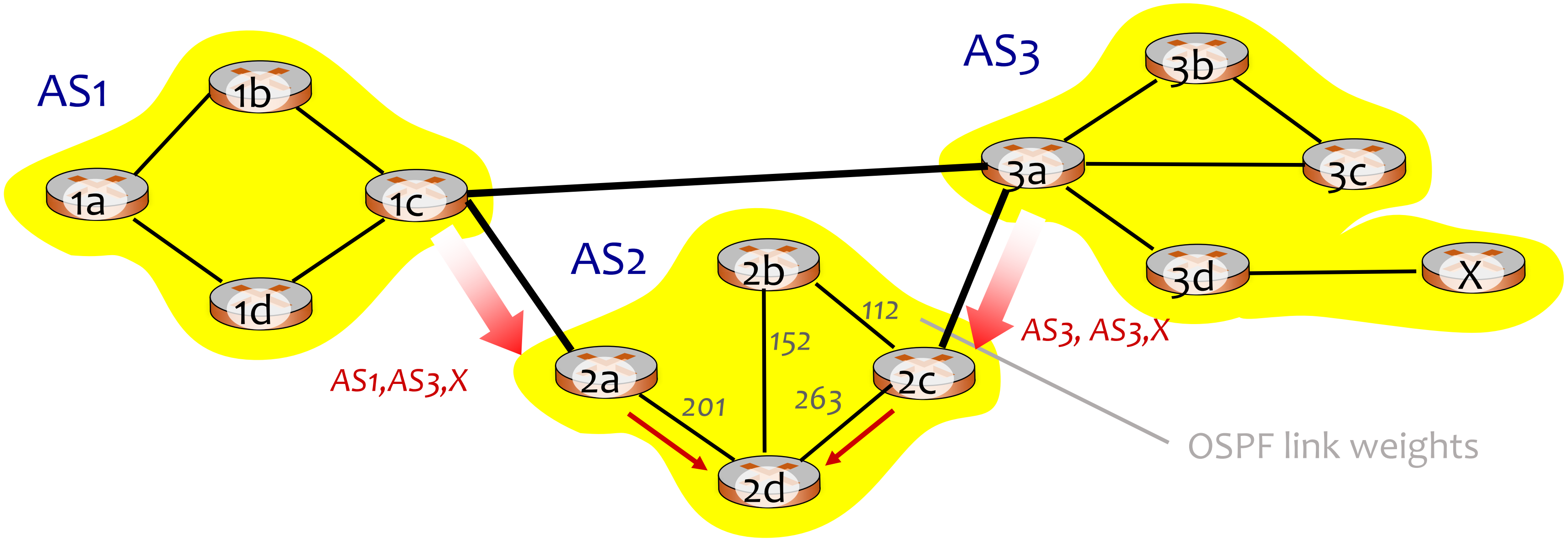
- When advertising a prefix, advert includes BGP attributes
  - *Prefix + other attributes = "route"*

- When gateway router receives route advertisement, uses *ingress filters* to accept/decline
  - Can make decision based on ASes on path, e.g., to avoid loops

BGP route selection

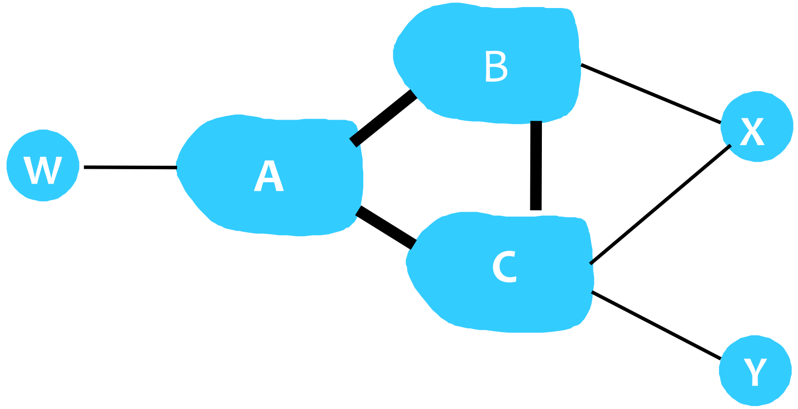
- Router learn *more than 1 route* to some prefix
- Router **must** select *best route*
- ***Elimination rules***:
    - Local preference value attribute: Policy decision
    - Shortest AS-PATH
    - Best MED (multi-exit-discriminator)
    - Closest NEXT-HOP router: Hot potato routing
    - Additional criteria
    - IP address of peer

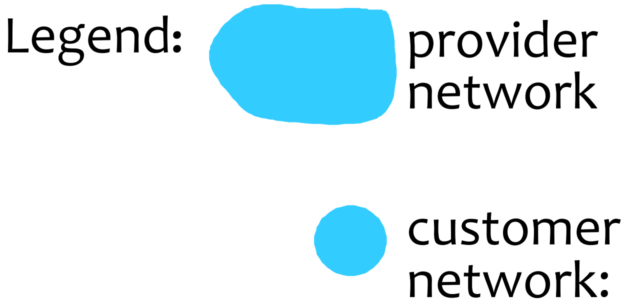
Hot Potato Routing



- 2d learns (via iBGP) it can route to X via 2a or 2c
- *Hot potato routing:* Choose local gateway that has least intra-domain cost (e.g., 2d chooses 2a, even though more AS hops to *X*): don't worry about inter-domain cost!

BGP: Achieving policy via advertisements

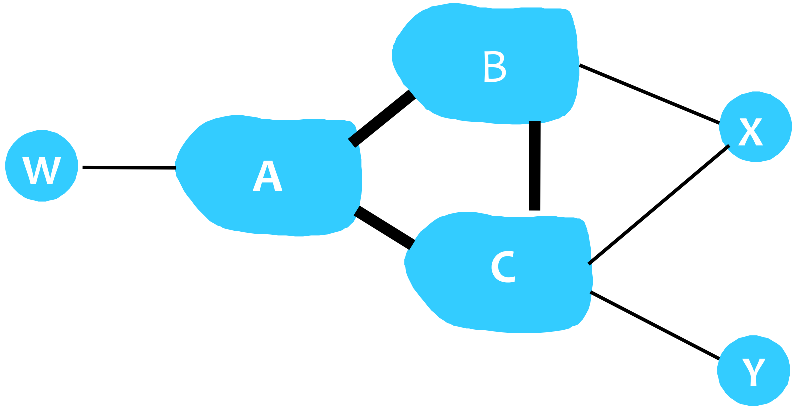




Suppose an ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs)

- A advertises path Aw to B and to C
- B *chooses not to advertise* BAw to C:
  - B gets no "revenue" for routing CBAw, since none of C, A, w are B's customers
  - C does not learn about CBAw path
- C will route CAw (not using B) to get to w

BGP: Achieving policy via advertisements

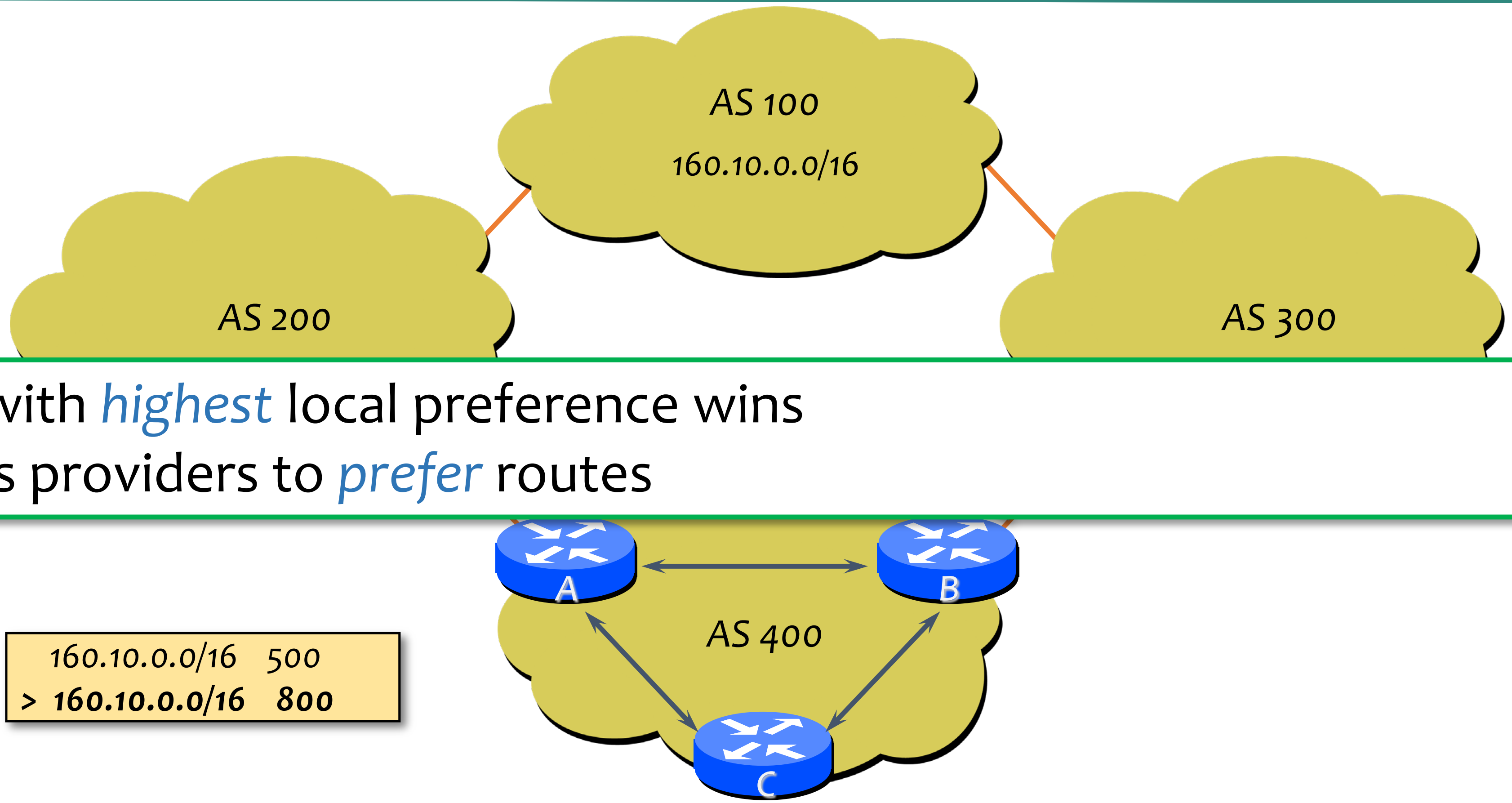


Legend: provider network; customer network:

Suppose an ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs)

- A,B,C are *provider networks*
- X,W,Y are customer (of provider networks)
- X is *dual-homed:* attached to two networks
- *Policy to enforce:* X does not want to route from B to C via X
  - .. so X will not advertise to B a route to C

BGP: Local preference



```
  160.10.0.0/16   500
> 160.10.0.0/16   800
```

- Path with *highest* local preference wins
- Allows providers to *prefer* routes

Local Preference – common uses

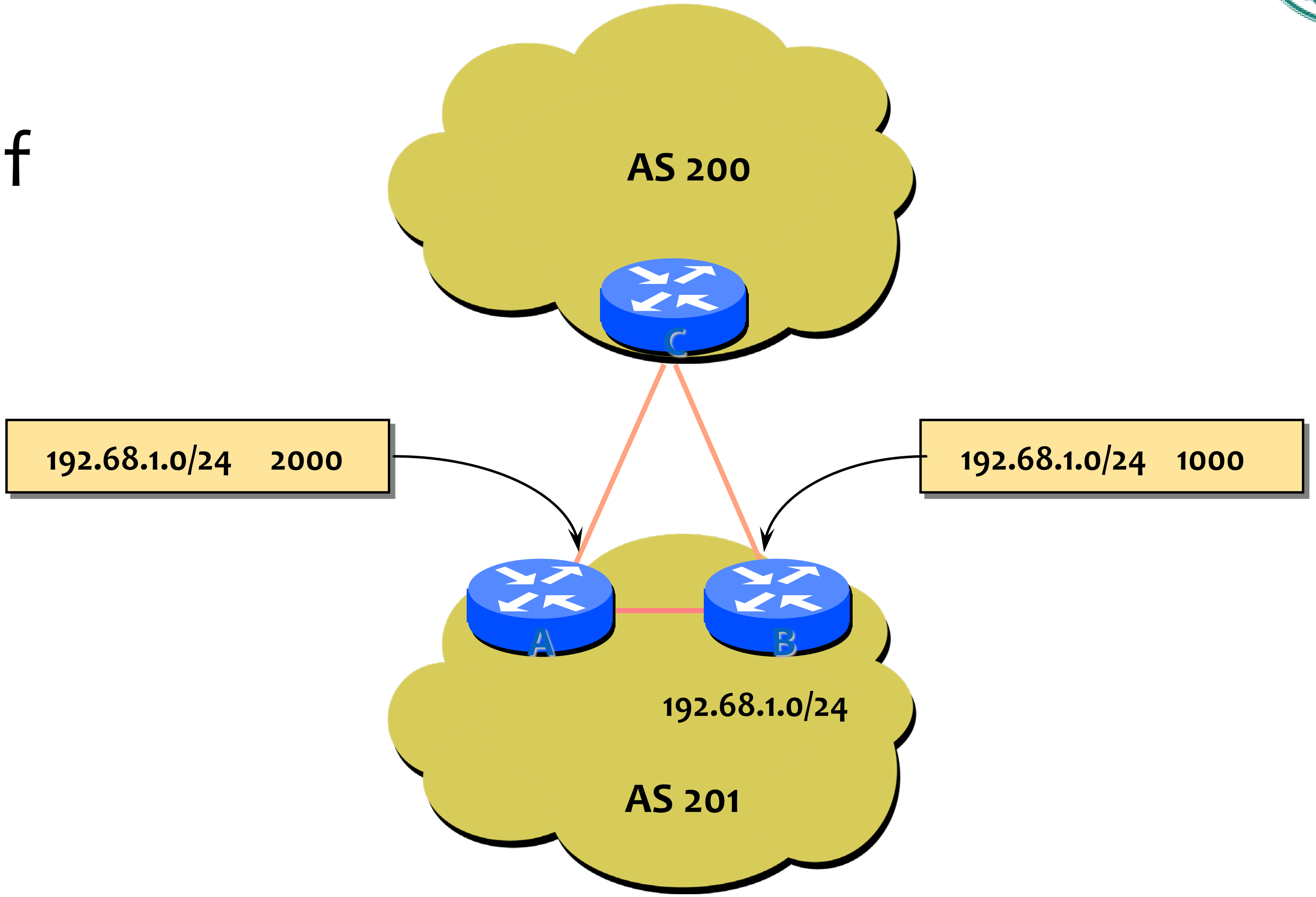
- Handle traffic directed to multi-homed transit customers
  - Allows providers to prefer a route
- Peering vs. transit
  - Prefer to use peering connection
  - Customer > peer > provider

Multi-Exit Discriminator (MED)

- Non-transitive
- Used to convey the relative preference of entry points
- Influences best path selection
- Comparable if paths are from same AS
- IGP metric can be conveyed as MED

BGP: MED attribute

- Used to convey the relative preference of entry points
- Comparable if paths are from same AS
- IGP metric can be conveyed as MED



Communities

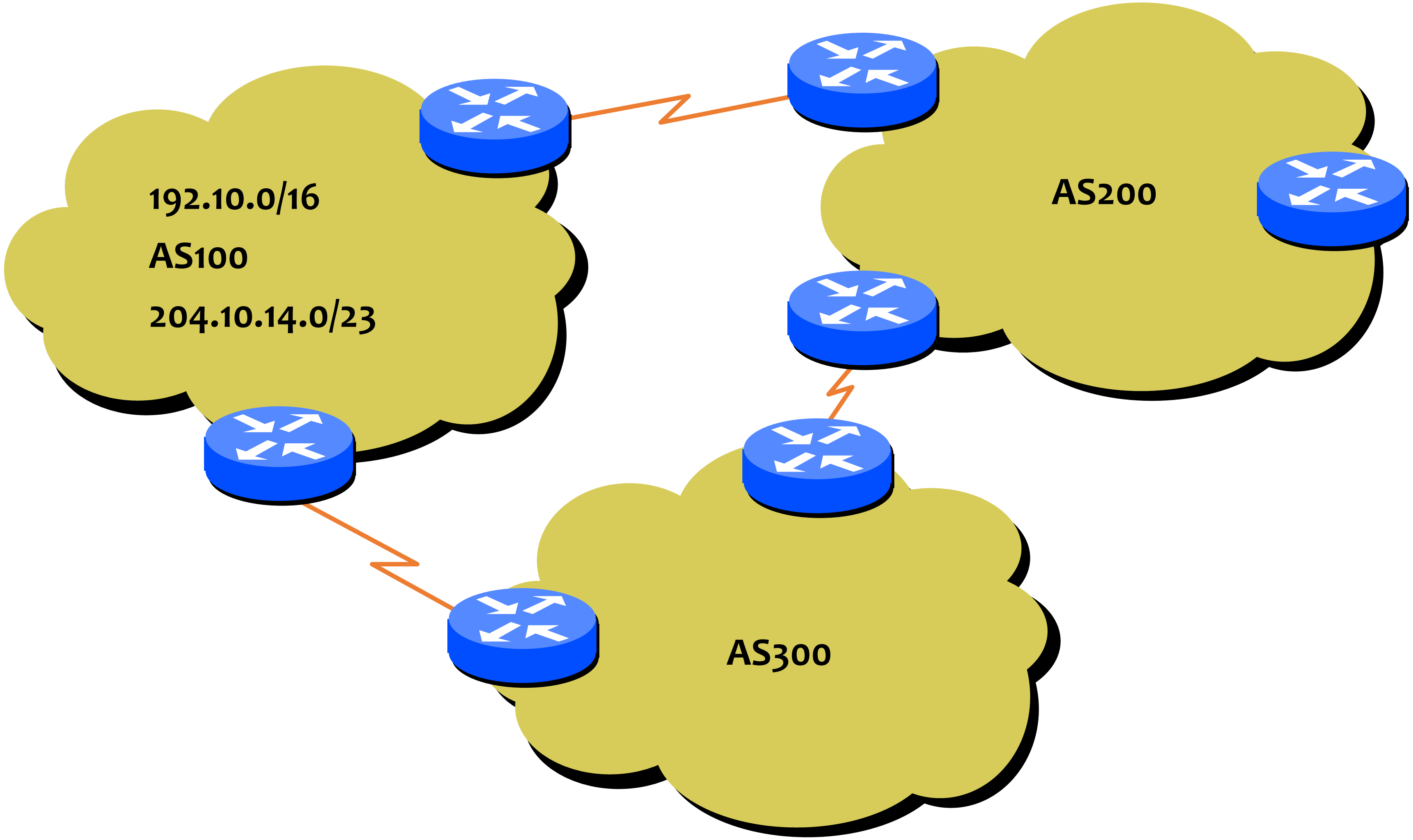
- Used to group prefixes and influence routing decisions (accept, prefer, redistribute, etc.), e.g., via route-maps to realize routing policies
- Represented as an integer Range: 0 to 4,294,901,760
- Each destination can have multiple communities
- Community attribute carried across AS's
- RFC1997, RFC1998

Load balancing

- BGP does not load-balance traffic; it chooses & installs a “best” route.

> “Since BGP picks a ‘best’ route based upon most specific prefix and shortest AS_PATH, it becomes non-trivial to figure out how to manually direct specific portions of internal traffic (prefixes) in a distributed fashion across multiple external gateways.”

Difficulties in load balancing



Multi-homing

Multi-homing:

  - Network has several connections to the Internet.

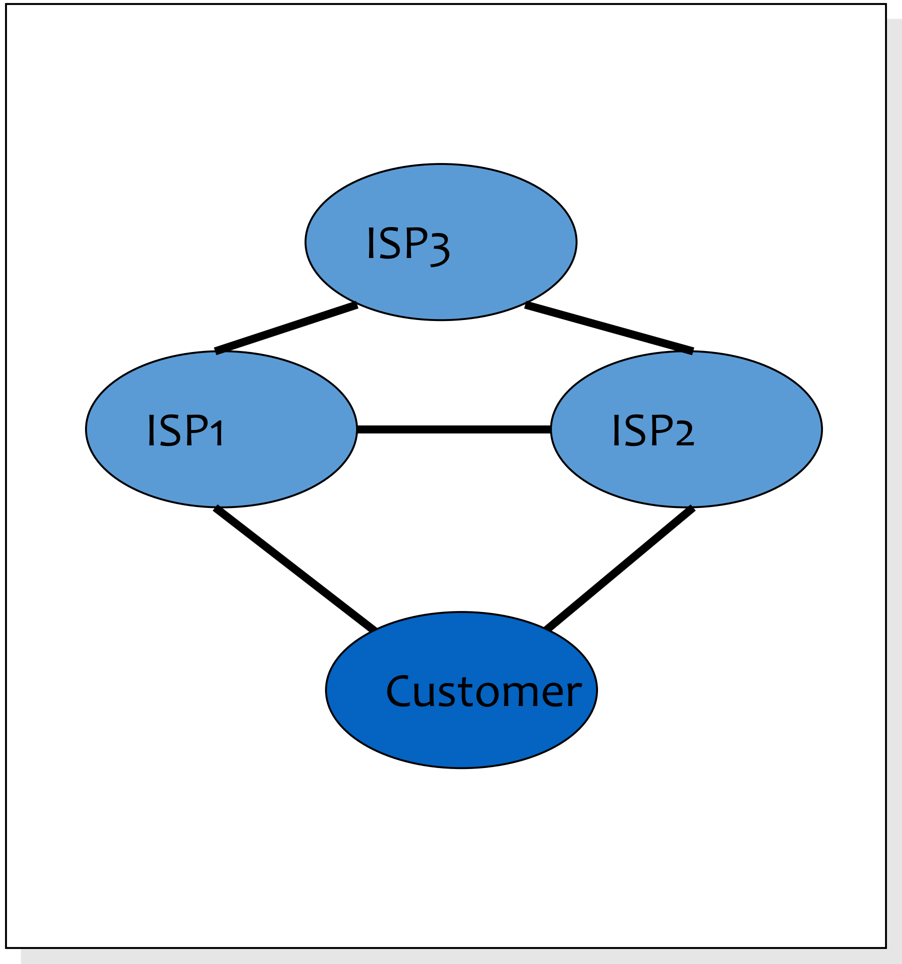
- Improves reliability and performance:
  - Can accommodate link failure
  - Bandwidth is sum of links to Internet
- Challenges
  - Getting policy right (MED, etc..)
  - Addressing

Multi-homing with multiple providers

Major issues:

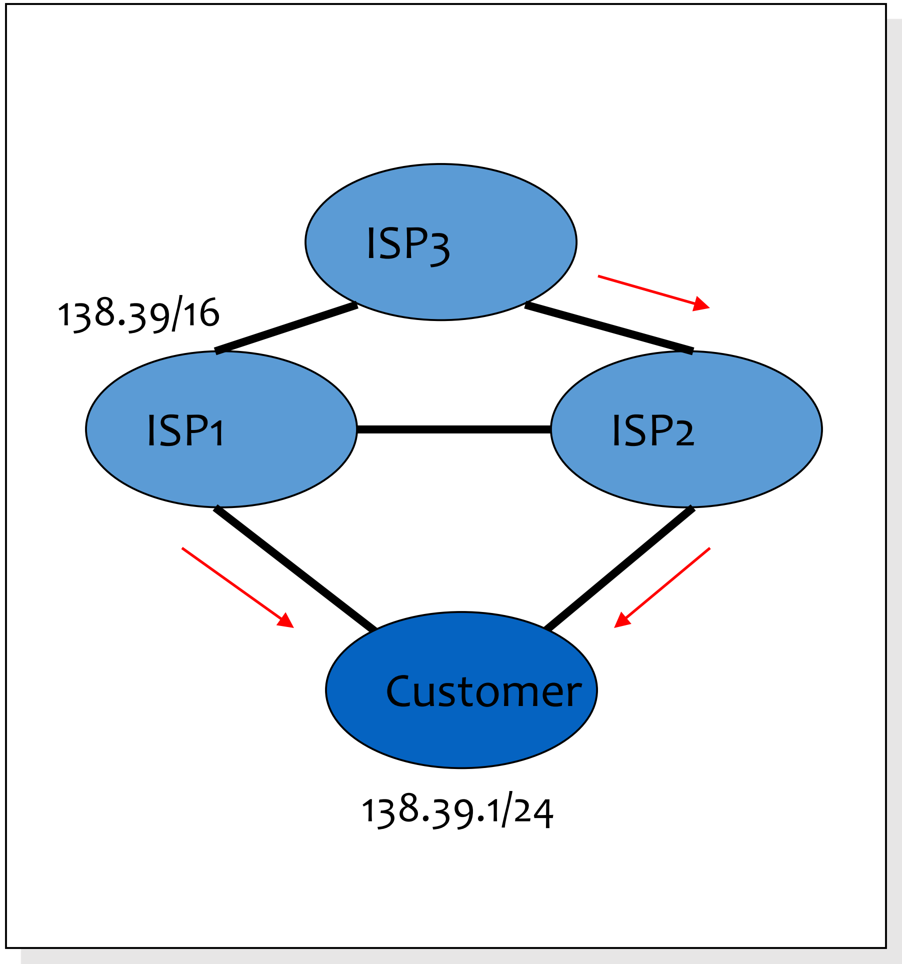
- Addressing
- Aggregation

- Customer address space:
  - Delegated by ISP1
  - Delegated by ISP2
  - Delegated by ISP1 and ISP2
  - Obtained independently



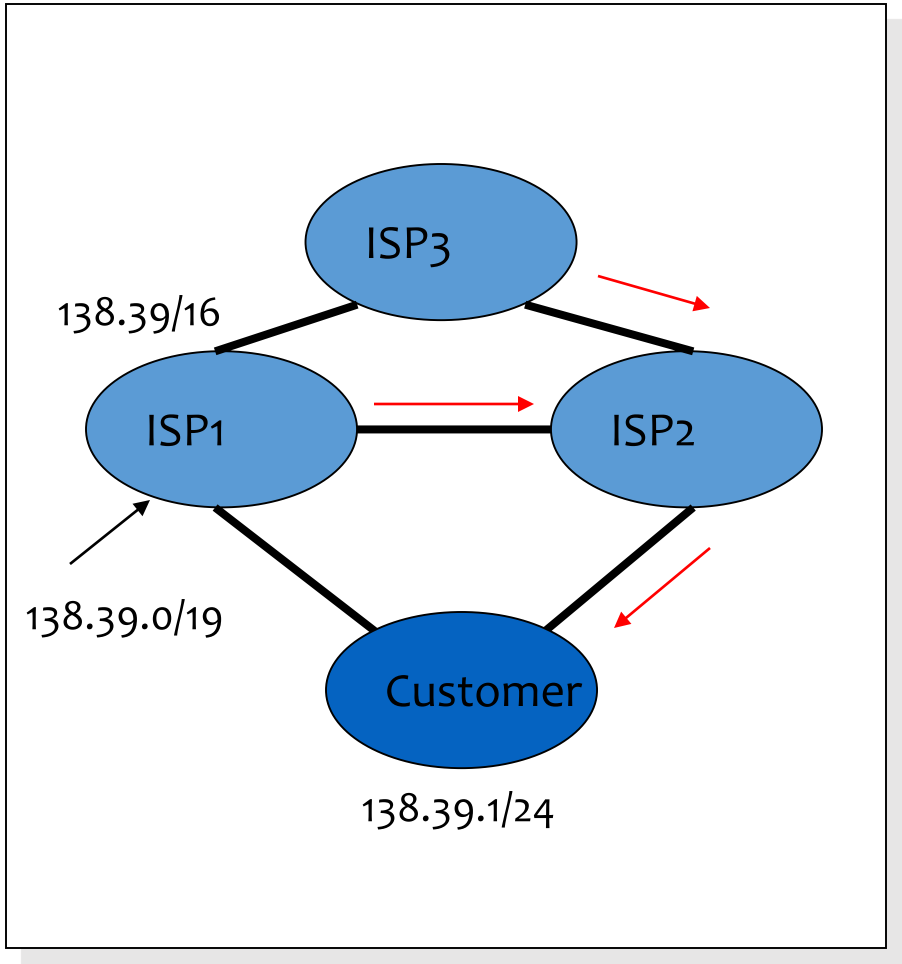
Multi-Homing: Address space from one ISP

- Customer uses address space from ISP1
- ISP1 advertises /16 aggregate
- Customer advertises /24 route to ISP2
- ISP2 relays route to ISP1 and ISP3
- ISP2-3 use /24 route
- ISP1 routes directly
- Problems with traffic load?



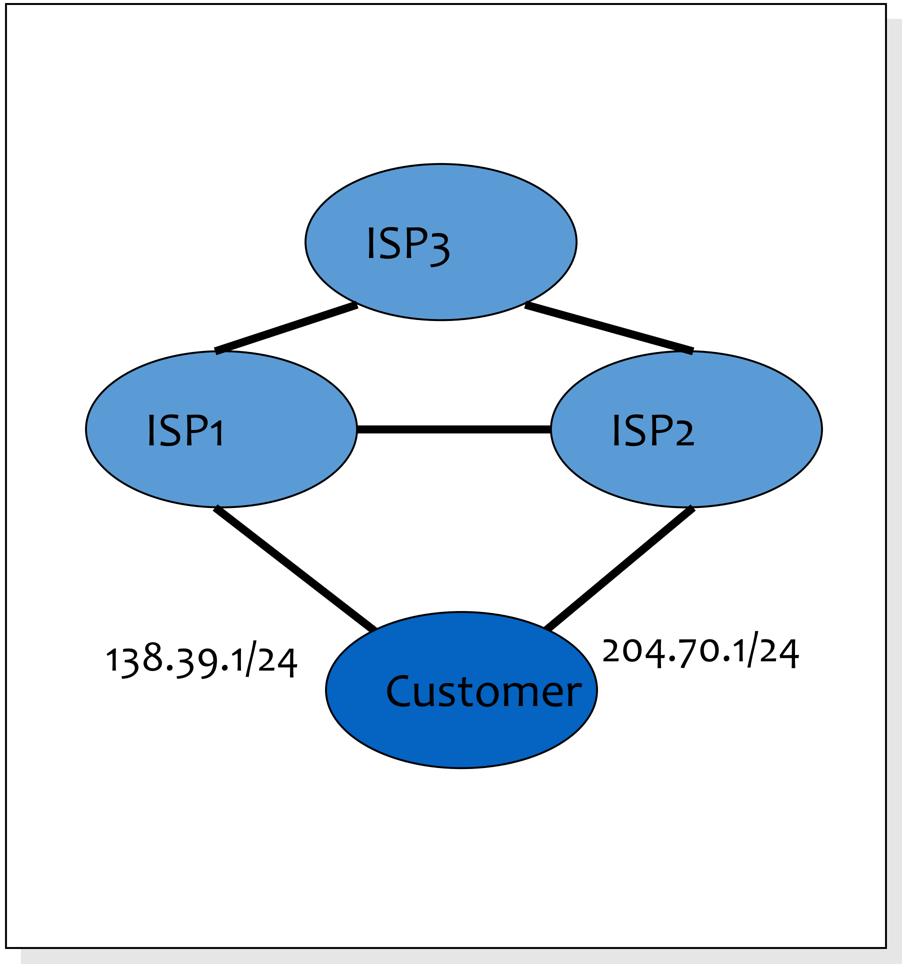
Multi-Homing: Pitfalls

- ISP1 aggregates to a /19 at border router to reduce internal tables.
- ISP1 still announces /16.
- ISP1 hears /24 from ISP2
- ISP1 routes packets for customer to ISP2!
- Workaround:
  ISP1 must inject /24 in I-BGP



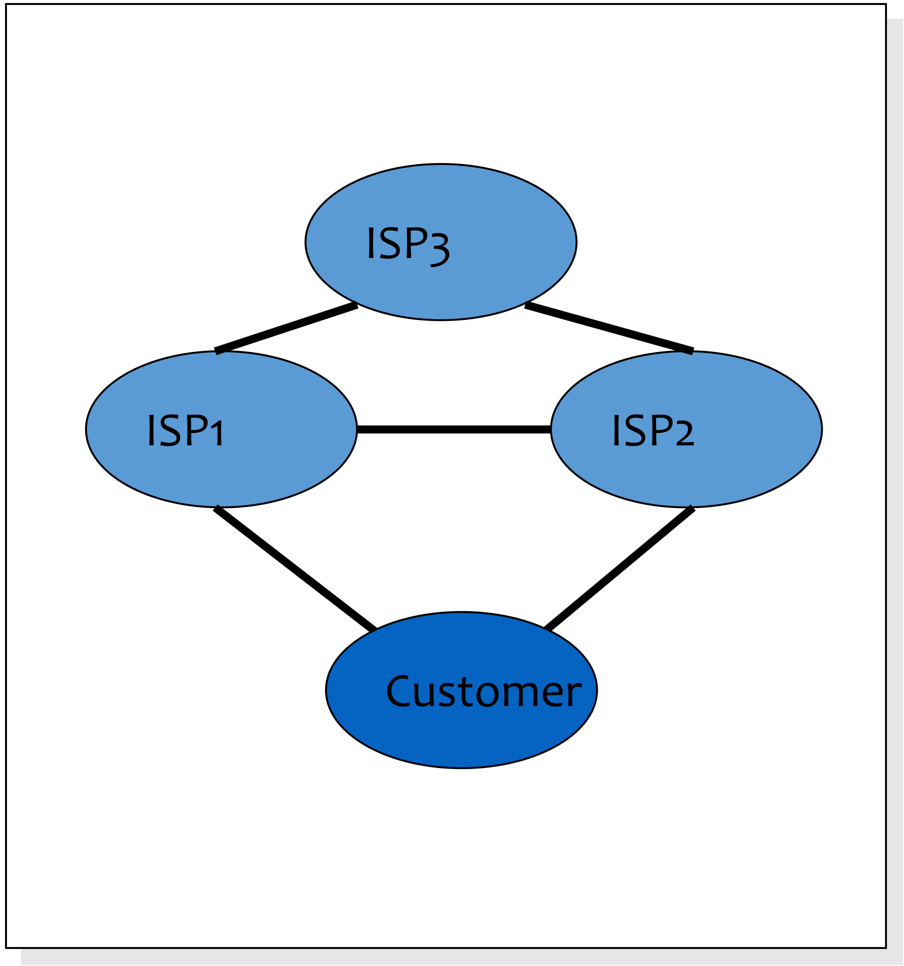
Multi-Homing: Address space from both ISPs

- ISP1 and ISP2 continue to announce aggregates
- Load sharing depends on traffic to two prefixes
- Lack of reliability: If ISP1 link goes down, part of customer becomes inaccessible.
- Customer may announce prefixes to both ISPs, but still problems with longest match as in case 1.



Multi-Homing: Independent address space

- Offers the most control, but at the cost of aggregation.
- Still need to control paths
- Many ISP's ignore advertisements of less than /19



BGP: A path-vector protocol

Distance vector algorithm with extra information

- When advertising a prefix, advert includes BGP attributes
  - *Prefix + other attributes = "route"*
- When gateway router receives route advertisement, uses *ingress filters* to accept/decline
- Before gateway router announces a route advertisement, uses *egress filters*