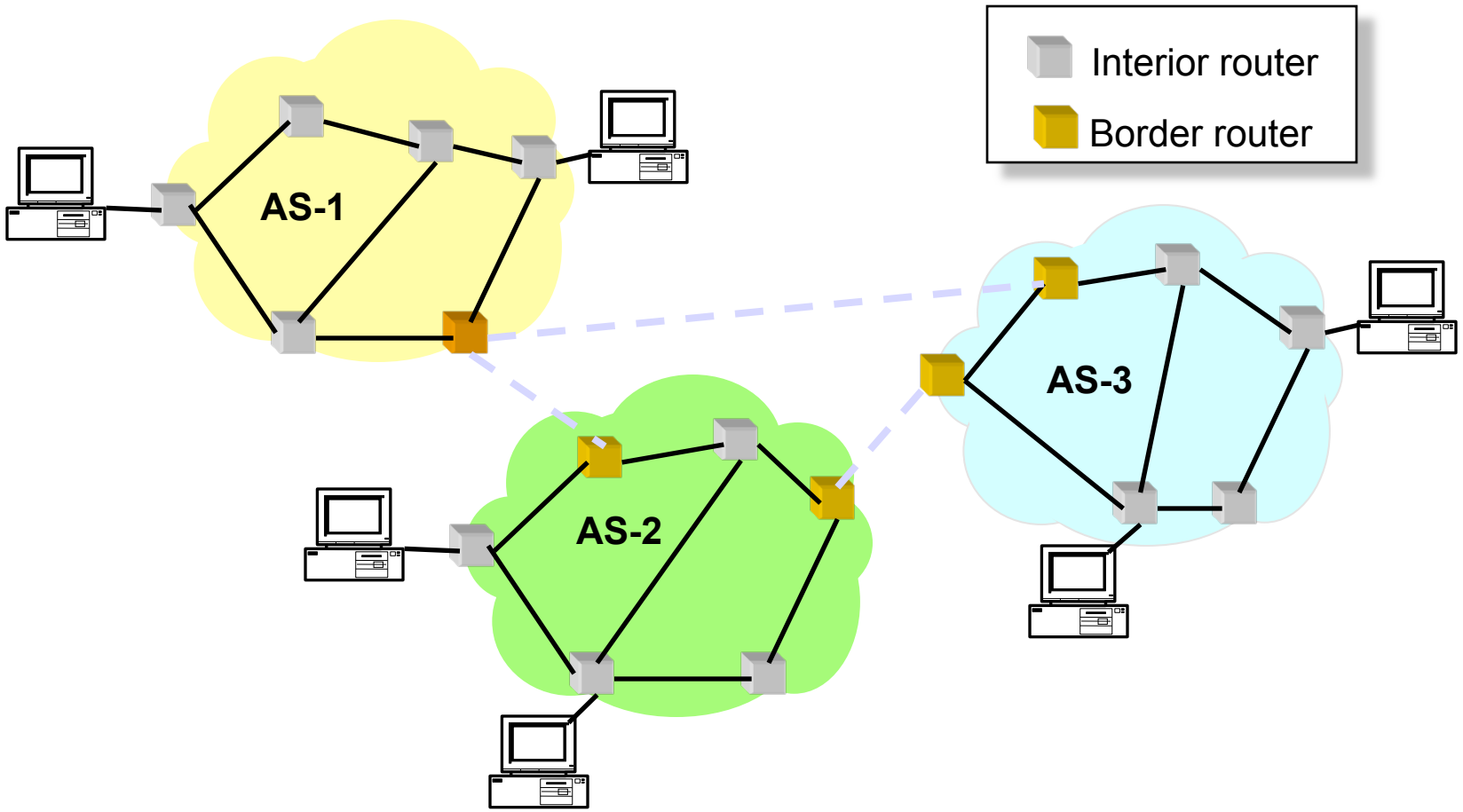
COMP/ELEC 429/556
Introduction to Computer Networks

Inter-domain routing

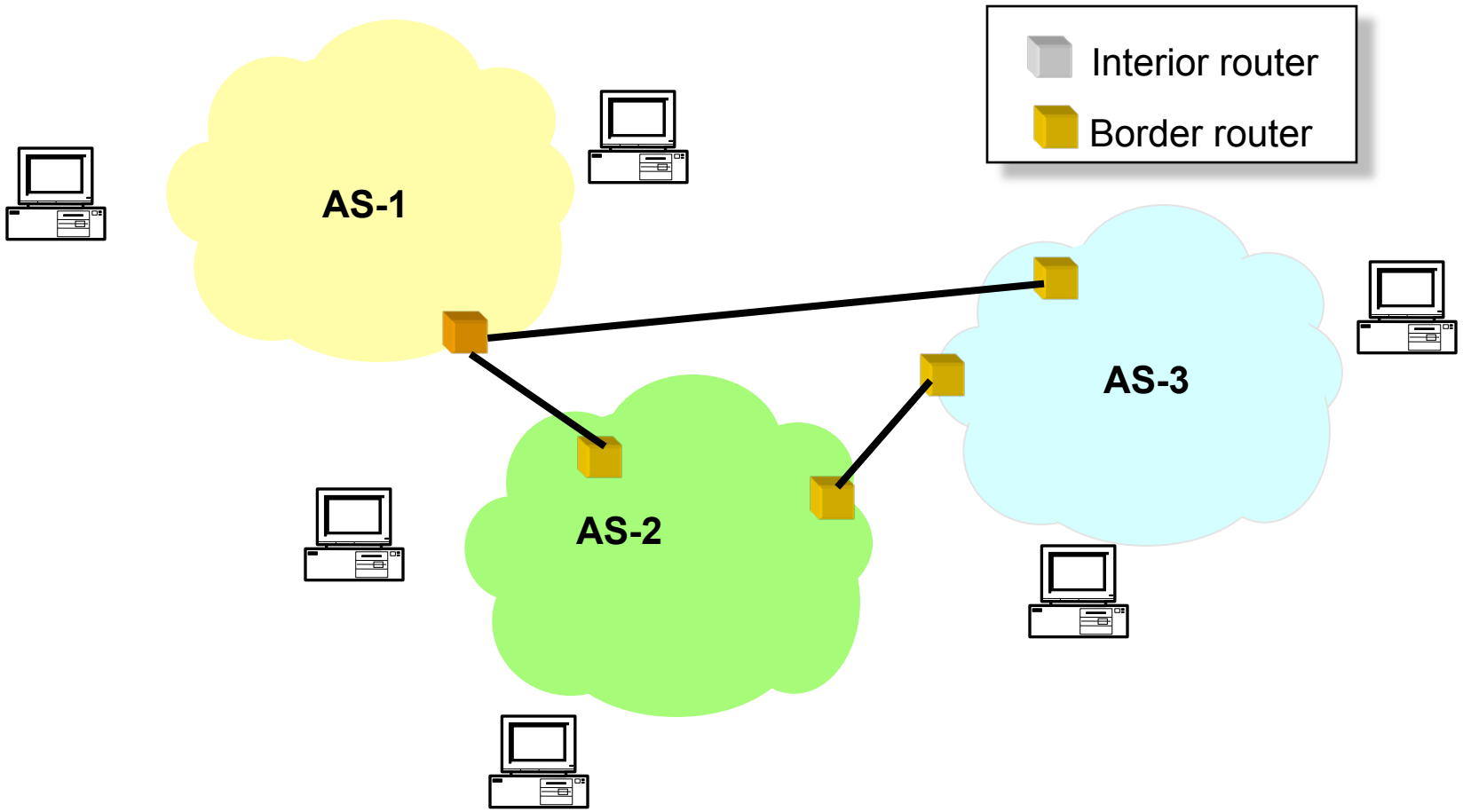
Some slides used with permissions from Edward W. Knightly, T. S. Eugene Ng, Ion Stoica, Hui Zhang

Previous focus: Intra-Domain Routing



Intra-domain routing protocol aka Interior Gateway Protocol (IGP)

Today's focus: Inter-Domain Routing

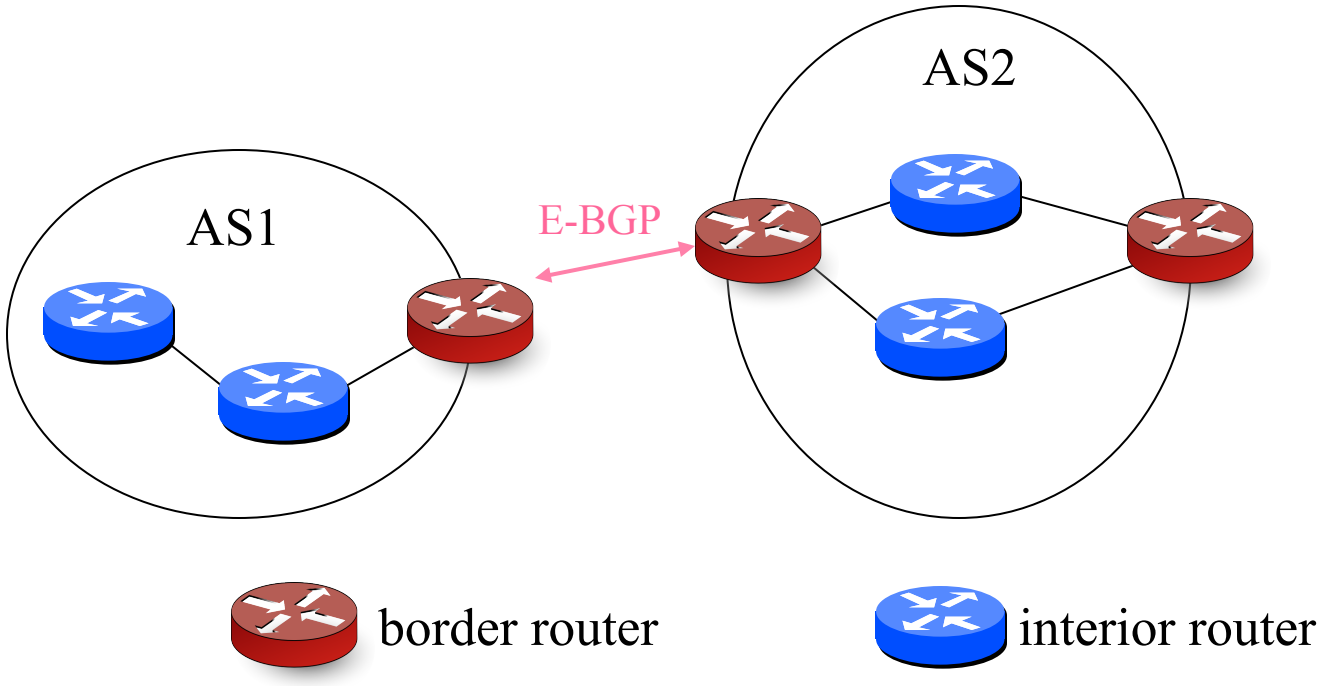


Inter-Domain Routing Considerations

- Global connectivity is at stake
- Inevitably leads to one single protocol that everyone must speak
  - Unlike many choices in intra-domain routing
- What are the requirements?
  - Scalability
  - Flexibility in choosing routes
- If you were to choose, link state based or distance vector based?

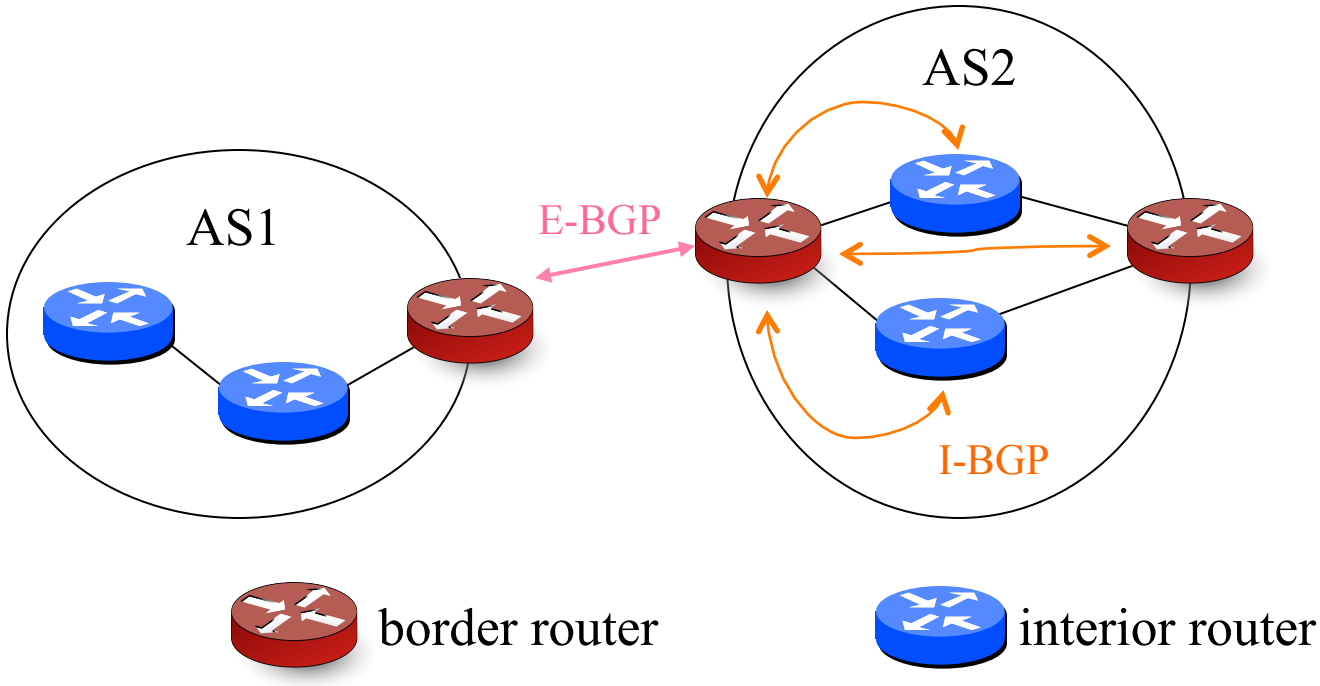
- Border Gateway Protocol (BGP)
  - A hybrid between link state and distance vector
  - “Path vector”

Border Gateway Protocol Part I: E-BGP



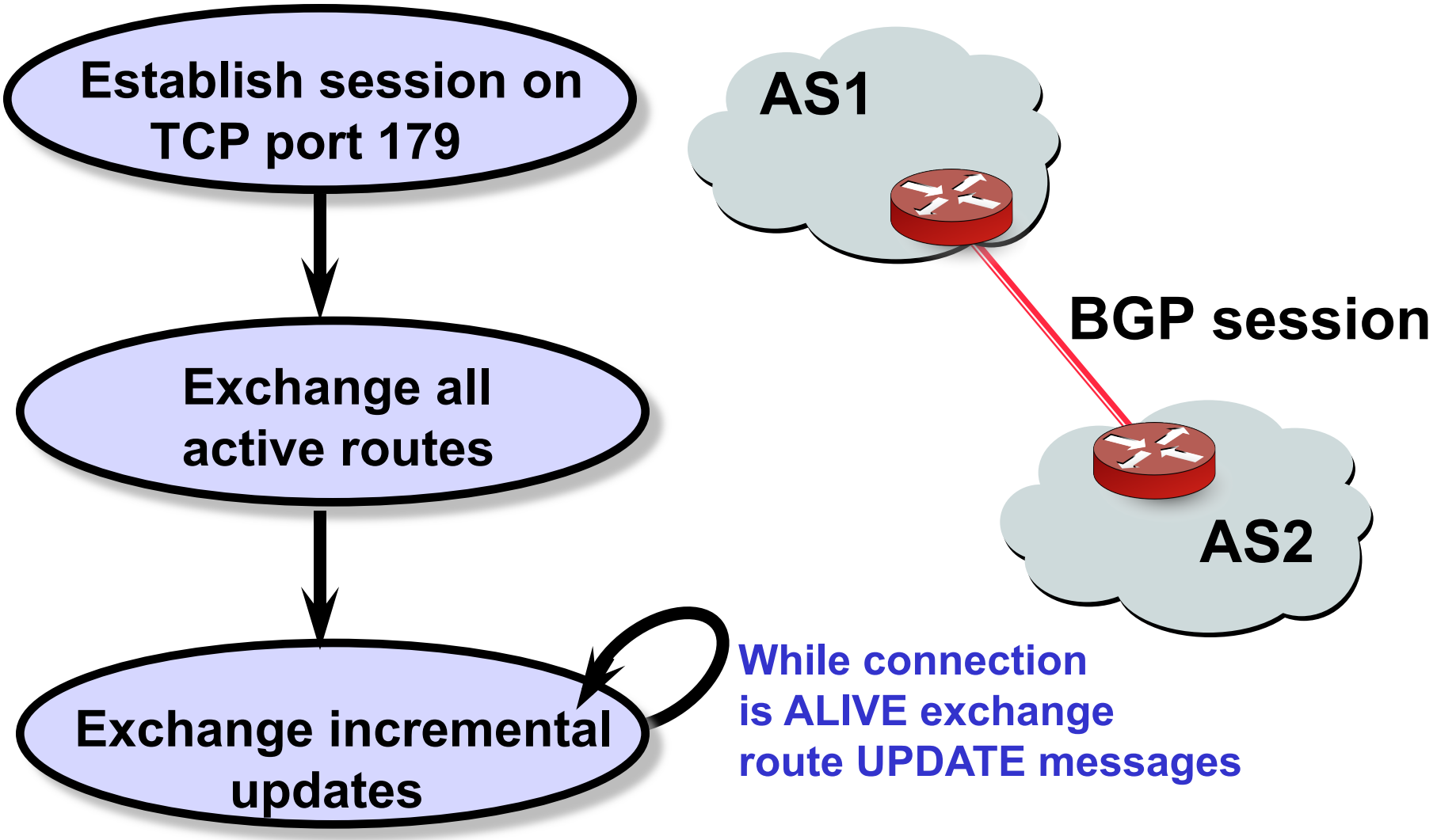
- Two types of routers
  - Border router, Interior router

Border Gateway Protocol Part II: I-BGP



- Two types of routers
  - Border router, Interior router

BGP Operations (Simplified)

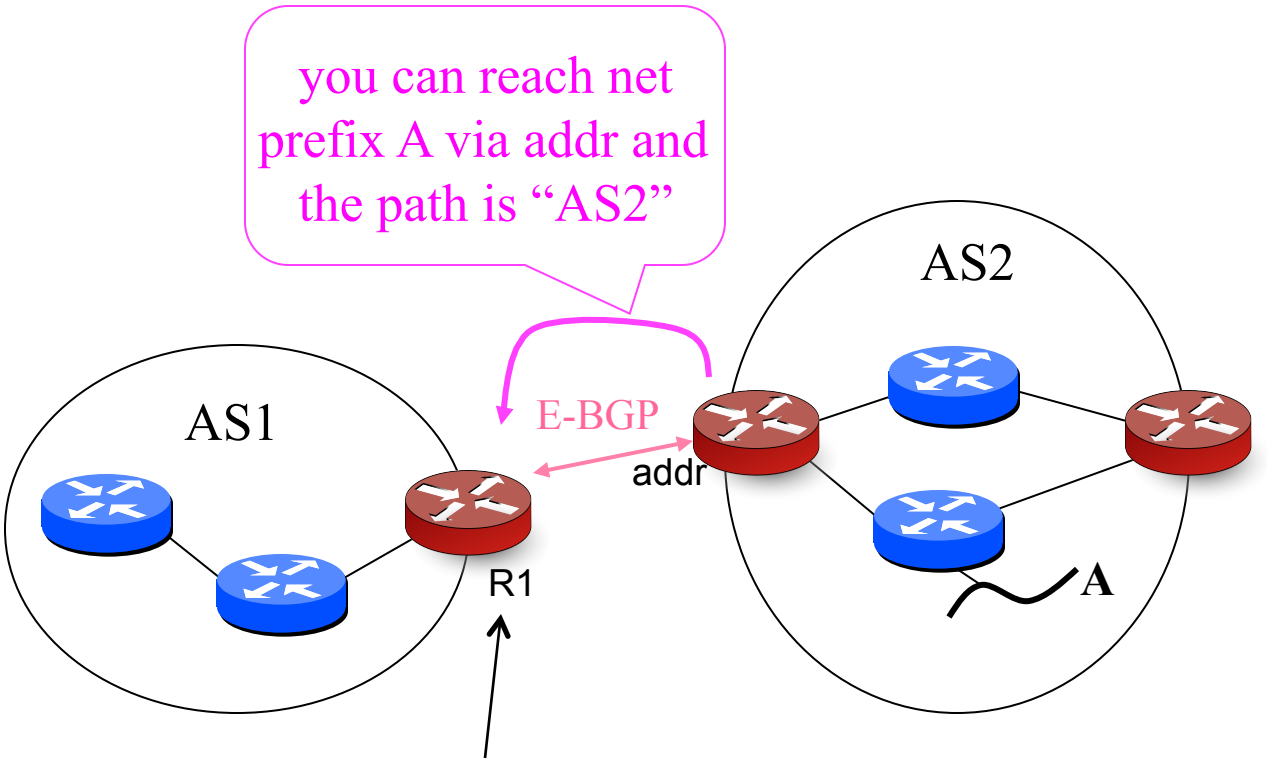


BGP Update Messages

- **Update** : Announcing new routes or withdrawing previously announced routes.

Update
=
Destination IP address prefix + attributes values
(e.g. a routing path)

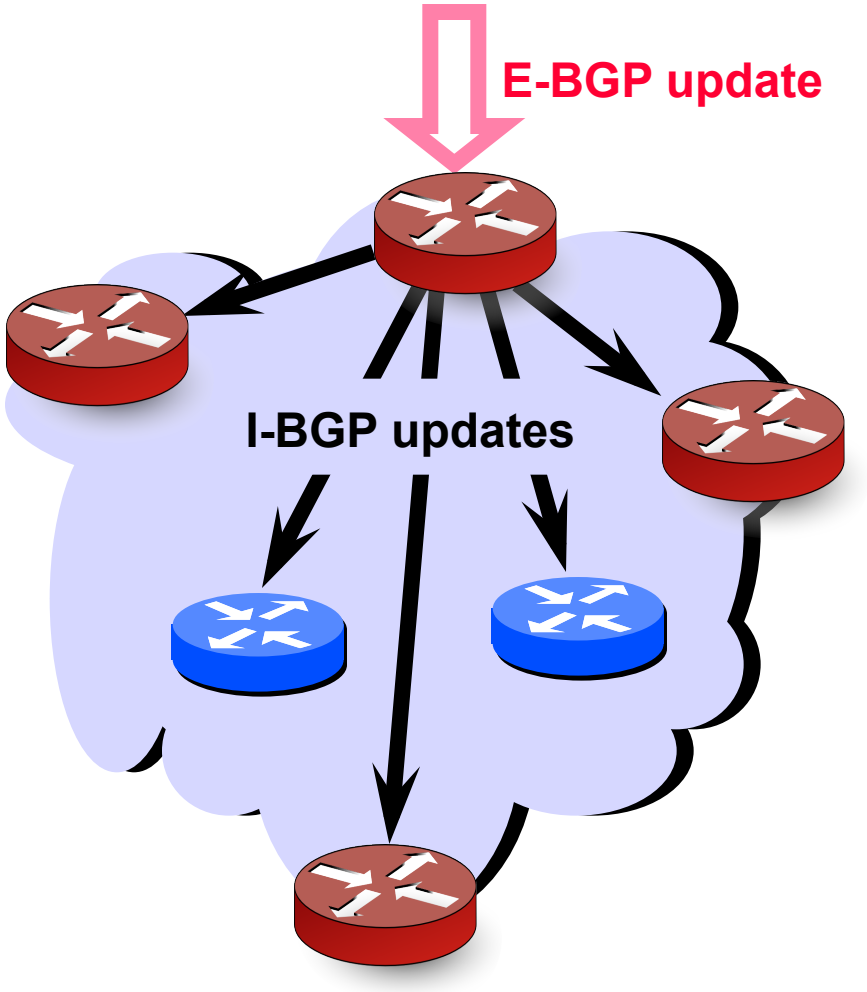
Part I: E-BGP, Share connectivity information across ASs



BGP table at R1:

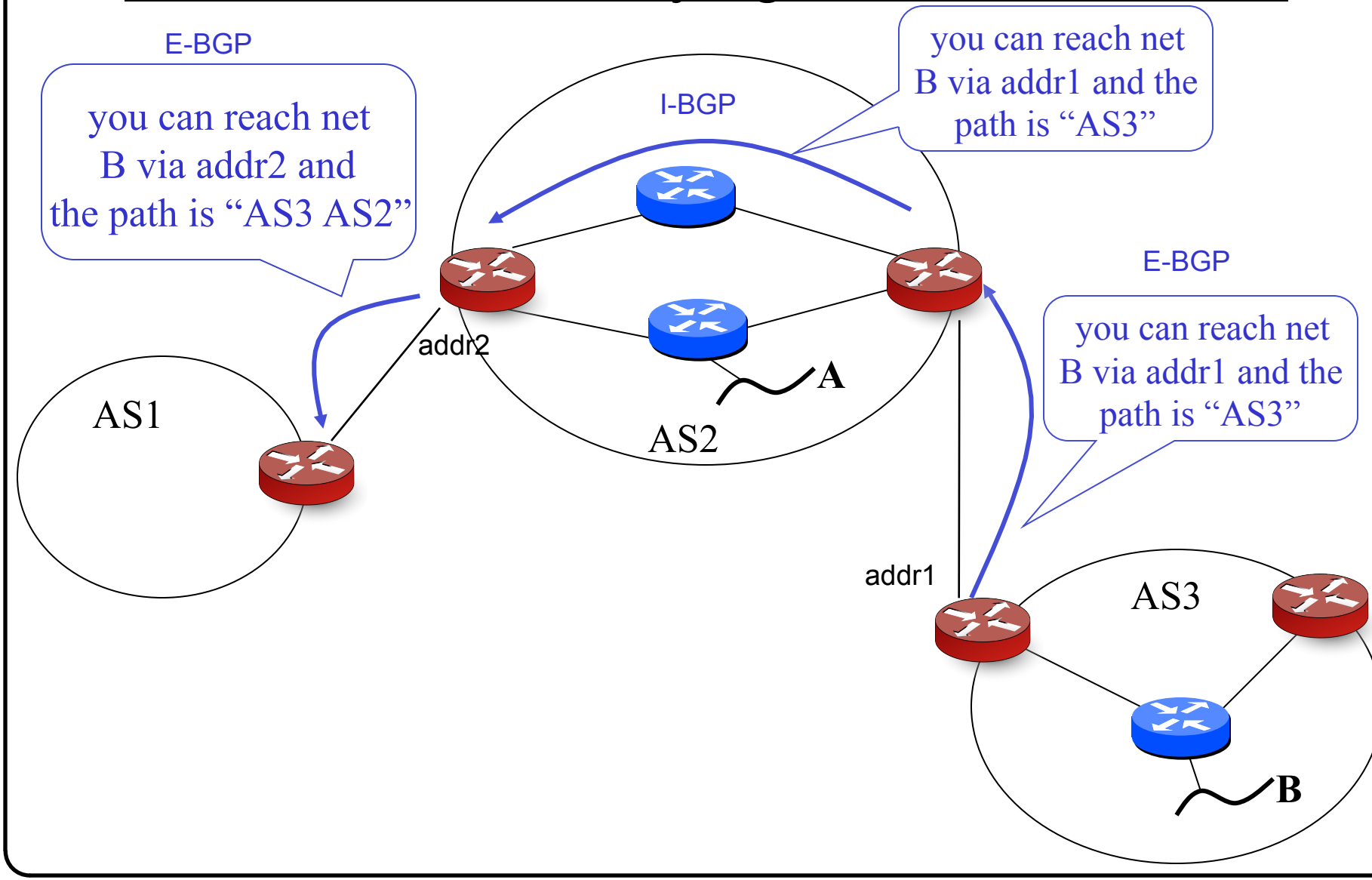
dest network prefix	AS path	next hop
A	AS2	addr

Part II: I-BGP, Carrying Info within an AS

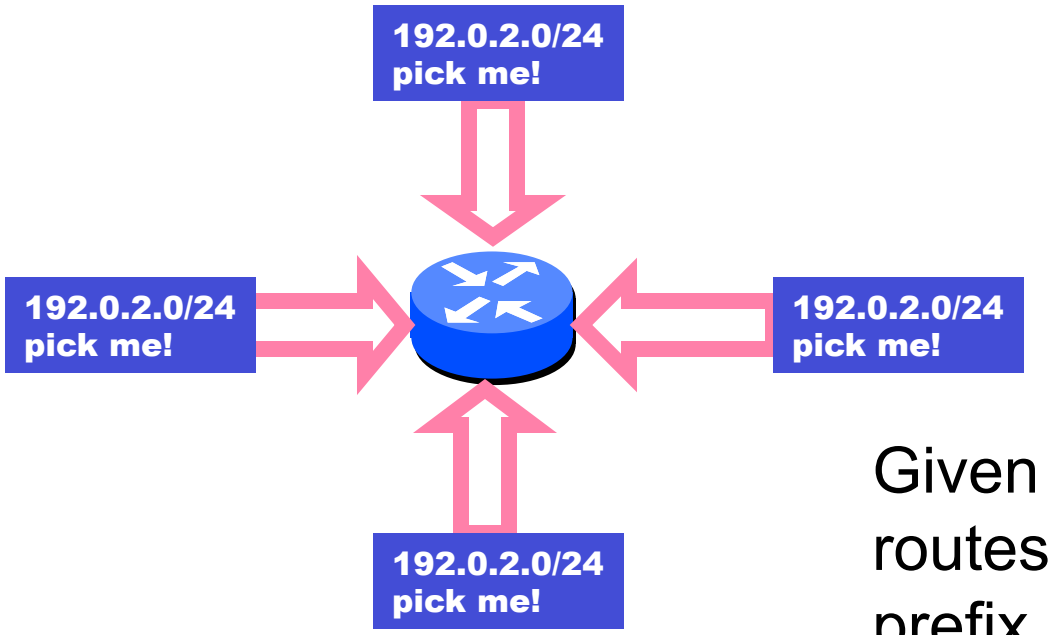


- I-BGP used to disseminate learned routes to all routers in AS

Part II: I-BGP, Carrying Info within an AS

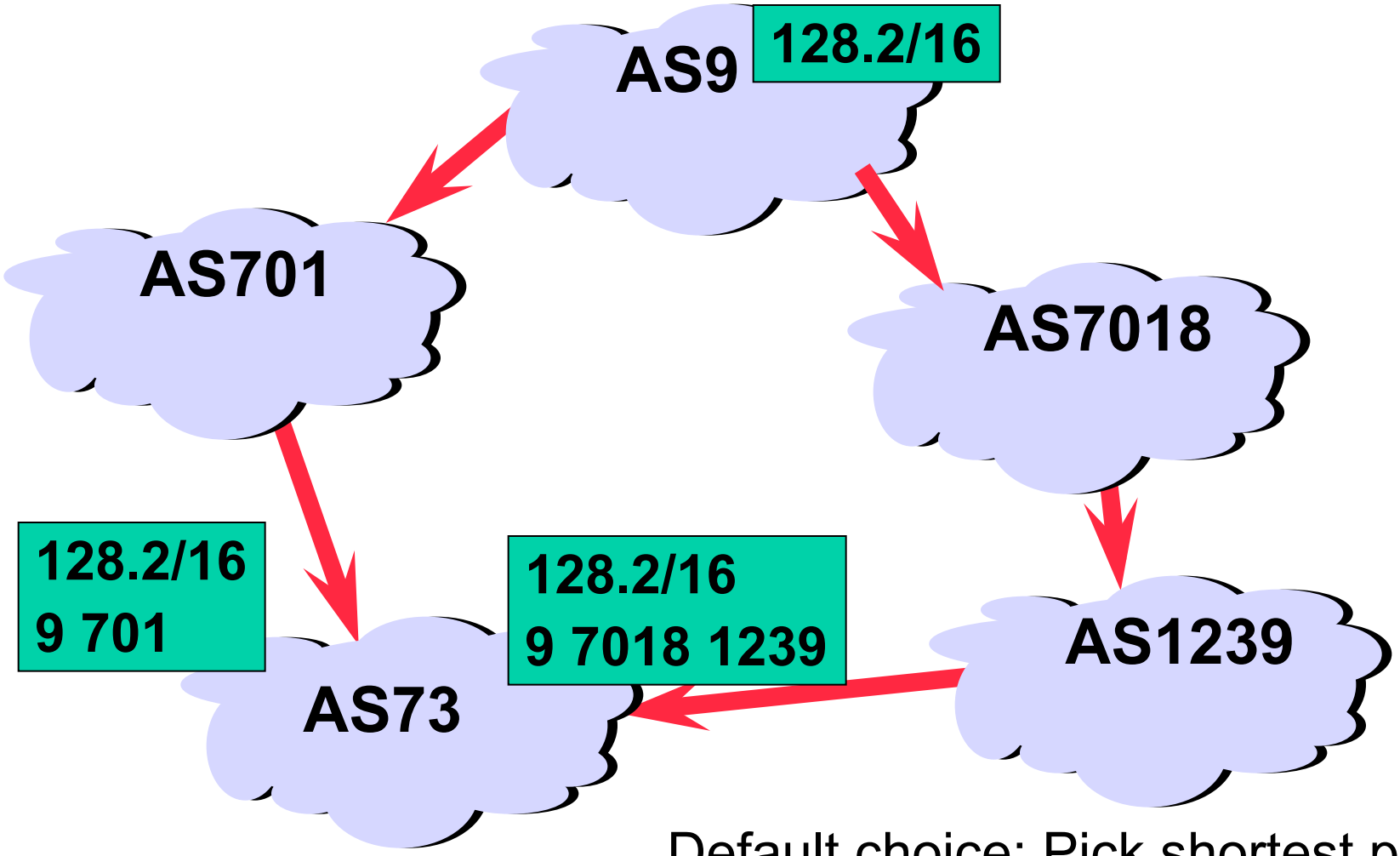


Attributes are Used to Select Best Routes



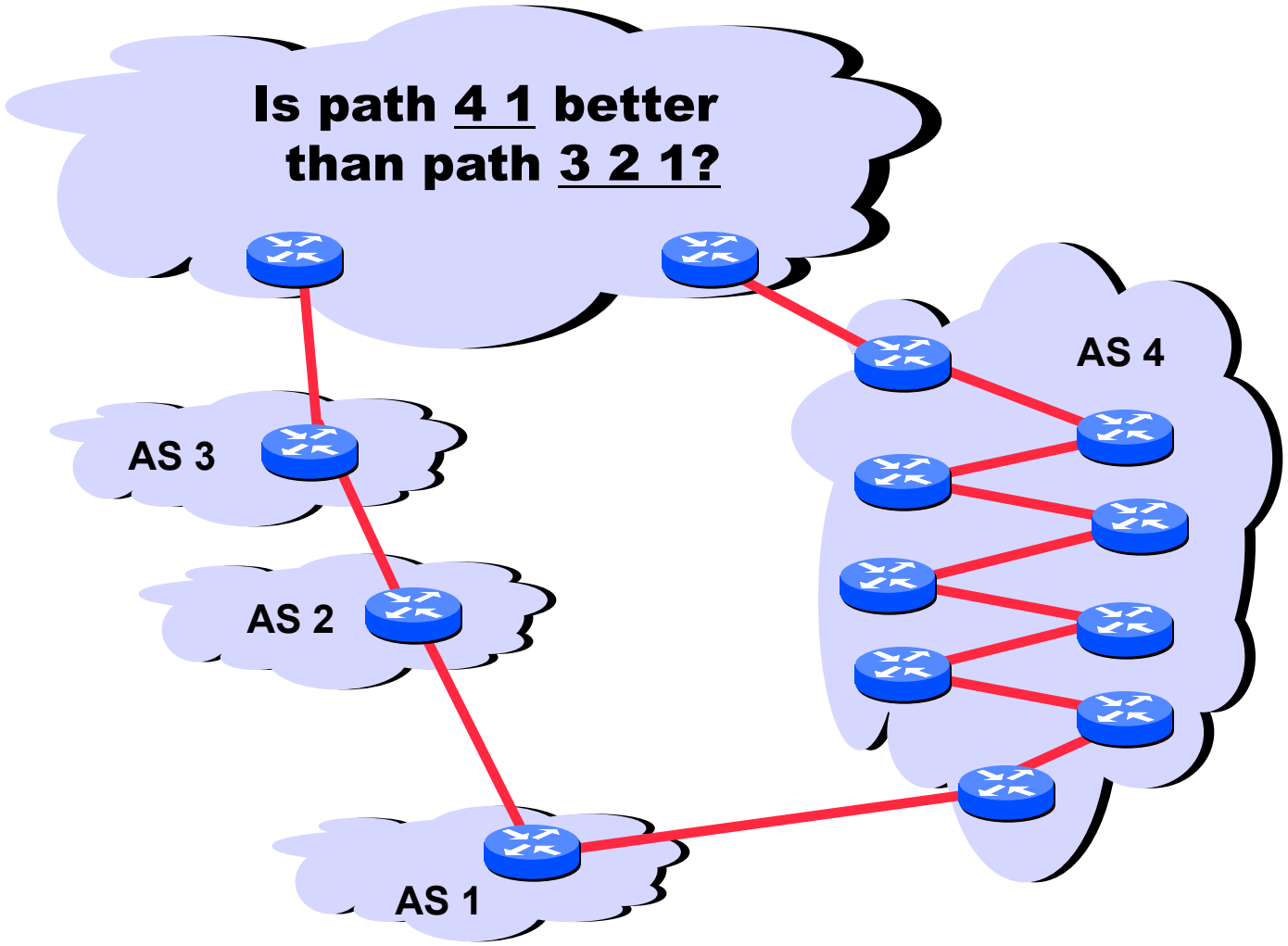
Given multiple routes to the same prefix, a BGP speaker must pick at most one best route

Example: Multiple AS Paths



Default choice: Pick shortest path

Shorter Doesn't Always Mean better



AS can use custom policies other than shortest path

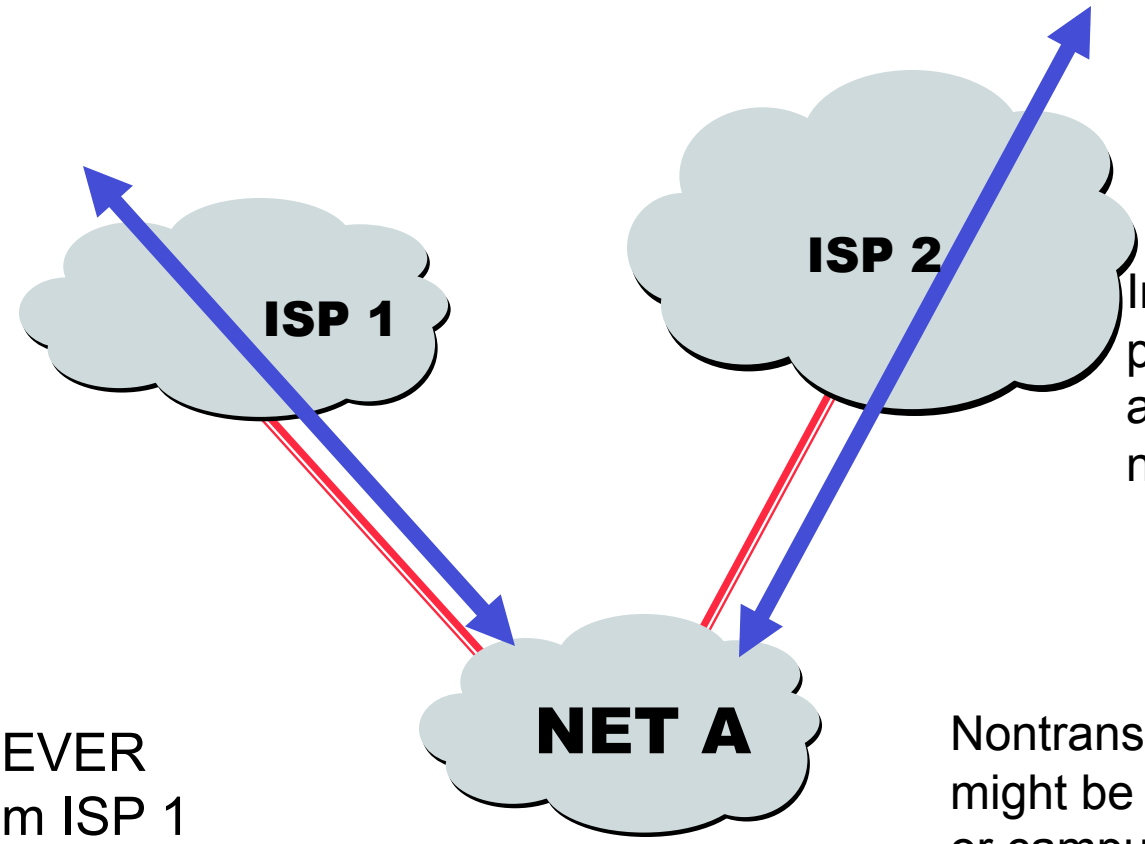
Benefits of BGP Design

- Path Vector style routing
  - Distance vector algorithm with extra information
  - For each route, store the complete path (ASs)
- Advantages:
  - can make policy choices (choose among many possible learned paths) based on set of ASs in path
  - can easily avoid loops

Announcing and Choosing Routes

- BGP may learn many different paths for a destination network
- Learns only reachability information, no performance metrics
  - Not about optimizing anything
  - All about policy (business and politics)
- What a BGP speaker announces or not announces to a neighbor determines what routes may get used by that neighbor
- Router chooses among paths based on policy

Nontransit vs. Transit ASes



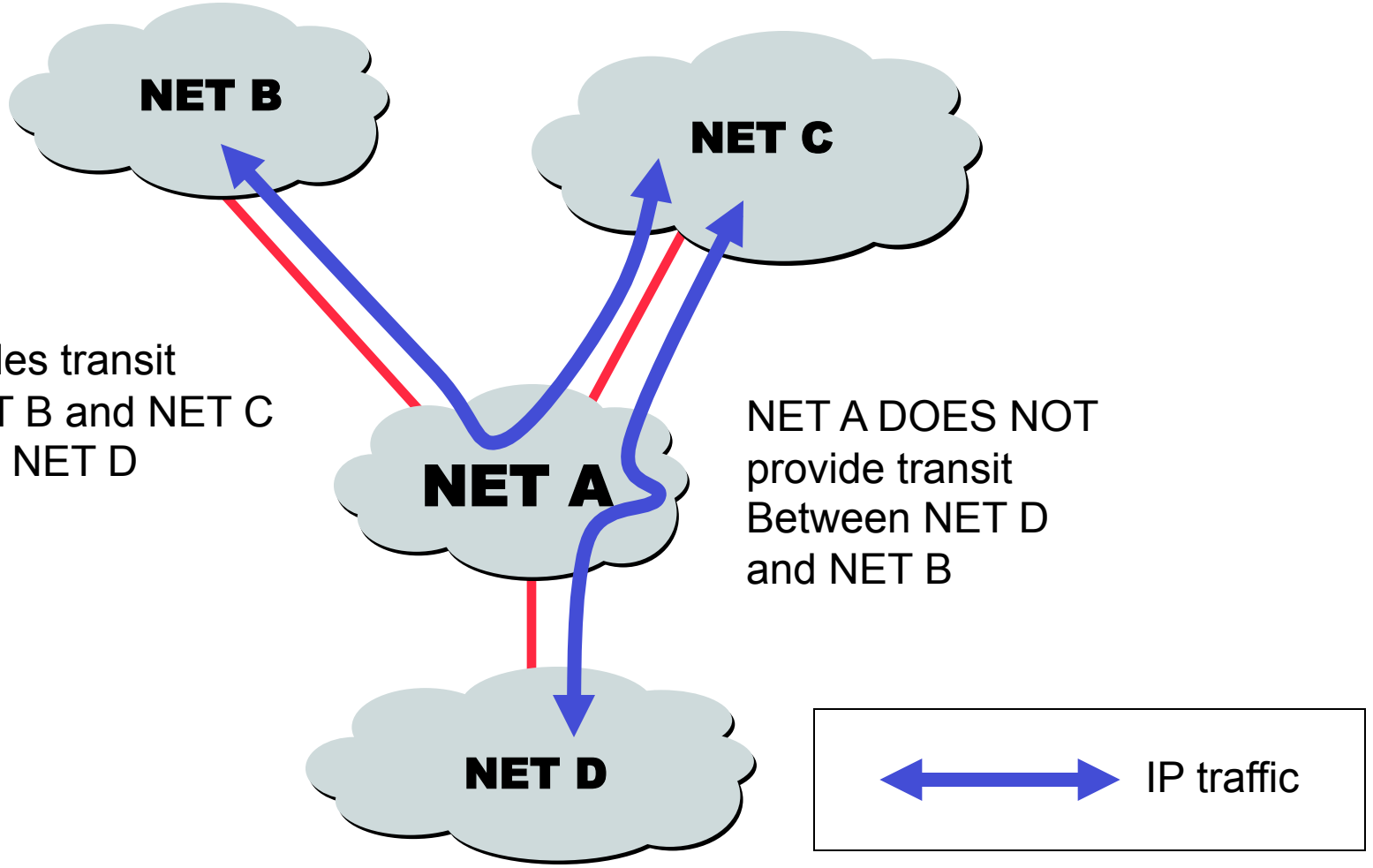
Internet Service providers (often) are transit networks

Traffic NEVER flows from ISP 1 through NET A to ISP 2 (At least not intentionally!)

Nontransit AS might be a corporate or campus network.

IP traffic

Selective Transit

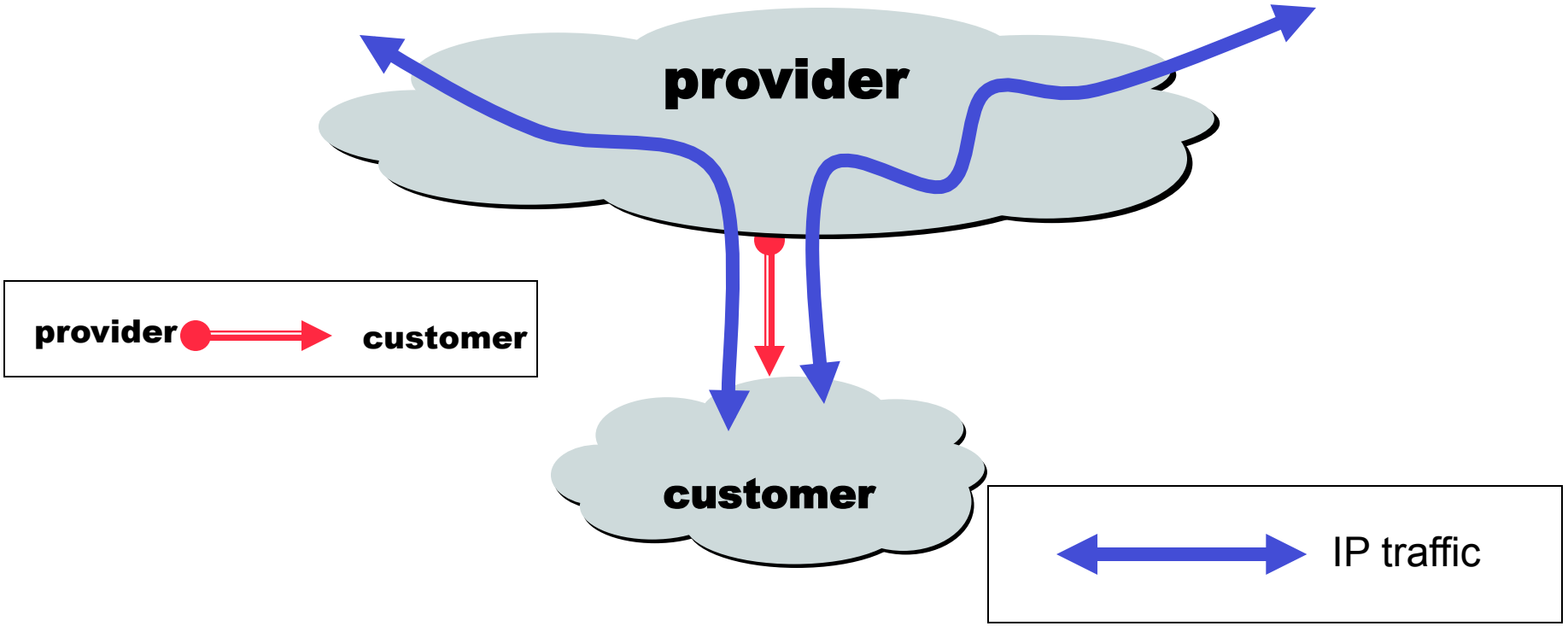


NET A provides transit between NET B and NET C and between NET D and NET C

NET A DOES NOT provide transit Between NET D and NET B

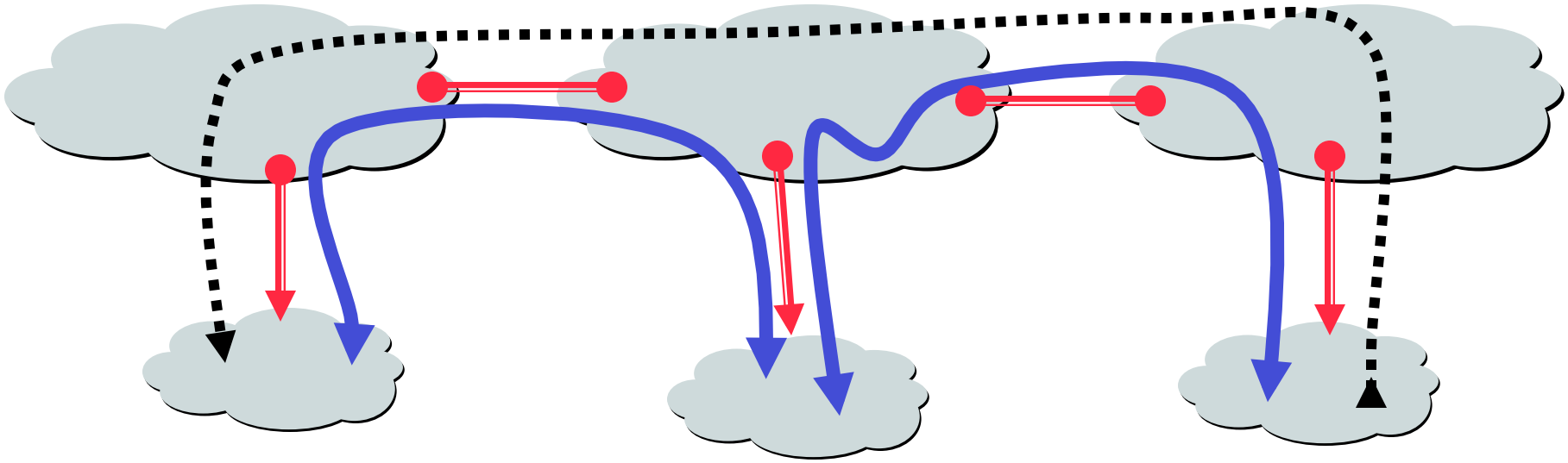
Most transit networks transit in a selective manner…

Customers and Providers



Customer pays provider for access to the Internet

The Peering Relationship



peer — peer

provider → customer

traffic allowed

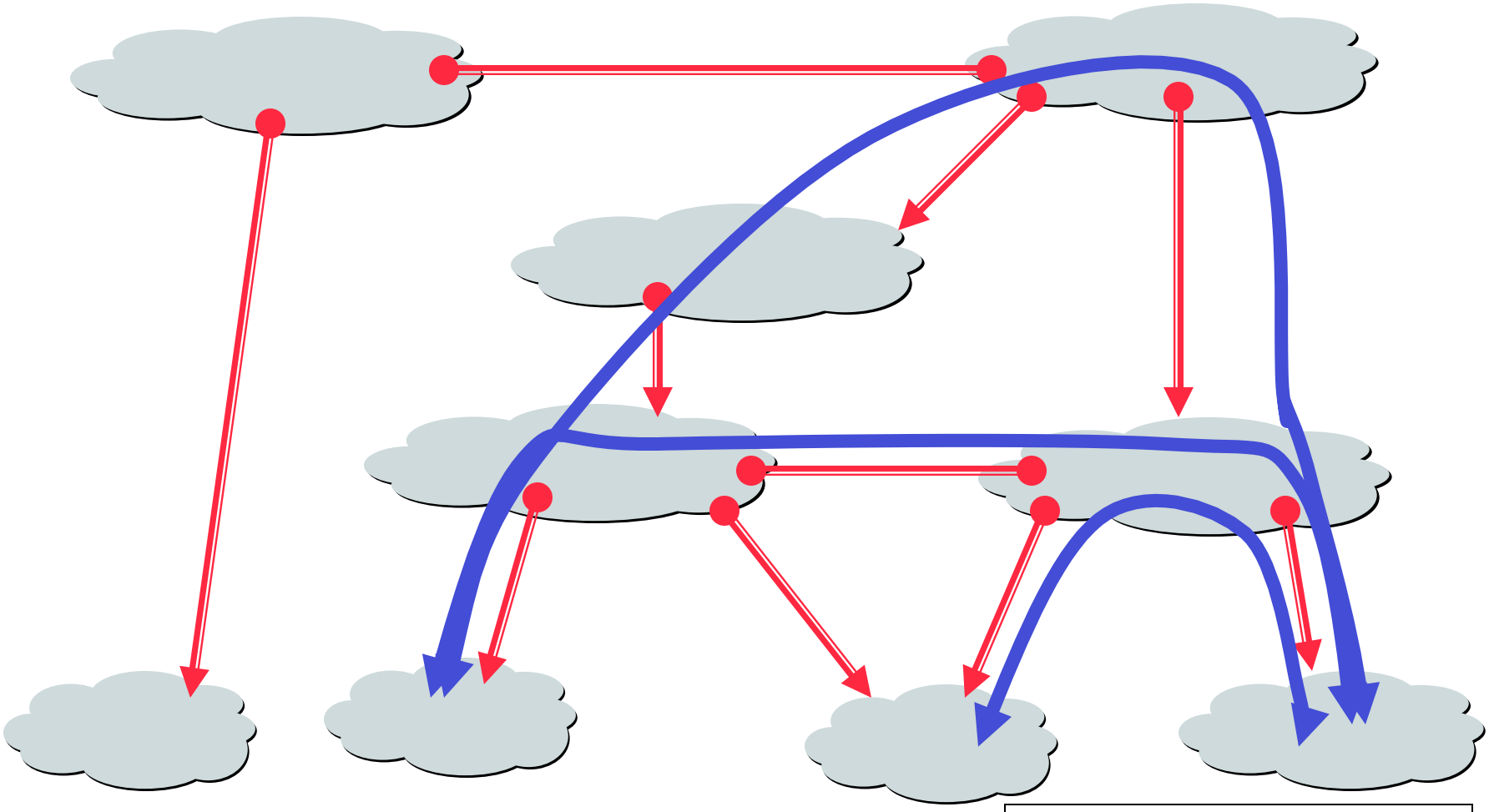
traffic NOT allowed

Peers provide transit between their respective customers

Peers do not provide transit between peers

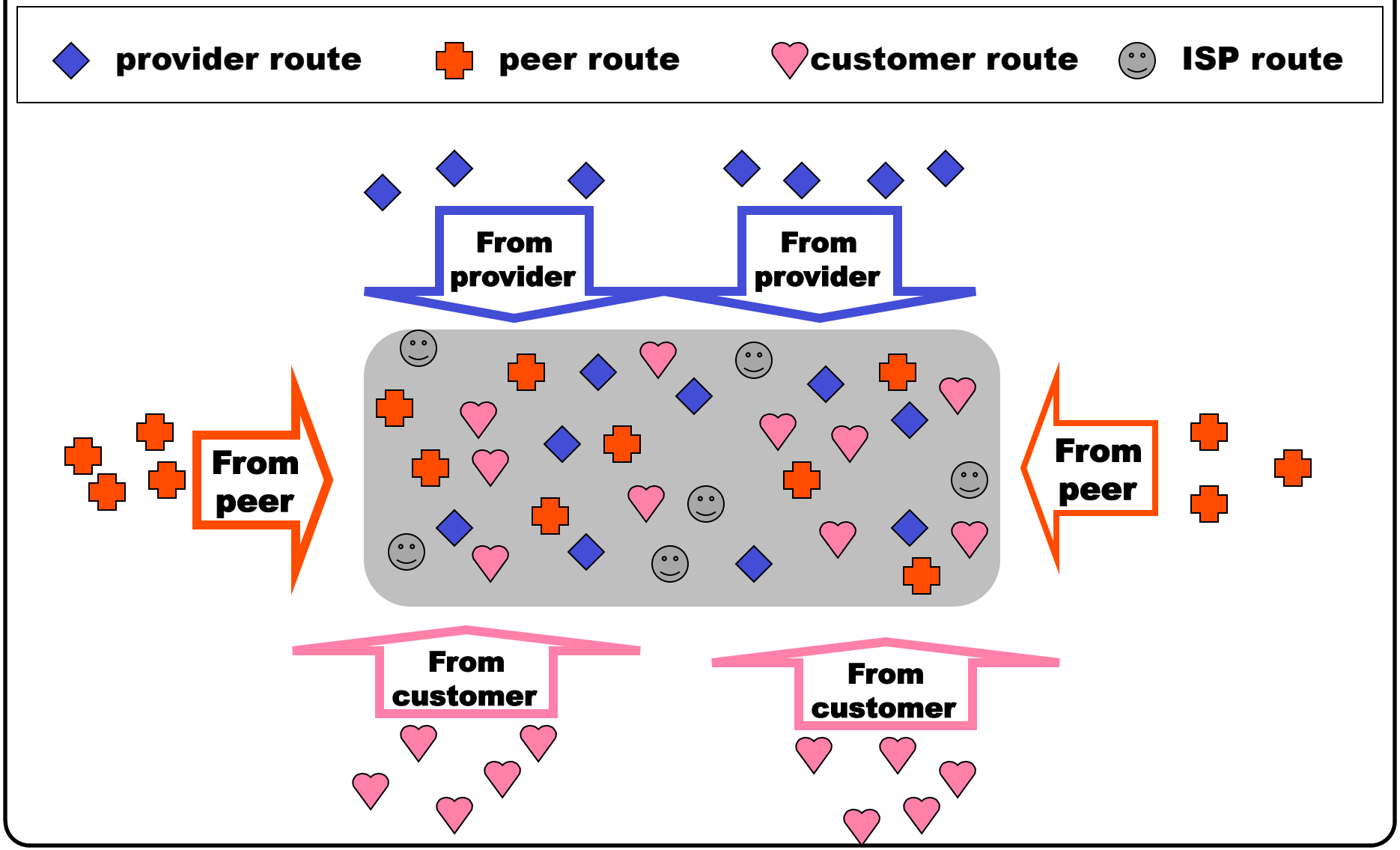
Peers (often) do not exchange $$$

Peering Provides Shortcuts

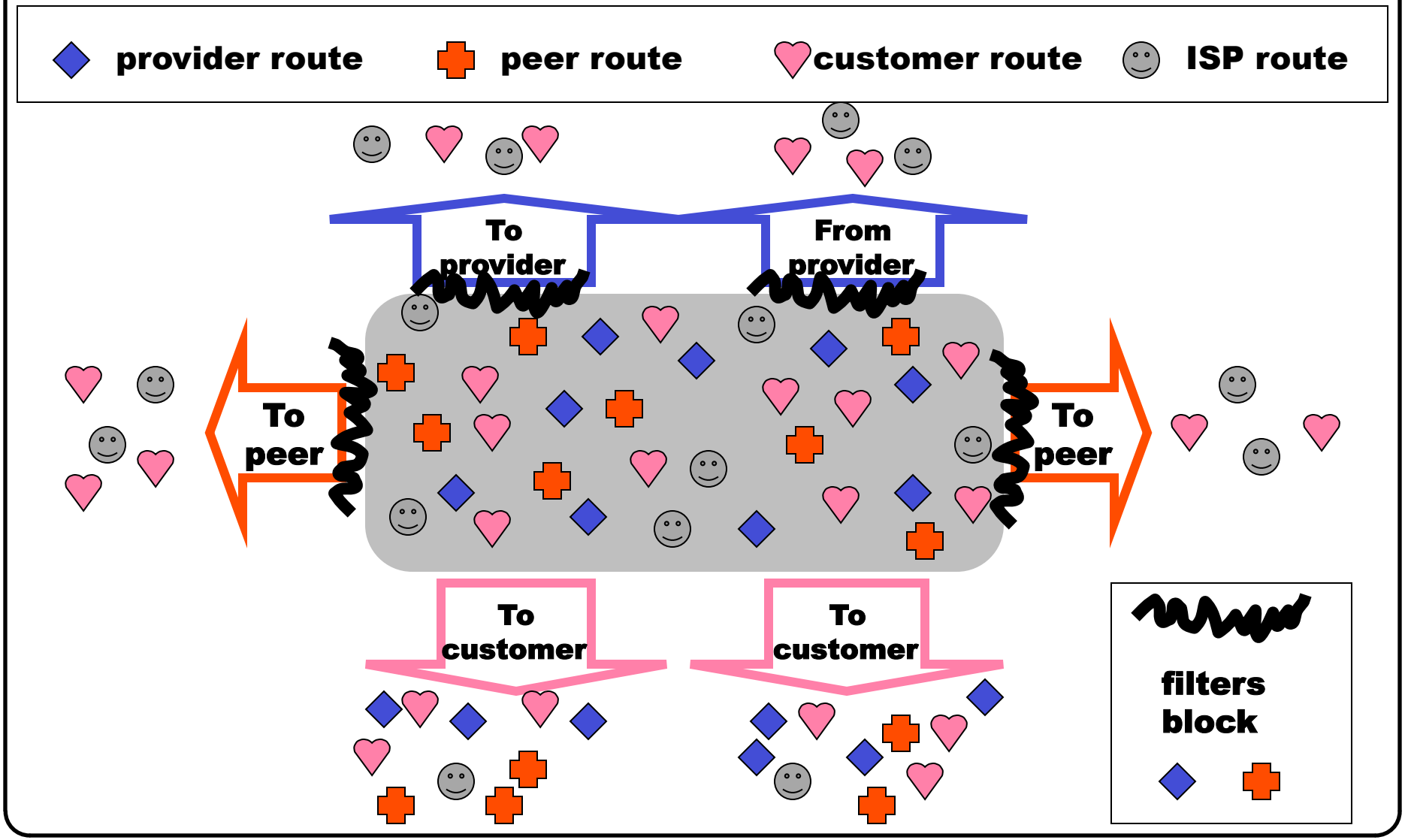


Peering also allows connectivity between the customers of "Tier 1" providers.

Import Routes



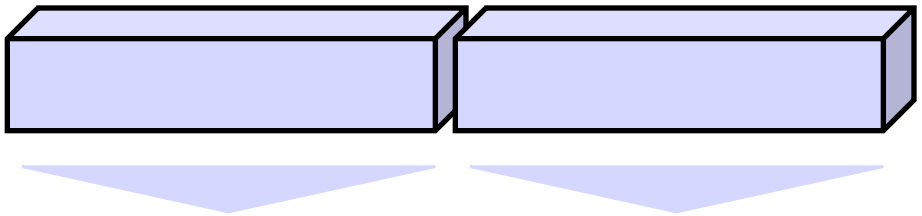
Export Routes



How can routes be marked as “provider”, “peer”, “customer”, “isp” ?

Use “Community Attribute” in route announcement

A community attribute is 32 bits



By convention, first 16 bits is ASN indicating who is giving it an interpretation

community number

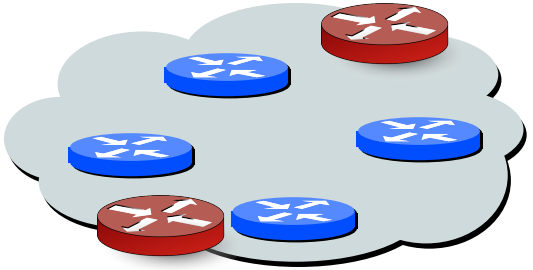
Used for signaling within and between ASs

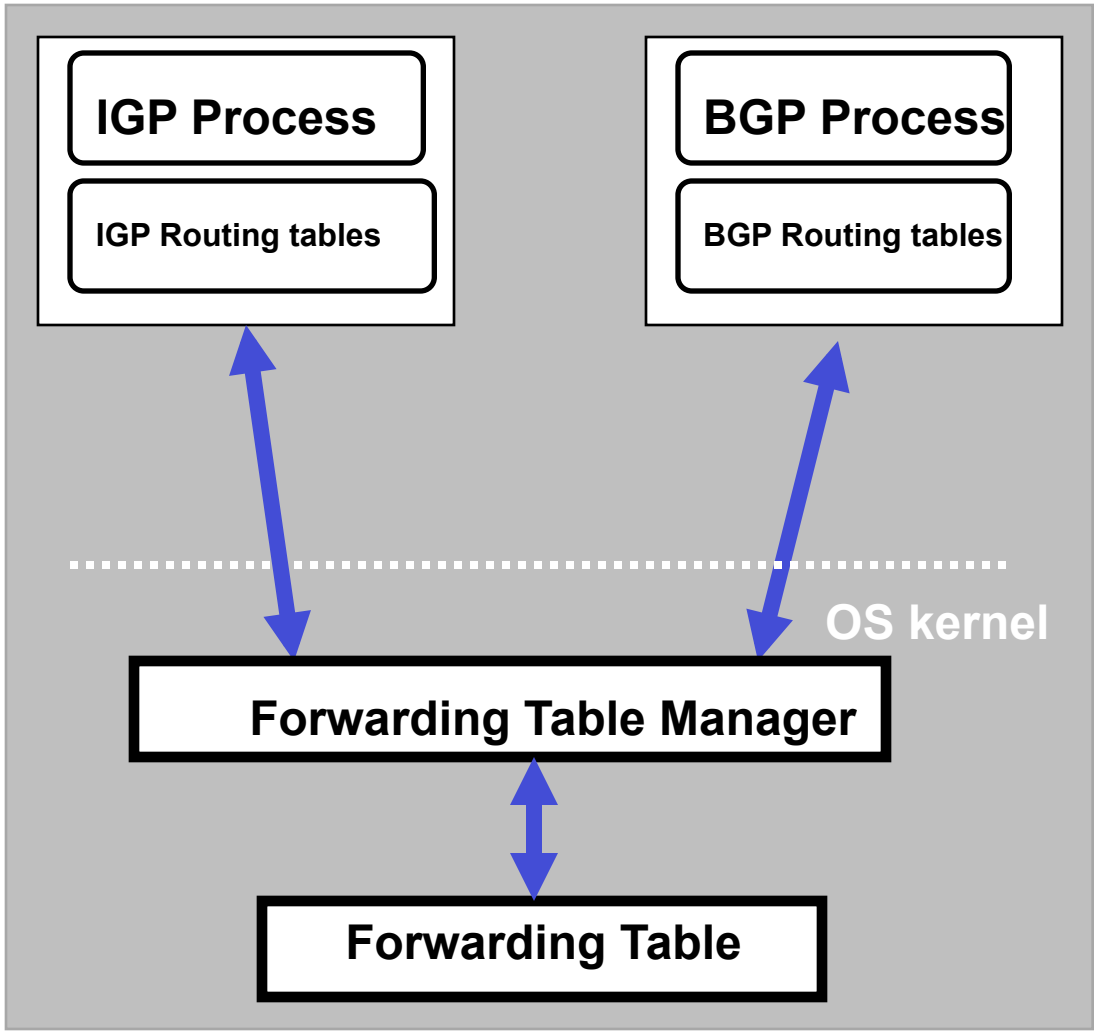
Very flexible BECAUSE it has no predefined meaning

BGP Issues

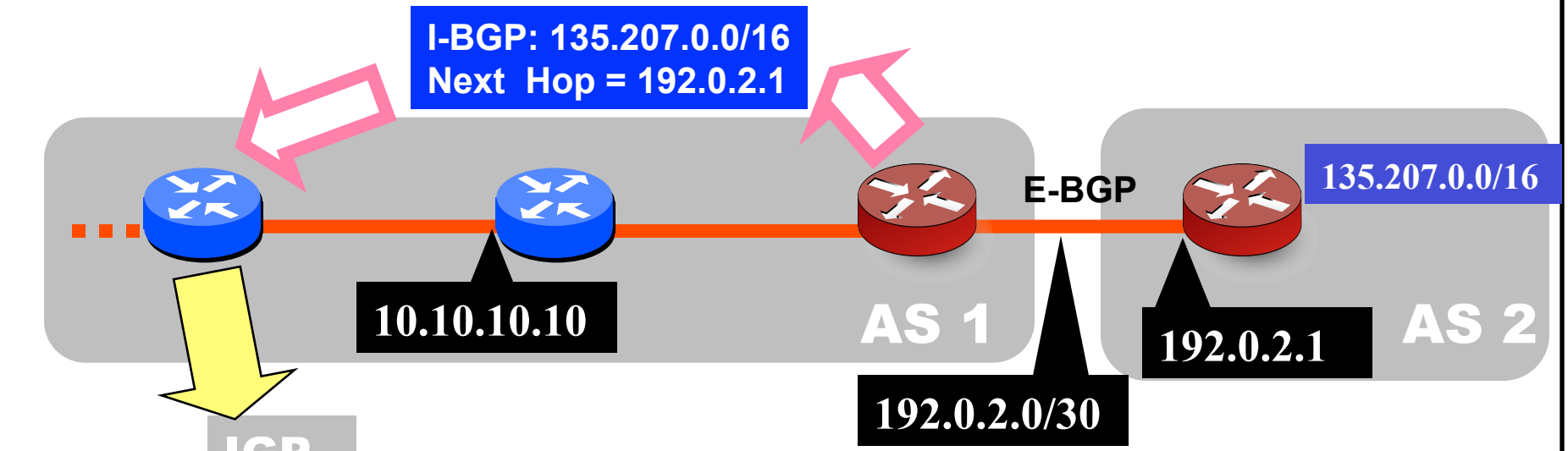
- BGP designed for policy not performance
- Susceptible to router misconfiguration
  - Blackholes: announce a route you cannot reach
- Slow convergence time
  - Rate limiting and route flap dampening

Combining IGP and BGP





Combine BGP and IGP Tables to Create Forwarding Table



IGP

destination	next hop
192.0.2.0/30	10.10.10.10

+

BGP

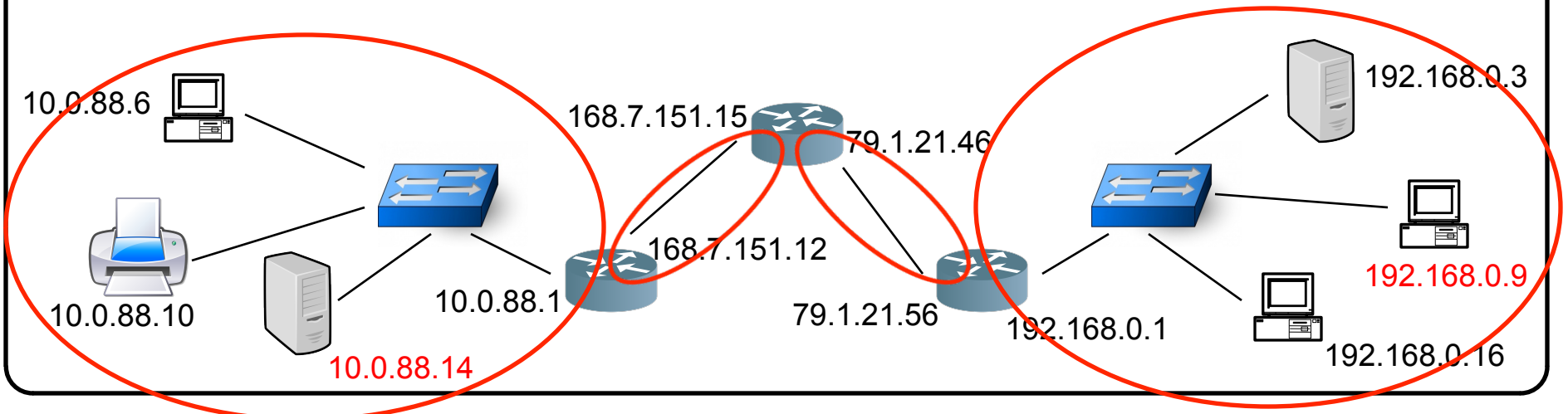
destination	next hop
135.207.0.0/16	192.0.2.1

Forwarding Table

destination	next hop
135.207.0.0/16	10.10.10.10
192.0.2.0/30	10.10.10.10

Life Cycle of a Packet in the Internet

- Address Resolution Protocol (ARP)
  - On the same subnet, need to map IP address to MAC (e.g. Ethernet) address
  - Host and router have ARP cache to store the IP-MAC pairs
  - In case of no match in ARP cache, broadcast an ARP request with the IP address in question and the device with the IP address will reply with its MAC address



Life Cycle of a Packet in the Internet

For each hop in the network, do the following steps:

1. Decapsulate the Ethernet frame to get the IP header (except no need to do this at the source)
2. Check routing table by the destination IP address, get the next-hop IP address and the network interface
3. Learn the MAC address of the next hop (look up in ARP cache or broadcast an ARP request)
4. Encapsulate the IP packet into an Ethernet frame with the destination MAC address
5. Send the Ethernet frame out from the next-hop network interface