CS 43: Computer Networks

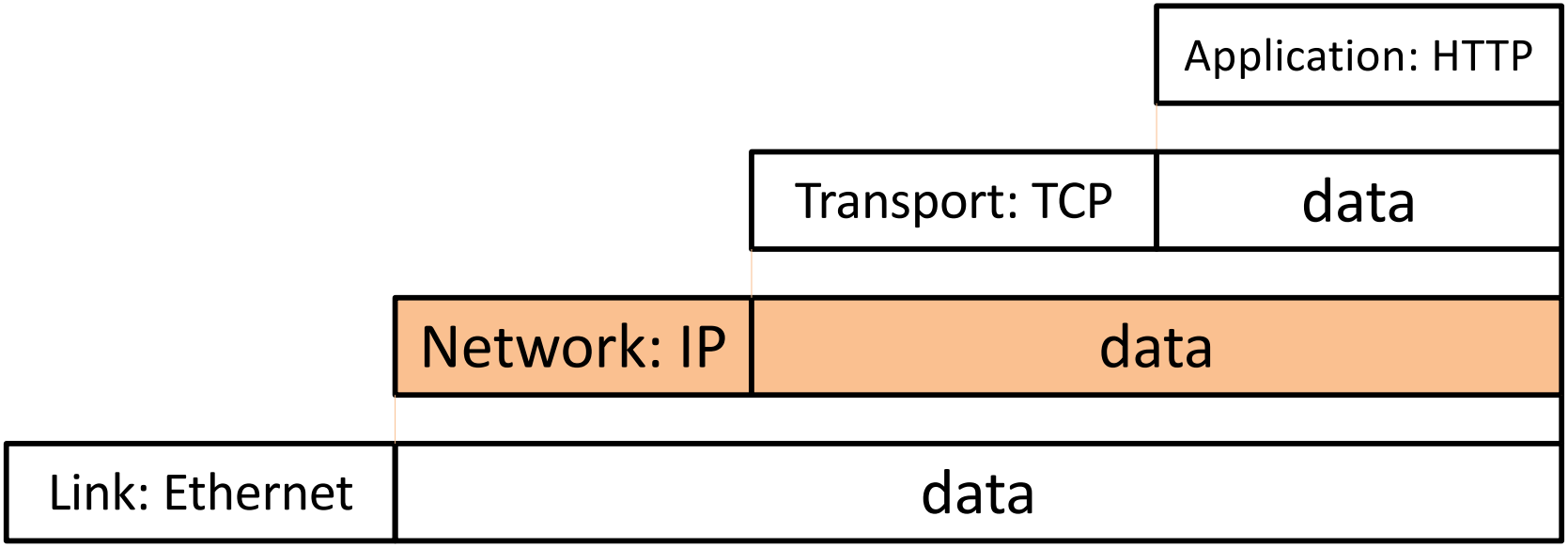
22: Interdomain Routing-Traffic Management
November 26, 2019

Adapted from Slides by: J.Kurose, D. Choffnes, K. Webb

Reading Quiz

Network Layer

- Function: Route packets end-to-end on a network, through multiple hops



Network Layer Functions

- Forwarding*:* move packets from router's input to appropriate router output
  - Look up in a table
- Routing: determine route taken by packets from source to destination.
  - Populating the table

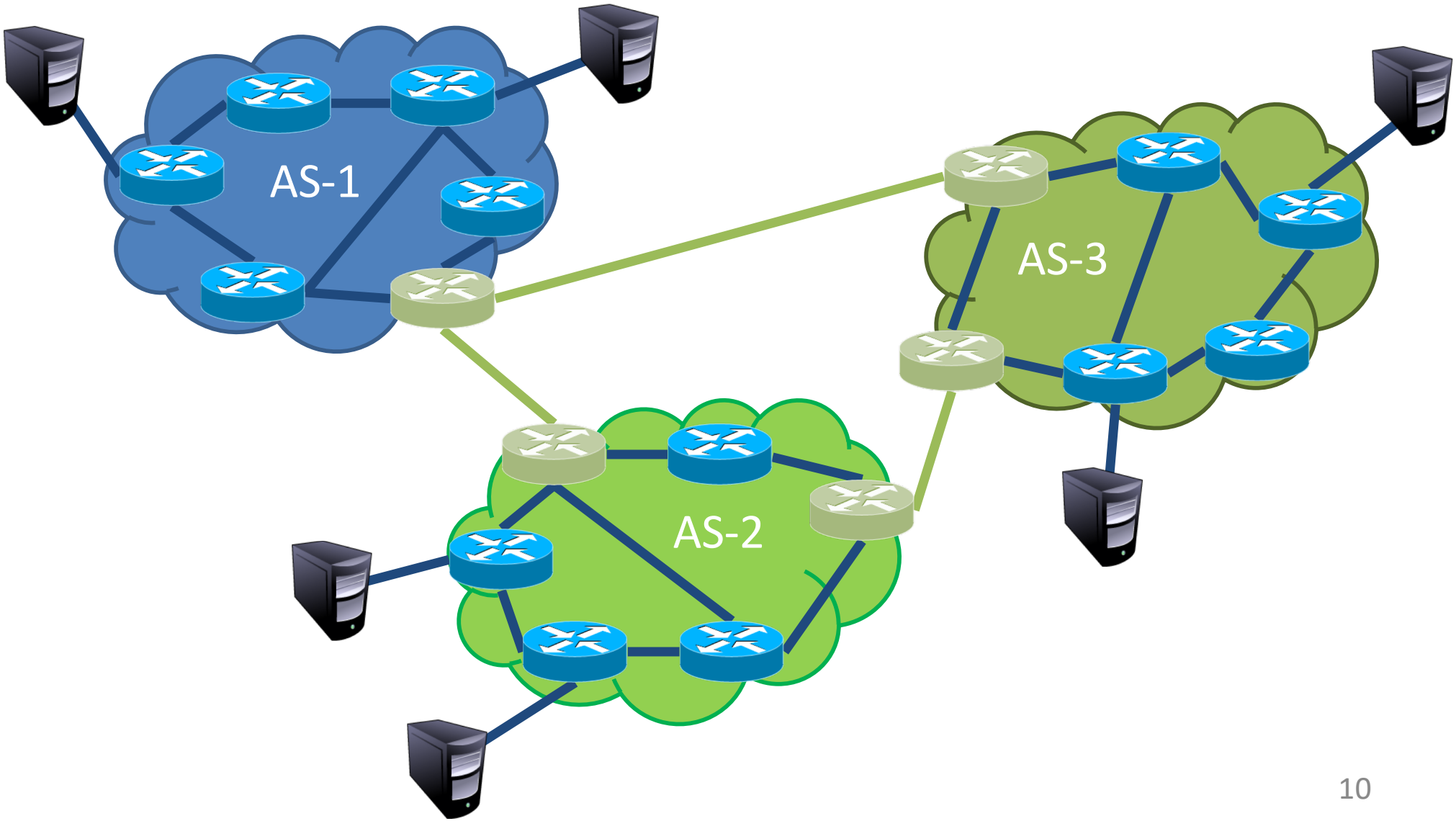
Why do we need different Intra and Interdomain AS routing ?

A. Scalability
B. Performance
C. A and B
D. More than just A and B

Internet/inter-AS Routing

Goal:
Get traffic from one AS to another.

Hierarchical routing: Autonomous Systems



The Inter-domain routing protocol, needs to be an agreed upon protocol across all Autonomous Systems

A. Yes, for inter-operability
B. Not necessarily, but reduces overhead
C. No, each AS can have its own inter-domain routing protocol of choice.

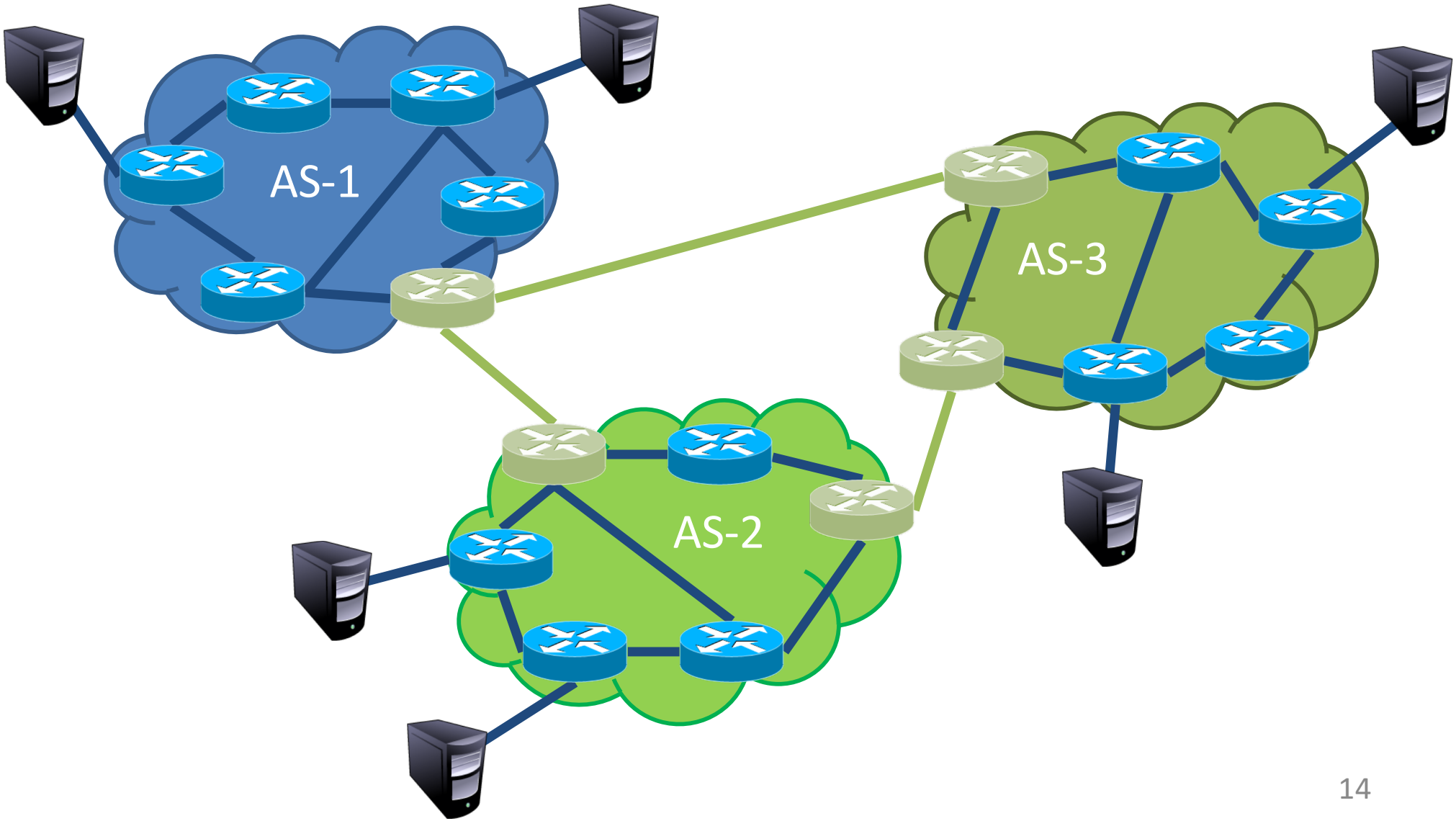
Hierarchical routing

- We aggregate routers into regions, "autonomous systems" (AS)
- Routers in same AS run same routing protocol
  - "intra-AS" or "interior" routing protocol
  - routers in different AS can run different intra-AS routing protocol

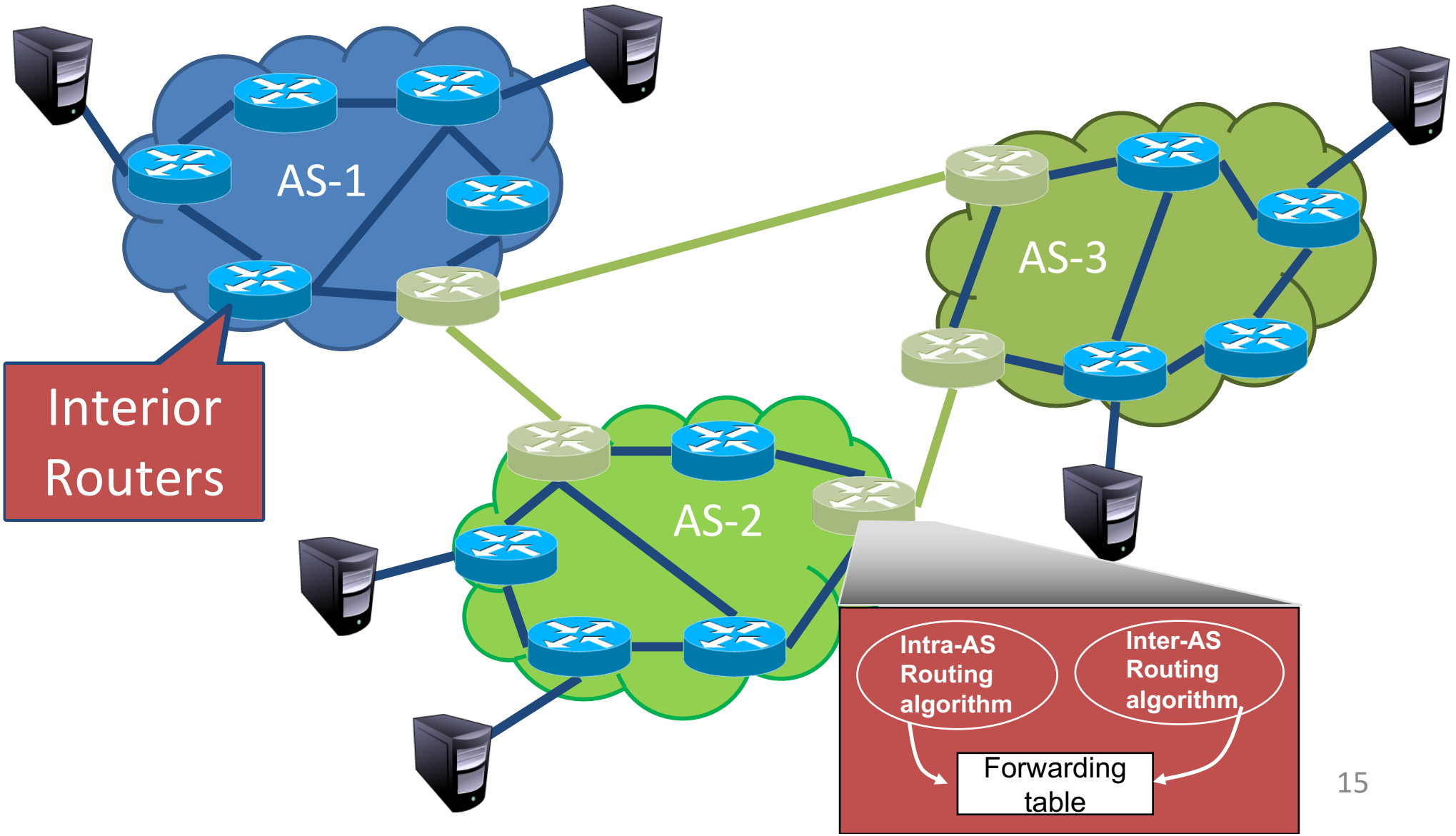
Gateway (or border) router:

- at "edge" of its own AS
- has link to router in another AS

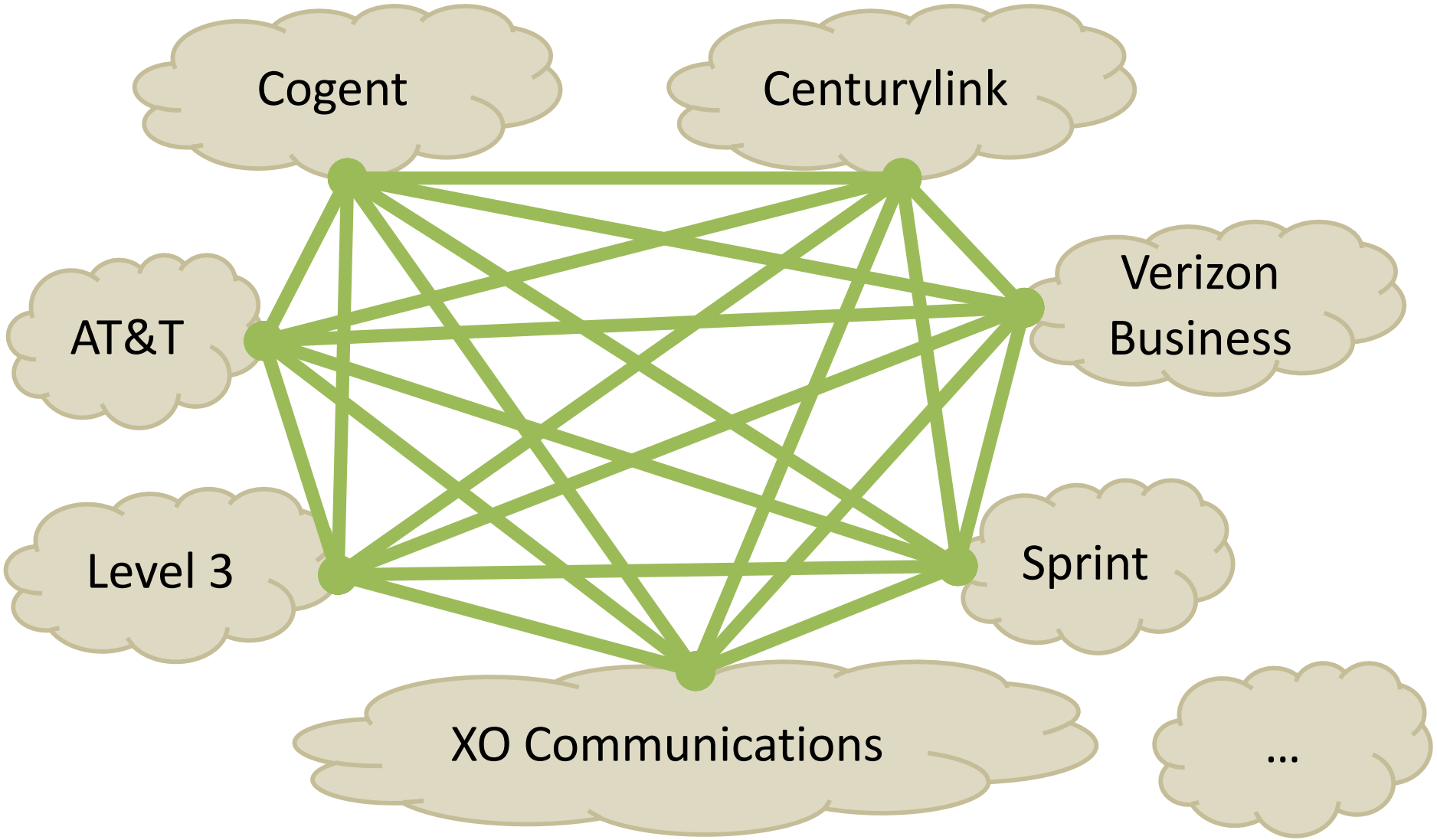
Hierarchical routing: Autonomous Systems



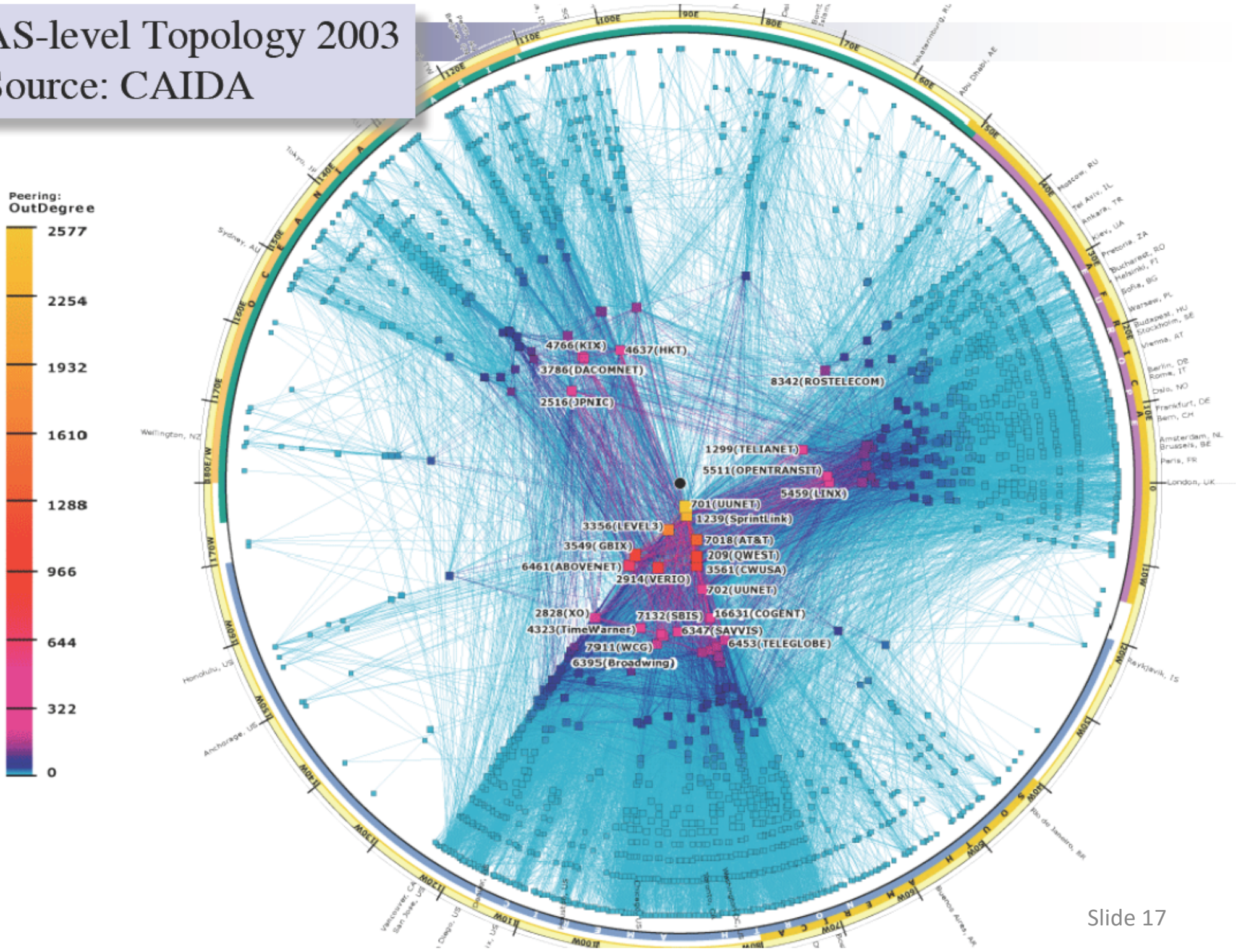
Hierarchical routing: Interconnected ASes



Tier-1 ISP Peering

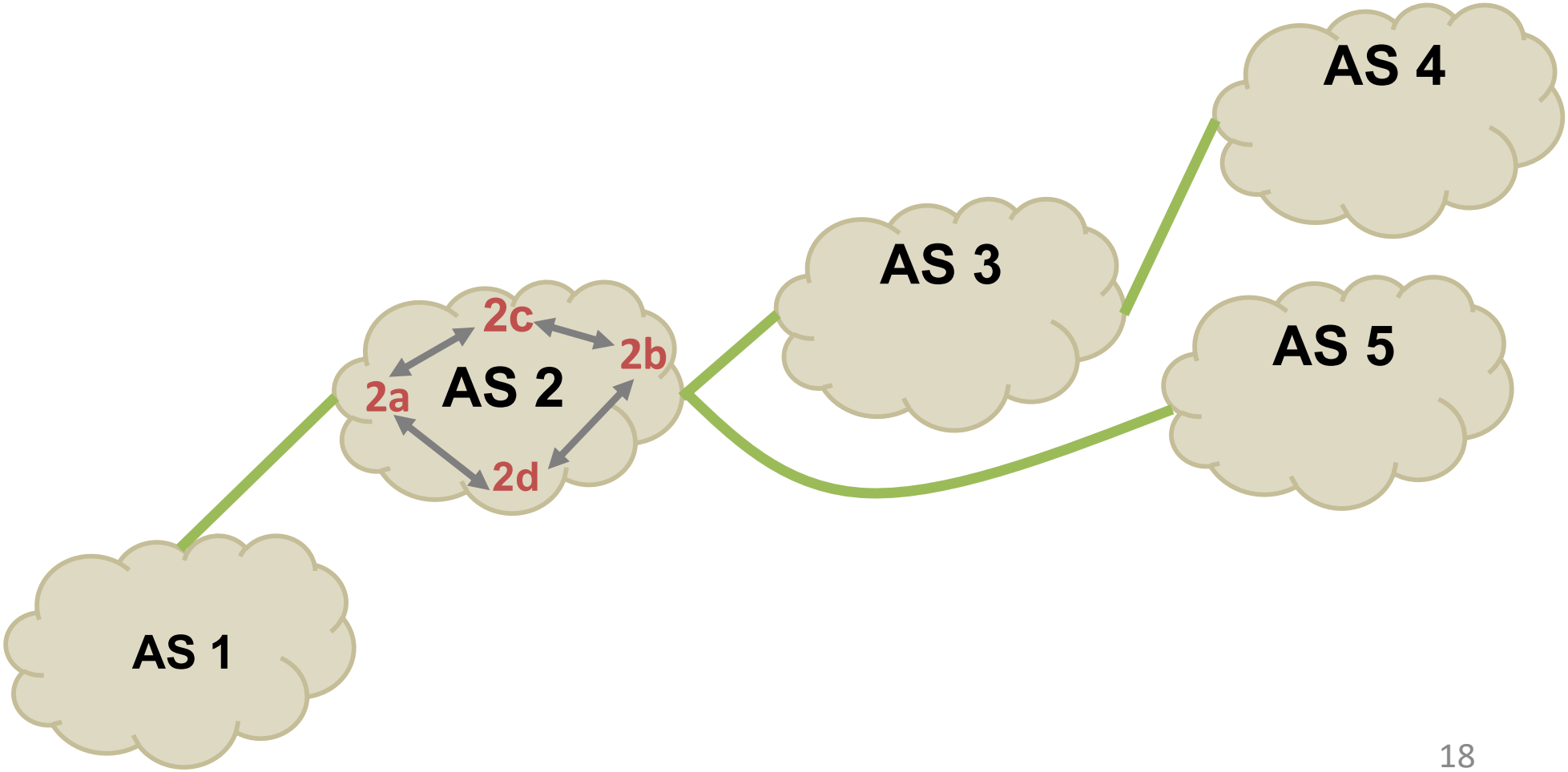


AS-level Topology 2003
Source: CAIDA



Path Vector Protocol

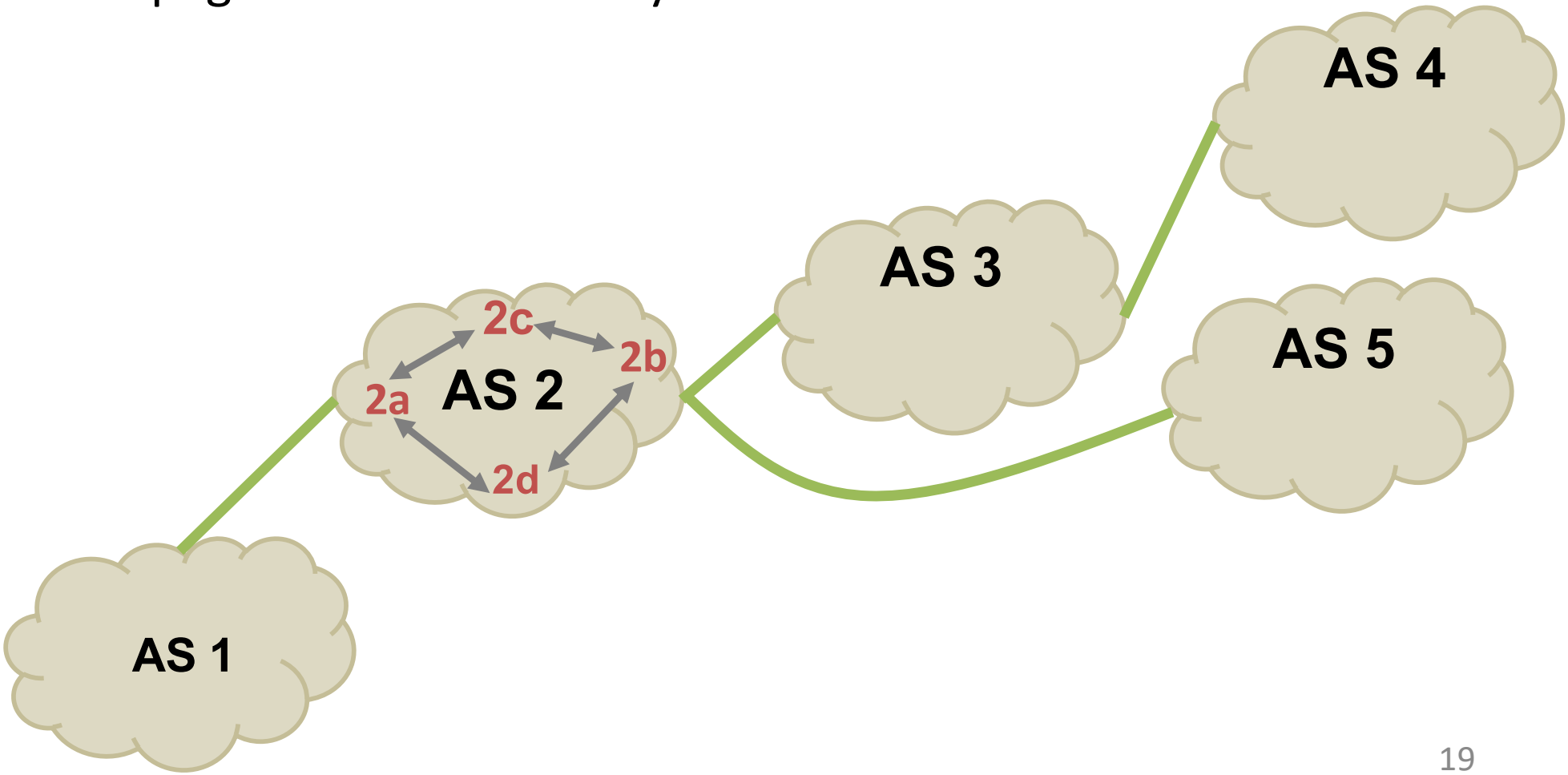
- Key idea: advertise the entire path
  - Distance vector: send *distance metric* per dest d
  - Path vector: send the *entire path* for each dest d



Inter-domain (Inter-ISP) Routing

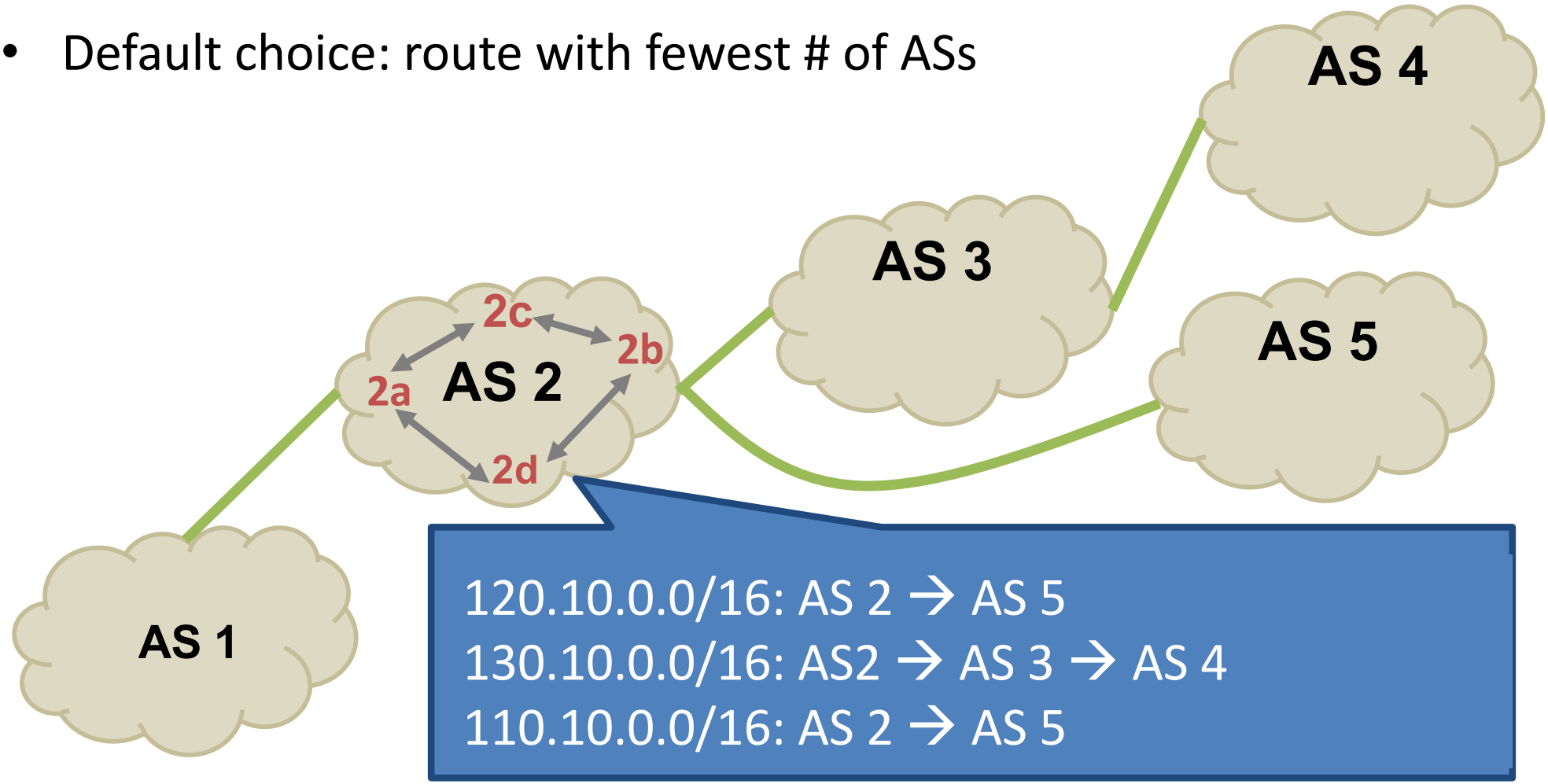
AS2 must:

1. Learn destinations reachable through AS2
2. Propagate this reachability info to all routers in AS2



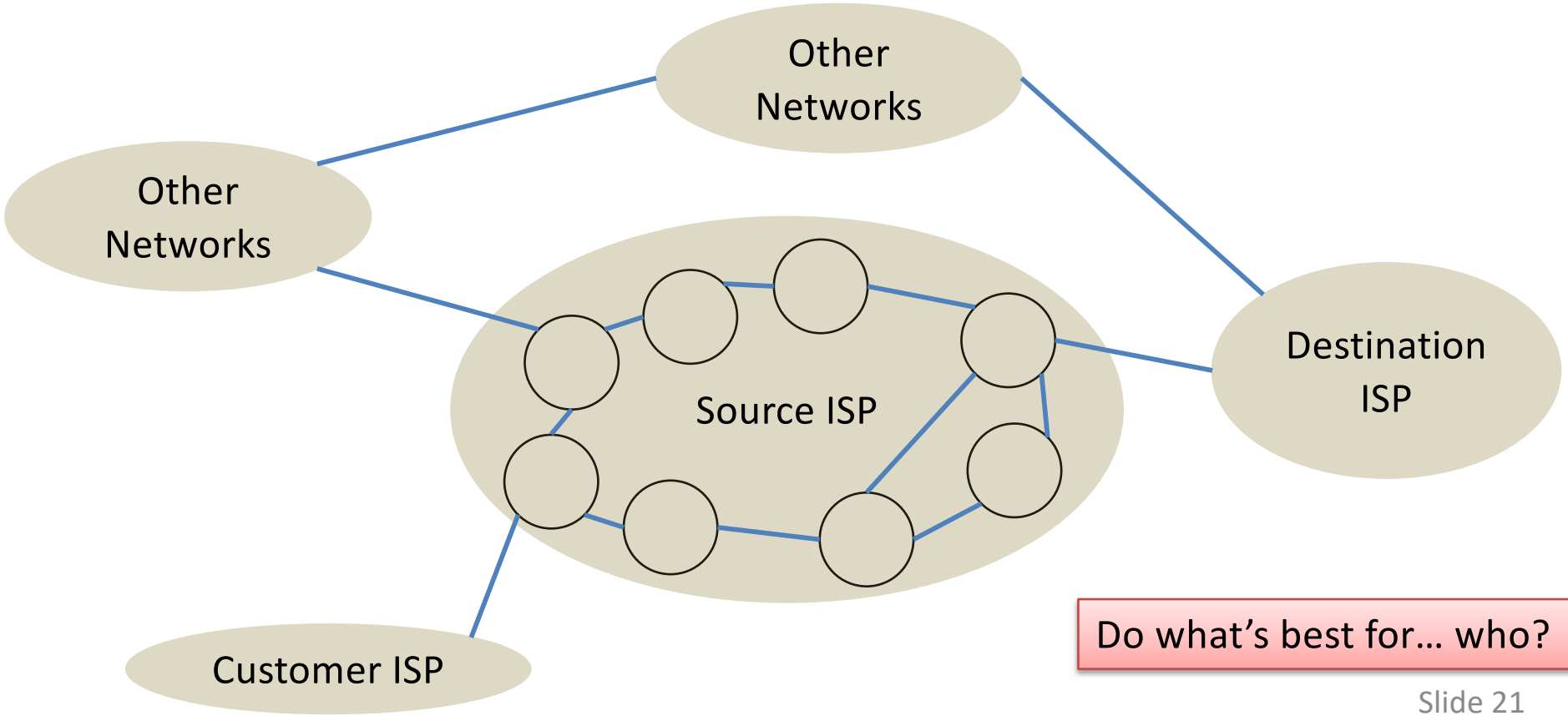
Path Vector Protocol

- AS-path: sequence of ASs a route traverses
  - Like distance vector, plus additional information
- Used for loop detection and to apply policy
- Default choice: route with fewest # of ASs



Routing Policy

- How should the ISP route the customer's traffic to the destination?



Which routes a BGP router advertises will depend on…

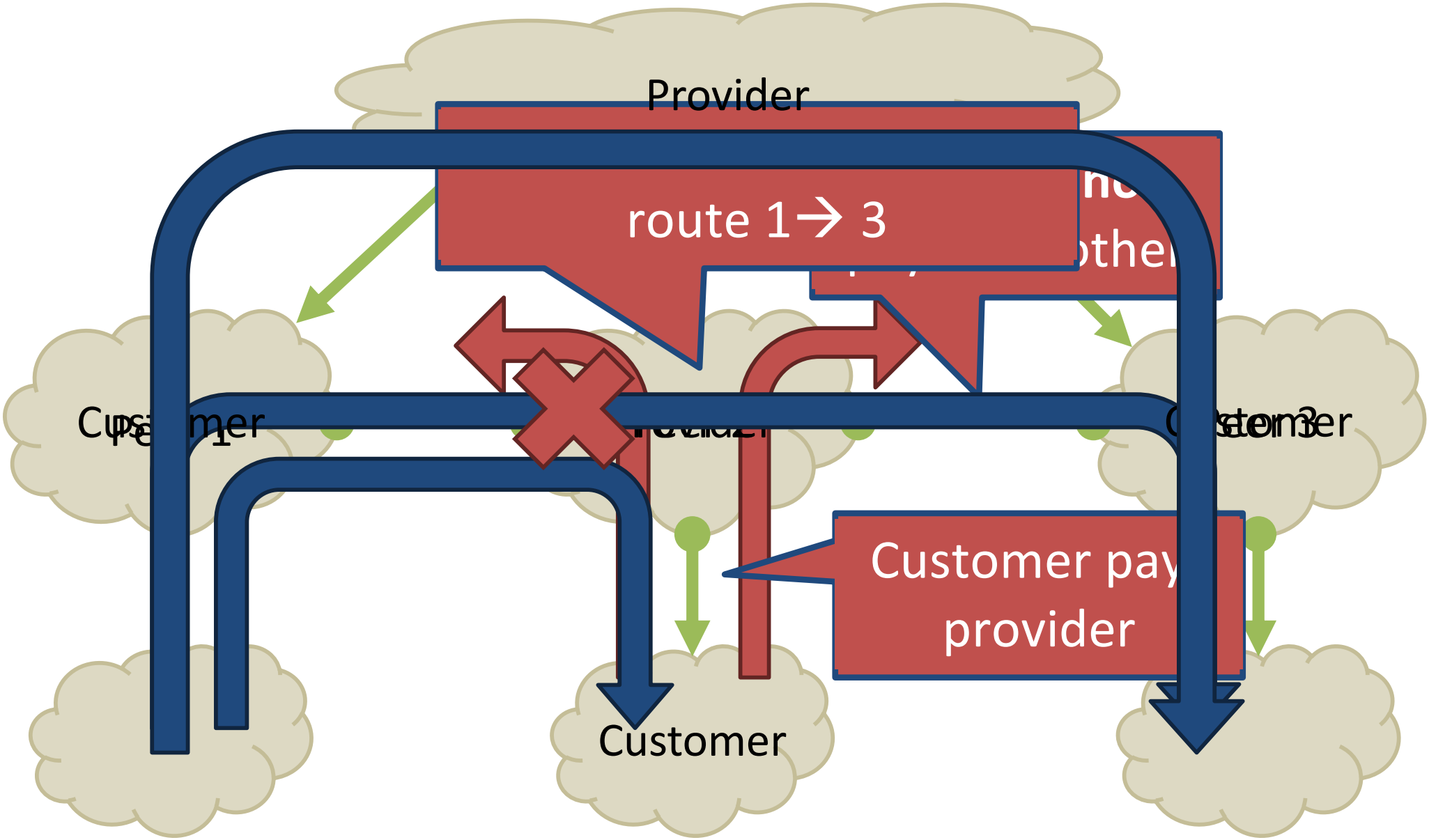
A. which ISPs have contractual agreements.

B. the shortest path to a subnet/prefix.

C. which subnets are customers of an ISP.

D. More than one of the above. (which?)

BGP Relationships



Peering/Interconnection Wars

- Peer

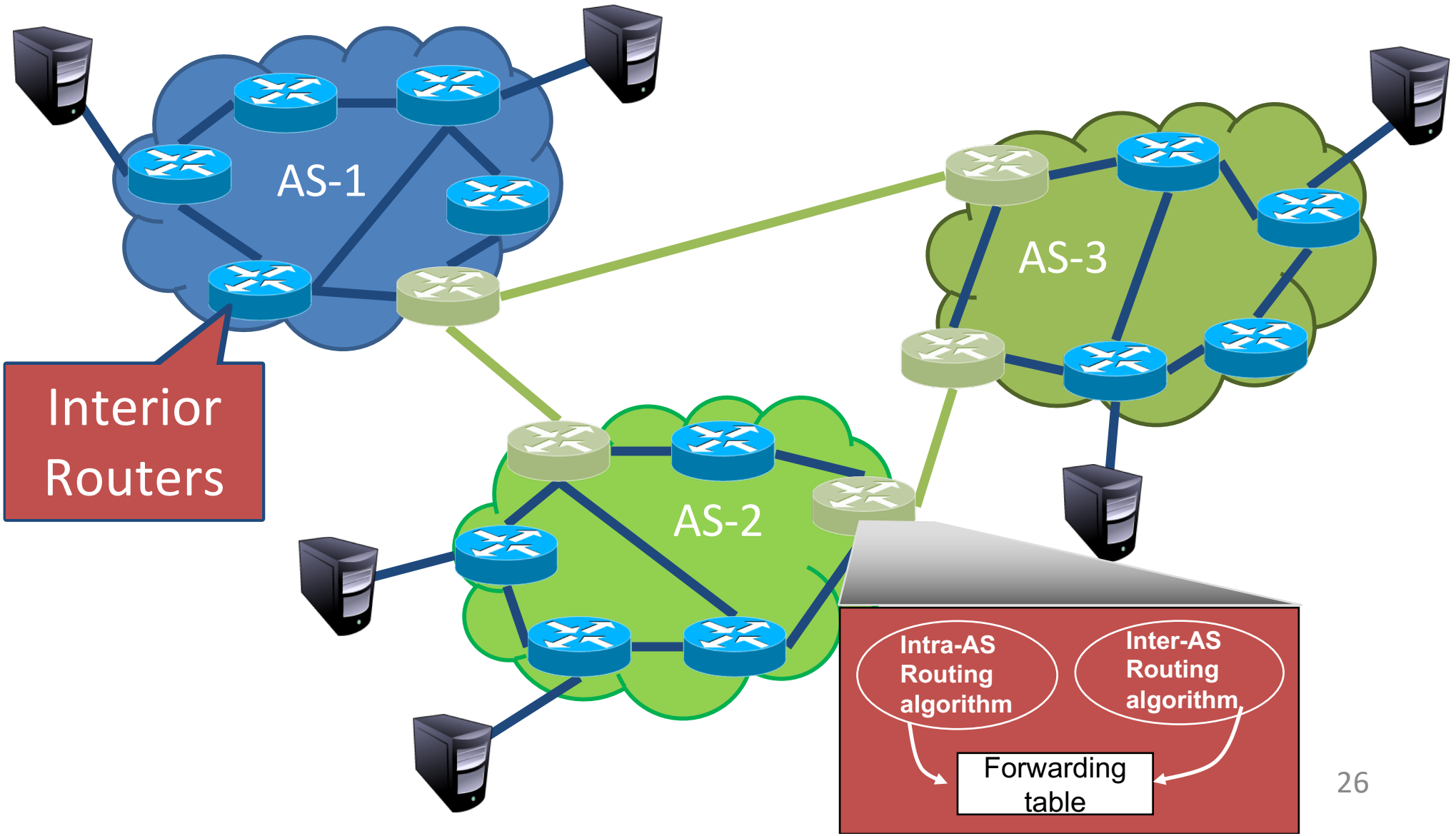
- Reduce upstream costs
- Improve end-to-end performance
- May be the only way to connect to parts of the Internet

- Don't Peer

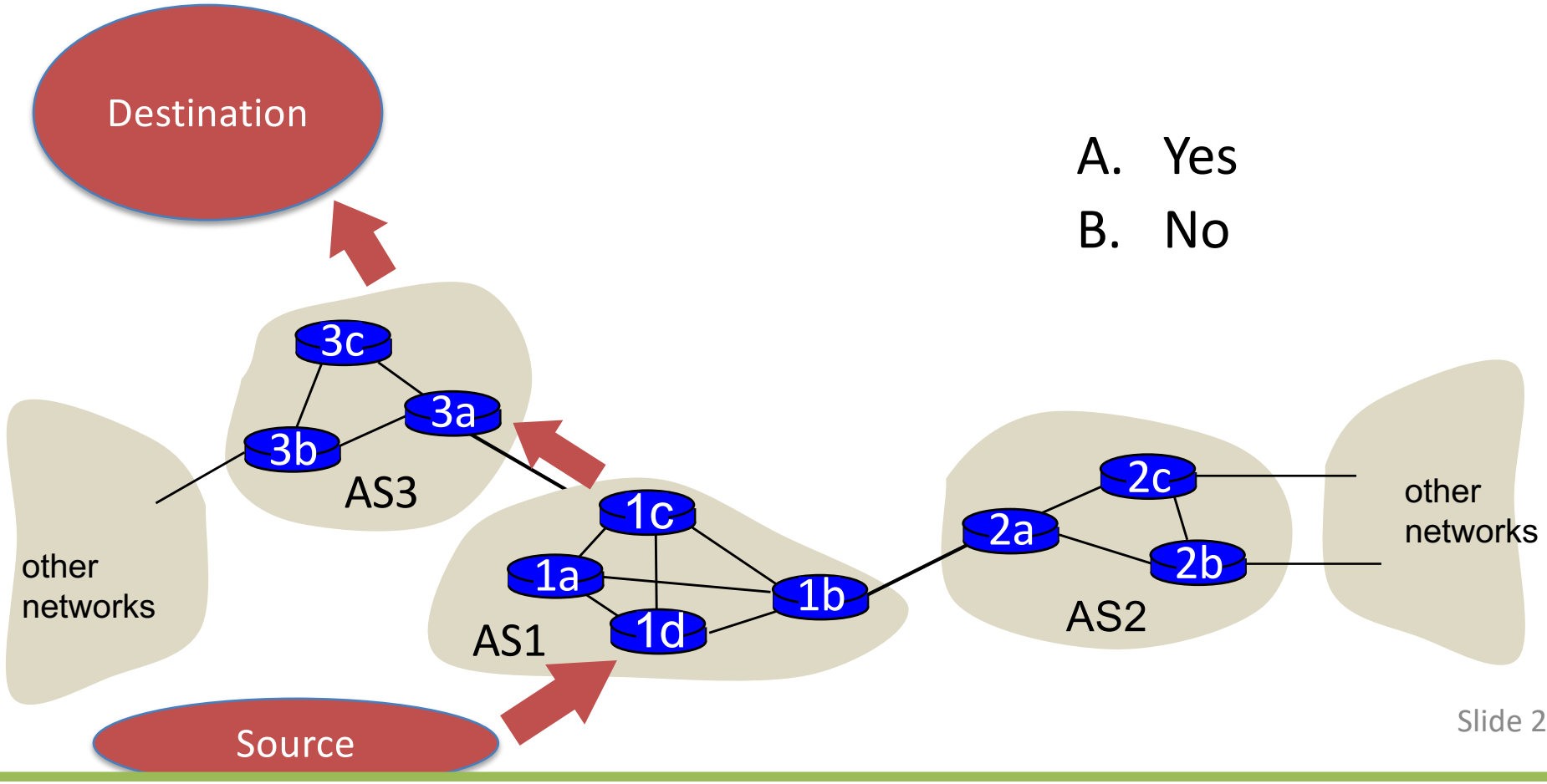
- You would rather have customers
- Peers are often competitors
- Peering agreements require periodic renegotiation

Peering struggles in the ISP world are extremely contentious, agreements are usually confidential

Hierarchical routing: Interconnected ASes

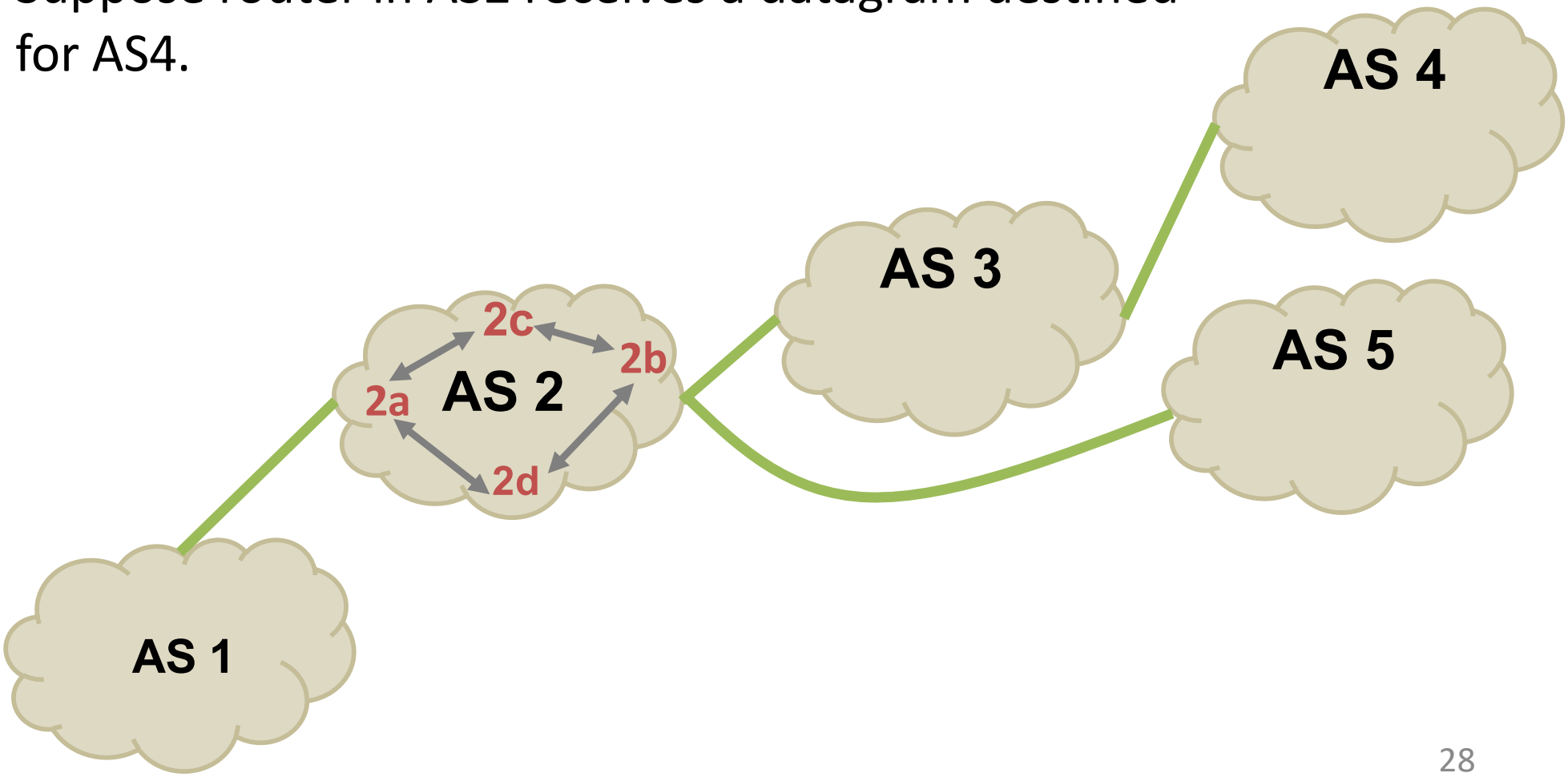


Border routers: exchange AS reachability, Internal routers: exchange intra-AS reachability., Is this sufficient to route from source to destination?



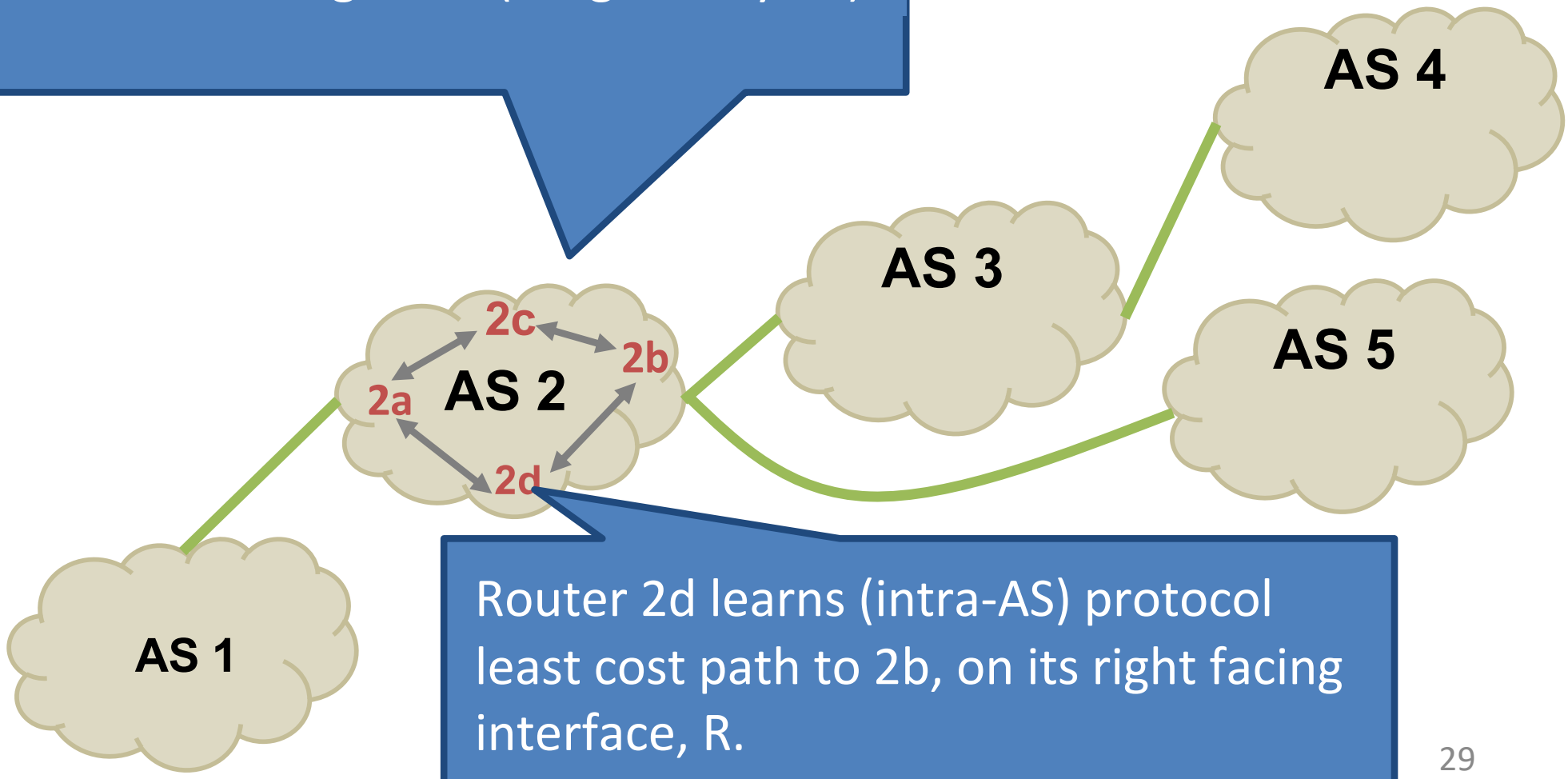
Building the forwarding table in router 2d, for path to AS4

Suppose router in AS2 receives a datagram destined for AS4.



Building the forwarding table in router 2d, for path to AS4

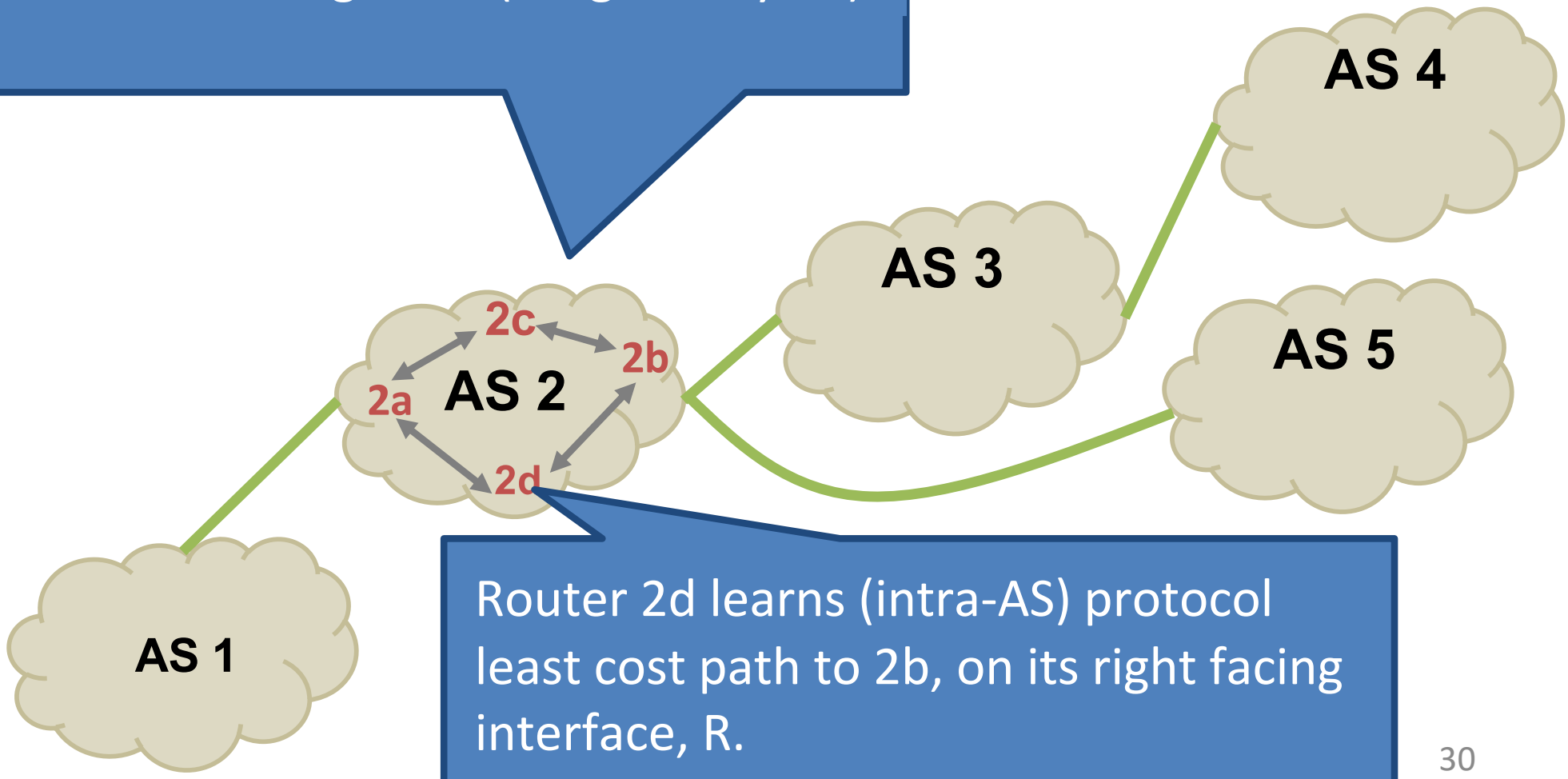
AS2 learns (Inter-AS) protocol that AS4 is reachable through AS3 (via gateway 2b).



Router 2d learns (intra-AS) protocol least cost path to 2b, on its right facing interface, R.

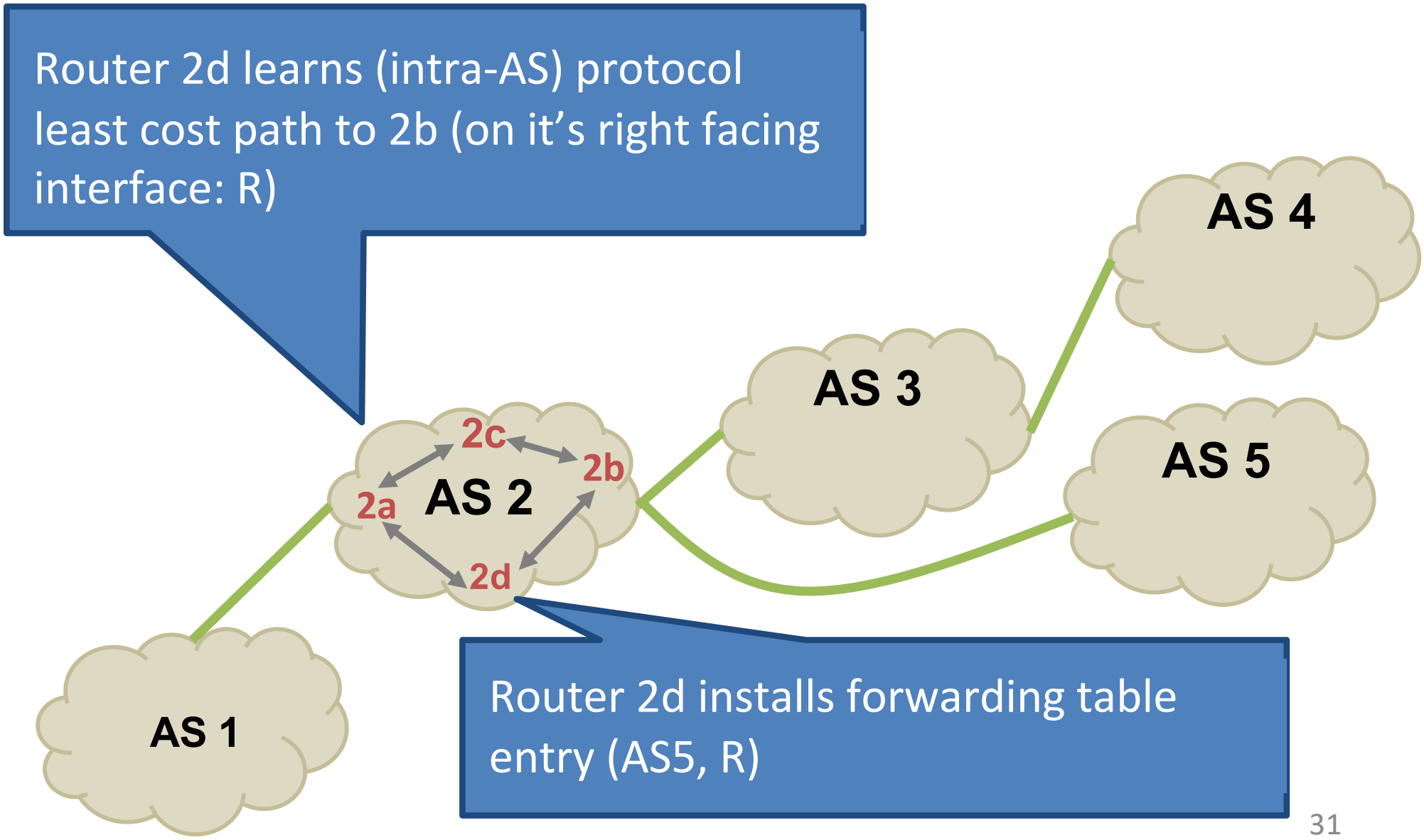
Building the forwarding table in router 2d, for path to AS4

AS2 learns (Inter-AS) protocol that AS4 is reachable through AS3 (via gateway 2b).

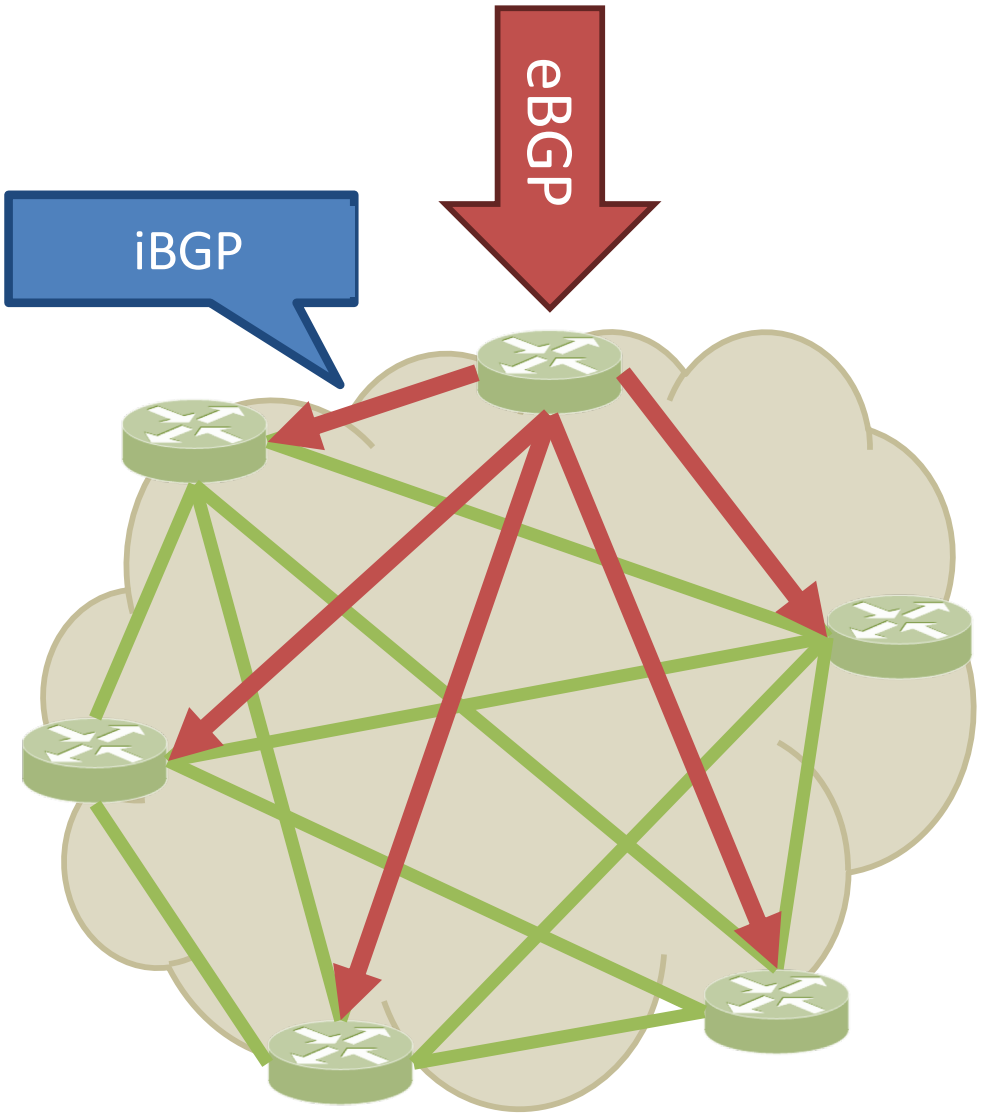


Router 2d learns (intra-AS) protocol least cost path to 2b, on its right facing interface, R.

Building the forwarding table in router 2d, for path to AS 5

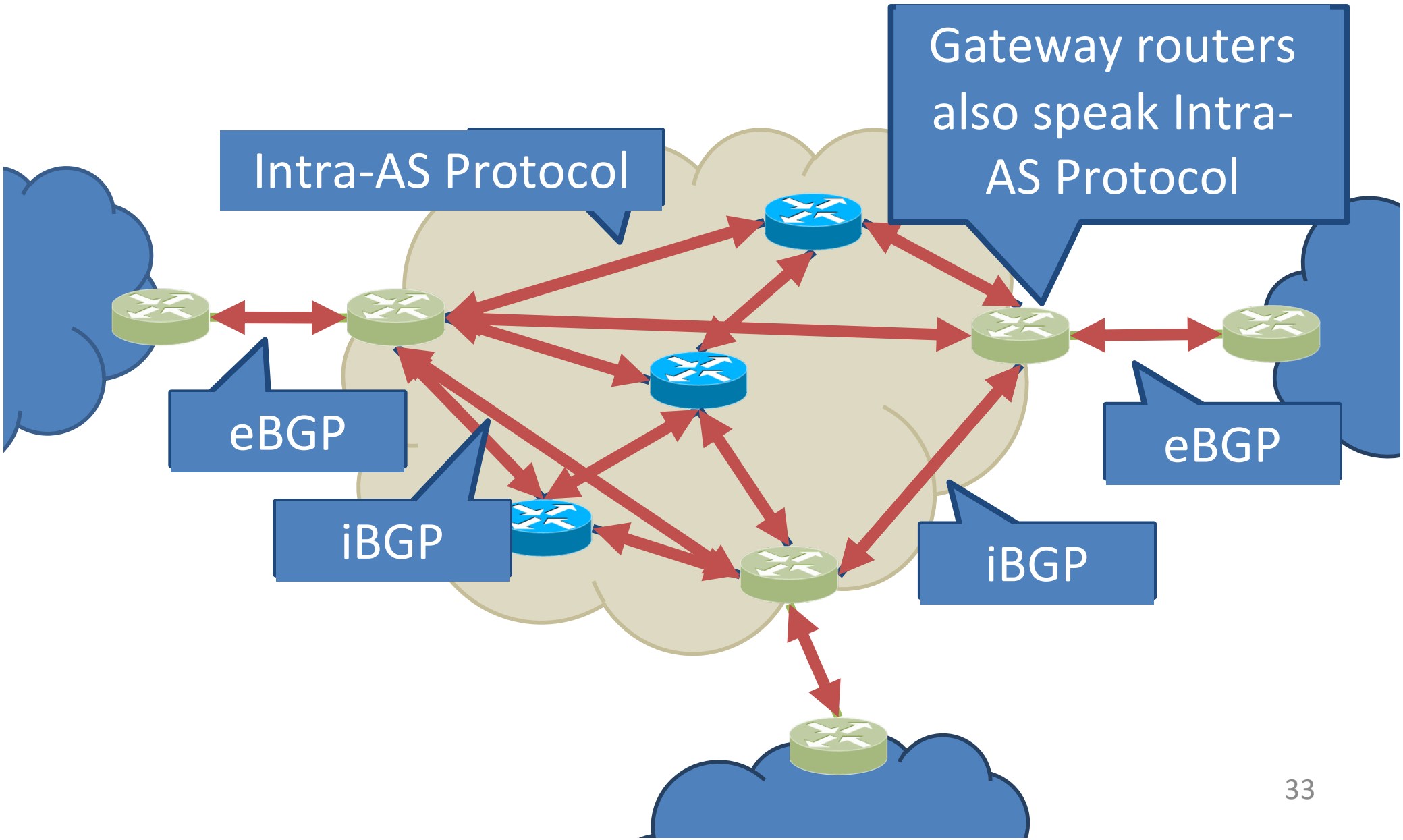


Internet inter-AS routing: BGP



- Question: why do we need iBGP?
  - OSPF does not include BGP policy info
  - Prevents routing loops within the AS
- iBGP updates do not trigger announcements

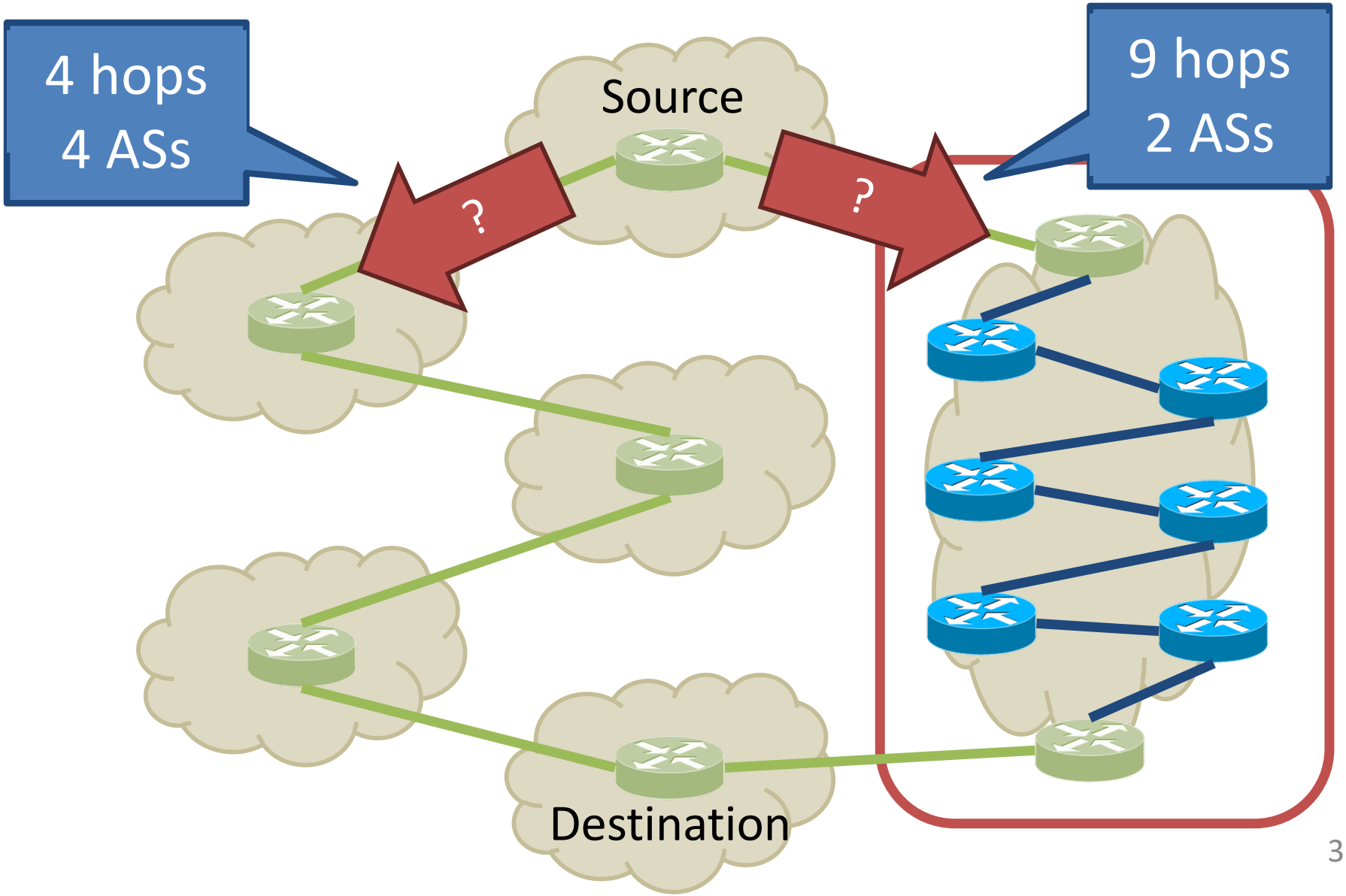
Internet inter-AS routing: BGP



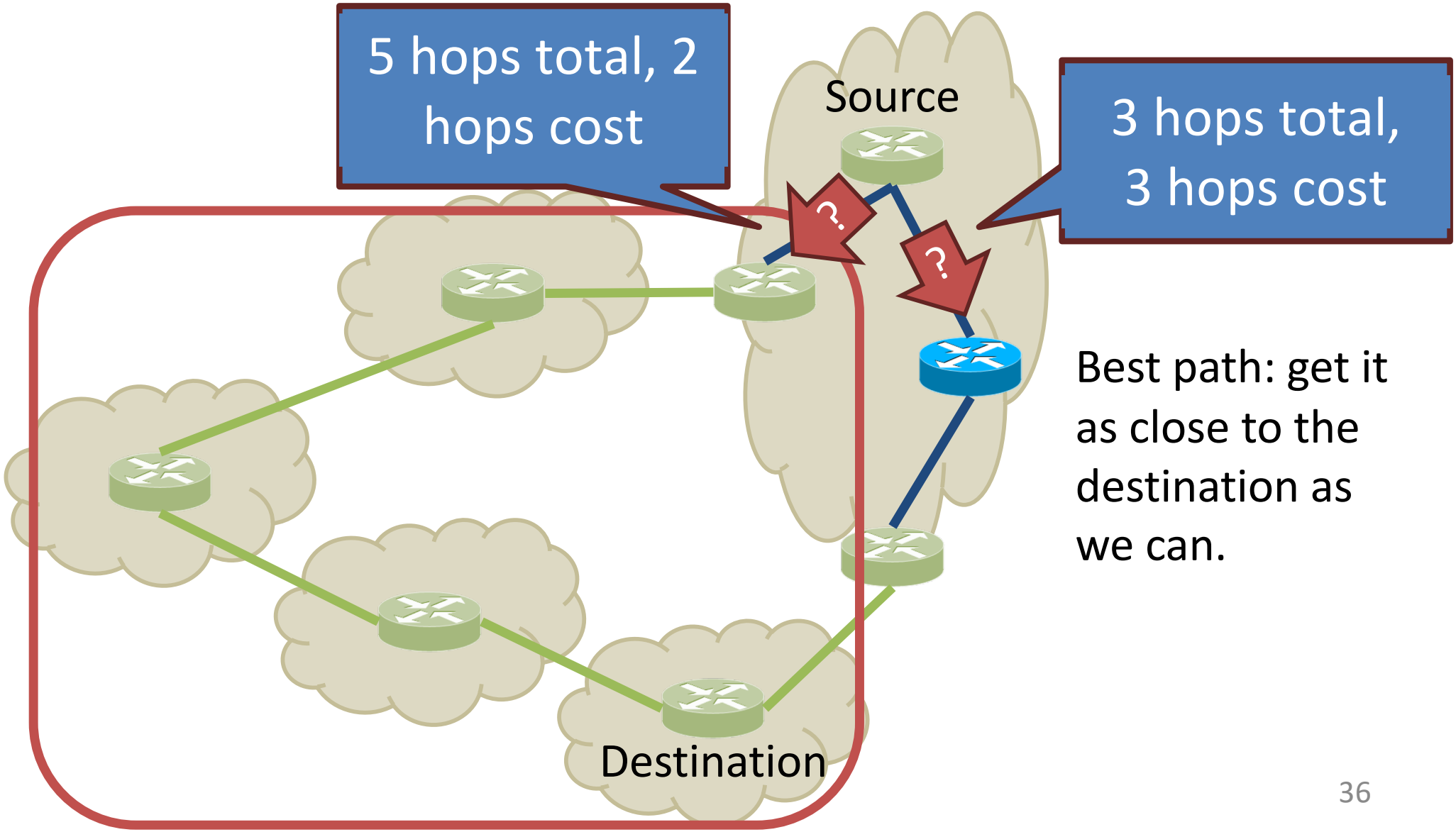
Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol):
  The de facto inter-domain routing protocol
- BGP provides each AS a means to:
  - external BGP: obtain subnet reachability information from neighboring ASs.
  - internal BGP: propagate reachability information to all AS-internal routers.
  - determine "good" routes to other networks based on reachability information and policy.
- Allows a subnet to advertise its prefix to the rest of the Internet

Shortest AS Path != Shortest Path

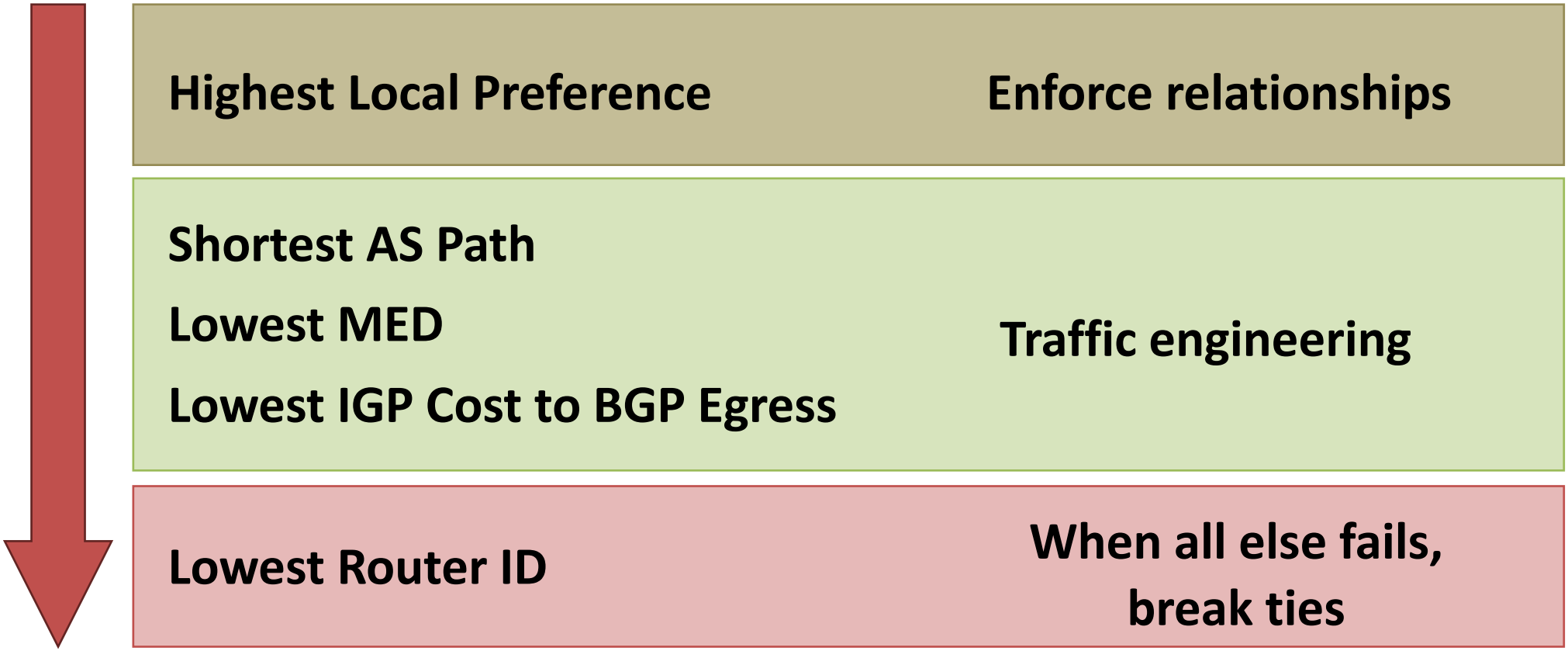


Hot Potato Routing: get rid of packets ASAP!

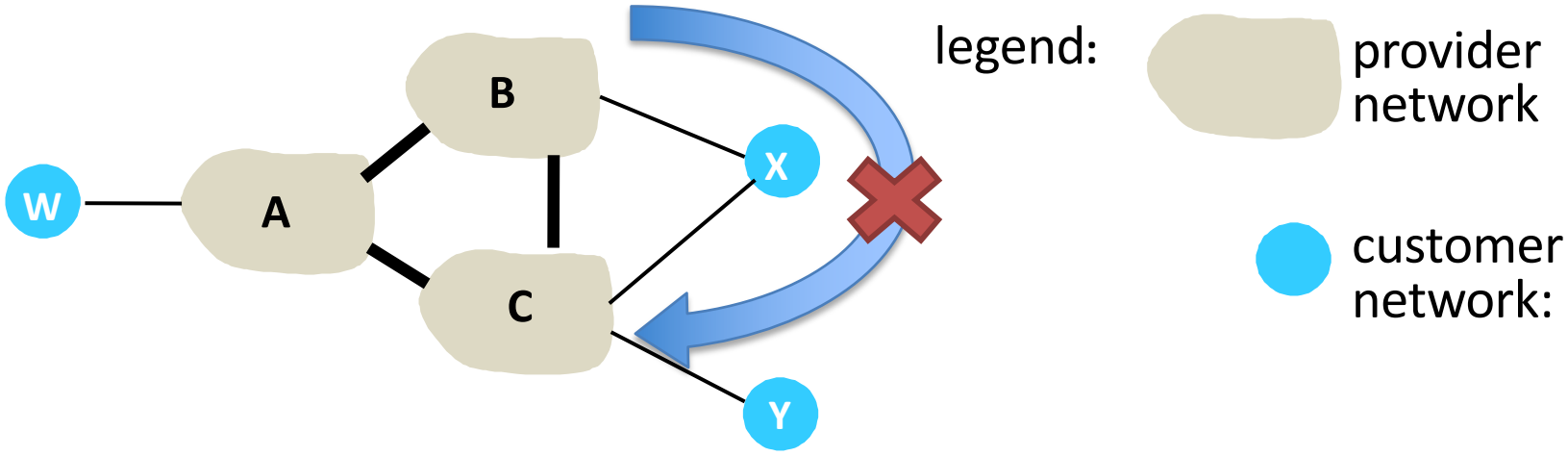


Best path: get it as close to the destination as we can.

Route Selection Summary

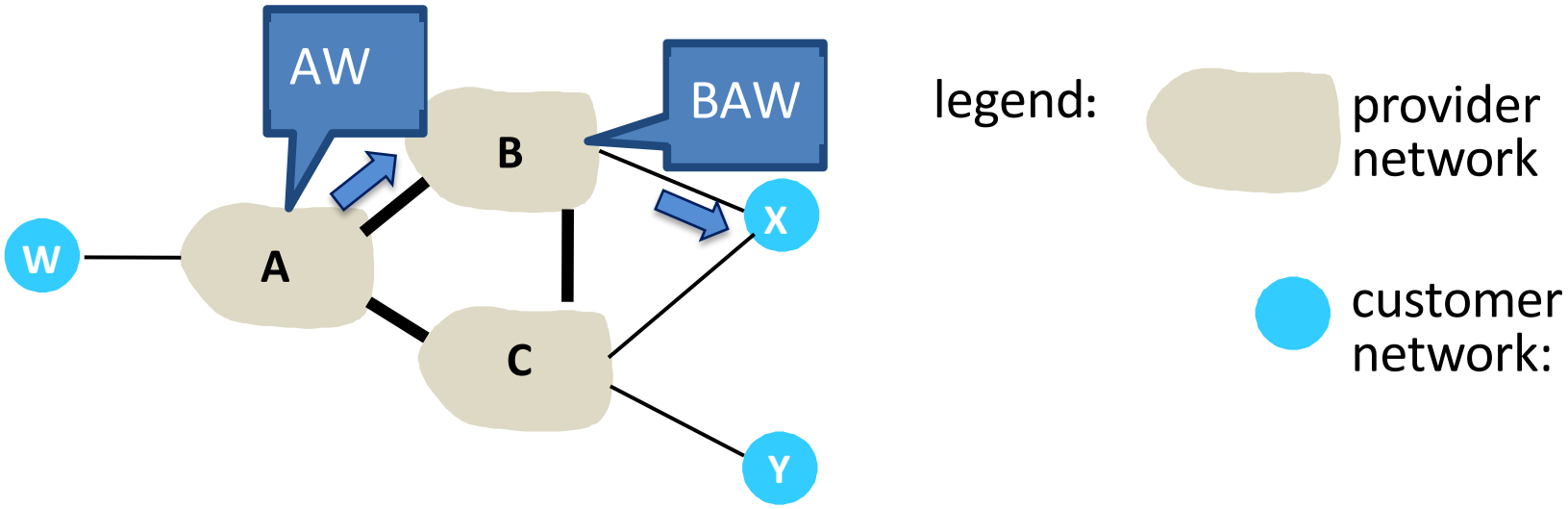


BGP routing policy



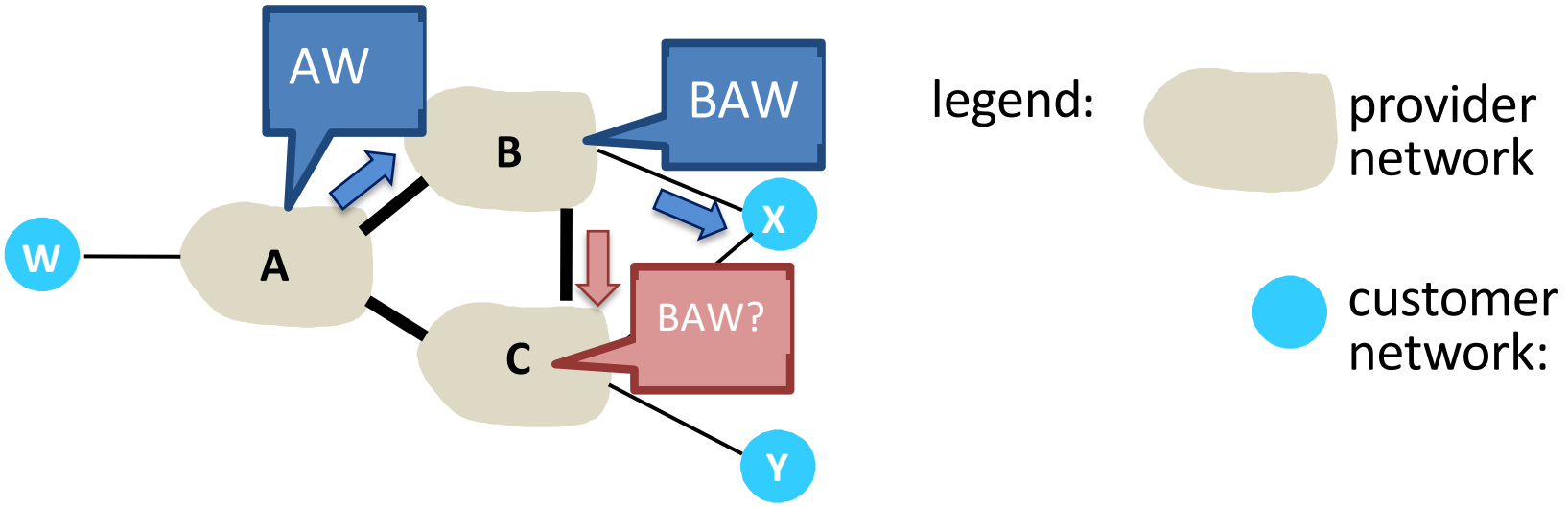
- A,B,C are provider networks
- X,W,Y are customers of the providers
- X is dual-homed: attached to two networks (B and C)
  - X does not want to route from B via X to C
  - .. so X will not advertise to B a route to C

BGP routing policy



- A advertises path AW to B
- B advertises path BAW to X

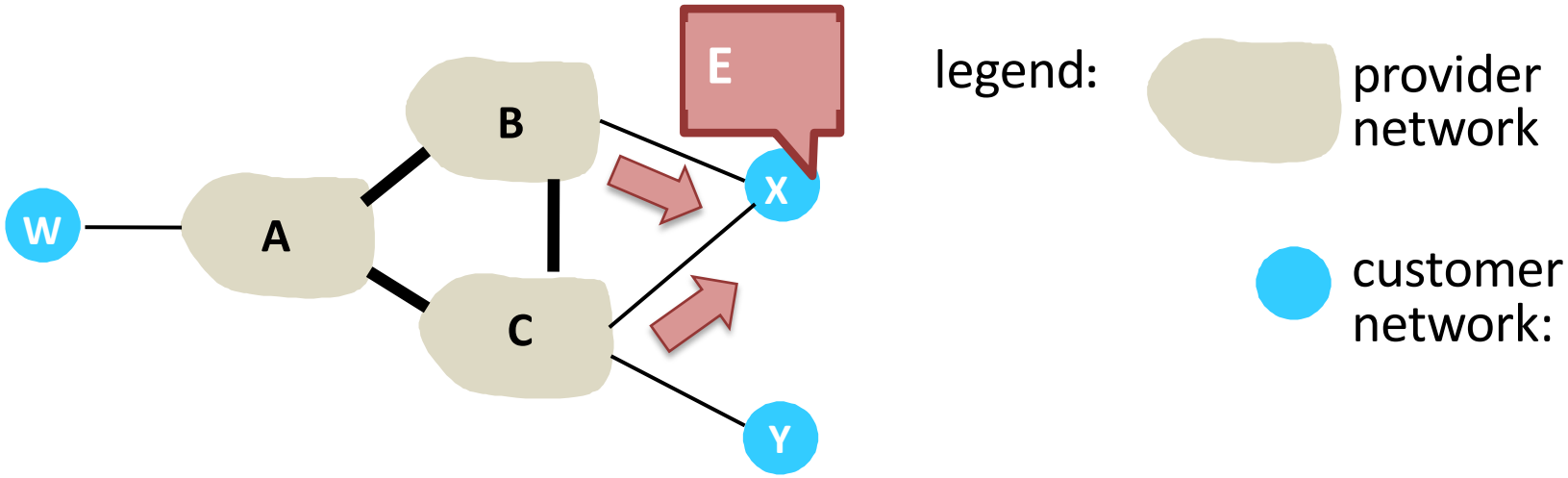
BGP routing policy: Should B advertise path BAW to C?



Should B advertise path BAW to C?

- B gets no "revenue" for routing CBAW since neither W nor C are B's customers
- B wants to force C to route to w via A
- B wants to route *only* to/from its customers!

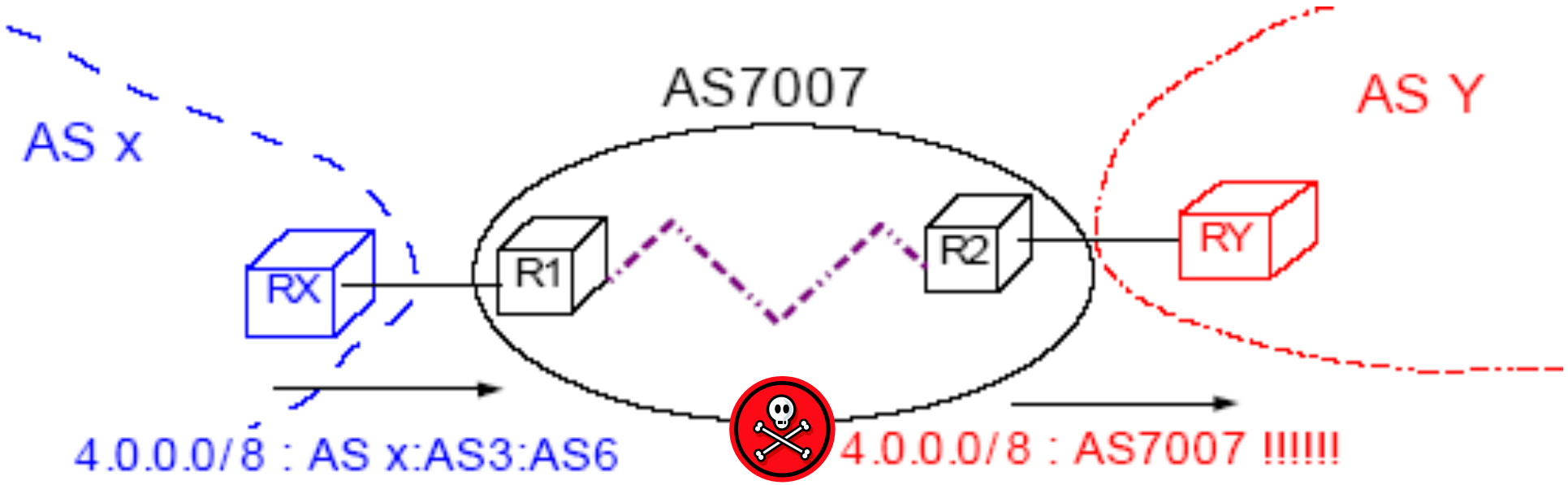
BGP routing policy gone wrong



- x advertises a path to E (that it is not connected to).
- all traffic starts to flow into x from B and C!

Faulty redistribution can be dangerous!

- AS7007 incident (April, 1997):



Hijack large part of the Internet

Summary

- As we've seen before (DNS), a hierarchy can help manage state storage constraints.
  - intra-AS routing: lots of info about local routes
  - inter-AS routing: less info about far away routes

- BGP: the inter-AS routing protocol for the Internet
  - Decisions often contractual

- BGP advertises AS prefixes, including:
  - entire path of ASes along the way
  - which border router heard the advertisement (Next Hop)

Additional Info: Inter-Domain Routing Challenges

- BGP4 is the only inter-domain routing protocol currently in use world-wide
- Issues?
  - Lack of security
  - Ease of misconfiguration
  - Poorly understood interaction between local policies
  - Poor convergence
  - Lack of appropriate information hiding
  - Non-determinism
  - Poor overload behavior

Additional Info: Lots of research into how to fix this

- Security
  - BGPSEC, RPKI
- Misconfigurations, inflexible policy
  - SDN
- Policy Interactions
  - PoiRoot (root cause analysis)
- Convergence
  - Consensus Routing
- Inconsistent behavior
  - LIFEGUARD, among others

Additional Info
Why are these still issues?

- Backward compatibility
- Buy-in / incentives for operators
- Stubbornness

Very similar issues to IPv6 deployment

Additional Info: Why Network Reliability Remains Hard

- Visibility
  - IP provides no built-in monitoring
  - Economic disincentives to share information publicly
- Control
  - Routing protocols optimize for policy, not reliability
  - Outage affecting your traffic may be caused by distant network
- Detecting, isolating and repairing network problems for Internet paths remains largely a slow, manual process

Net Neutrality

- how an ISP should share/allocation its resources
  - protecting innovation, free speech, and competition on the Internet

- Example: Comcast didn't like BitTorrent, started injecting RSTs into user TCP streams.

- Scarier example: You like Netflix, but your ISP has their own video service. They degrade (or block) Netflix service unless you pay $$$.

Net Neutrality

Cases for:

- End to end principle
- Prevent customer extortion
- Allow for innovation

Google, Microsoft, Yahoo, Amazon, eBay

Cases against:

- ISP owns their network
- Asymmetric application bandwidth usage
- We shouldn't legislate the Internet, it moves too fast

Cisco, many ISPs

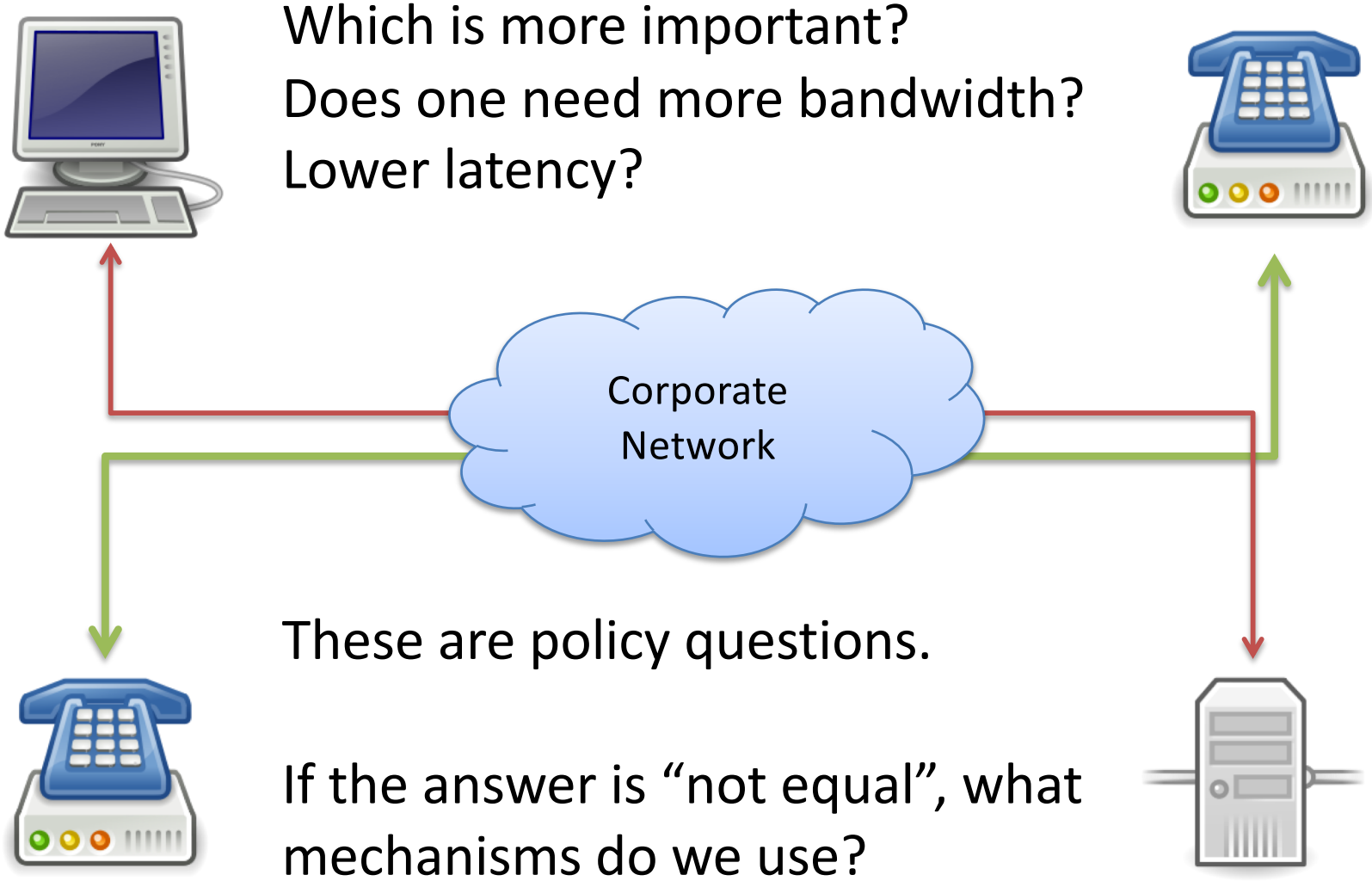
Today

- We've seen the behavior of TCP/IP, and routers
- We've joked about the option of marking packets as "urgent"
  - As a lone user, your cries for urgency will likely be ignored by one or more ISPs on the Internet
- False implication: All traffic is treated equally.

Scenarios

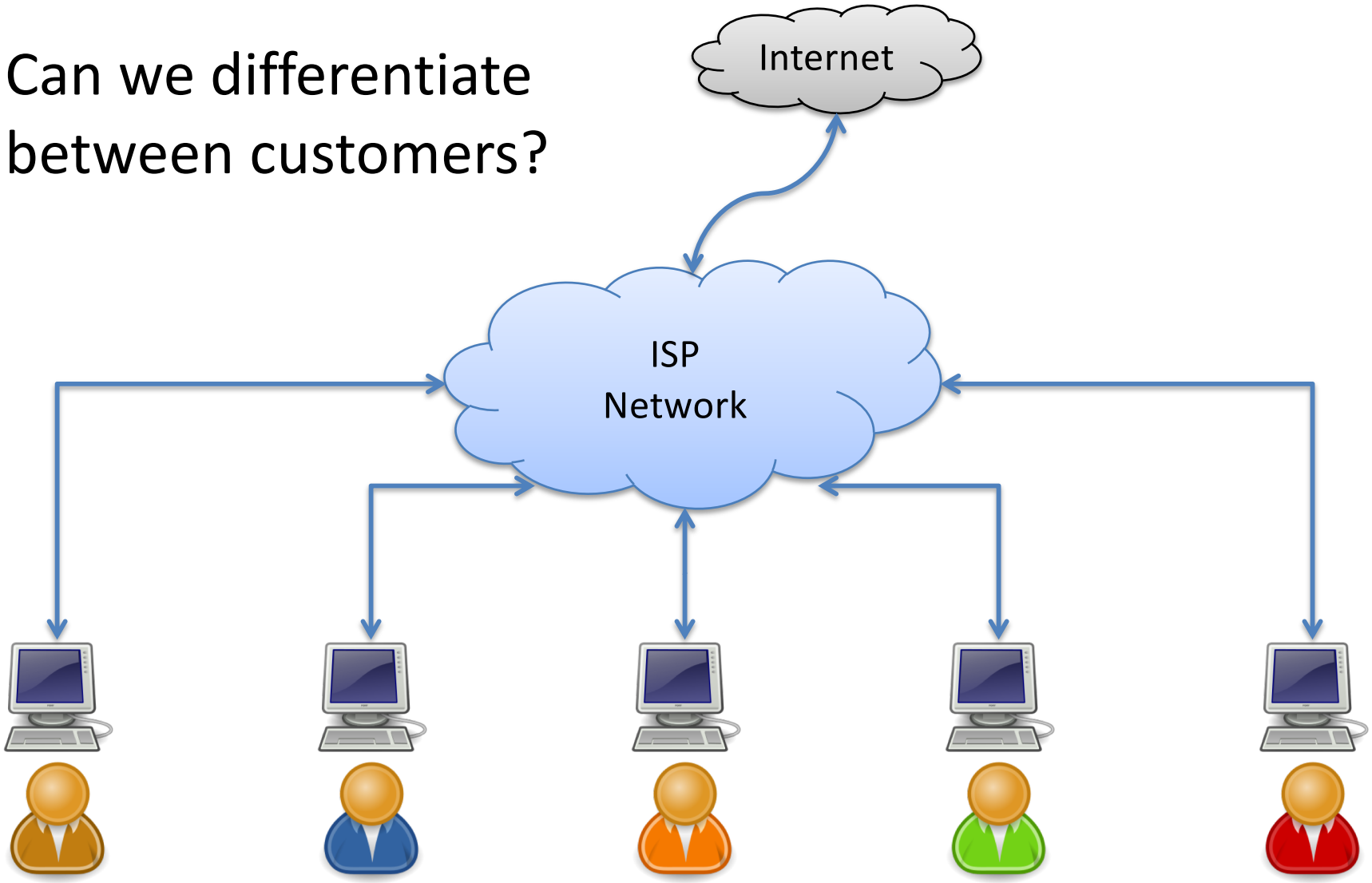
- Things we can do at the network layer to:
  - Treat traffic differently
  - Improve congestion control
- You own a private network
  - Corporate network
  - Data center
  - ISP
- You want to provide better performance to:
  - More important services
  - Customers who pay more

Example 1: Corporate Phones



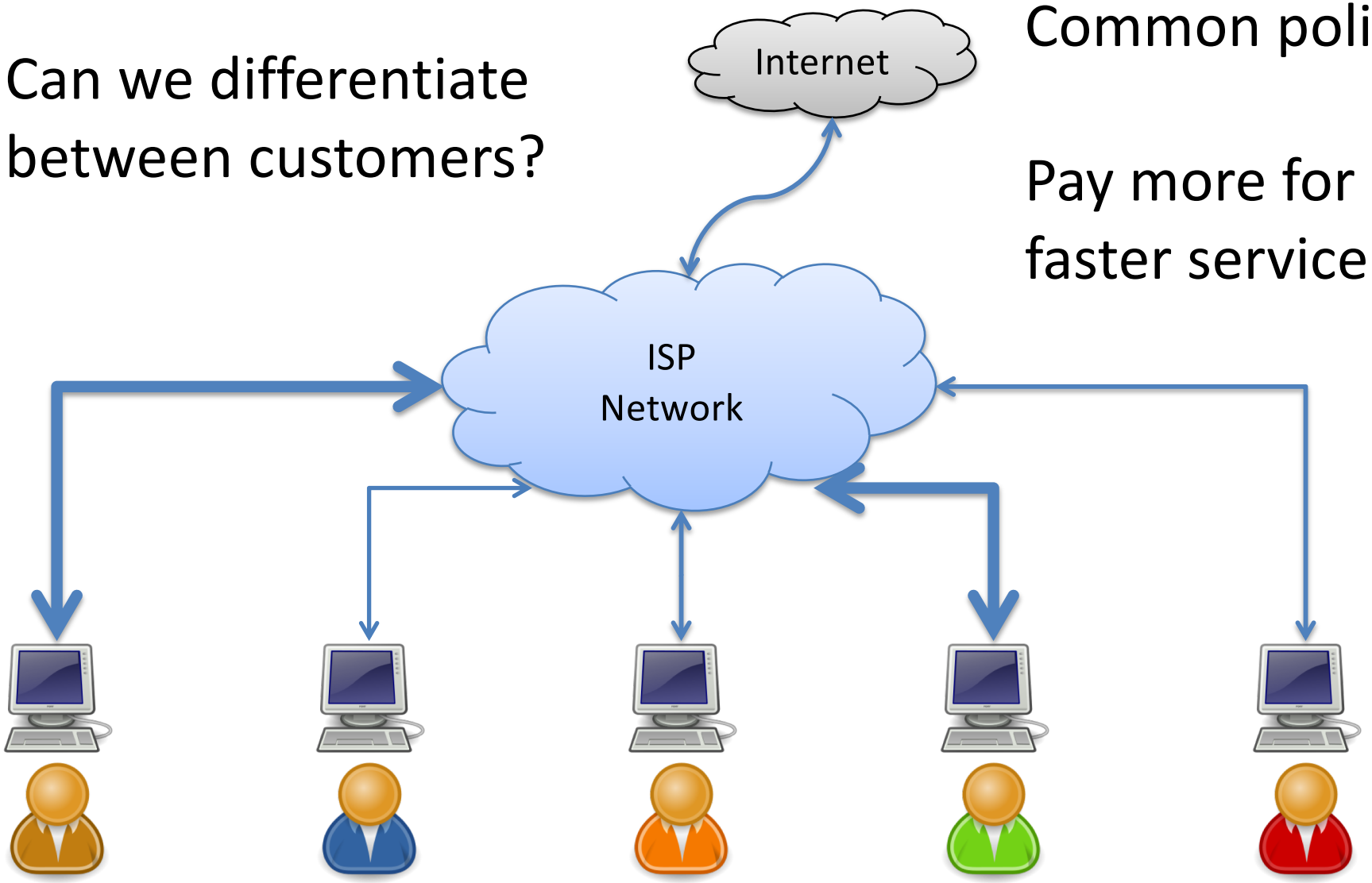
Example 2: ISP Customers

Can we differentiate between customers?



Example 2: ISP Customers

Can we differentiate between customers?



Common policy:

Pay more for faster service!

How might we enforce these types of policies?

A. Require that end-hosts police their traffic.

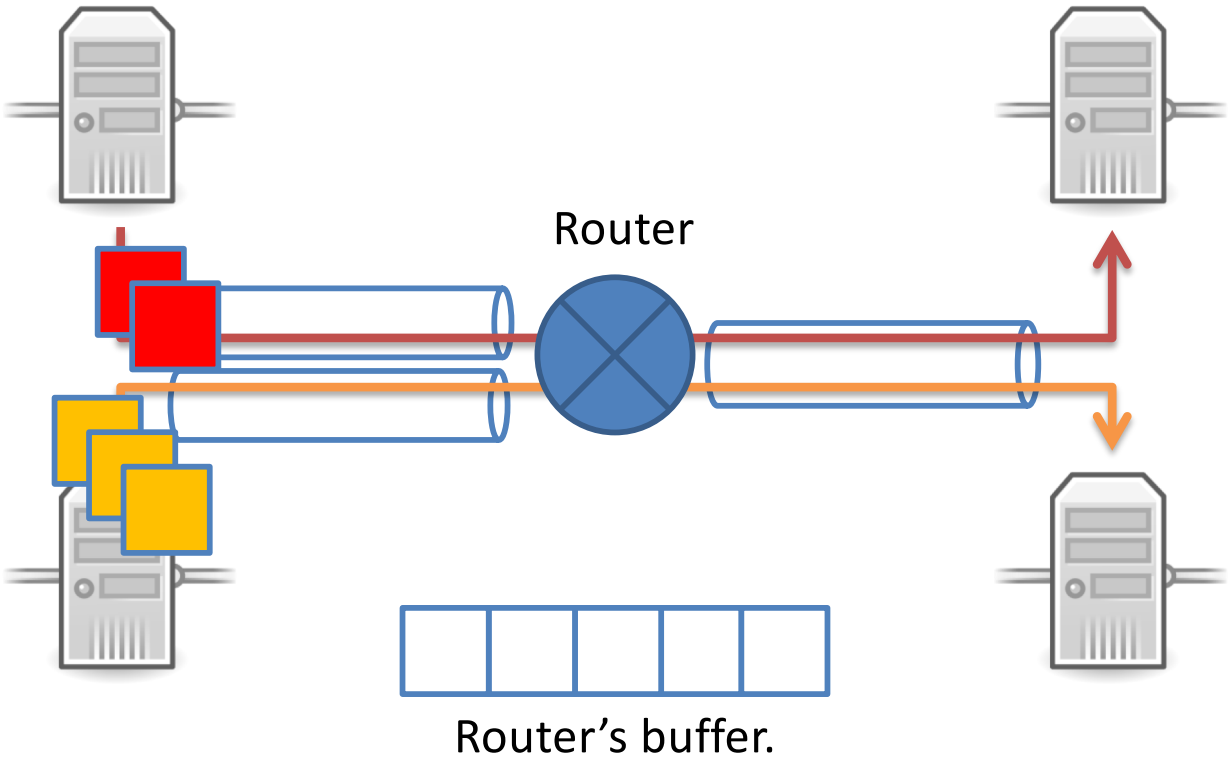
B. Change how routers queue traffic.

C. Ask users nicely to comply with policy.

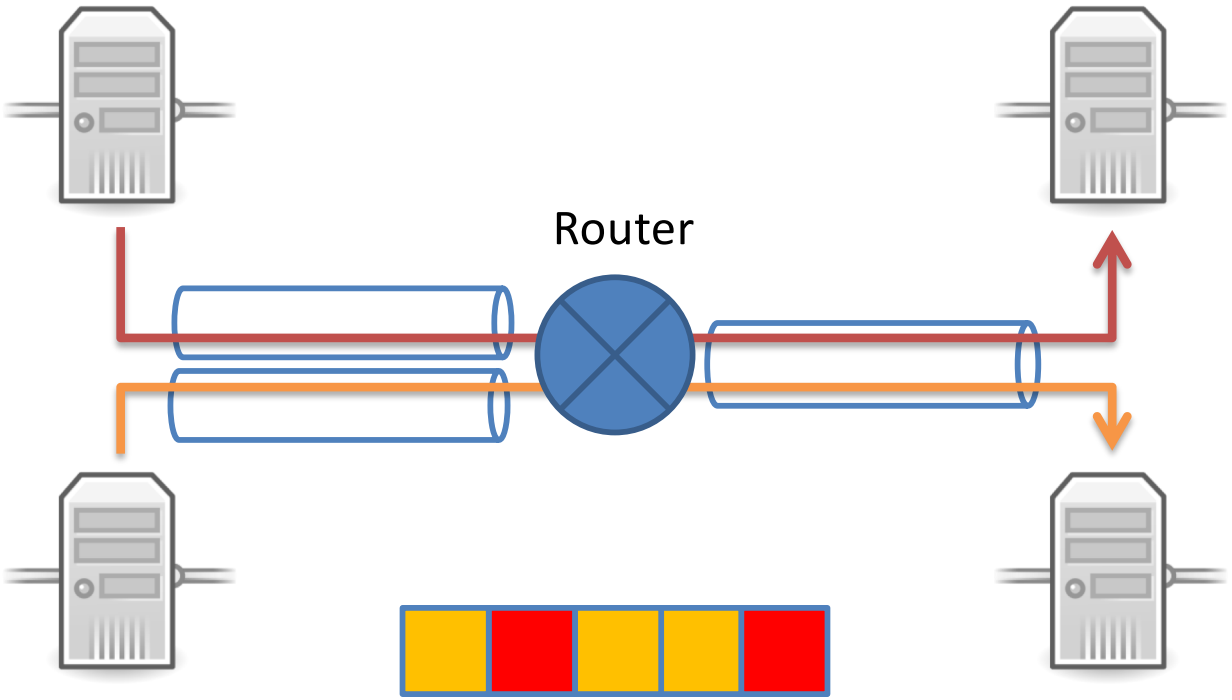
D. Enforce policies some other way.

E. There is nothing we can do.

Recall Queueing



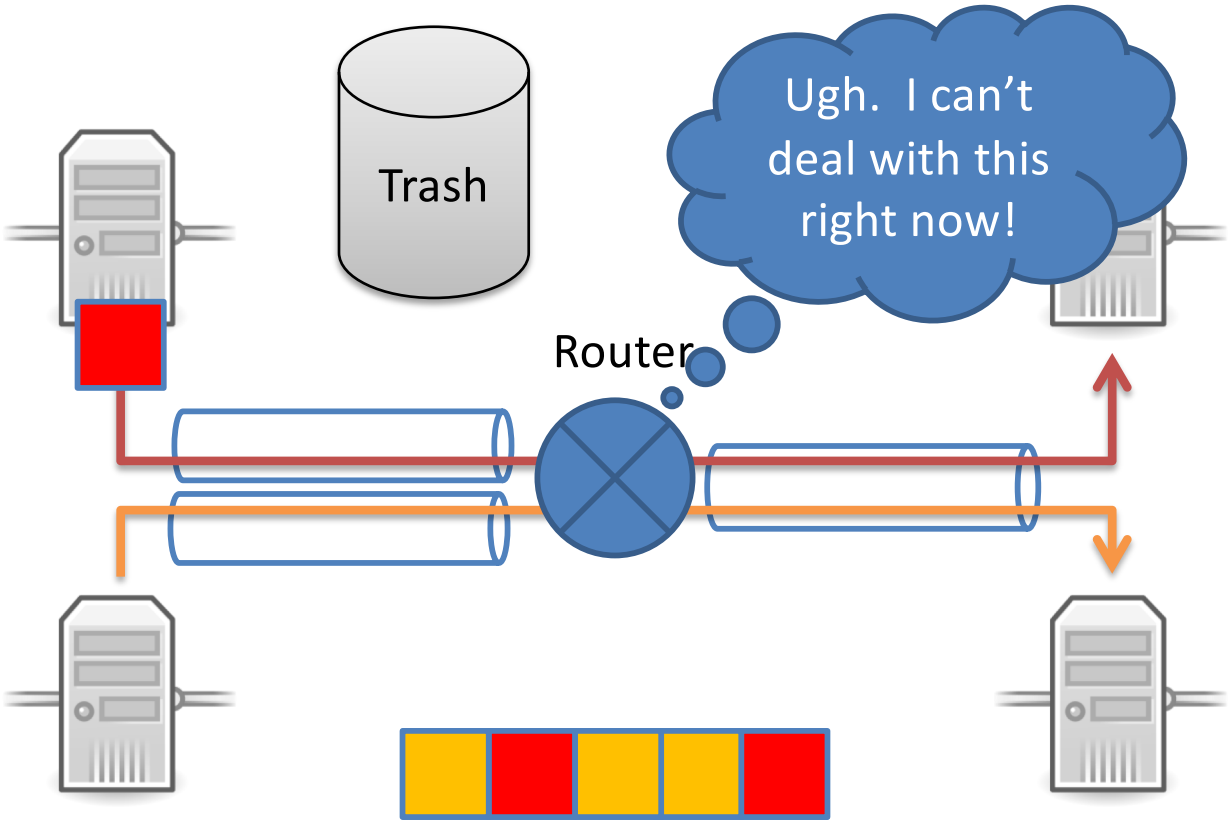
Recall Queueing



Router's buffer.

Incoming rate is faster than outgoing link can support.

Recall Queueing



Router's buffer.

Incoming rate is faster than outgoing link can support.

Basic Buffer Management

- FIFO + drop-tail
  - Simplest choice
  - Used widely in the Internet
- FIFO (first-in-first-out)
  - Traffic queued in first-come, first-served fashion
- Drop-tail
  - Arriving packets get dropped when queue is full
- Important distinction:
  - FIFO: queueing (scheduling) discipline
  - Drop-tail: drop policy

FIFO/Drop-Tail Problems

- Doesn't differentiate between flows/users
- No policing: send more, get more service
- Leaves responsibility of congestion control completely to the edges (e.g., TCP)
- Synchronization: hosts react to same events

FIFO/Drop-Tail Problems

- Doesn't differentiate between flows/users
- No policing: send more, get more service

QoS

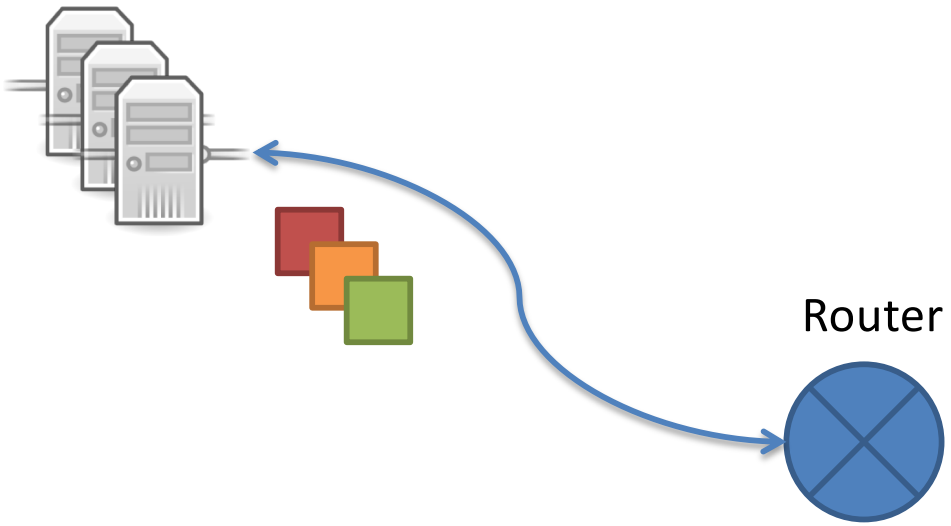
- Leaves responsibility of congestion control completely to the edges (e.g., TCP)
- Synchronization: hosts react to same events

AQM

Quality of Service (QoS)

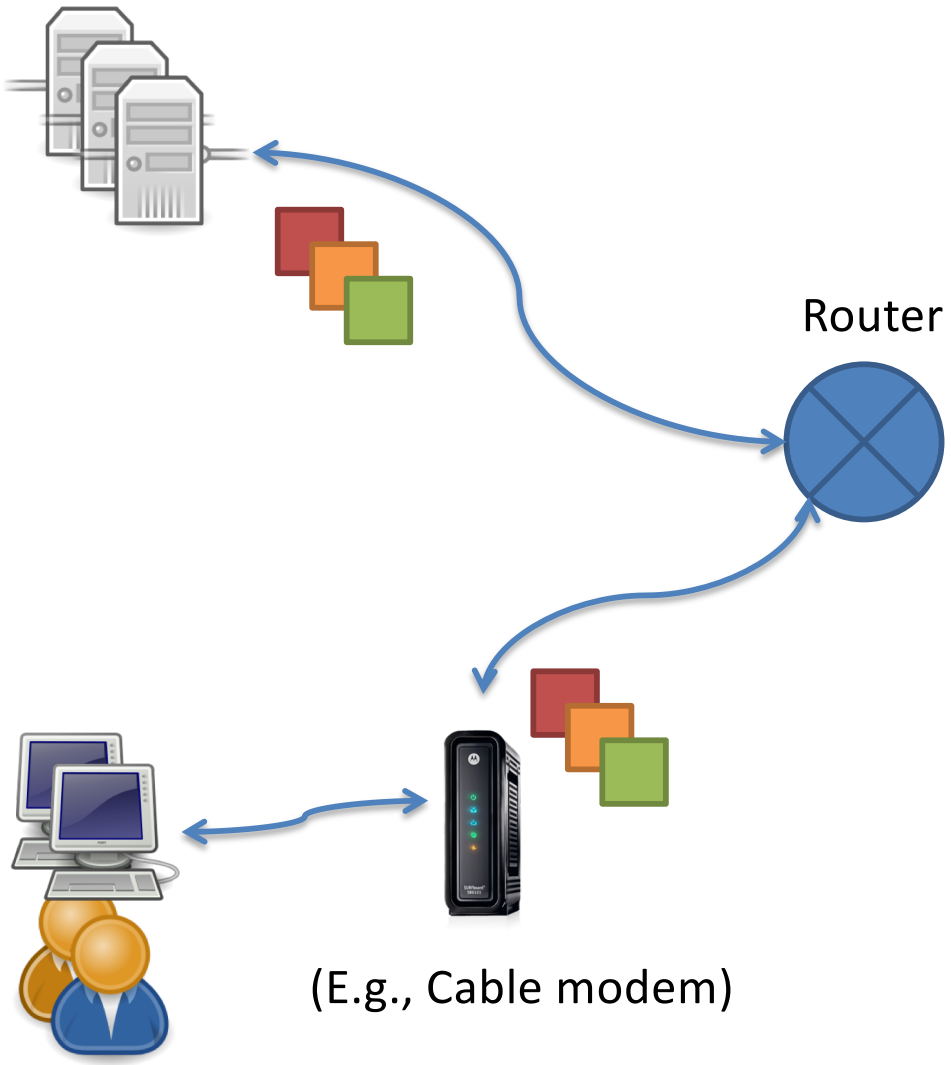
- QoS is a broad topic! We're going to discuss:
  - Mechanism for differentiating users/flows
  - Mechanism for enforcing rate limits
  - Mechanism for prioritizing traffic

Differentiating Users



- If you control end hosts:
  - Mark packets in OS according to policy.
- Take advantage of IP's class of service or options header fields

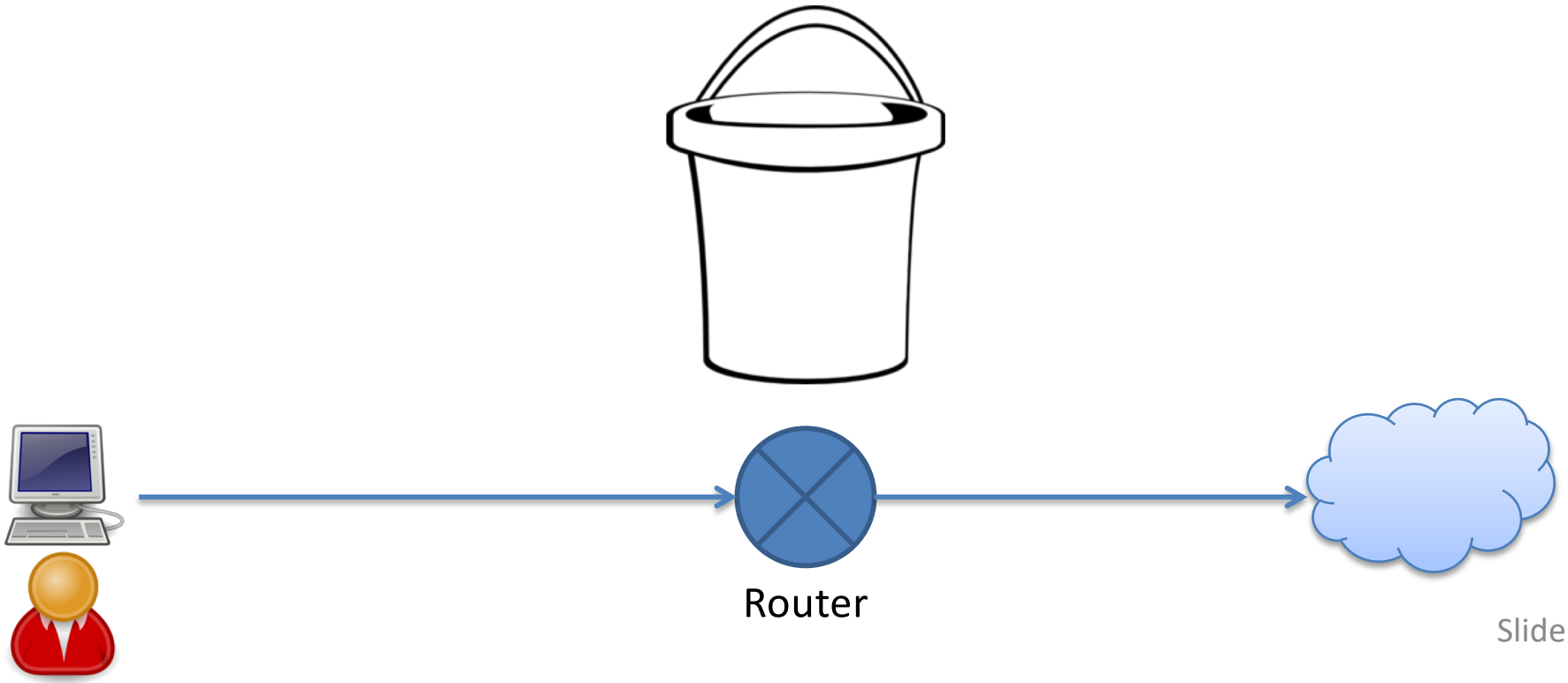
Differentiating Users



- If you control end hosts:
  - Mark packets in OS according to policy.
- Take advantage of IP's class of service or options header fields
- Otherwise:
  - Introduce an intermediate device you trust.

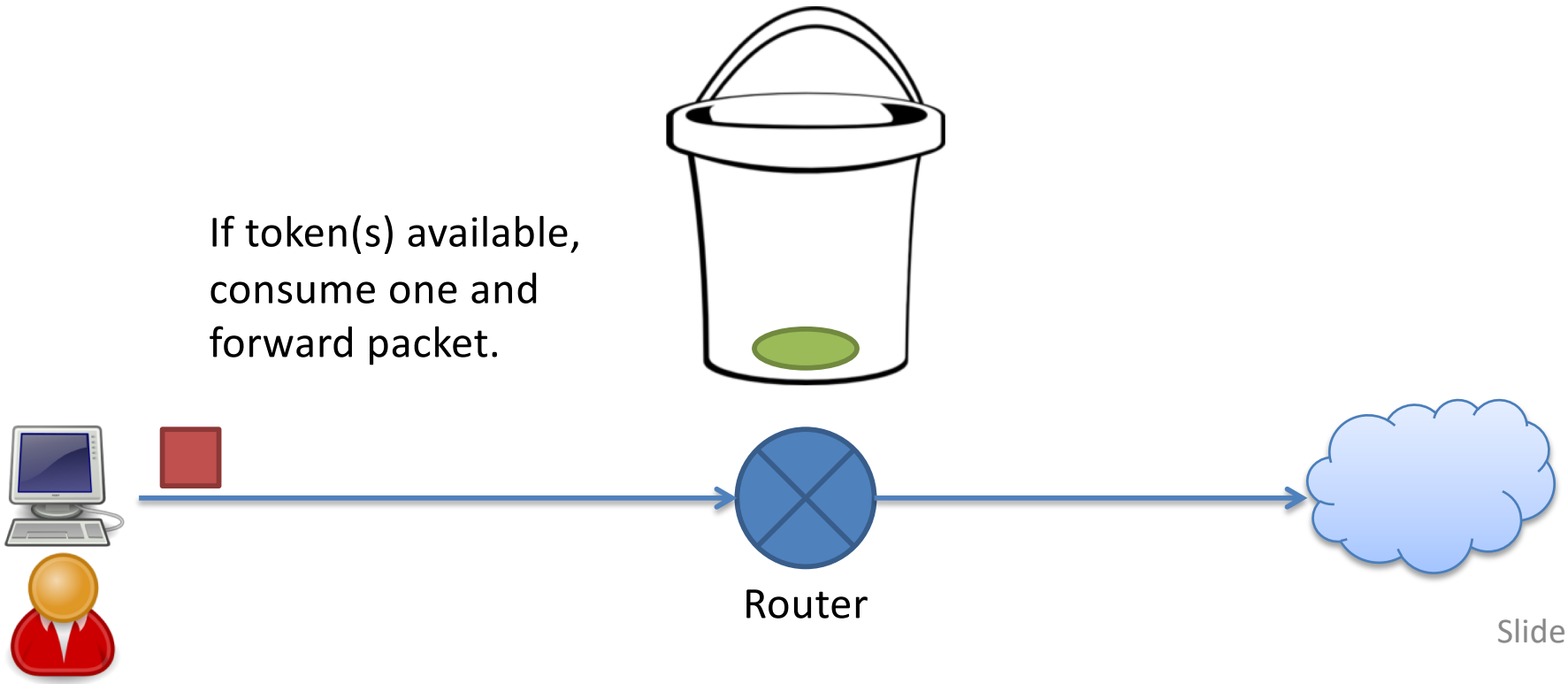
Enforcing (Policing) Rate Limits

- Example: the red user gets at most 10 Mbps
- Solution: Token bucket



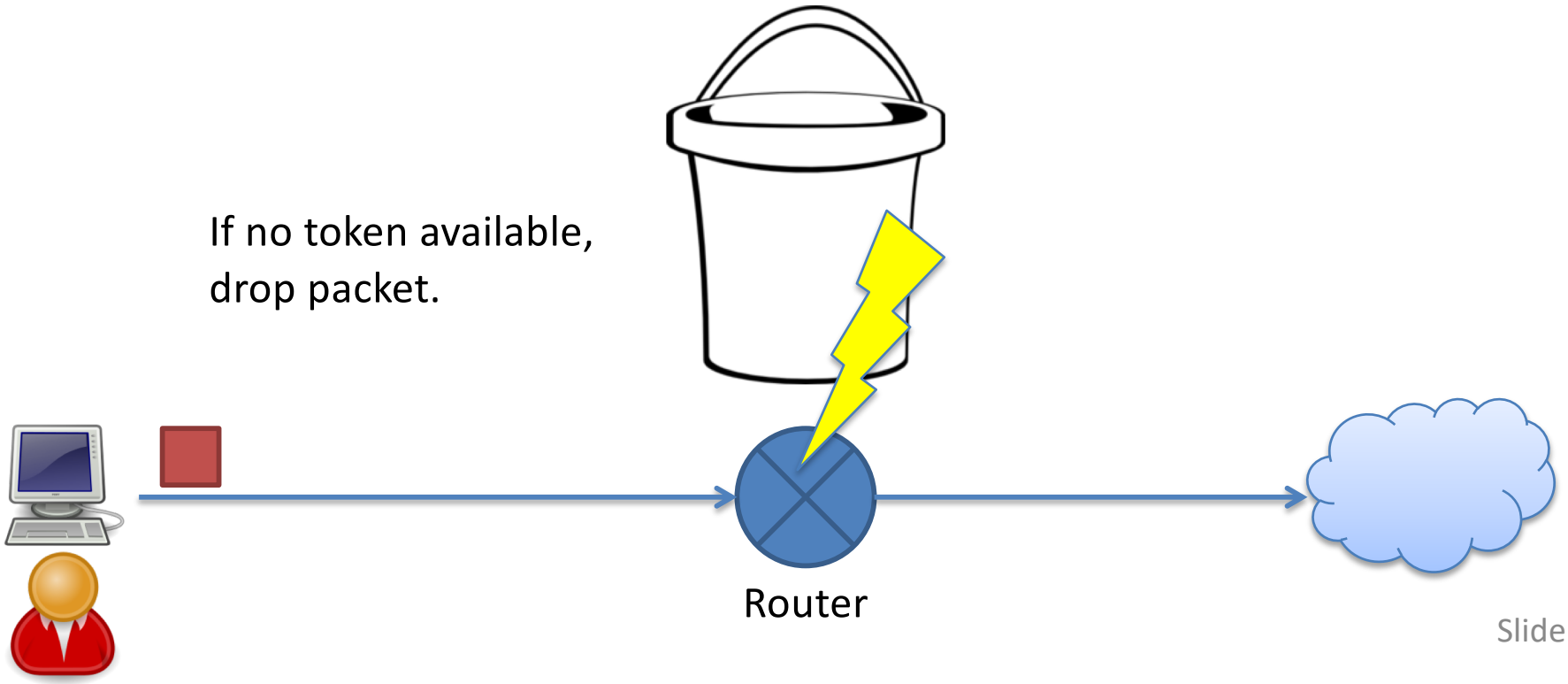
Enforcing (Policing) Rate Limits

- Example: the red user gets at most 10 Mbps
- Solution: Token bucket



Enforcing (Policing) Rate Limits

- Example: the red user gets at most 10 Mbps
- Solution: Token bucket

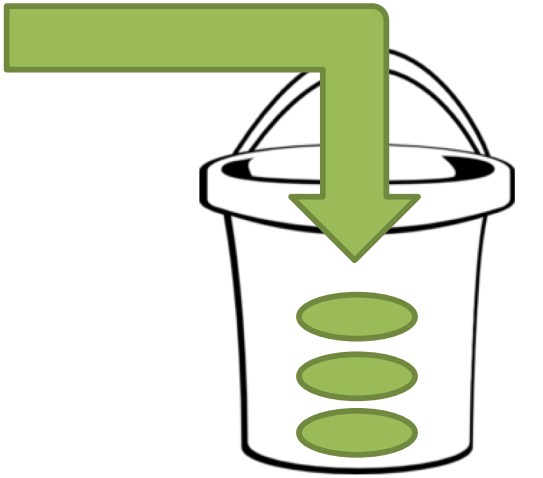


Enforcing (Policing) Rate Limits

- Example: the red user gets at most 10 Mbps
- Solution: Token bucket

No matter how fast user sends, limited by number of tokens, which replenish at controlled rate!

Router adds tokens at specified rate. (10 Mbps)



Bucket depth determines burst size.