Interdomain Routing Policy

Reading: Sections 4.3.3 plus optional reading

Acknowledgments: Lecture slides are from Computer networks course thought by Jennifer Rexford at Princeton University. When slides are obtained from other sources, a reference will be noted on the bottom of that slide and full reference details on the last slide.

Goals of Today's Lecture

- Business relationships between ASes
  - Customer-provider: customer pays provider
  - Peer-peer: typically settlement-free
- Realizing routing policies
  - Import and export filtering
  - Assigning preferences to routes
- Multiple routers within an AS
  - Disseminated BGP information within the AS
  - Combining with intradomain routing information

Business Relationships

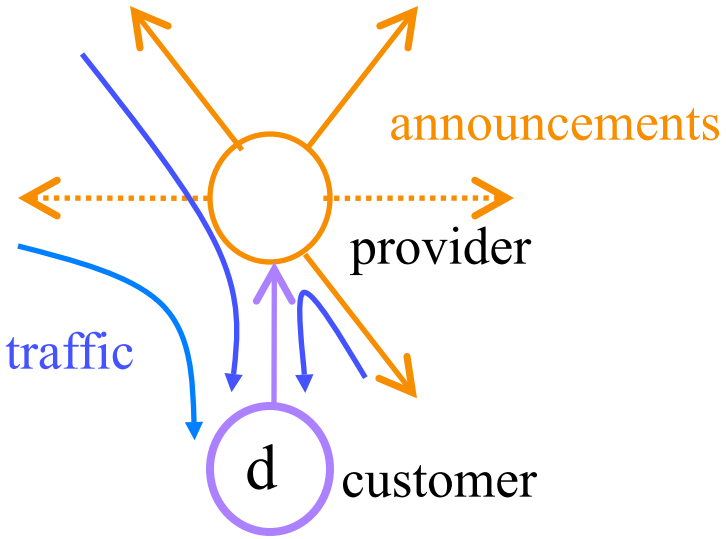
Business Relationships

- Neighboring ASes have business contracts
  - How much traffic to carry
  - Which destinations to reach
  - How much money to pay
- Common business relationships
  - Customer-provider
    - E.g., Princeton is a customer of USLEC
    - E.g., MIT is a customer of Level3
  - Peer-peer
    - E.g., UUNET is a peer of Sprint
    - E.g., Harvard is a peer of Harvard Business School

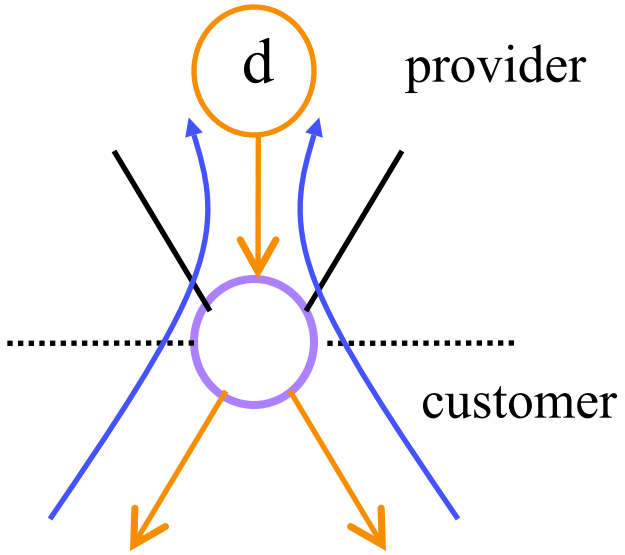
Customer-Provider Relationship

- Customer needs to be reachable from everyone
  - Provider tells all neighbors how to reach the customer
- Customer does not want to provide transit service
  - Customer does not let its providers route through it

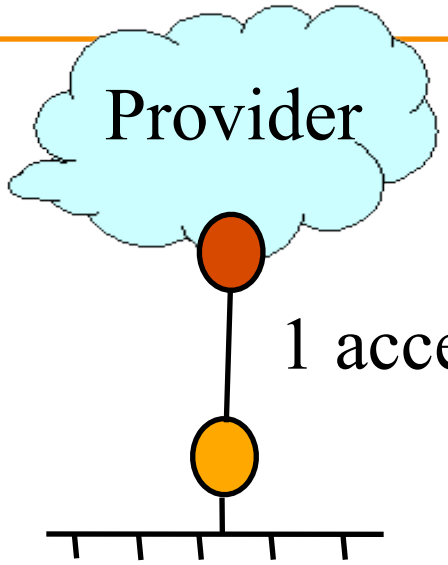
Traffic **to** the customer



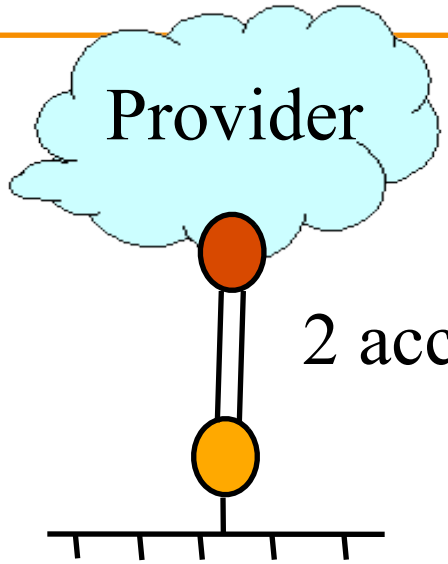
Traffic **from** the customer



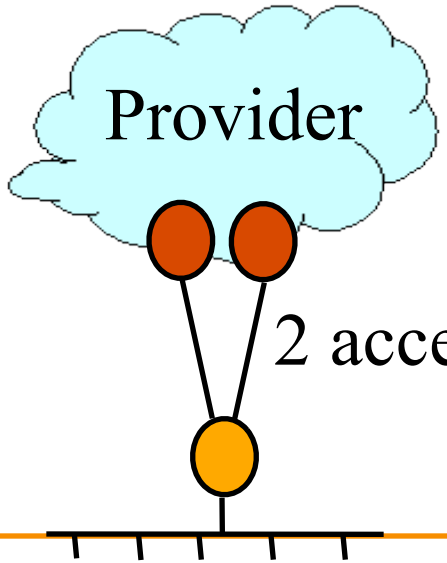
Customer Connecting to a Provider



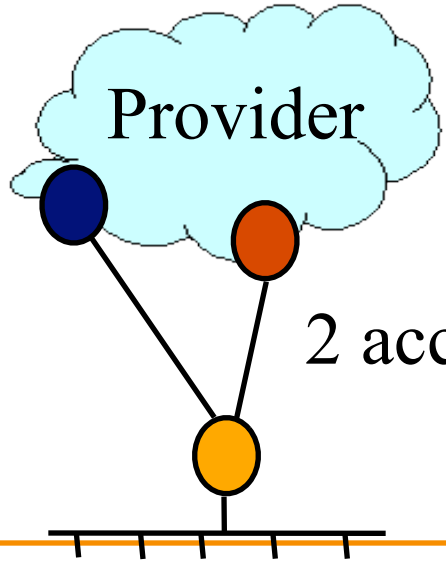
1 access link



2 access links



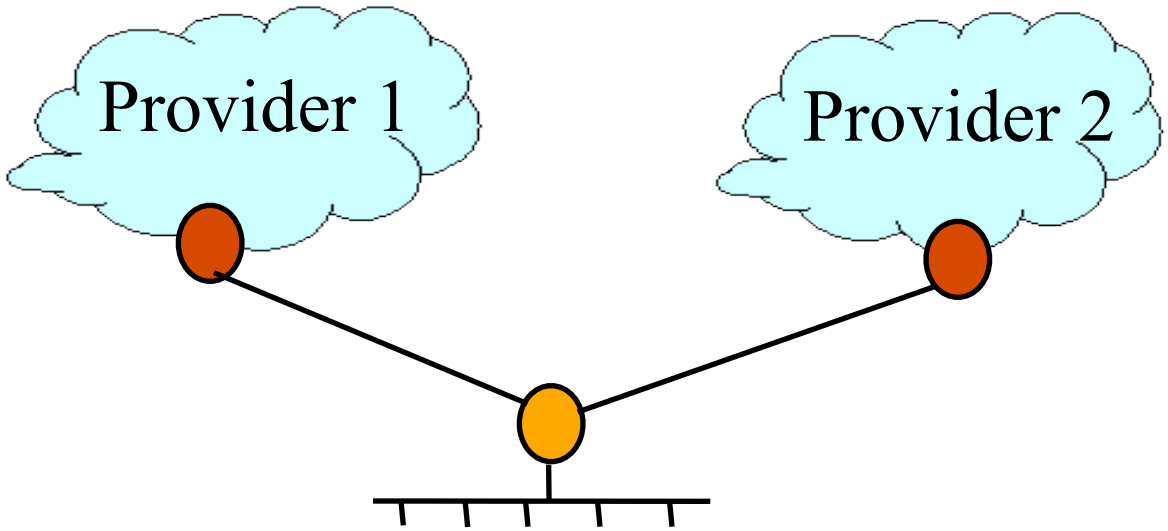
2 access routers



2 access PoPs

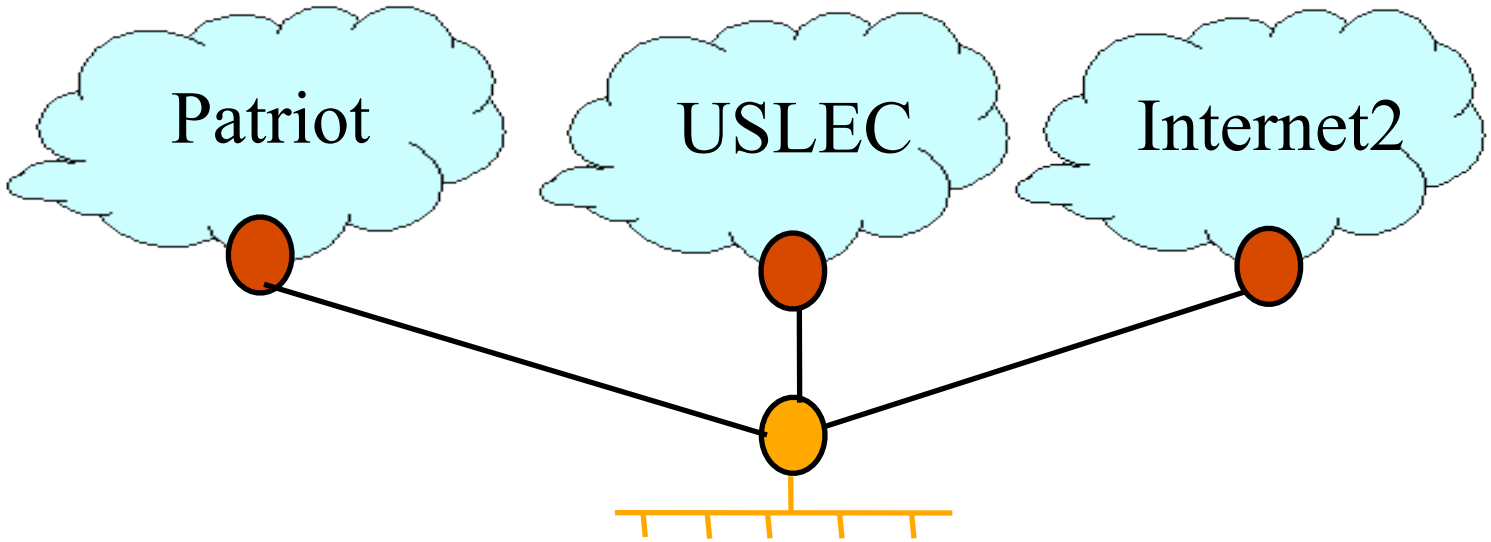
Multi-Homing: Two or More Providers

- Motivations for multi-homing
  - Extra reliability, survive single ISP failure
  - Financial leverage through competition
  - Better performance by selecting better path



Princeton Example

- Internet: customer of USLEC and Patriot
- Research universities/labs: customer of Internet2
- Local non-profits: provider for several non-profits



AS12660 SHARIF-EDU-NET Sharif University of Technology, Tehran,Iran

12657

- 212.72.64.0/19
- 213.131.192.0/19

12660

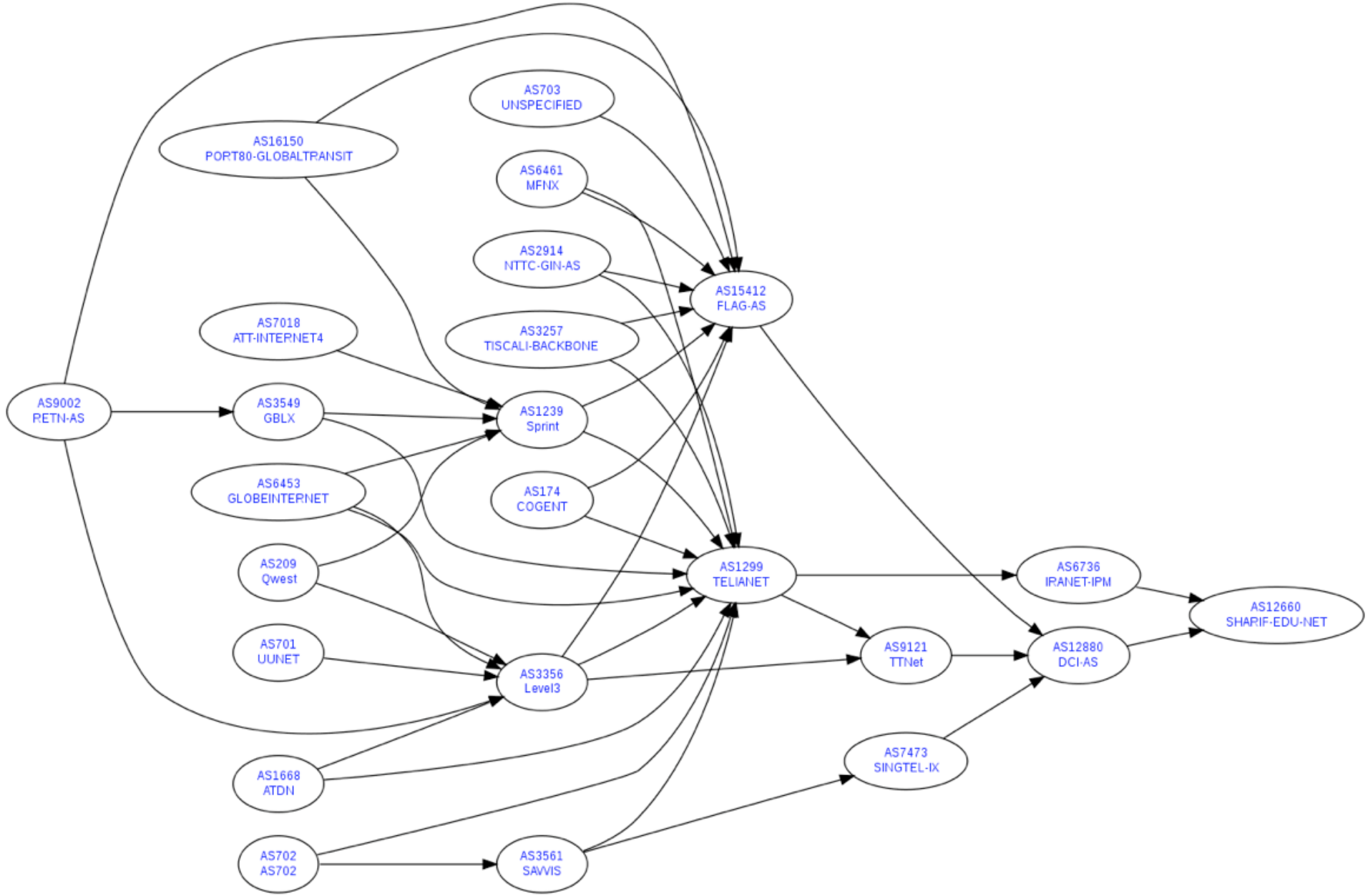
- 81.31.160.0/19
- 213.233.160.0/19

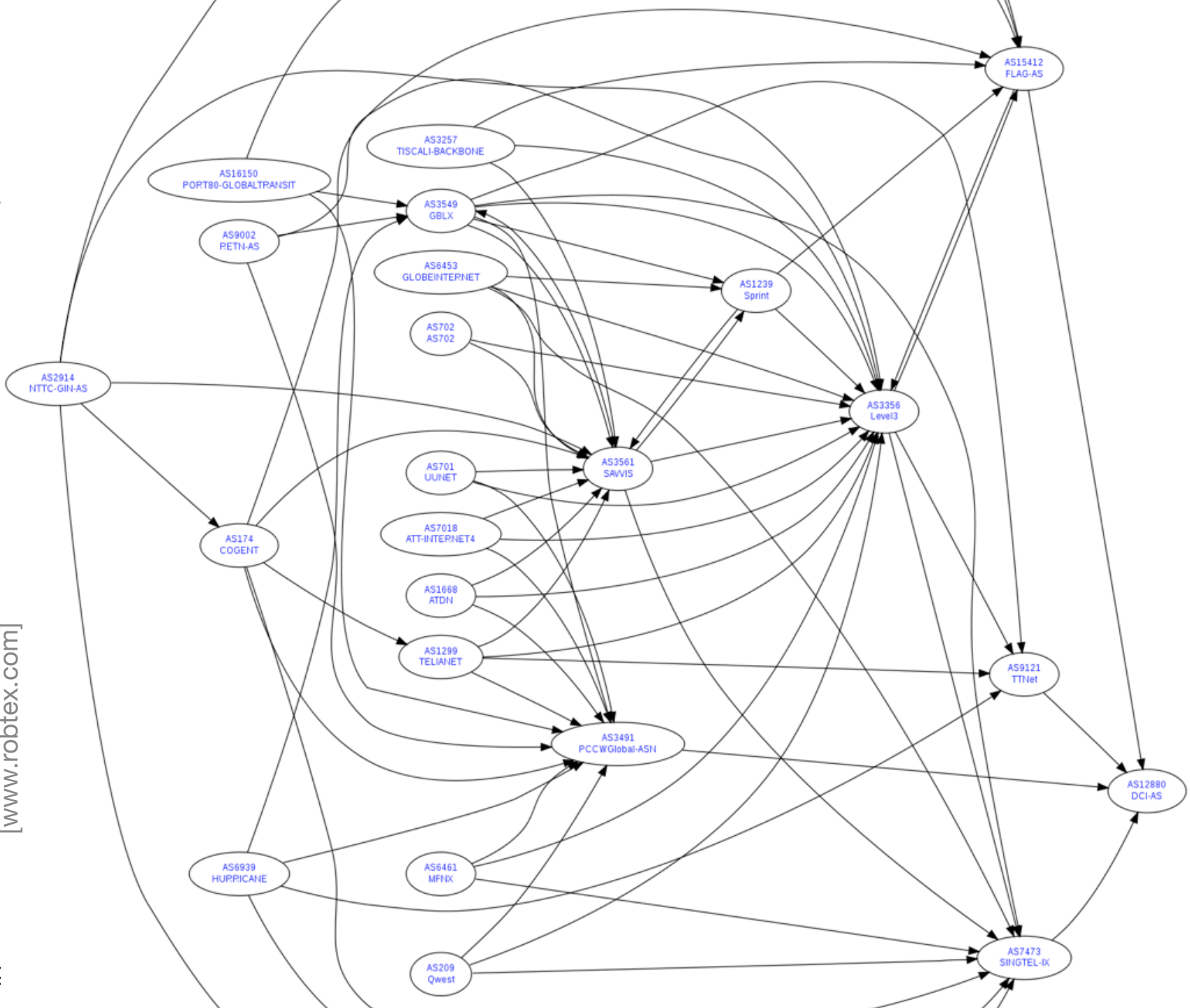
12692

- 140.204.0.0/16
- 161.71.0.0/16

12711

- 212.48.224.0/19
- 212.48.228.0/24


AS703
UNSPECIFIED
AS16150
PORT80-GLOBALTRANSIT
AS6461
MFNX
AS2914
NTTC-GIN-AS
AS15412
FLAG-AS
AS7018
ATT-INTERNET4
AS3257
TISCALI-BACKBONE
AS9002
RETN-AS
AS3549
GBLX
AS1239
Sprint
AS6453
GLOBEINTERNET
AS174
COGENT
AS209
Qwest
AS1299
TELIANET
AS6736
IRANET-IPM
AS12660
SHARIF-EDU-NET
AS701
UUNET
AS9121
TTNet
AS12880
DCI-AS
AS3356
Level3
AS7473
SINGTEL-IX
AS1668
ATDN
AS702
AS702
AS3561
SAVVIS



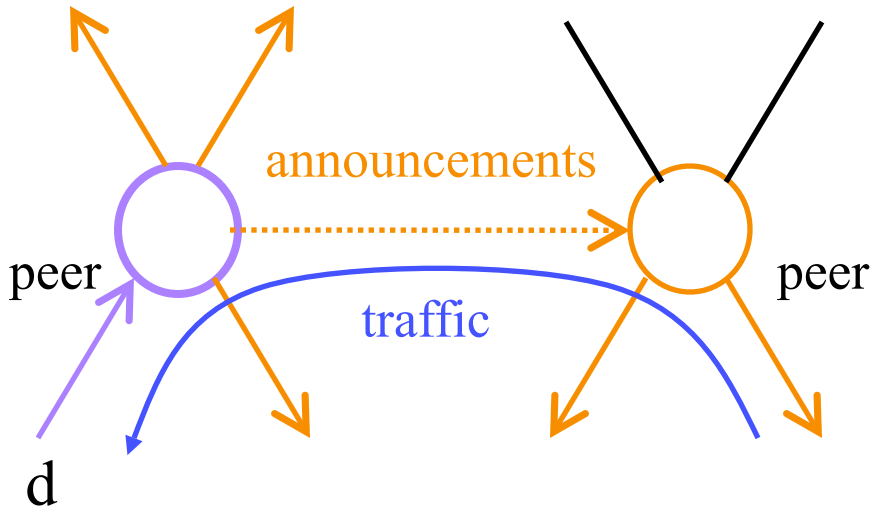
[www.robtex.com]

Spring

Peer-Peer Relationship

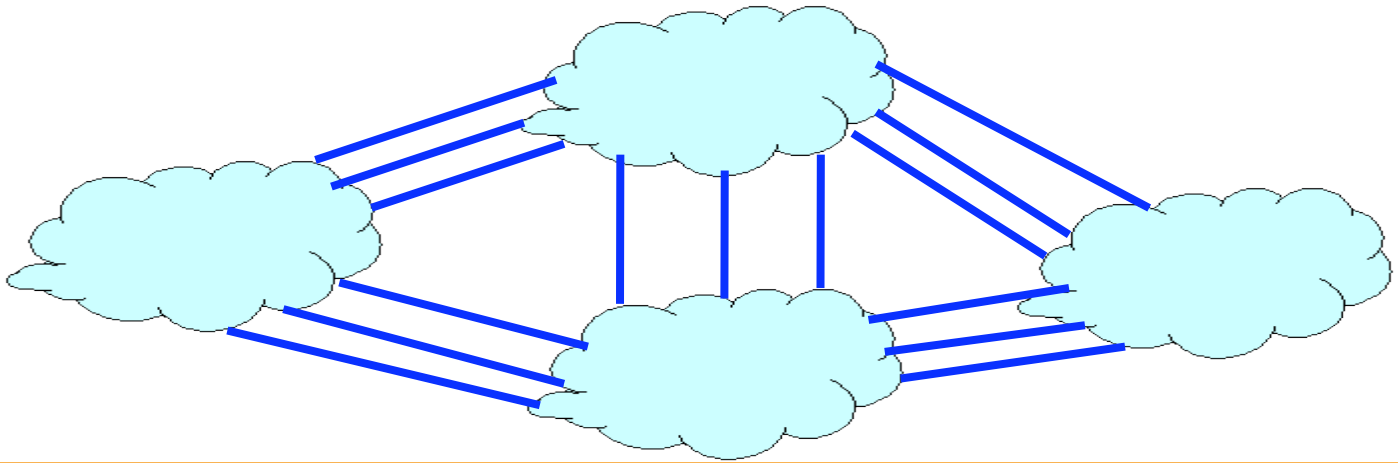
- Peers exchange traffic between customers
  - AS exports *only* customer routes to a peer
  - AS exports a peer's routes *only* to its customers
  - Often the relationship is settlement-free (i.e., no $$$)

Traffic to/from the peer and its customers



AS Structure: Tier-1 Providers

- Tier-1 provider
  - Has no upstream provider of its own
  - Typically has a national or international backbone
- Top of the Internet hierarchy of ~10 ASes
  - AOL, AT&T, Global Crossing, Level3, UUNET, NTT, Qwest, SAVVIS (formerly Cable & Wireless), and Sprint
  - Full peer-peer connections between tier-1 providers



AS Structure: Other ASes

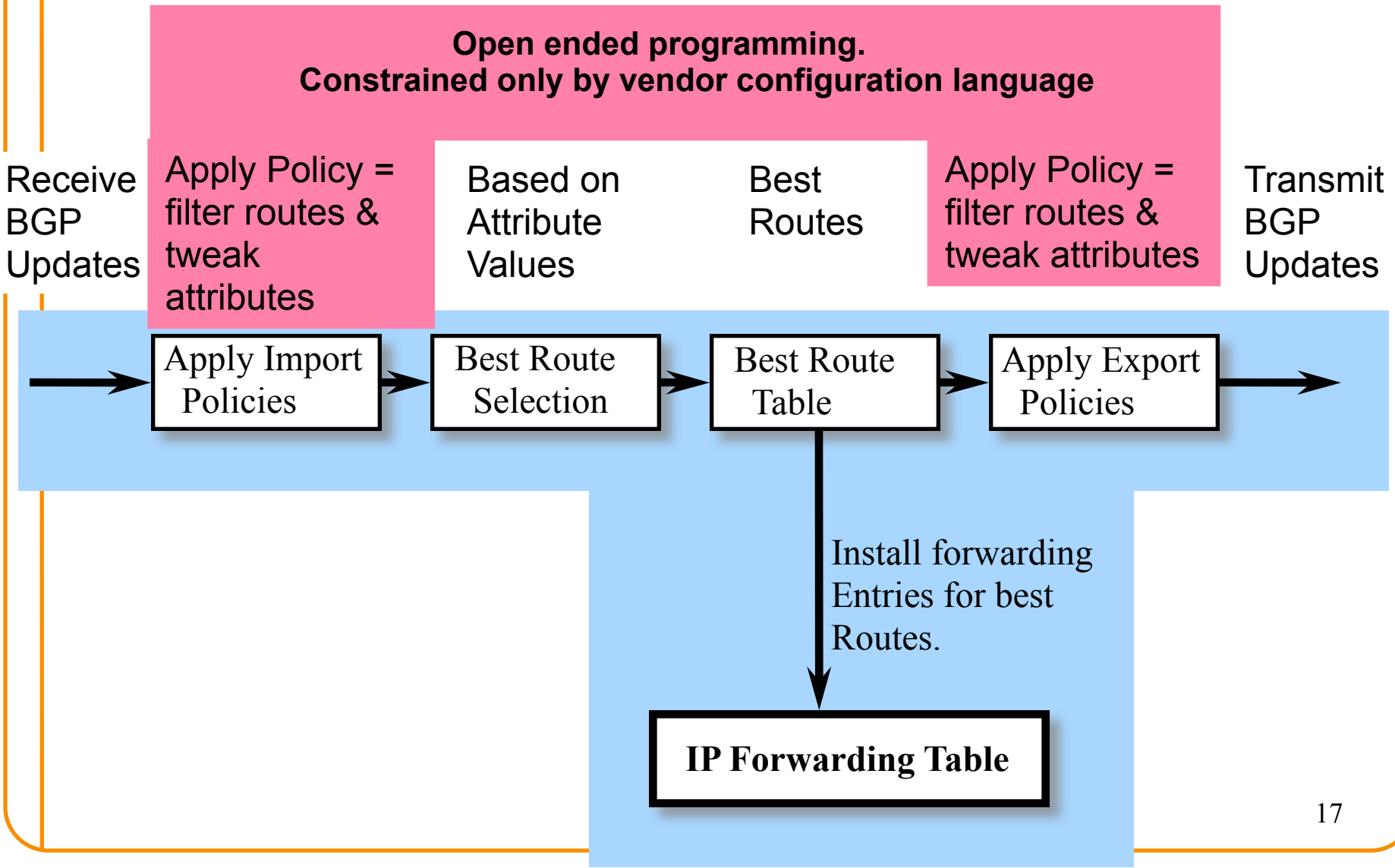
- Other providers
  - Provide transit service to downstream customers
  - … but, need at least one provider of their own
  - Typically have national or regional scope
  - Includes several thousand ASes
- Stub ASes
  - Do not provide transit service to others
  - Connect to one or more upstream providers
  - Includes the vast majority (e.g., 85-90%) of the ASes

Realizing BGP Routing Policy

BGP Policy: Applying Policy to Routes

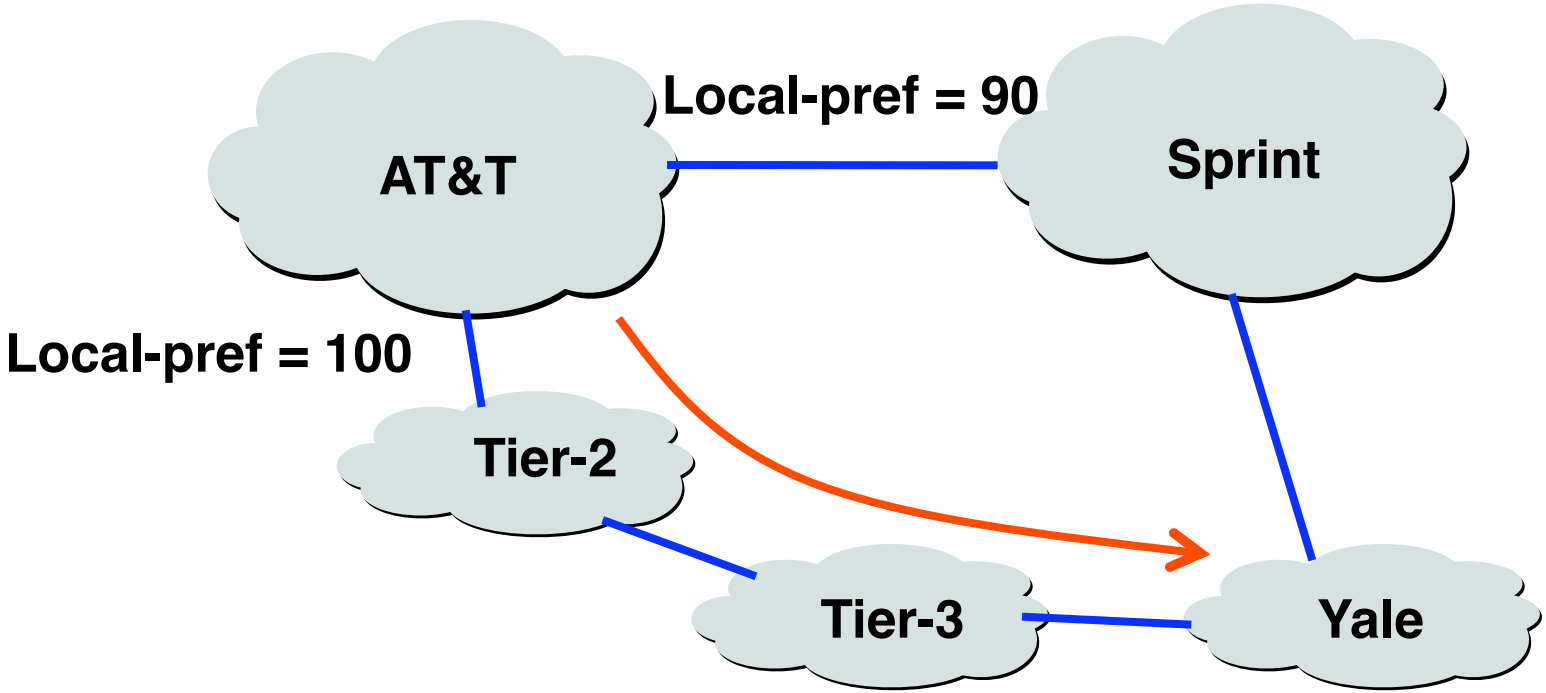
- Import policy
  - Select routes from neighbor
    - E.g. prefix that your customer doesn't own
  - Manipulate attributes to influence path selection
    - E.g., assign local preference to favored routes
- Export policy
  - Filter routes you don't want to tell your neighbor
    - E.g., don't tell a peer a route learned from other peer
  - Manipulate attributes to control what they see
    - E.g., make a path look artificially longer than it is

BGP Policy: Influencing Decisions



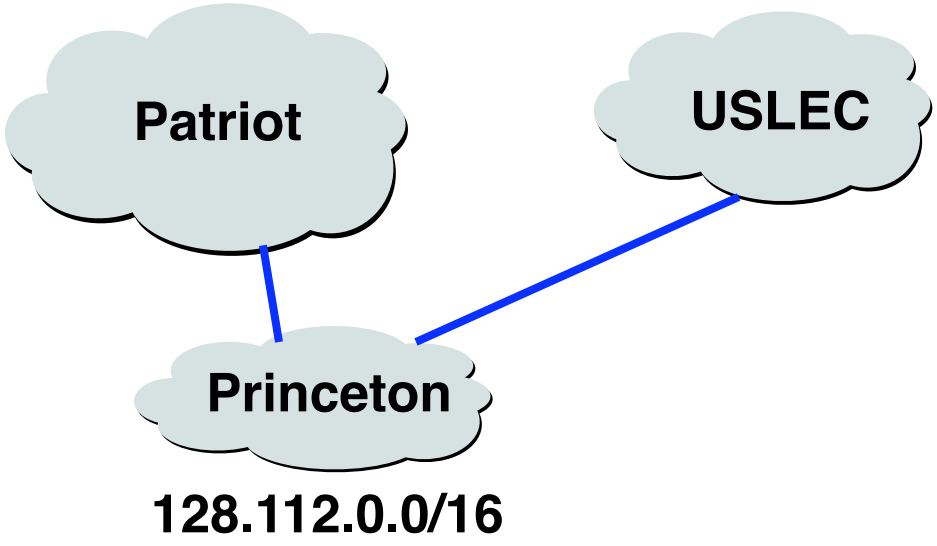
Import Policy: Local Preference

- Favor one path over another
  - Override the influence of AS path length
  - Apply local policies to prefer a path
- Example: prefer customer over peer



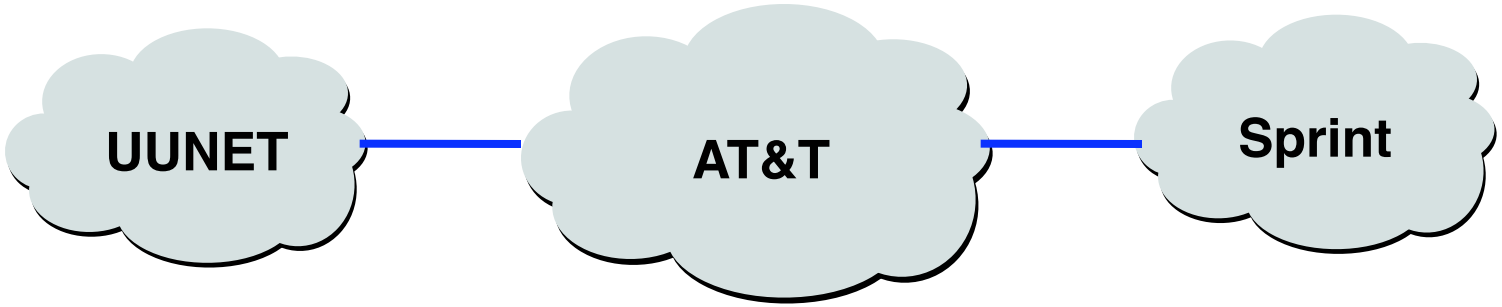
Import Policy: Filtering

- Discard some route announcements
  - Detect configuration mistakes and attacks
- Examples on session to a customer
  - Discard route if prefix not owned by the customer
  - Discard route that contains other large ISP in AS path



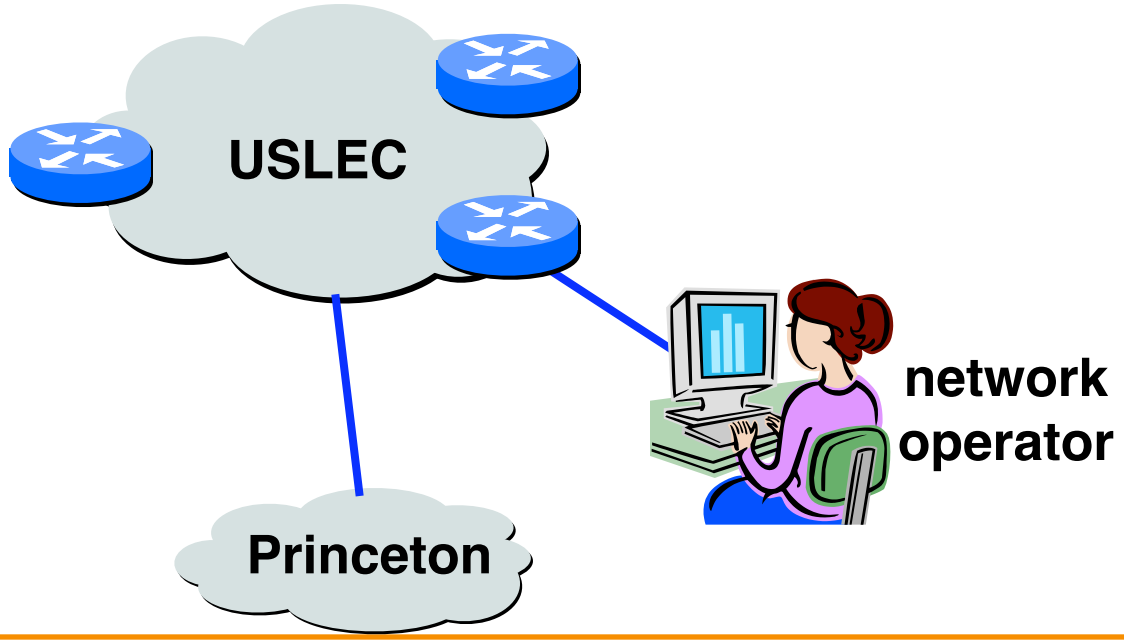
Export Policy: Filtering

- Discard some route announcements
  - Limit propagation of routing information
- Examples
  - Don't announce routes from one peer to another



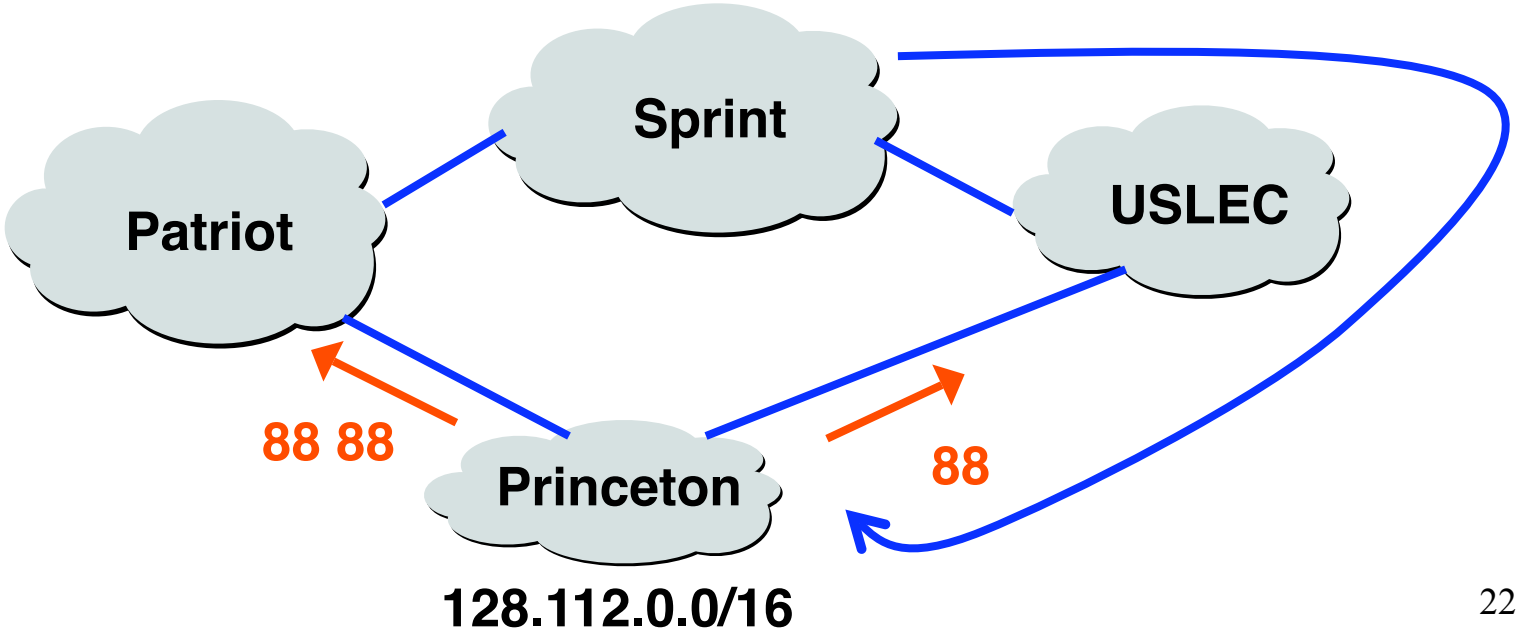
Export Policy: Filtering

- Discard some route announcements
  - Limit propagation of routing information
- Examples
  - Don't announce routes for network-management hosts or the underlying routers themselves



Export Policy: Attribute Manipulation

- Modify attributes of the active route
  - To influence the way other ASes behave
- Example: AS prepending
  - Artificially inflate the AS path length seen by others
  - To convince some ASes to send traffic another way



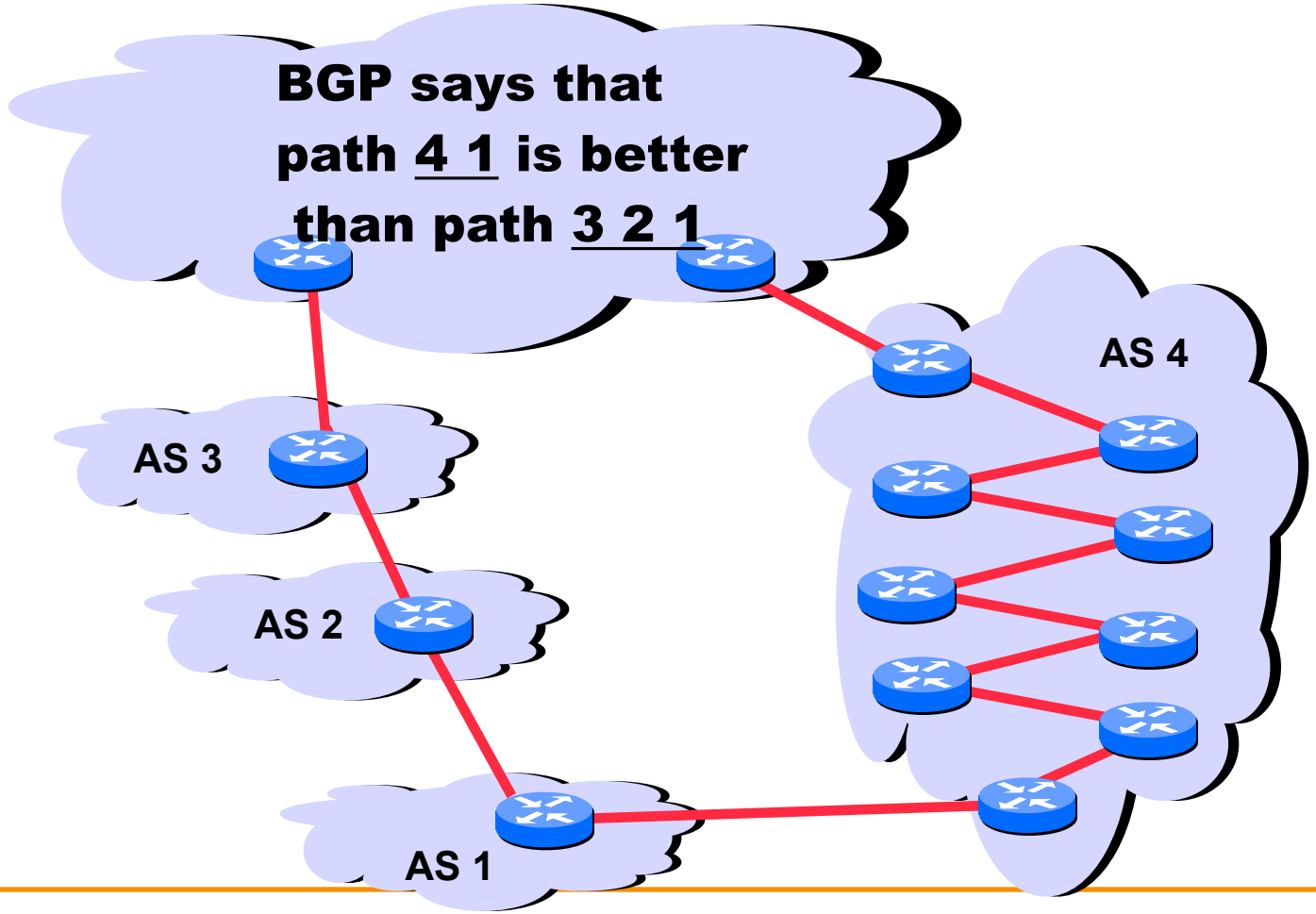
BGP Policy Configuration

- Routing policy languages are vendor-specific
  - Not part of the BGP protocol specification
  - Different languages for Cisco, Juniper, etc.
- Still, all languages have some key features
  - Policy as a list of clauses
  - Each clause matches on route attributes
  - … and either discards or modifies the matching routes
- Configuration done by human operators
  - Implementing the policies of their AS
  - Business relationships, traffic engineering, security, …
  - http://www.cs.princeton.edu/~jrex/papers/policies.pdf

Multiple Routers in an AS

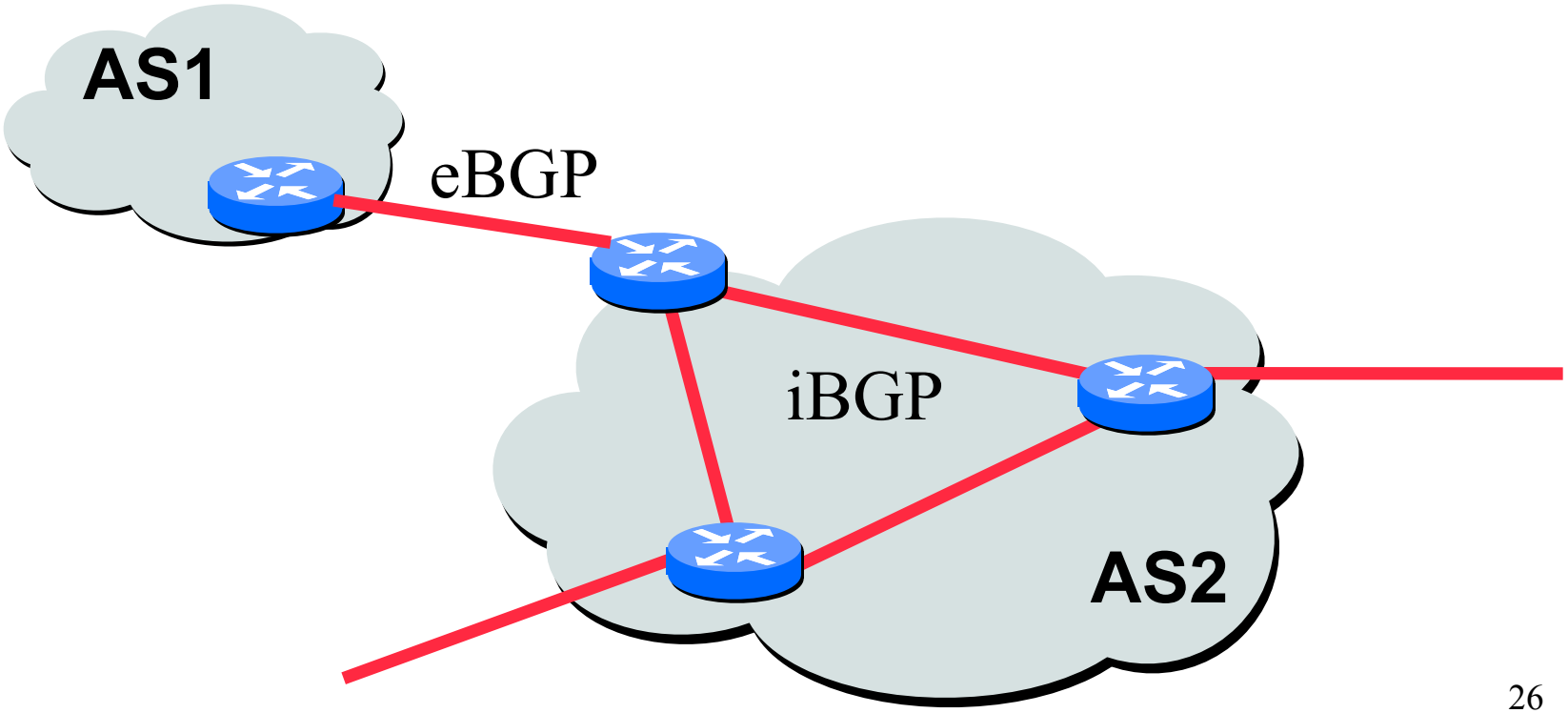
AS is Not a Single Node

- AS path length can be misleading
  - An AS may have many router-level hops



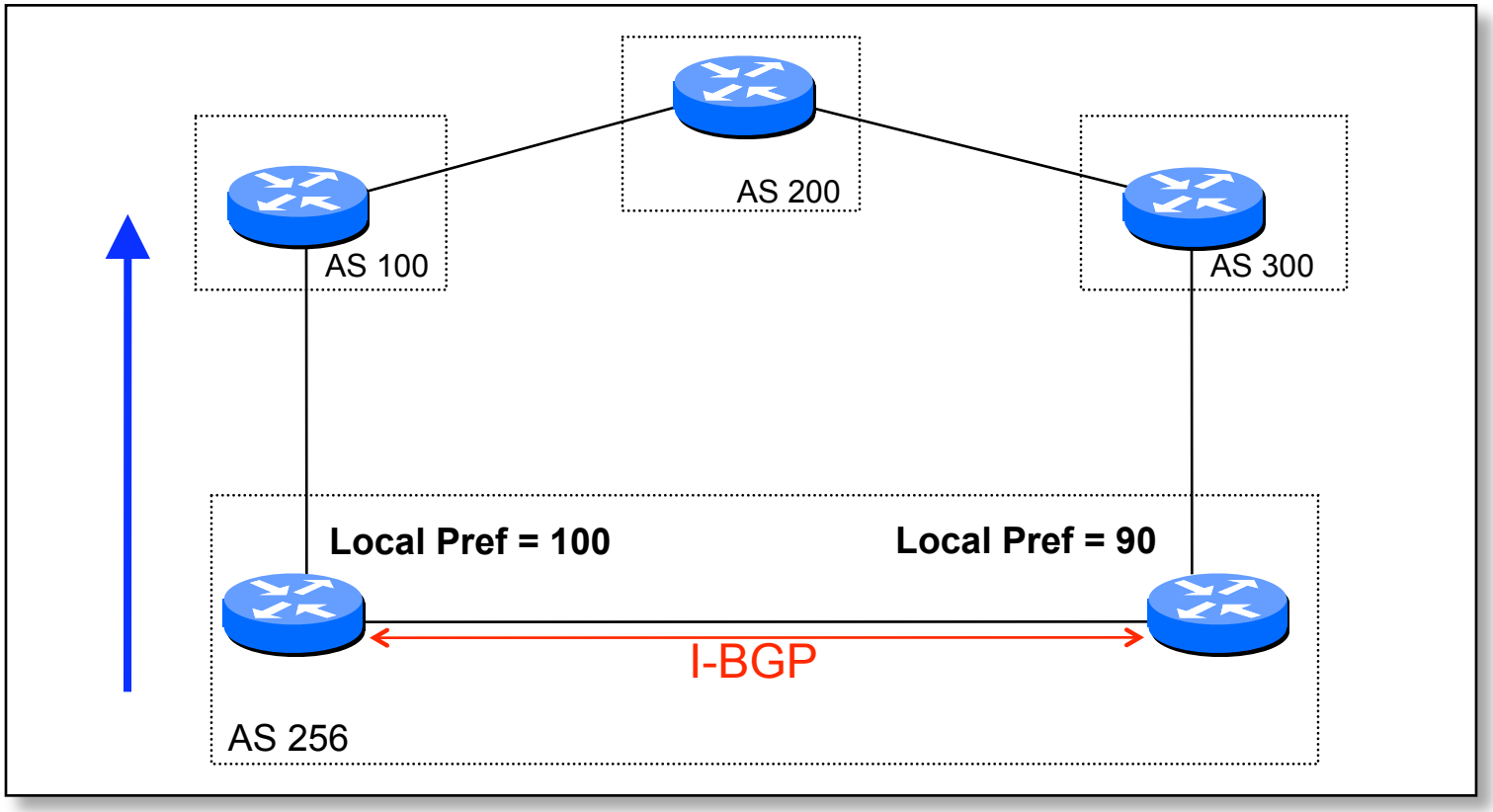
An AS is Not a Single Node

- Multiple routers in an AS
  - Need to distribute BGP information within the AS
  - Internal BGP (iBGP) sessions between routers

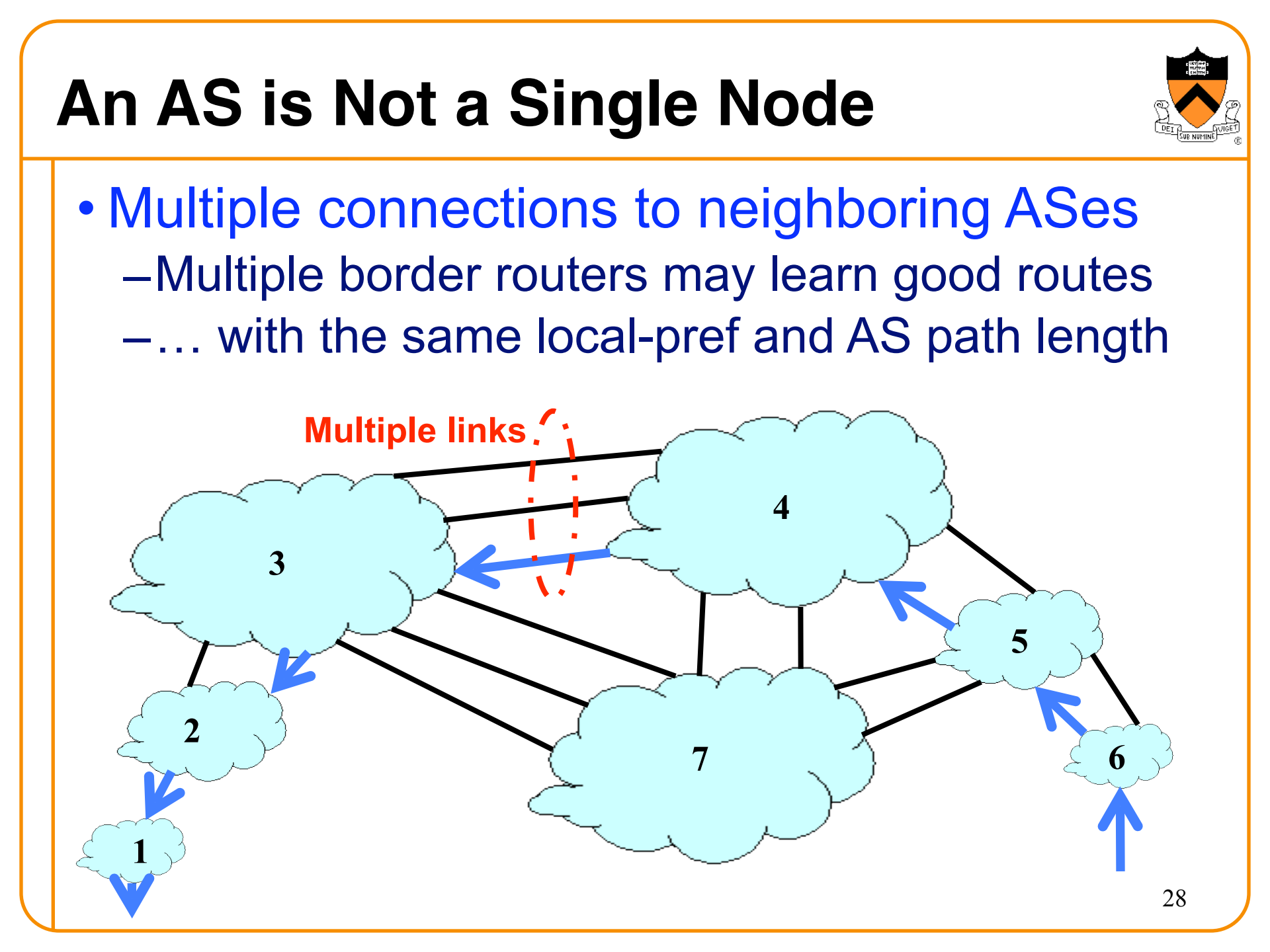


Internal BGP and Local Preference

- Example
  - Both routers prefer the path through AS 100 on the left
  - … even though the right router learns an external path

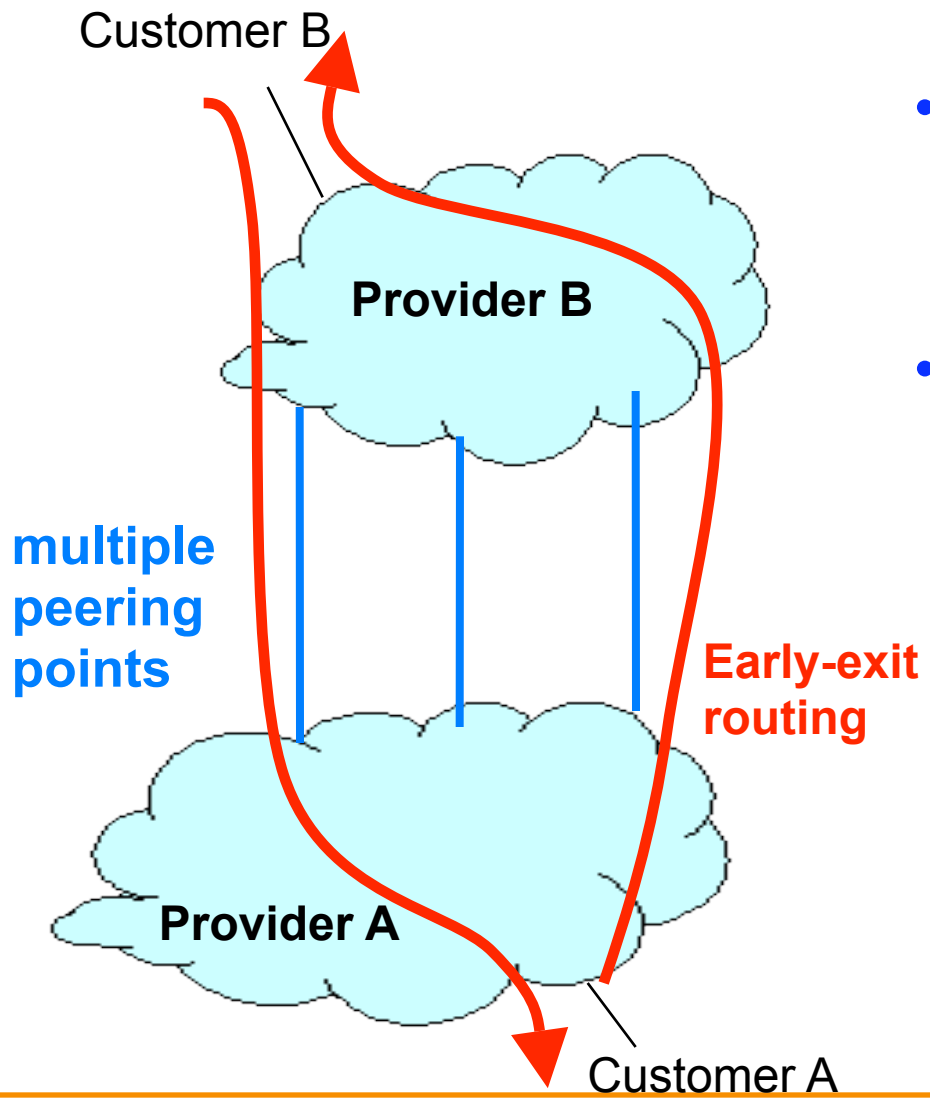


An AS is Not a Single Node



- Multiple connections to neighboring ASes
  - Multiple border routers may learn good routes
  - … with the same local-pref and AS path length

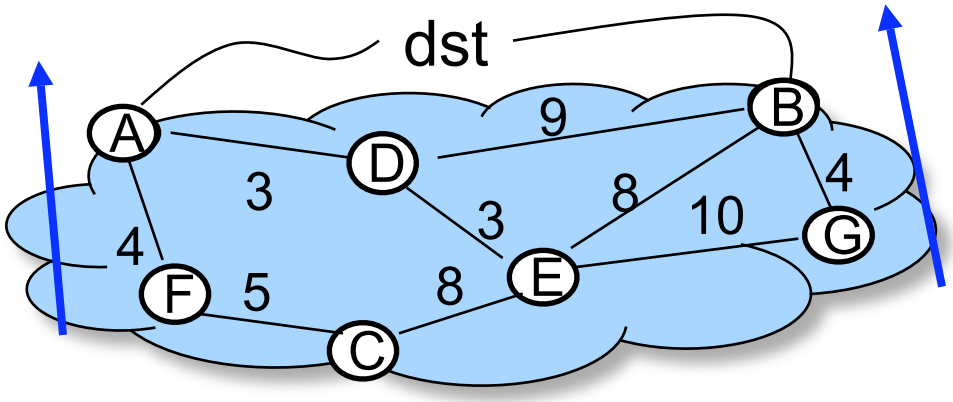
Early-Exit or Hot-Potato Routing



- Diverse peering locations
- Comparable capacity at all peering points
  - Can handle even load
- Consistent routes
  - Same destinations advertised at all points
  - Same AS path length for a destination at all points

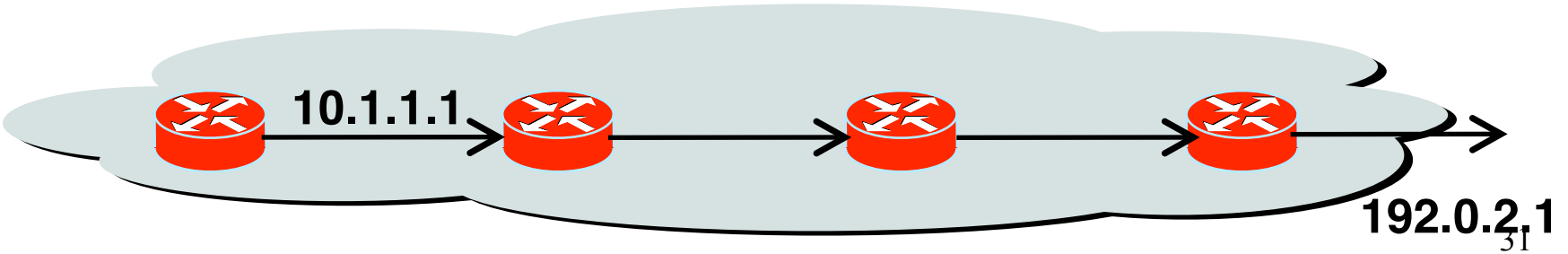
Realizing Hot-Potato Routing

- Hot-potato routing
  - Each router selects the closest egress point
  - … based on the path cost in intradomain protocol
- BGP decision process
  - Highest local preference
  - Shortest AS path
  - Closest egress point
  - Arbitrary tie break

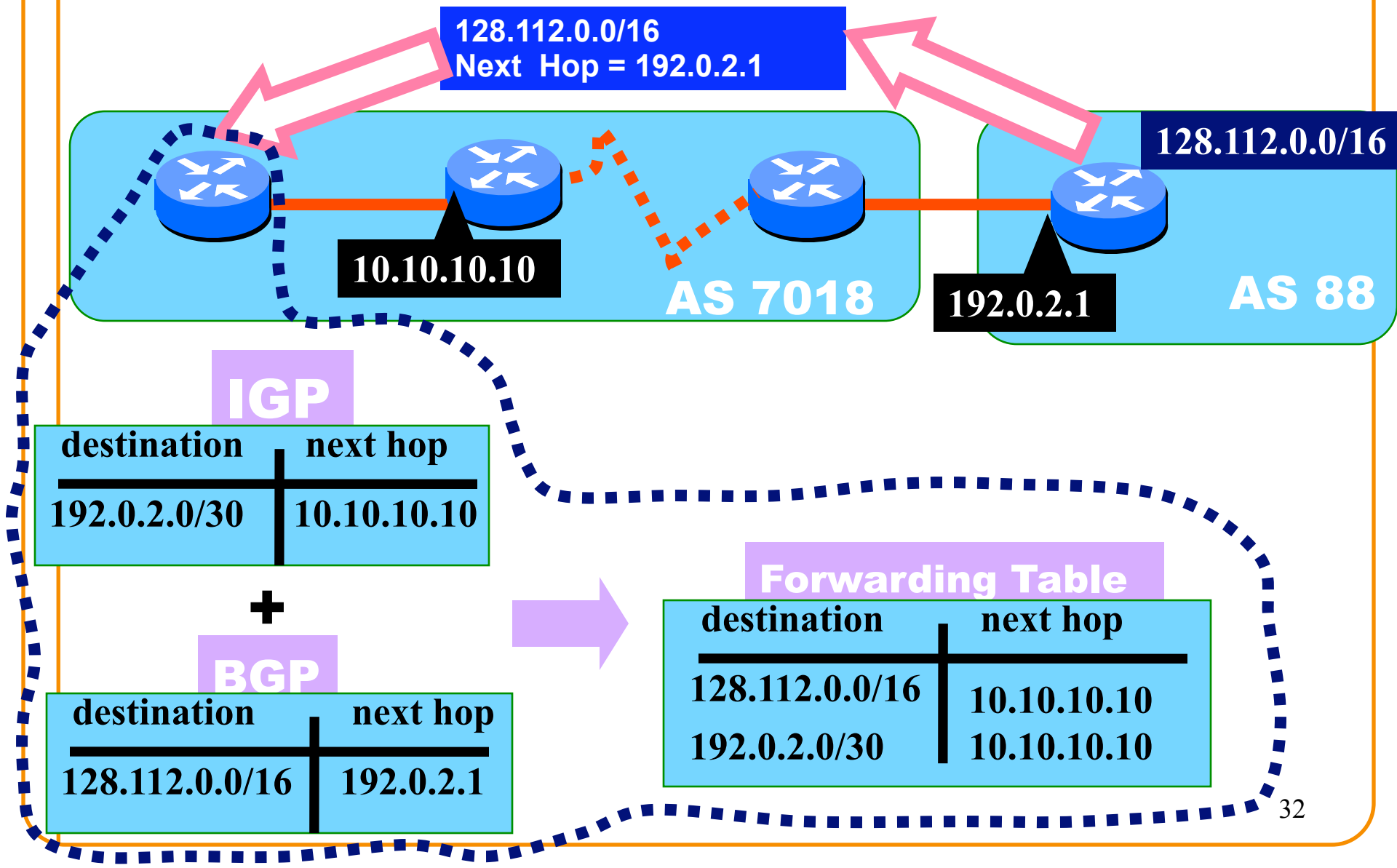


Joining BGP and IGP Information

- Border Gateway Protocol (BGP)
  - Announces reachability to external destinations
  - Maps a destination prefix to an egress point
    - 128.112.0.0/16 reached via 192.0.2.1
- Interior Gateway Protocol (IGP)
  - Used to compute paths within the AS
  - Maps an egress point to an outgoing link
    - 192.0.2.1 reached via 10.1.1.1

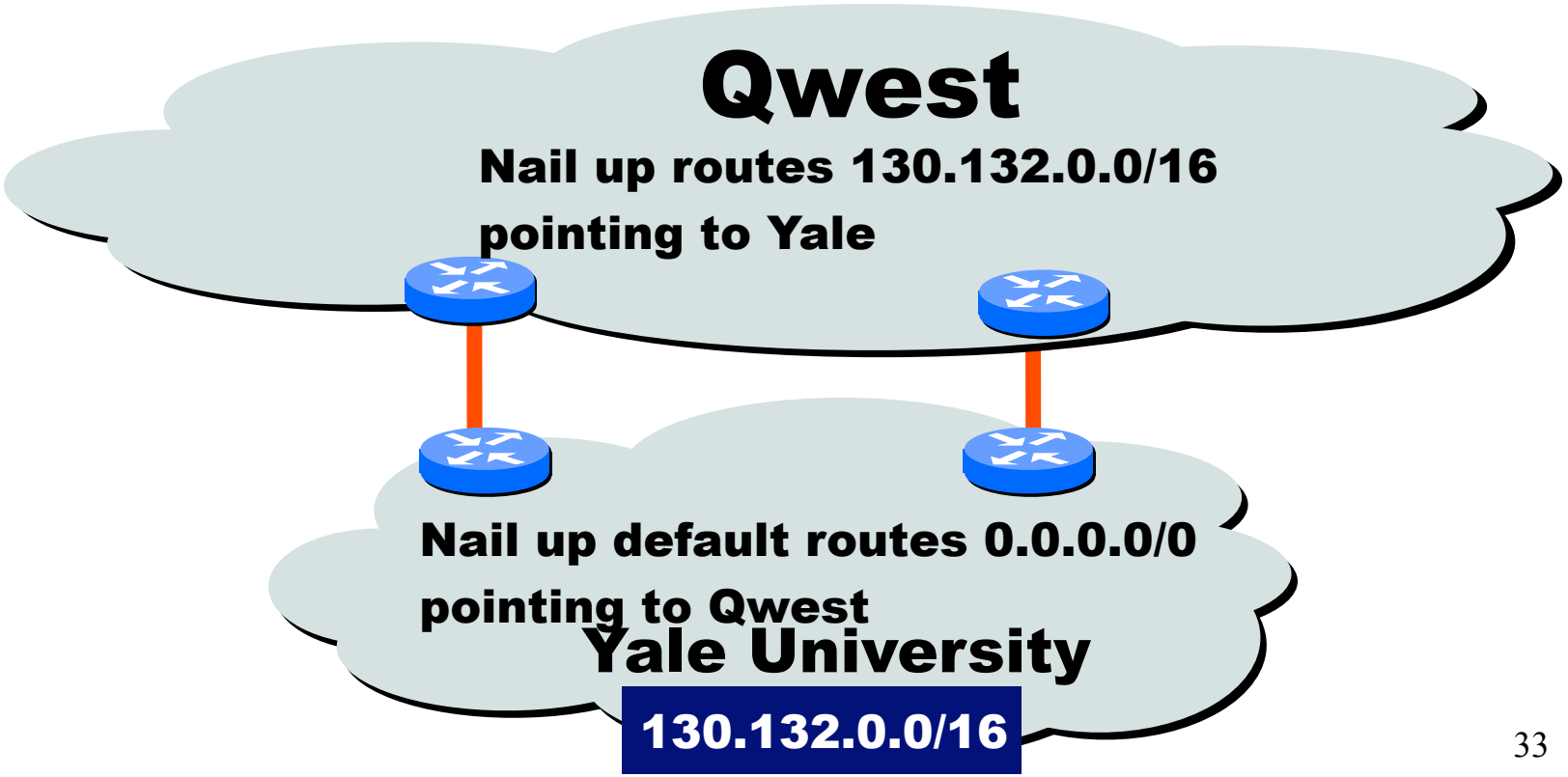


Joining BGP with IGP Information



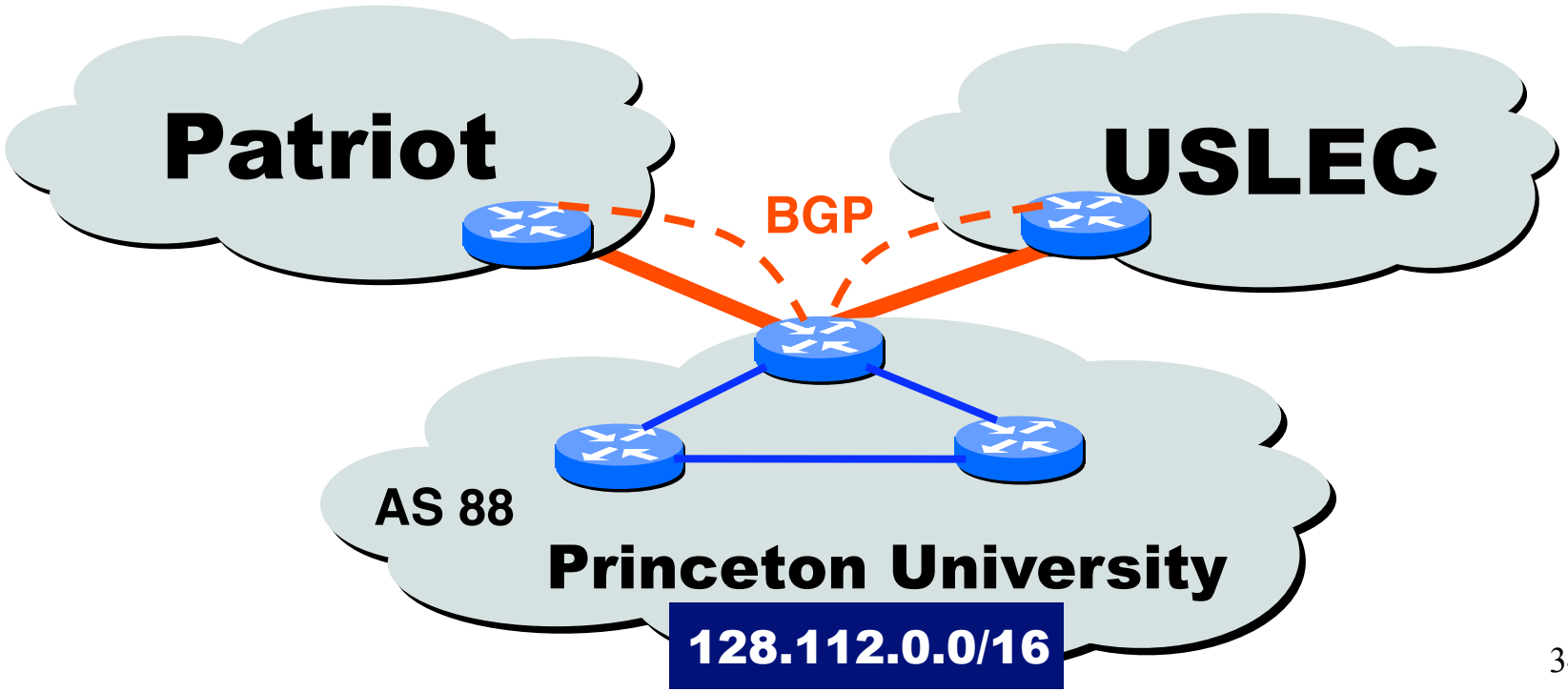
Some Routers Don't Need BGP

- Customer that connects to a single upstream ISP
  - The ISP can introduce the prefixes into BGP
  - … and the customer can simply default-route to the ISP



Some Routers Don't Need BGP

- Routers inside a "stub" network
  - Border router may speak BGP to upstream ISPs
  - But, internal routers can simply "default route"



Conclusions

- BGP is solving a hard problem
  - Routing protocol operating at a global scale
  - With tens of thousands of independent networks
  - That each have their own policy goals
  - And all want fast convergence
- Key features of BGP
  - Prefix-based path-vector protocol
  - Incremental updates (announcements and withdrawals)
  - Policies applied at import and export of routes
  - Internal BGP to distribute information within an AS
  - Interaction with the IGP to compute forwarding tables