Introduction to Computer Networks

CS640

Inter-domain Routing

https://pages.cs.wisc.edu/~mgliu/CS640/S25/index.html

Ming Liu
mgliu@cs.wisc.edu

Outline

- Last
  - Software-Defined Networking

- Today
  - Inter-domain Routing

- Announcements
  - Lab3 due on 04/01/2025 12:01PM

Intra-domain Routing

- Two routing schemes:
  - Distance vector routing
  - Link state routing

Intra-domain Routing

- Two routing schemes:
  - Distance vector routing
  - Link state routing
- What is a domain?
  - An organization under the same administrative control
  - E.g., university, company, etc.
  - Refer to as an **Autonomous System (AS)**
  - **An AS has a 16-bit unique number**

Intra-domain Routing

- Two routing schemes:
  - Distance vector routing
  - Link state routing
- What is a domain?
  - An organization under the same administrative control
  - E.g., university, company, etc.
  - Refer to as an **Autonomous System (AS)**
  - **An AS has a 16-bit unique number**

Is the intra-domain routing sufficient to support networks at scale?

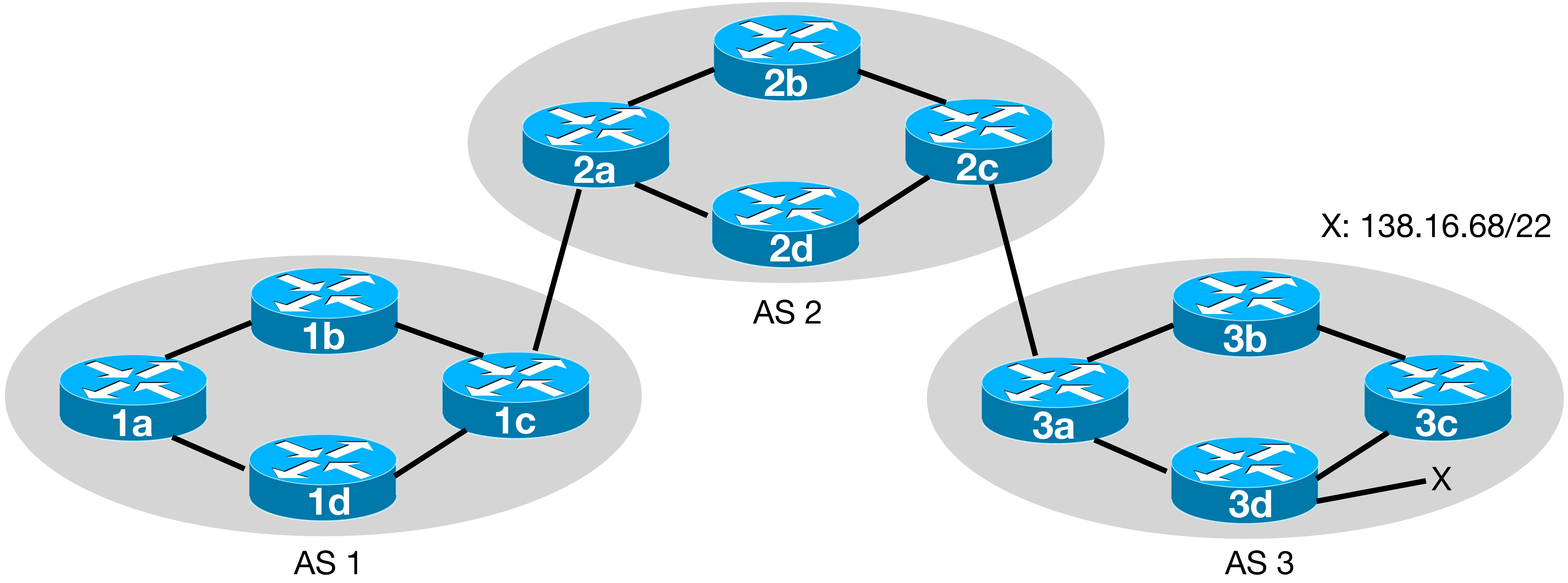
Intra-domain Routing

- Two routing schemes:
  - Distance vector routing
  - Link state routing
- What is a domain?
  - An organization under the same administrative control
  - E.g., university, company, etc.
  - Refer to as an **Autonomous System (AS)**
  - **An AS has a 16-bit unique number**

No. The number of hosts (subnets), the bandwidth requirement, and the device constraints limit its routing capability.

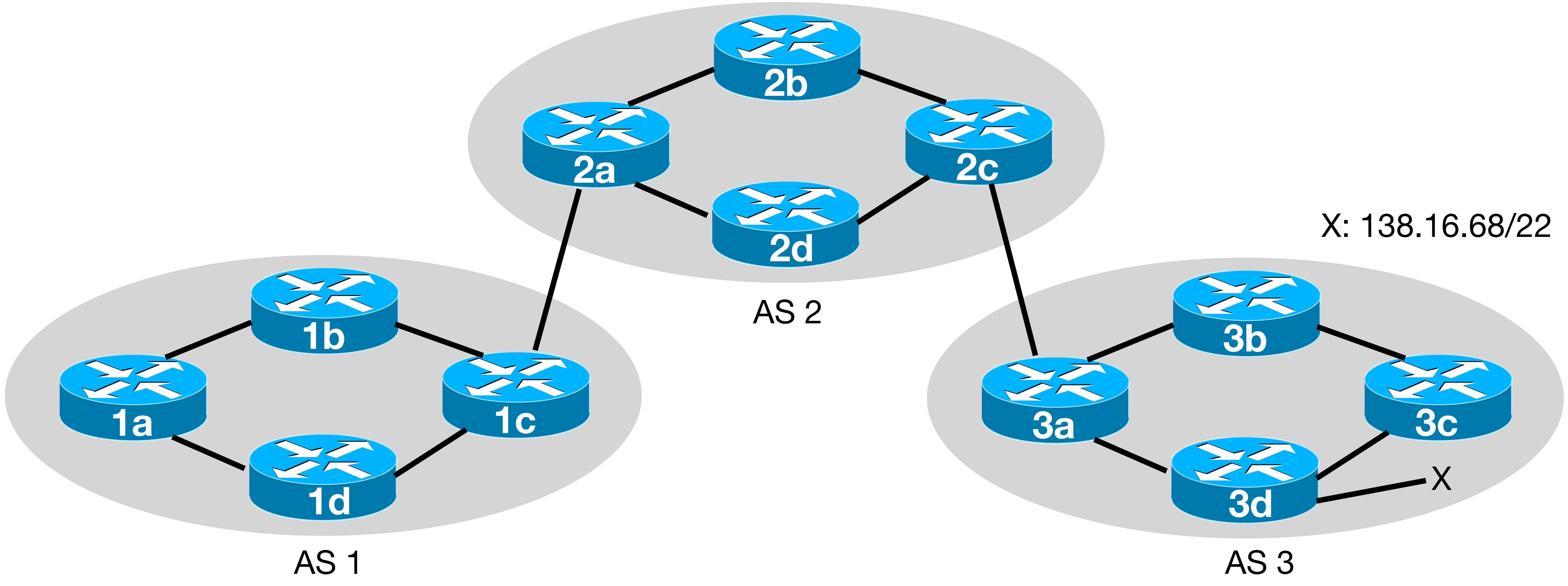
We need inter-domain routing to connect autonomous systems (domains).

Inter-domain Routing



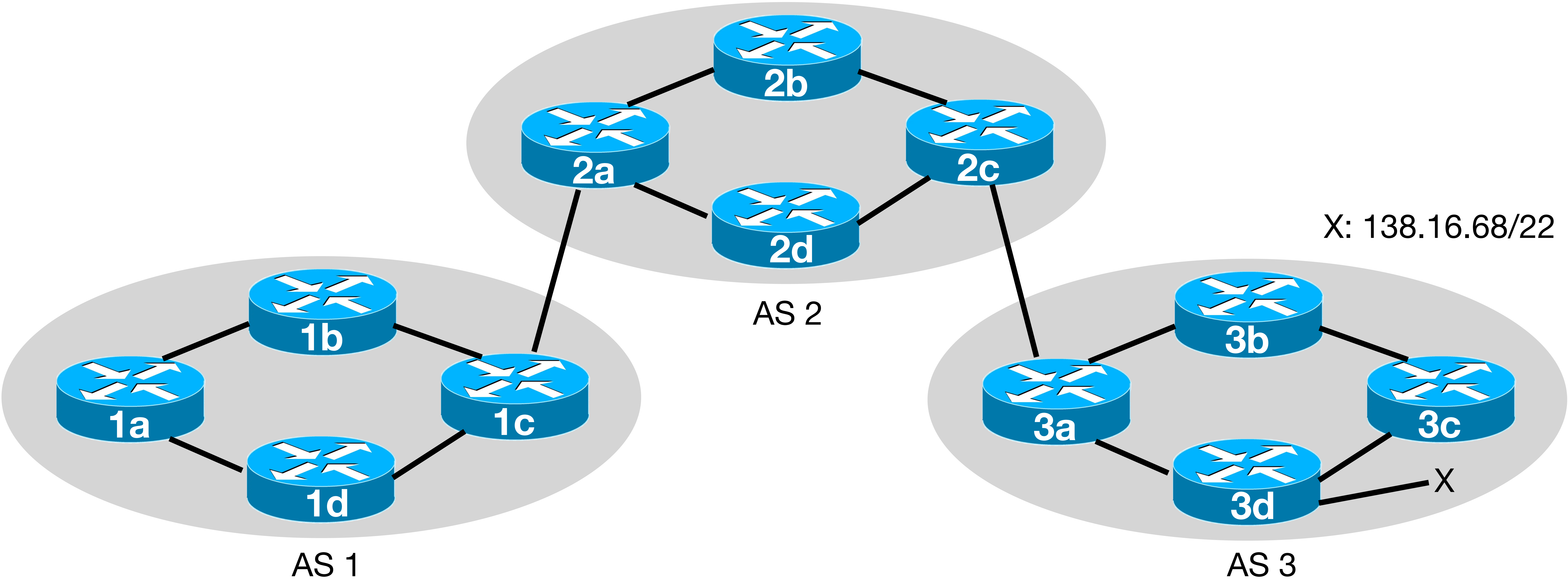
Inter-domain Routing

How can AS1/AS2 route packets to X?



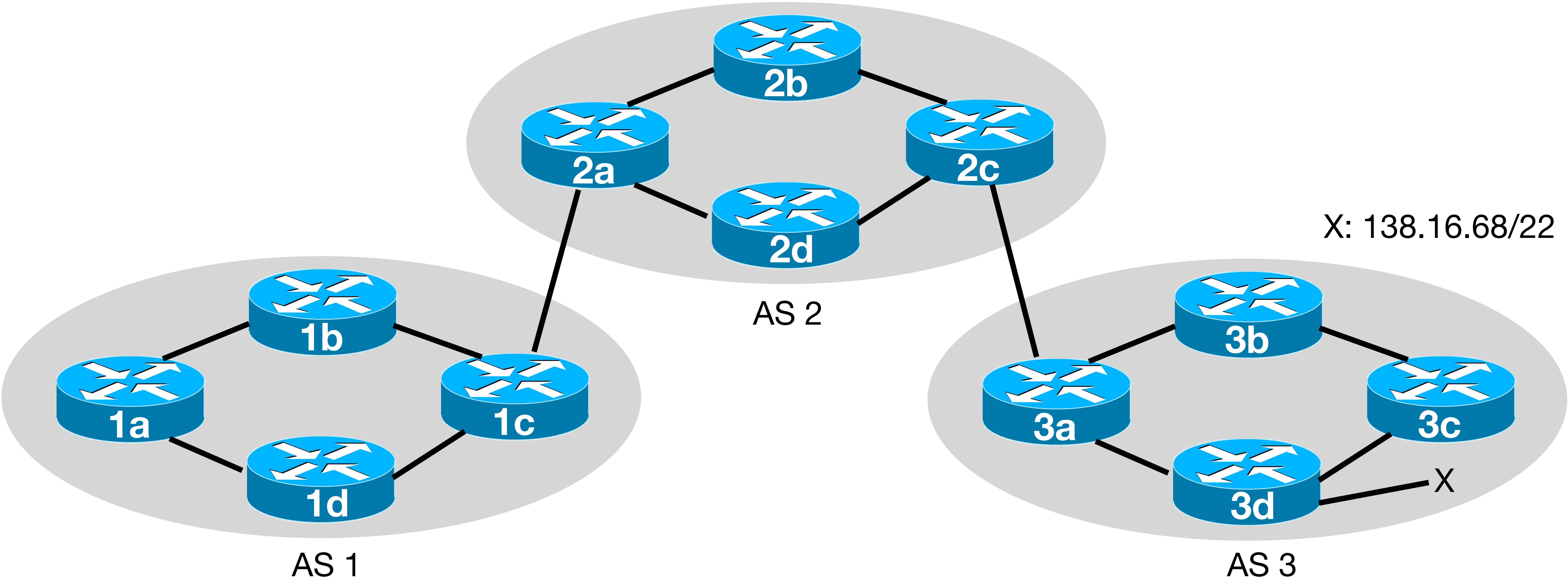
Two Types of Routers

- Gateway router: connect to one or more routers in other ASs
- Internal router: connect to hosts and routers within its own AS



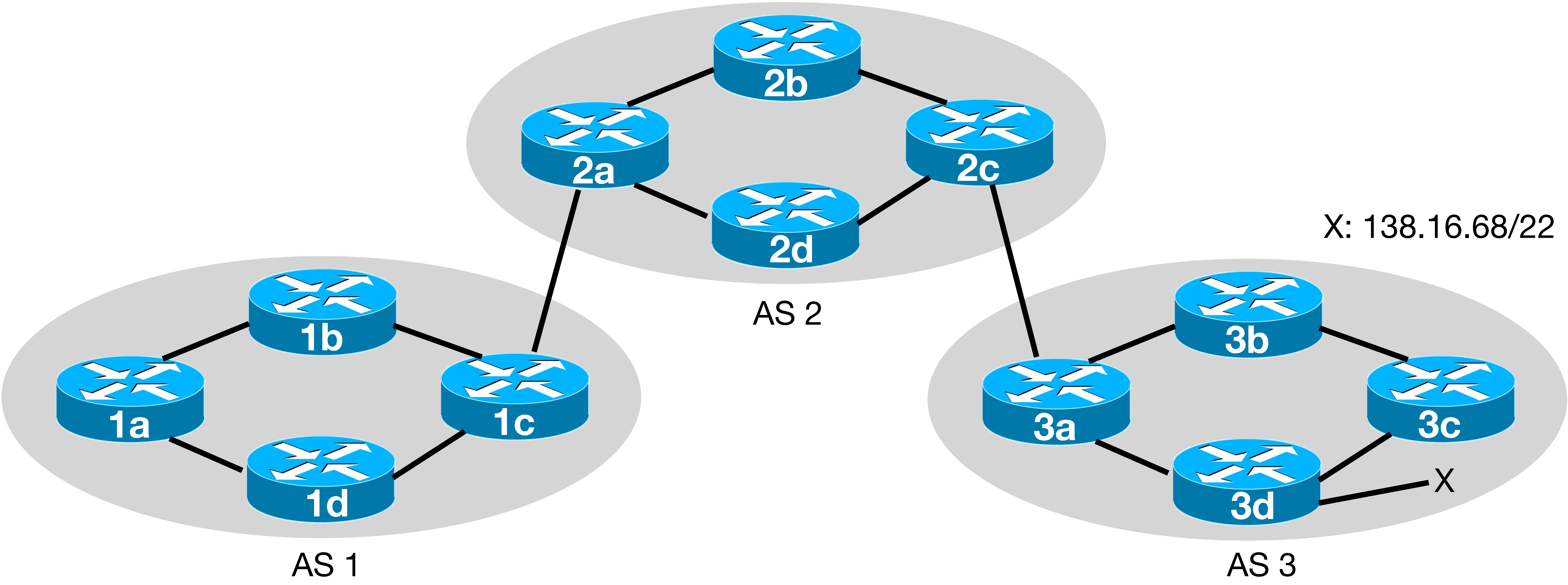
Route Advising

- An AS must tell others what prefix reachability it owns
  - AS 3 —> AS 2: X (138.16.68/22) exists (AS 3 X)
  - AS 2 —> AS 1: X (138.16.68/22) exists from me (AS 2 AS 3 X)



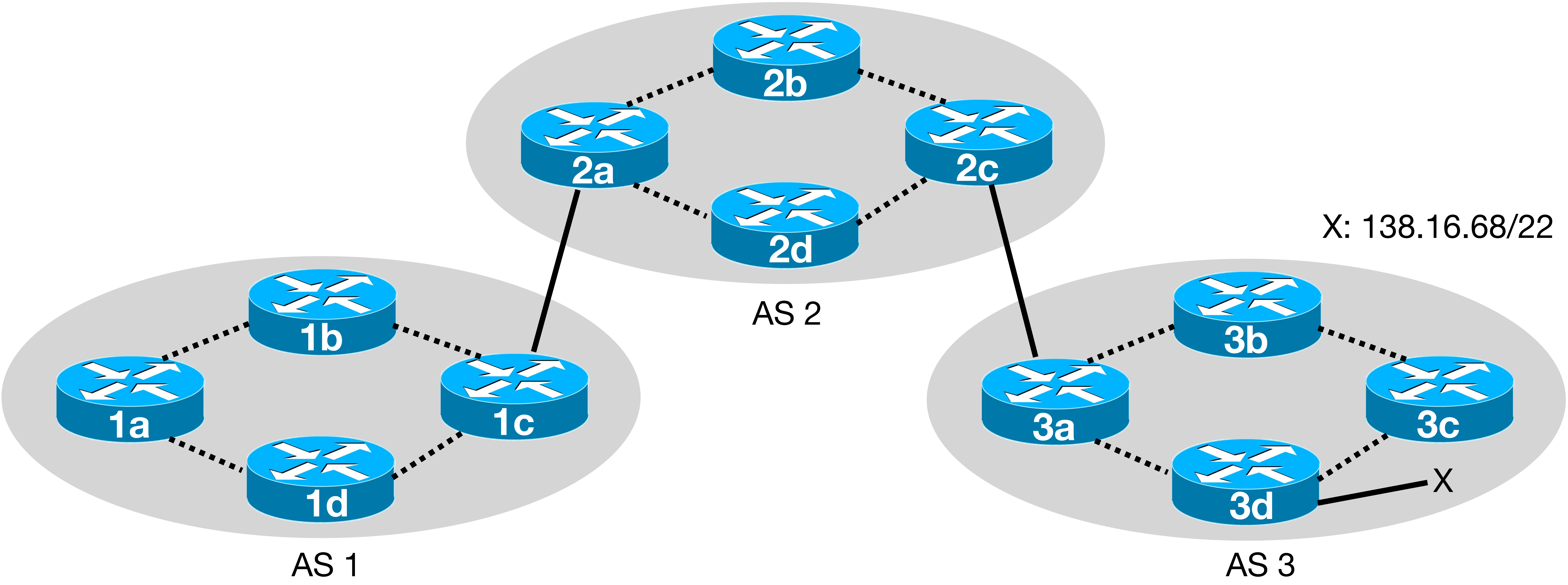
Change Route Advising to Protocol

- Using messages to drive the protocol execution



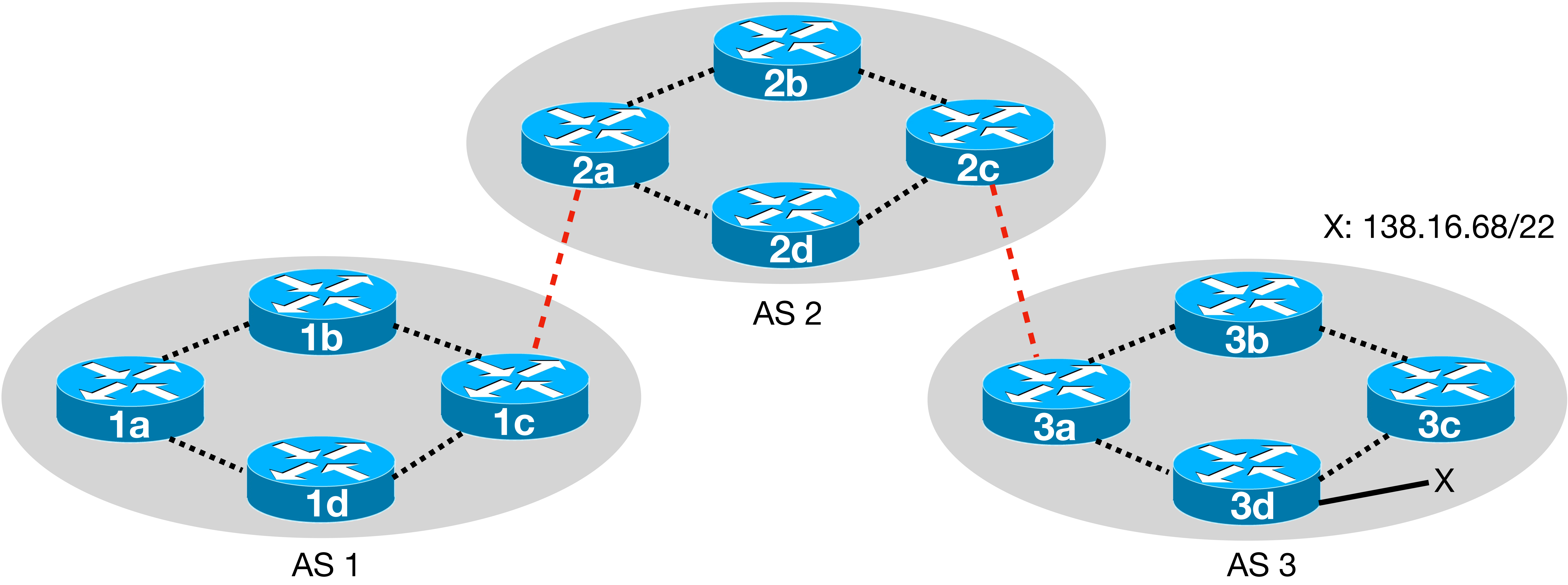
Change Route Advising to Protocol

- Using messages to drive the protocol execution
  - Internal message: advertise the prefix reachability within an AS



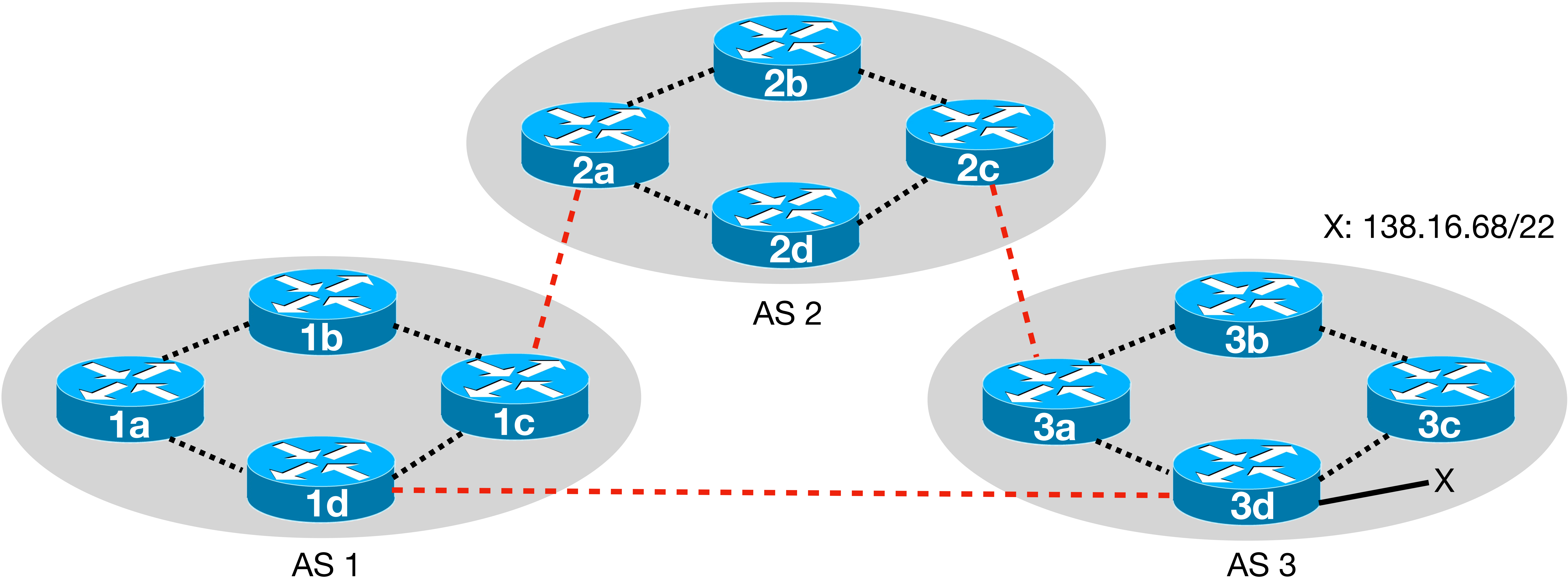
Change Route Advising to Protocol

- Using messages to drive the protocol execution
  - Internal message: advertise the prefix reachability within an AS
  - External message: advertise the prefix reachability across ASs



Determine the Routing Path

- Two paths from AS1 to AS3
  - AS 1 —> AS 3
  - AS 1 —> AS 2 —> AS 3

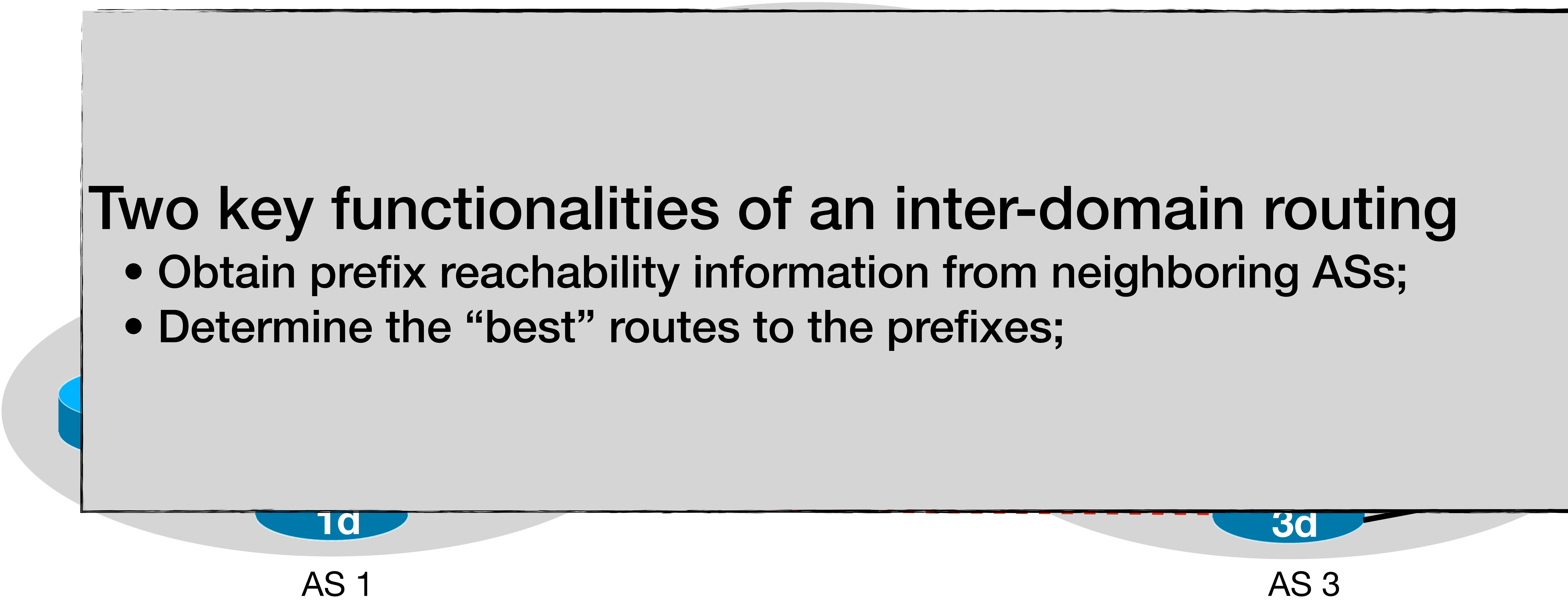


Determine the Routing Path

- Two paths from AS1 to AS3
  - AS 1 —> AS 3
  - AS 1 —> AS 2 —> AS 3

Two key functionalities of an inter-domain routing

- Obtain prefix reachability information from neighboring ASs;
- Determine the “best” routes to the prefixes;



BGP (Border Gateway Protocol)

- BGP was developed in 1989 (BGP-1)
  - Designed to address the problem of EGP (exterior gateway protocol)
  - RFC 4271
  - Current version: BGP-4

```
Network Working Group                                   Y. Rekhter, Ed.
Request for Comments: 4271                                    T. Li, Ed.
Obsoletes: 1771                                            S. Hares, Ed.
Category: Standards Track                                   January 2006
```

A Border Gateway Protocol 4 (BGP-4)

Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2006).

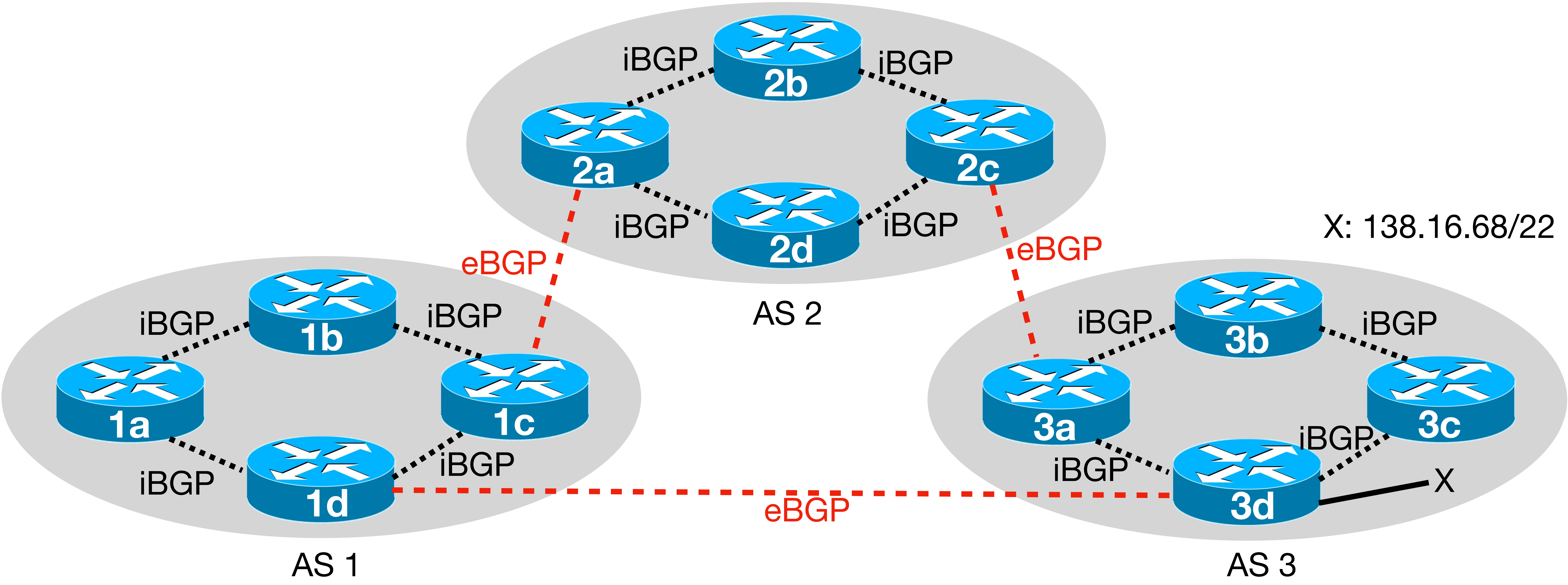
Abstract

This document discusses the Border Gateway Protocol (BGP), which is an inter-Autonomous System routing protocol.

The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. This network reachability information includes information on the list of Autonomous Systems (ASes) that reachability information traverses.

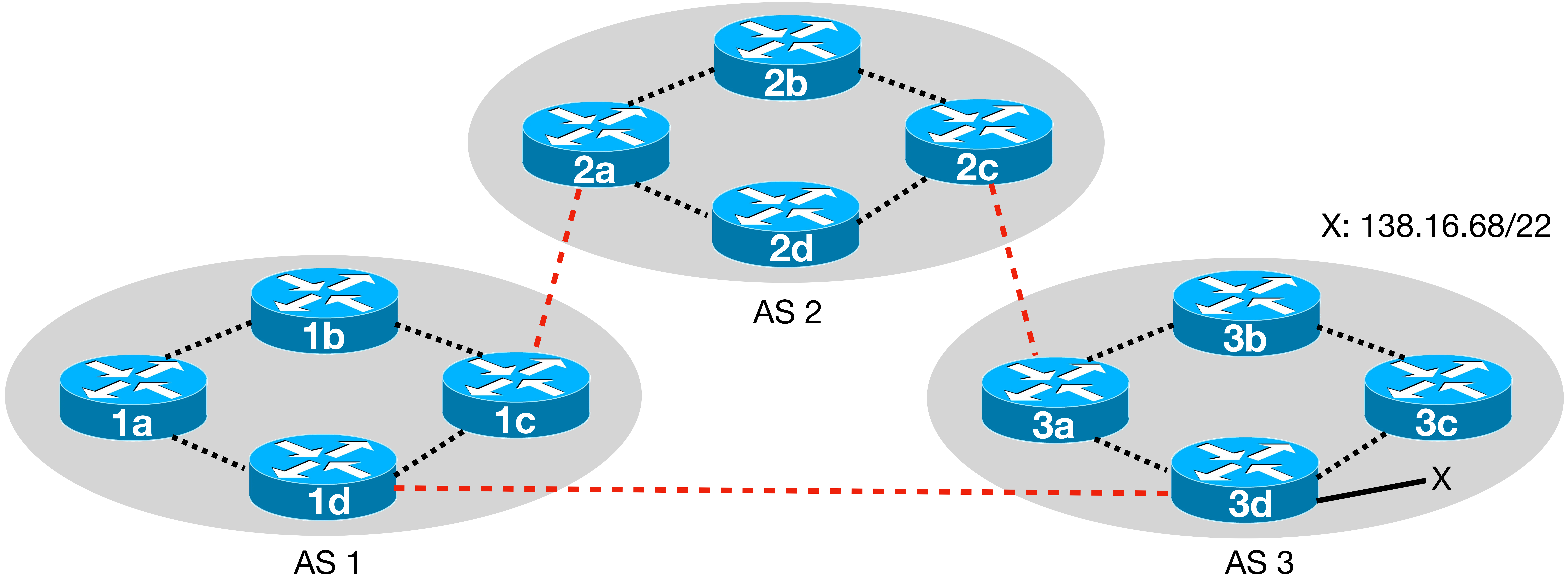
BGP Route Advertisement

- BGP runs over TCP (discussed later) using port 179
  - eBGP connection: span two ASs
  - iBGP connection: span routers within an AS



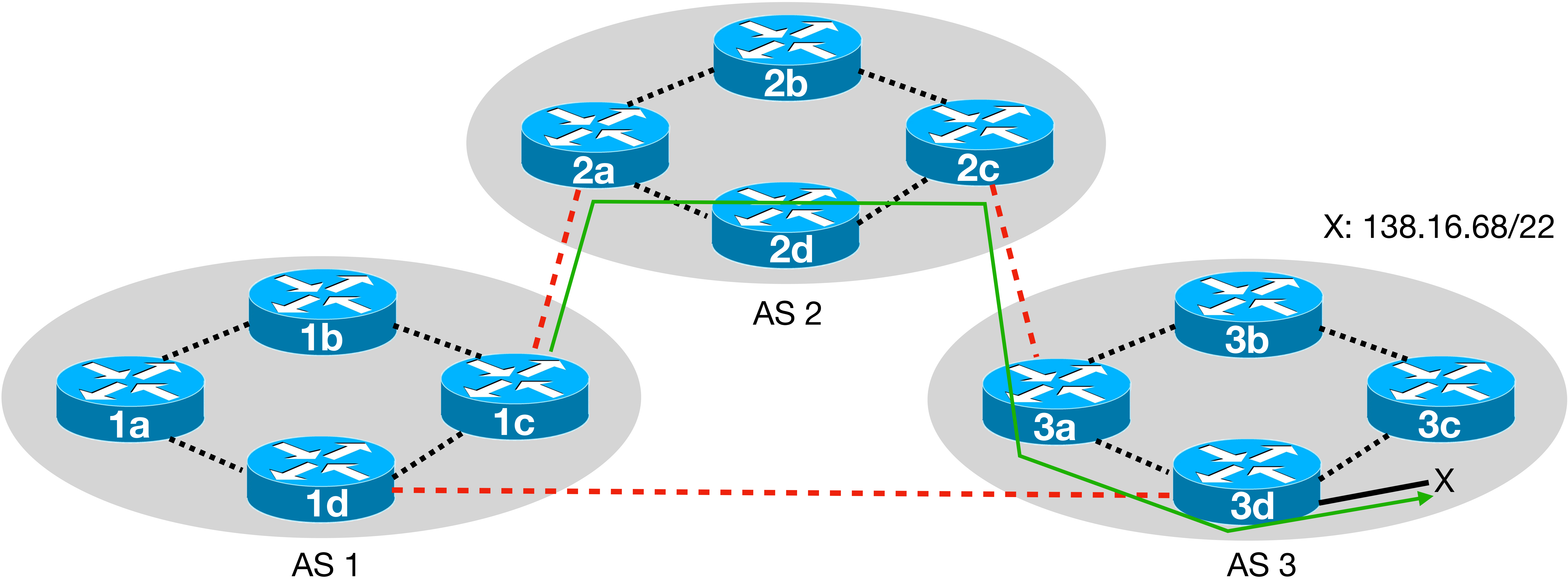
BGP Route Selection

- A BGP route: AS-PATH and NEXT-HOP
  - AS-PATH: the list of ASs through which the advertisement has passed
  - NEXT-HOP: the IP addr. of the router interface that begins the AS-PATH



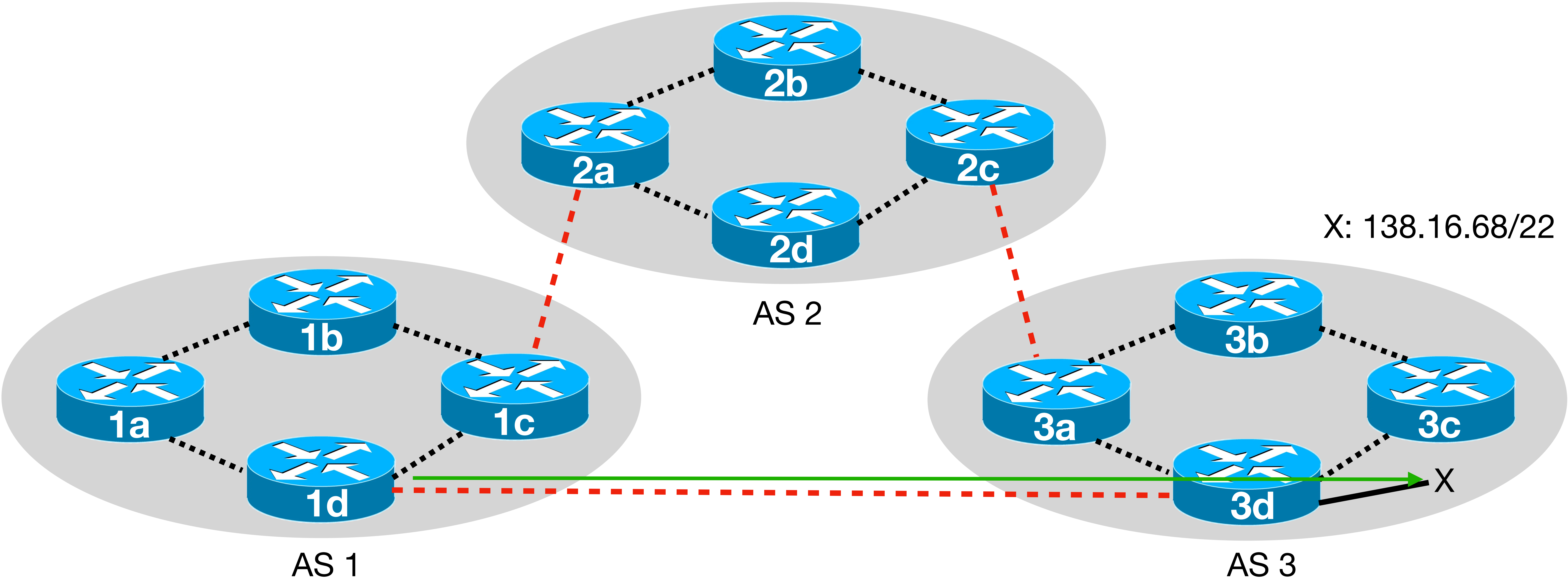
A BGP Route Example

- AS-PATH: AS2 AS 3; X
- NEXT-HOP: IP address of the leftmost interface for router 2a



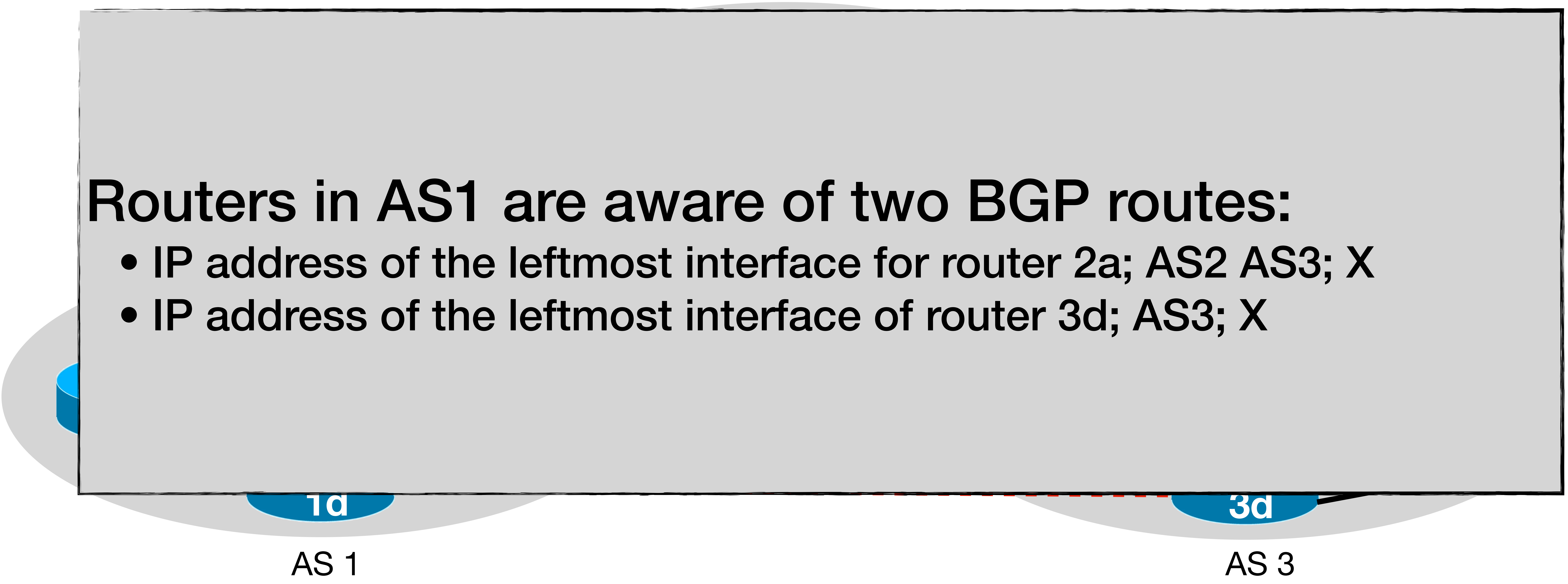
A BGP Route Example (cont'd)

- AS-PATH: AS 3; X
- NEXT-HOP: IP address of the leftmost interface of router 3d



A BGP Route Example (cont’d)

- AS-PATH: AS 3; X
- NEXT-HOP: IP address of the leftmost interface of router 3d

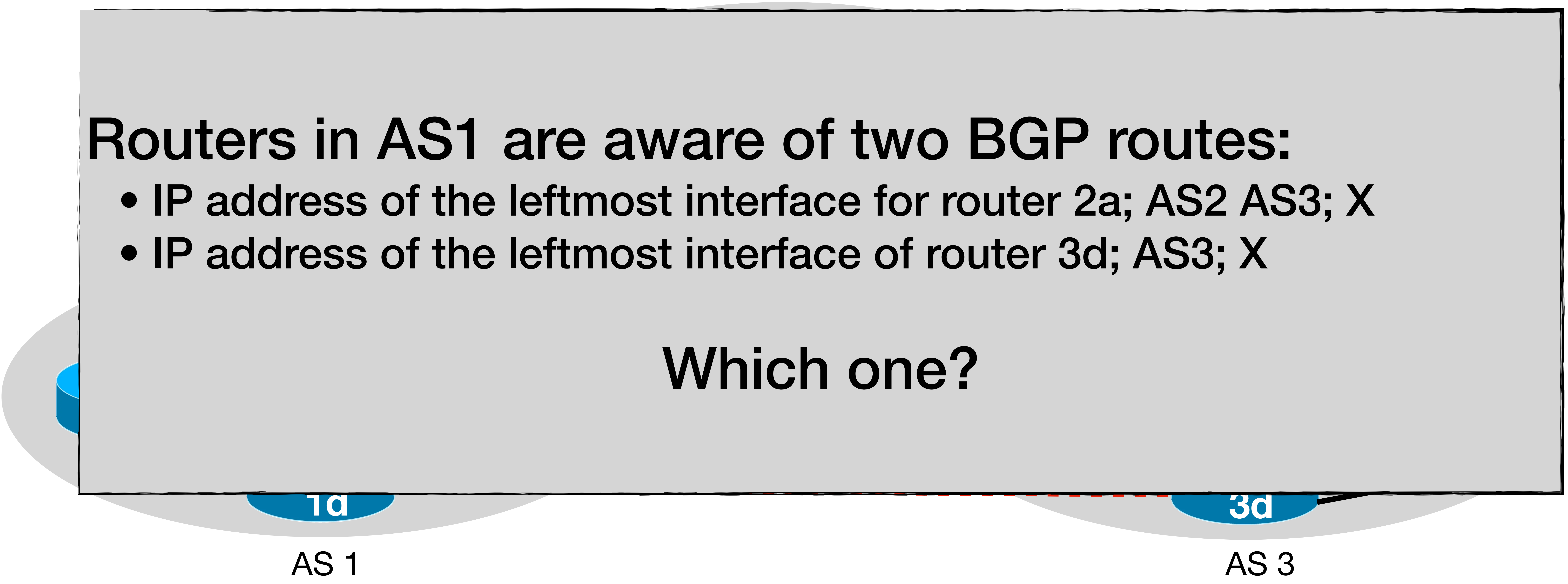


Routers in AS1 are aware of two BGP routes:

- IP address of the leftmost interface for router 2a; AS2 AS3; X
- IP address of the leftmost interface of router 3d; AS3; X

A BGP Route Example (cont'd)

- AS-PATH: AS 3; X
- NEXT-HOP: IP address of the leftmost interface of router 3d



Routers in AS1 are aware of two BGP routes:

- IP address of the leftmost interface for router 2a; AS2 AS3; X
- IP address of the leftmost interface of router 3d; AS3; X

Which one?

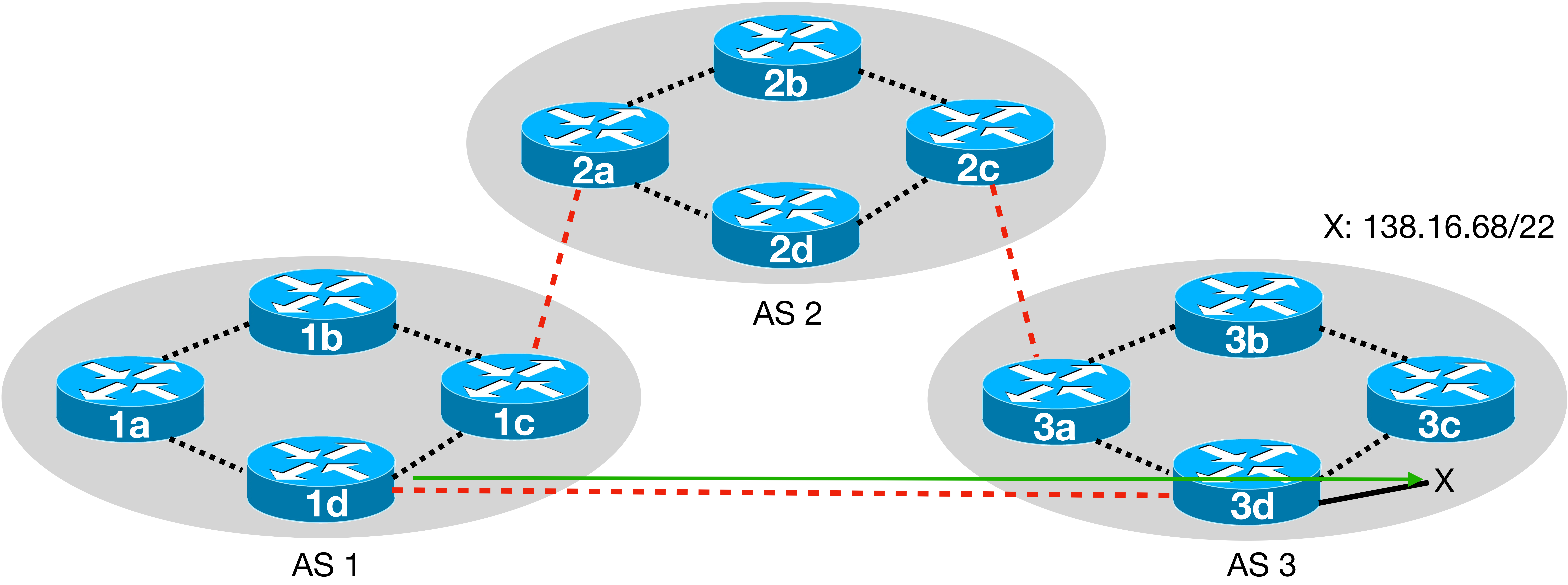
Let's Start with Simple

- Hot Potato Routing
  - The router chooses the least cost to the NEXT-HOP

Step 1: Learn from inter-AS protocol that subnet X is reachable via multiple gateways → **Step 2:** Use routing info from intra-AS protocol to determine cost of least-cost paths to each of the gateway → **Step 3:** Choose the gateway that has the smallest least cost → **Step 4:** Determine from the forwarding table the interface *I* that leads to the least-cost gateway

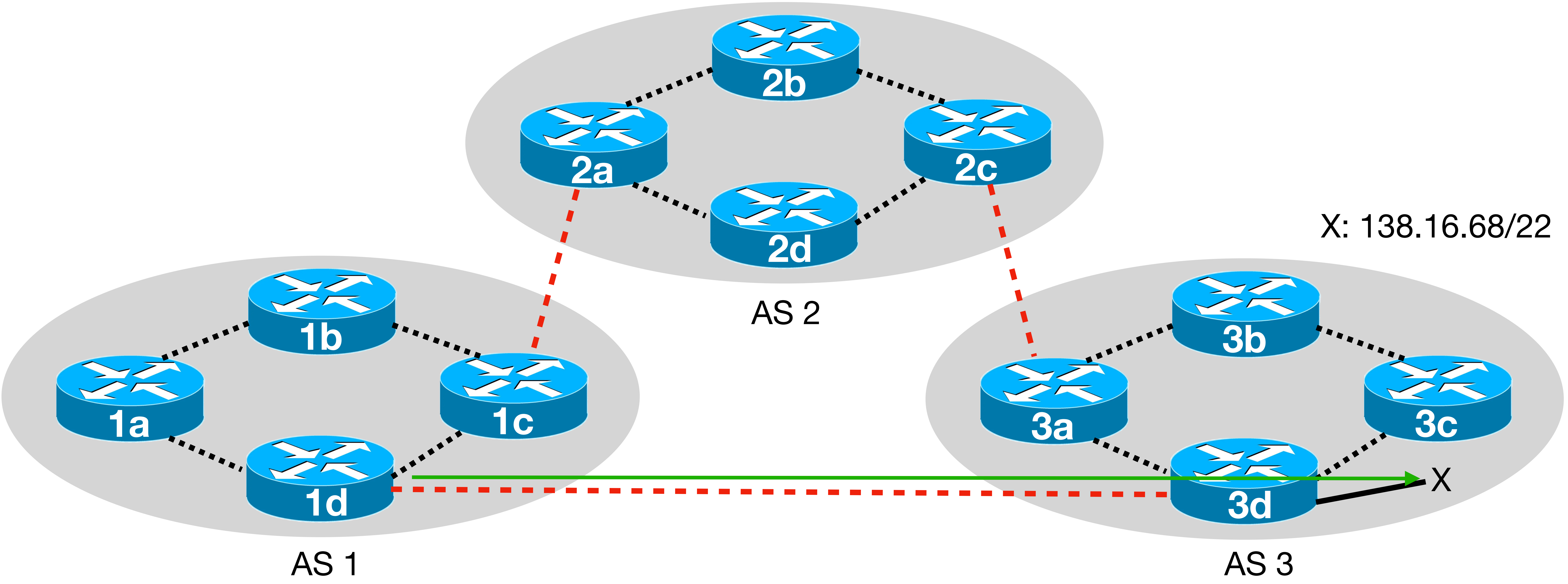
Hot Potato Routing: 1b—>X

- Cost is the number of hops
  - AS2->AS3->X: Cost (1b->2a) = 2
  - AS3-X: Cost (1b->3d) = 3



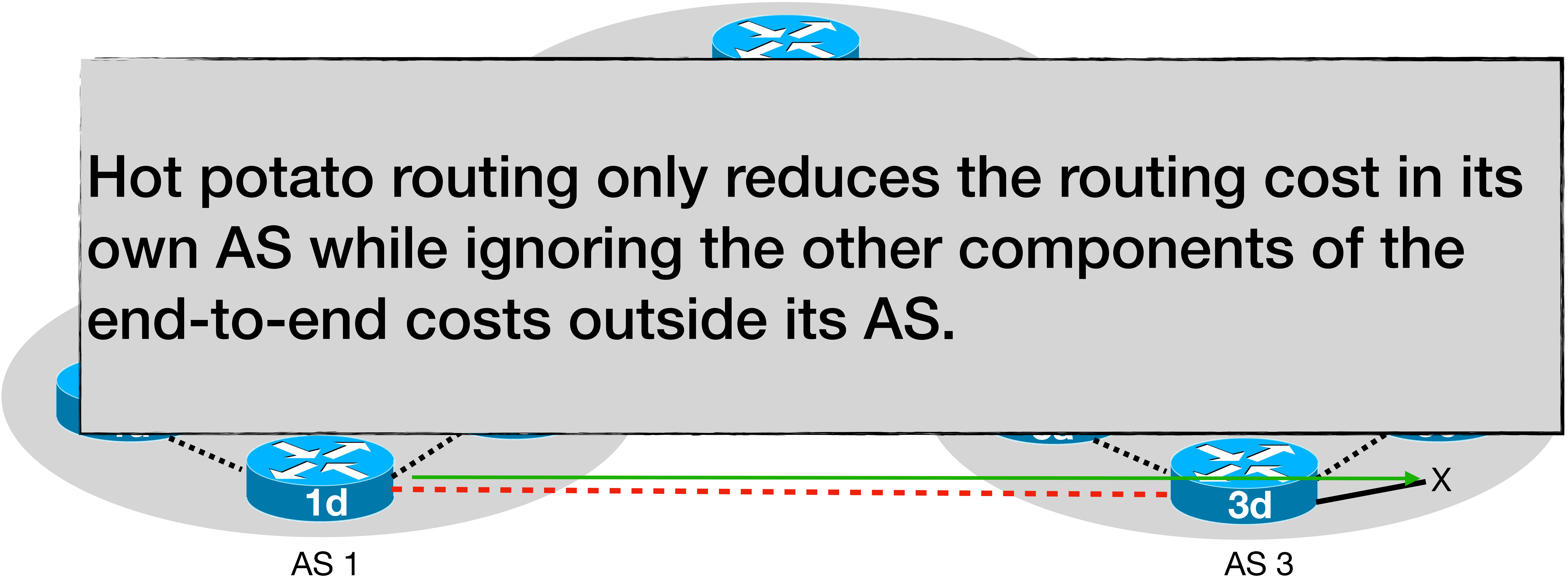
Hot Potato Routing: 1b—>X

- Cost is the number of hops
  - AS2->AS3->X: Cost (1b->2a) = 2, selected
  - AS3-X: Cost (1b->3d) = 3



Hot Potato Routing: 1b—>X

- Cost is the number of hops
  - AS2->AS3->X: Cost (1b->2a) = 2, selected
  - AS3-X: Cost (1b->3d) = 3

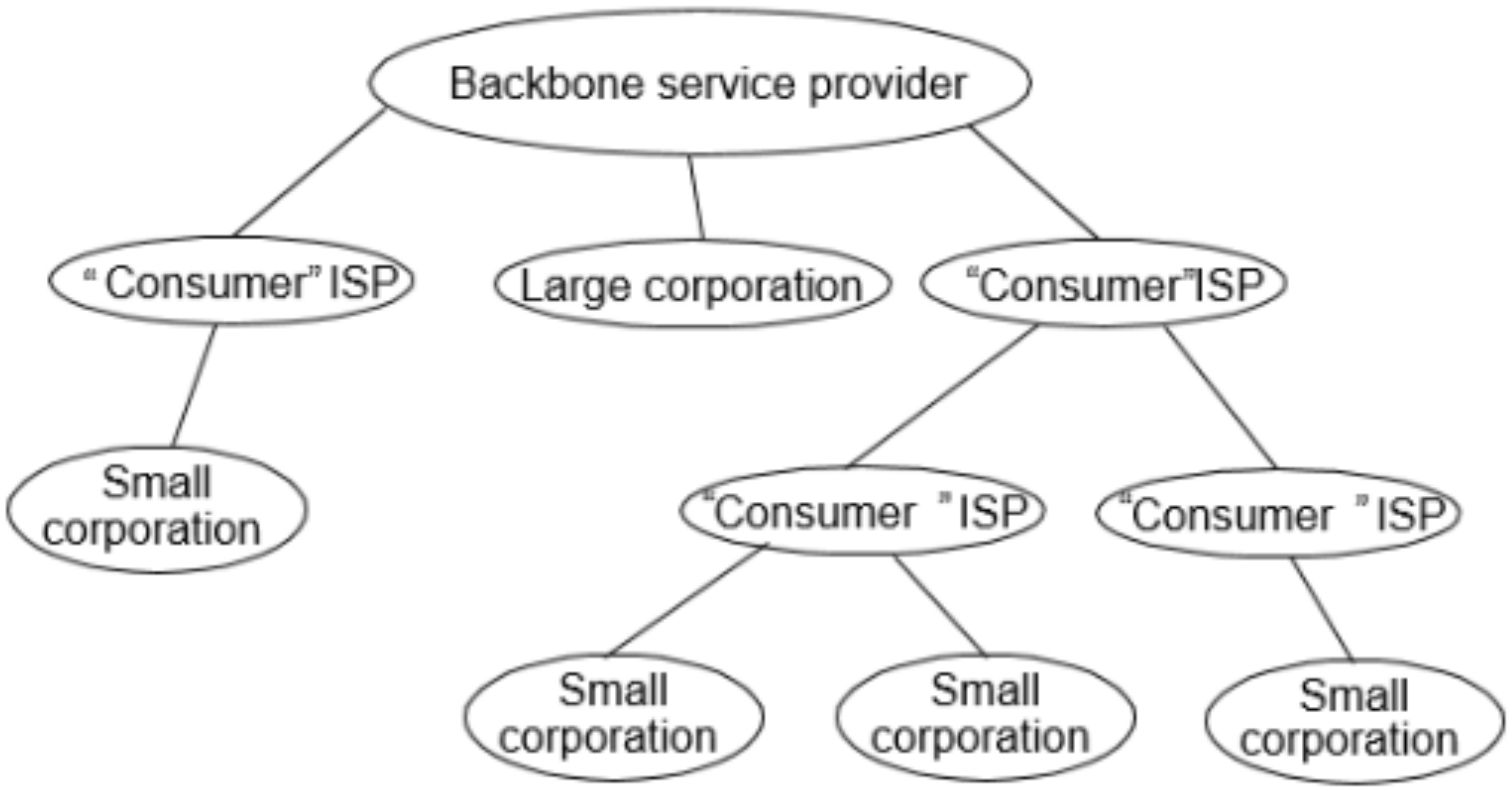


Hot potato routing only reduces the routing cost in its own AS while ignoring the other components of the end-to-end costs outside its AS.

BGP Routing in Reality: Very Complicated!

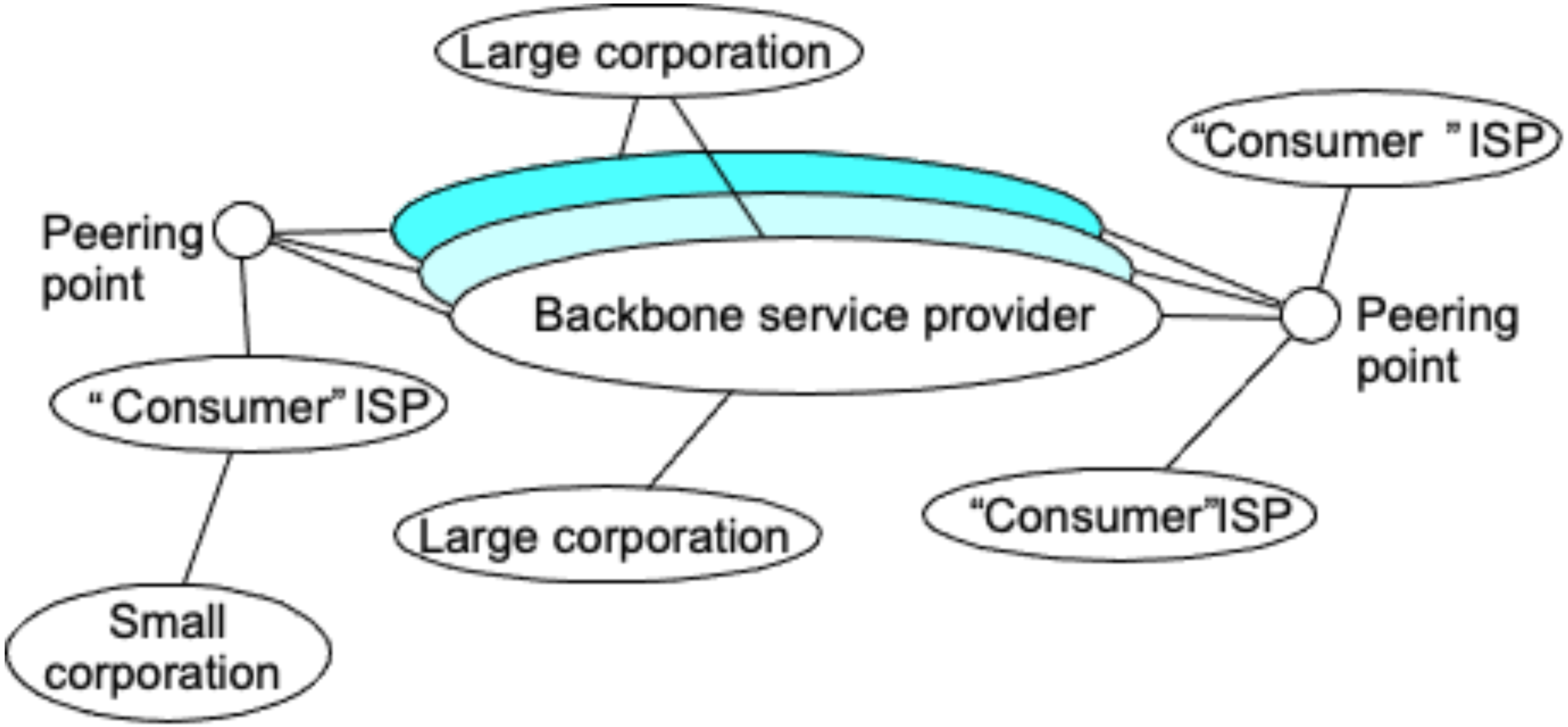
BGP Routing in Reality: Very Complicated!

- Internet structure (original idea)



BGP Routing in Reality: Very Complicated!

- Internet structure (today)



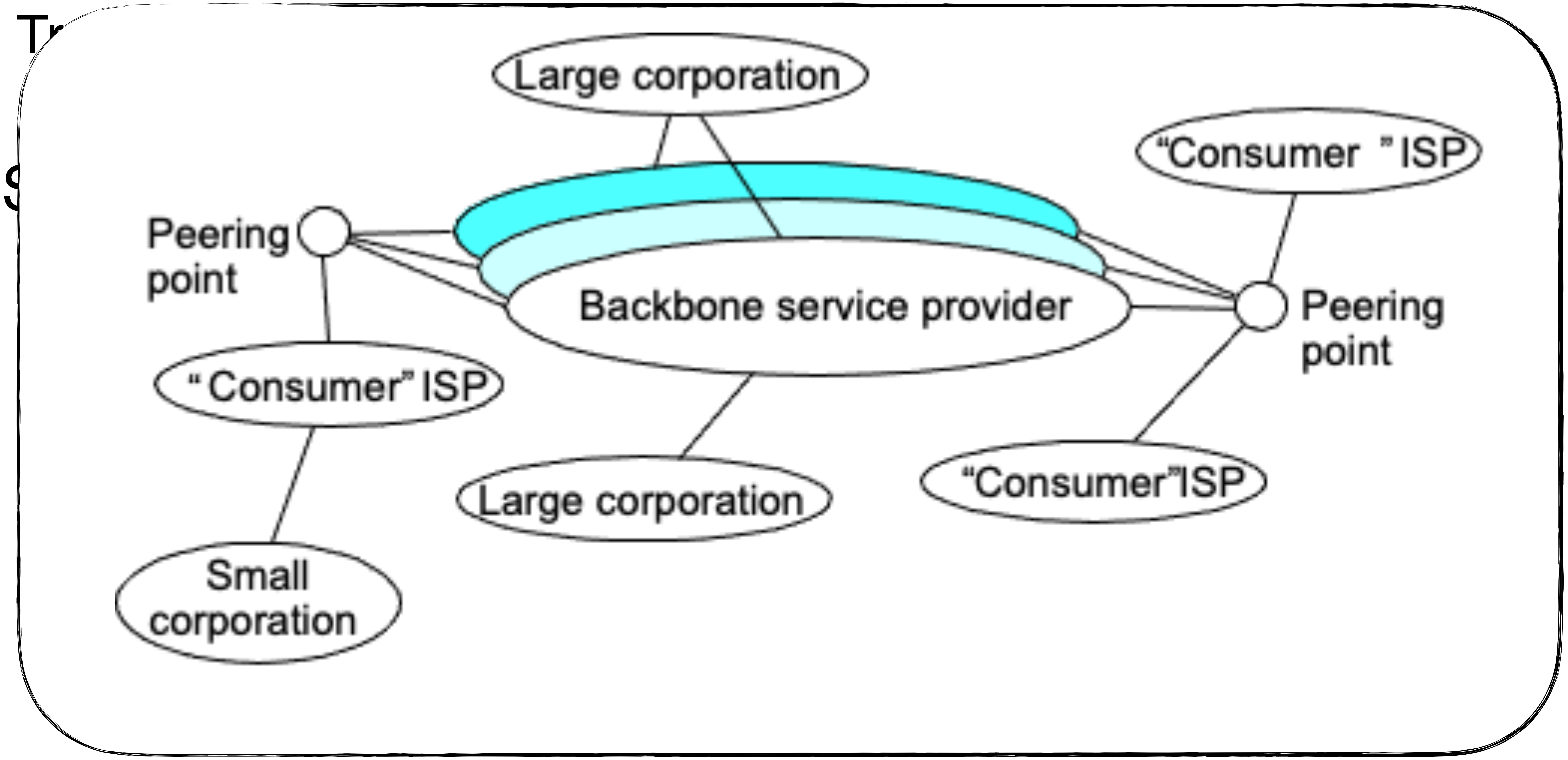
Autonomous System (AS)

- AS Traffic Types
  - Local: start or end within an AS
  - Transit: pass through an AS

- AS Types
  - sub AS: has a single connection to other AS
    - Carries local traffic only
  - multi-homed AS: has connections to more than one AS
    - Refuse to carry transit traffic
  - transmit AS: has connections to more than one AS
    - Carries both transmit and local traffic

Autonomous System (AS)

- AS Traffic Types
  - Local: start or end within an AS
  - Tr
- AS
  -
  -
  -



AS Characteristics

- #1: Each AS has one or more border routers
  - Handle inter-AS traffic

- #2: Each AS has at least one BGP speaker that participates in the routing process
  - Border routers might or might not be a BGP speaker

AS Characteristics (Cont'd)

- #3: A BGP speaker establishes BGP sessions with peers and advertises route information
  - Local network names and prefixes
  - Other reachable networks (transit AS only)
  - Provide AS-PATH
  - Withdraw routes

AS Characteristics (Cont'd)

- #3: A BGP speaker establishes BGP sessions with peers and advertises route information
  - Local network names and prefixes
  - Other reachable networks (transit AS only)
  - Provide AS-PATH
  - Withdraw routes

> - Peers: neighbor routers exchange routing information
> - Advertises: an AS publicizes its learned routing information

AS Characteristics (Cont'd)

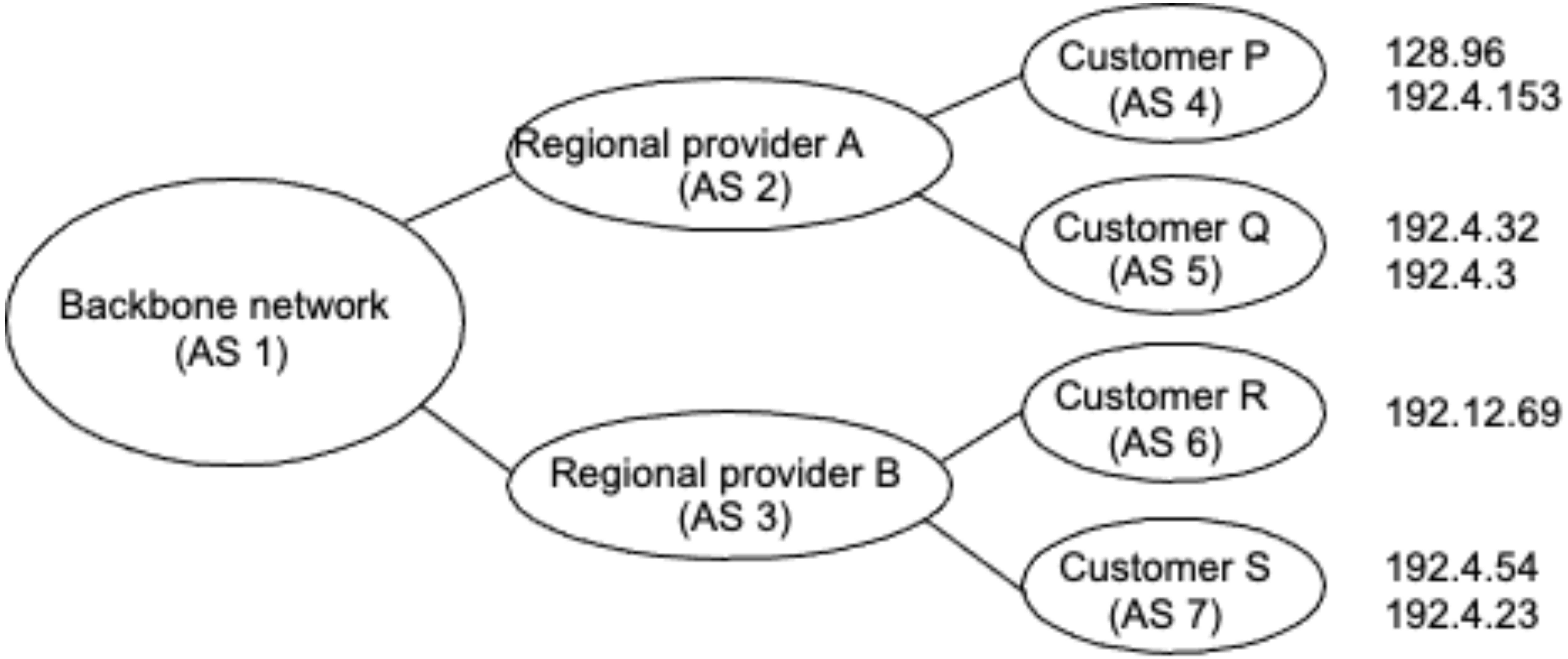
- #3: A BGP speaker establishes BGP sessions with peers and advertises route information
  - Local network names and prefixes
  - Other reachable networks (transit AS only)
  - Provide AS-PATH
  - Withdraw routes

> - Peers: neighbor routers exchange routing information
> - Advertises: an AS publicizes its learned routing information

> - Unlike RIP and OSPF, BGP advertises complete path as an enumerated list of autonomous systems to reach a particular network

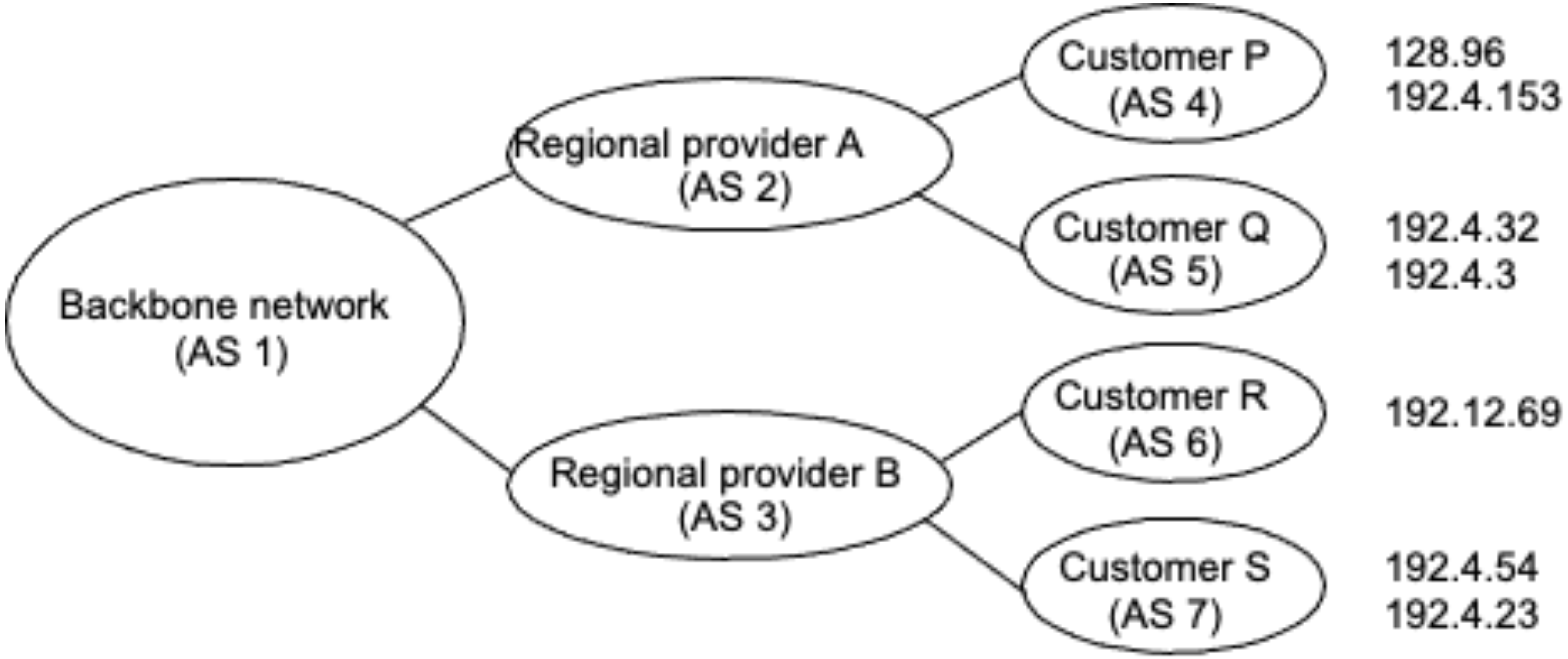
A BGP Example: AS2

- The speaker of AS2 advertises reachability to P and Q
  - Network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from AS2



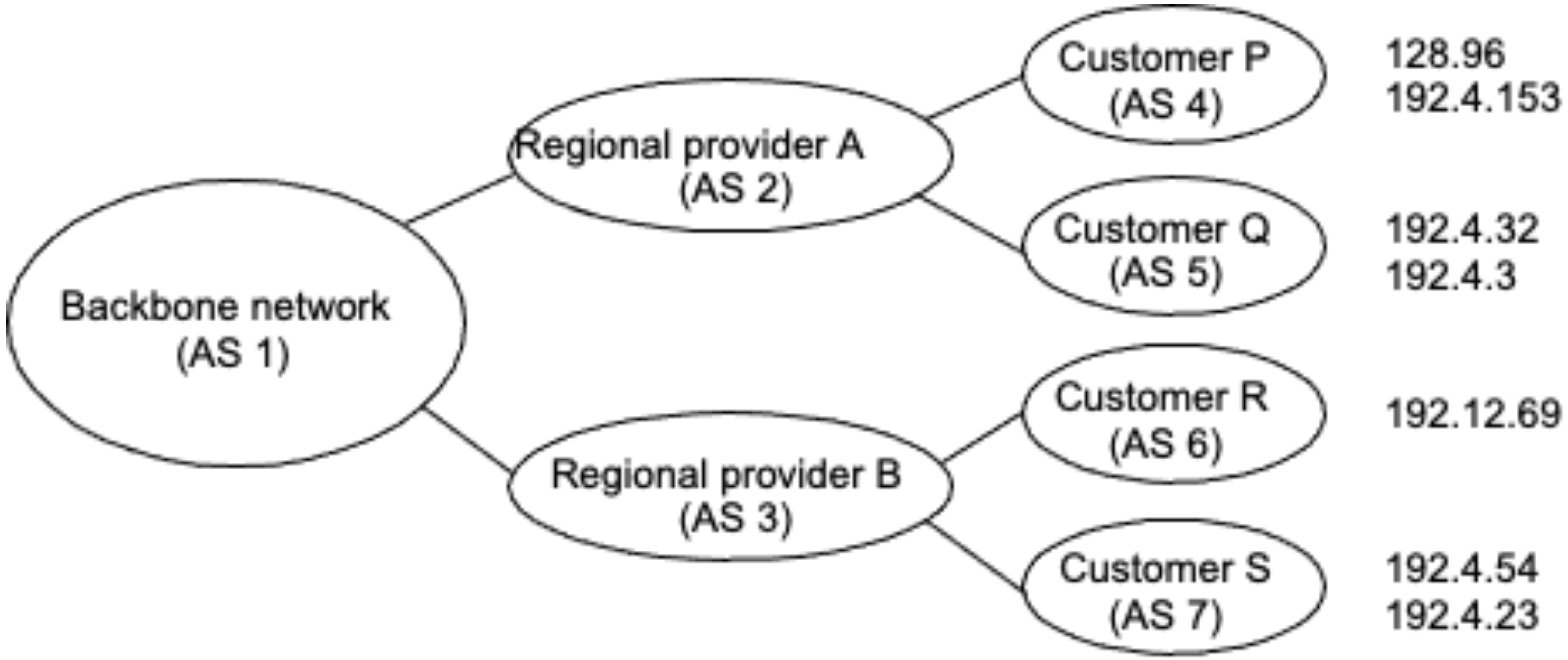
A BGP Example: AS1

- The speaker of the backbone (AS 1) advertises:
  - Network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from (AS1, AS2)



A BGP Example: Path Cancel

- The speaker can cancel previously advertised paths

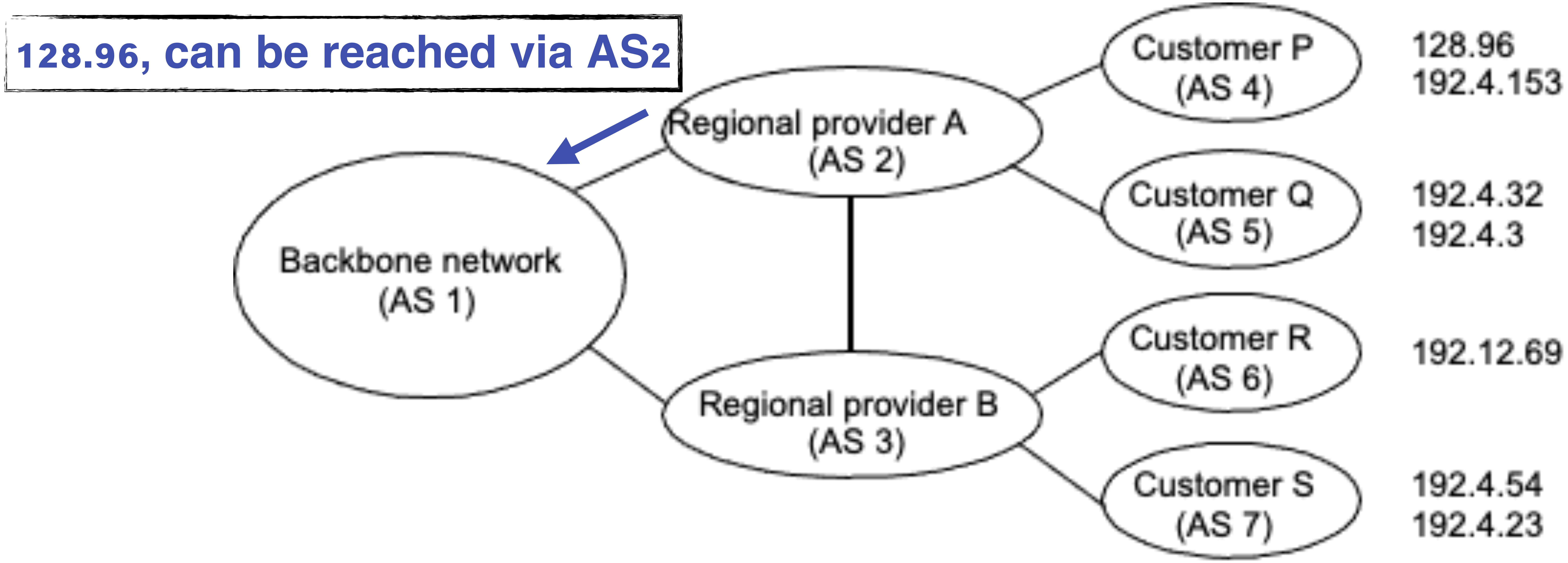


Two key functionalities of an inter-domain routing

- Obtain prefix reachability information from neighboring ASs;
- Determine the “best” routes to the prefixes;

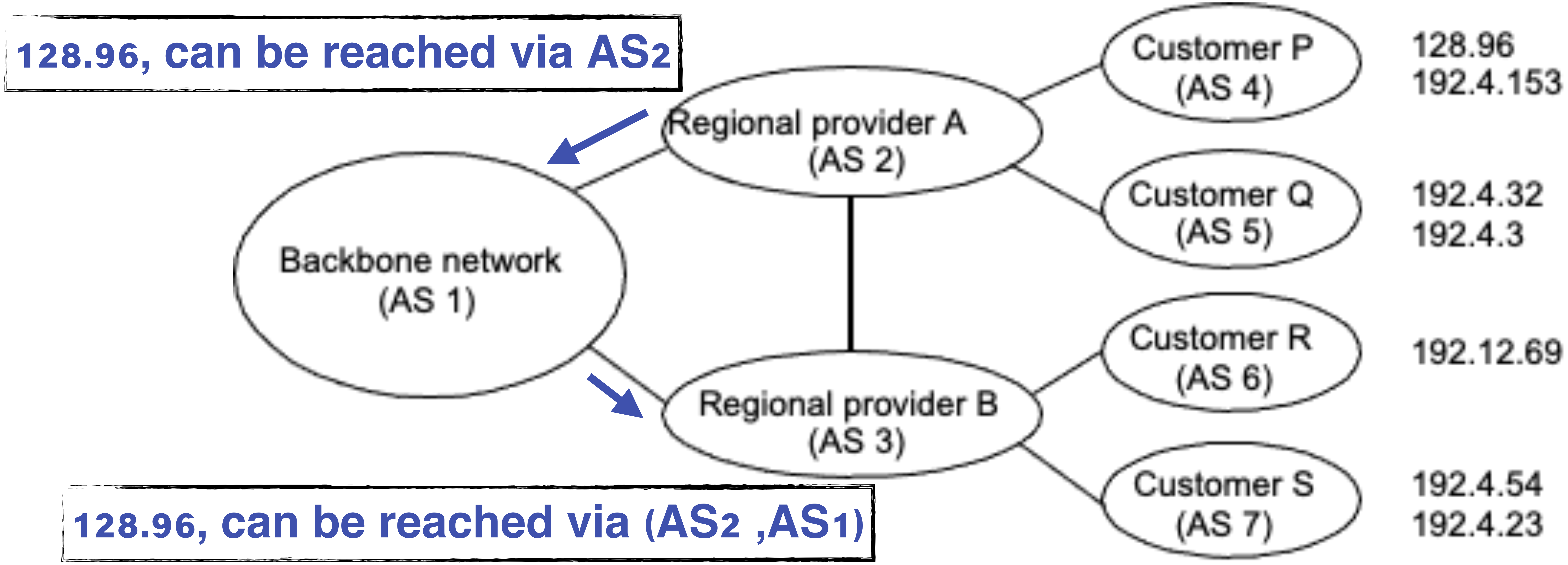
What is the Best Route?

- Loop-free paths among ASs
  - Optimality is a secondary goal



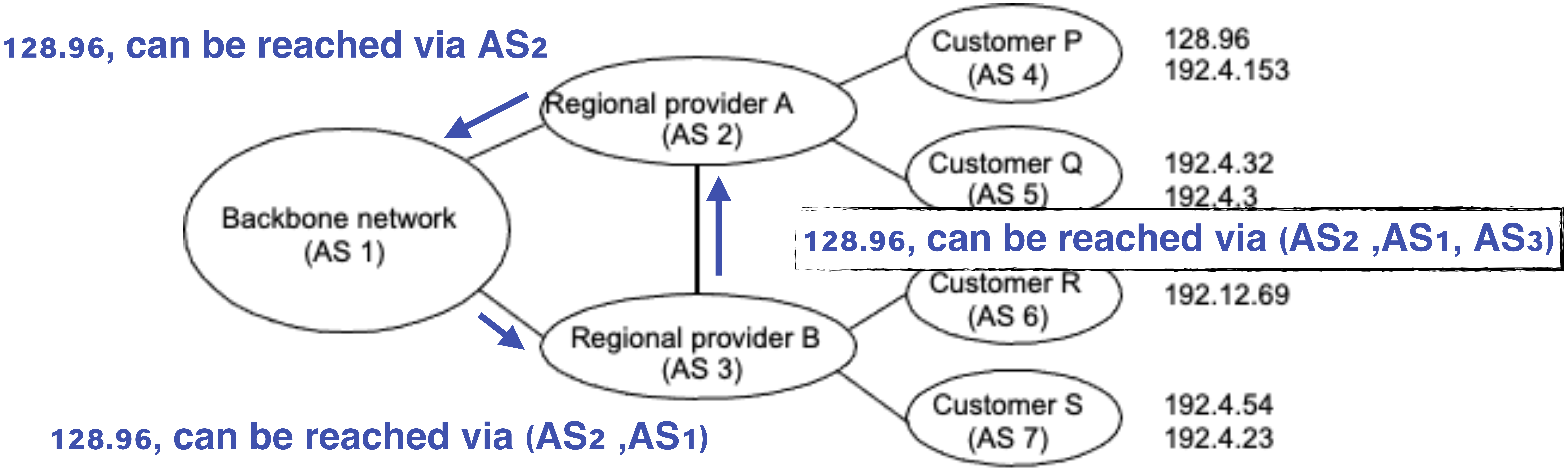
What is the Best Route?

- Loop-free paths among ASs
  - Optimality is a secondary goal



What is the Best Route?

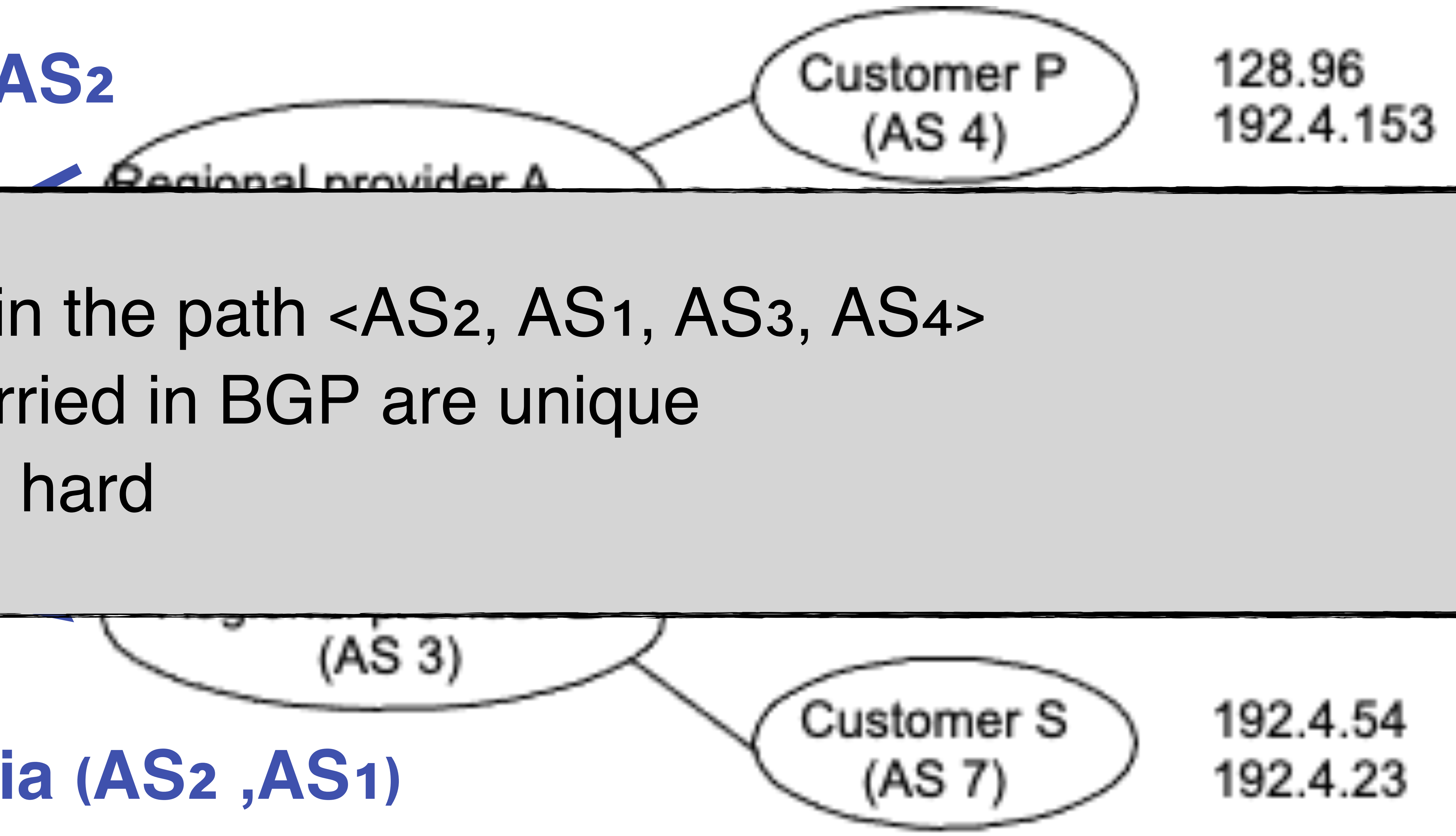
- Loop-free paths among ASs
  - Optimality is a secondary goal



What is the Best Route?

- Loop-free paths among ASs
  - Optimality is a secondary goal

128.96, can be reached via AS2



- AS2 sees itself in the path <AS2, AS1, AS3, AS4>
- AS numbers carried in BGP are unique
- Loop-free is not hard

S2 ,AS1, AS3)

128.96, can be reached via (AS2 ,AS1)

But, how to define "the best"?

- Loop-free paths among ASs
  - Optimality is a secondary goal

- Challenges
  - Internet size (~12K active ASs) mean large tables in BGP routers
  - Policy-compliant path (not just scalar cost of a path)
  - Different ASs use different path metrics
  - Not inherent trust among different ASs

How does BGP determine the path?

How does BGP determine the path?

Policy management

- #1: Learn — Import routing information from my neighbors
- #2: Speak — Export routing information to my neighbors

Policy in BGP

- BGP provides the capacity for enforcing policies
  - Policies are not part of BGP
  - They are provided to BGP for routing configuration

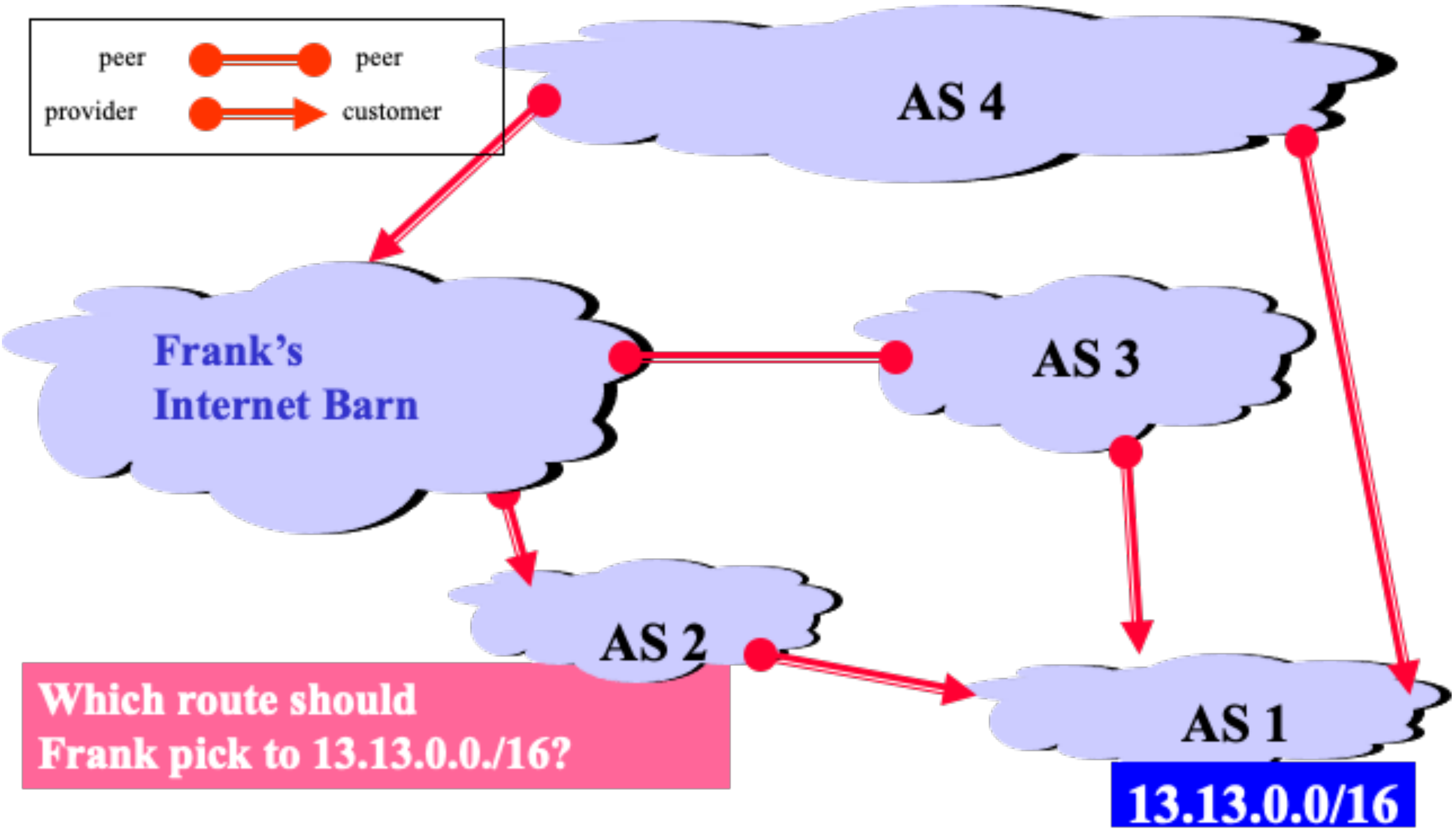
- Policy enforcement
  - Import: choose appropriate paths from multiple alternatives
  - Export: control advertisement to other ASs

Policy in BGP

- BGP provides the capacity for enforcing policies
  - Policies are not part of BGP
  - They are provided to BGP for routing configuration
- Policy enforcement
  - Import: choose appropriate paths from multiple alternatives
  - Export: control advertisement to other ASs

Policies can be arbitrarily complex. There are some common ones.

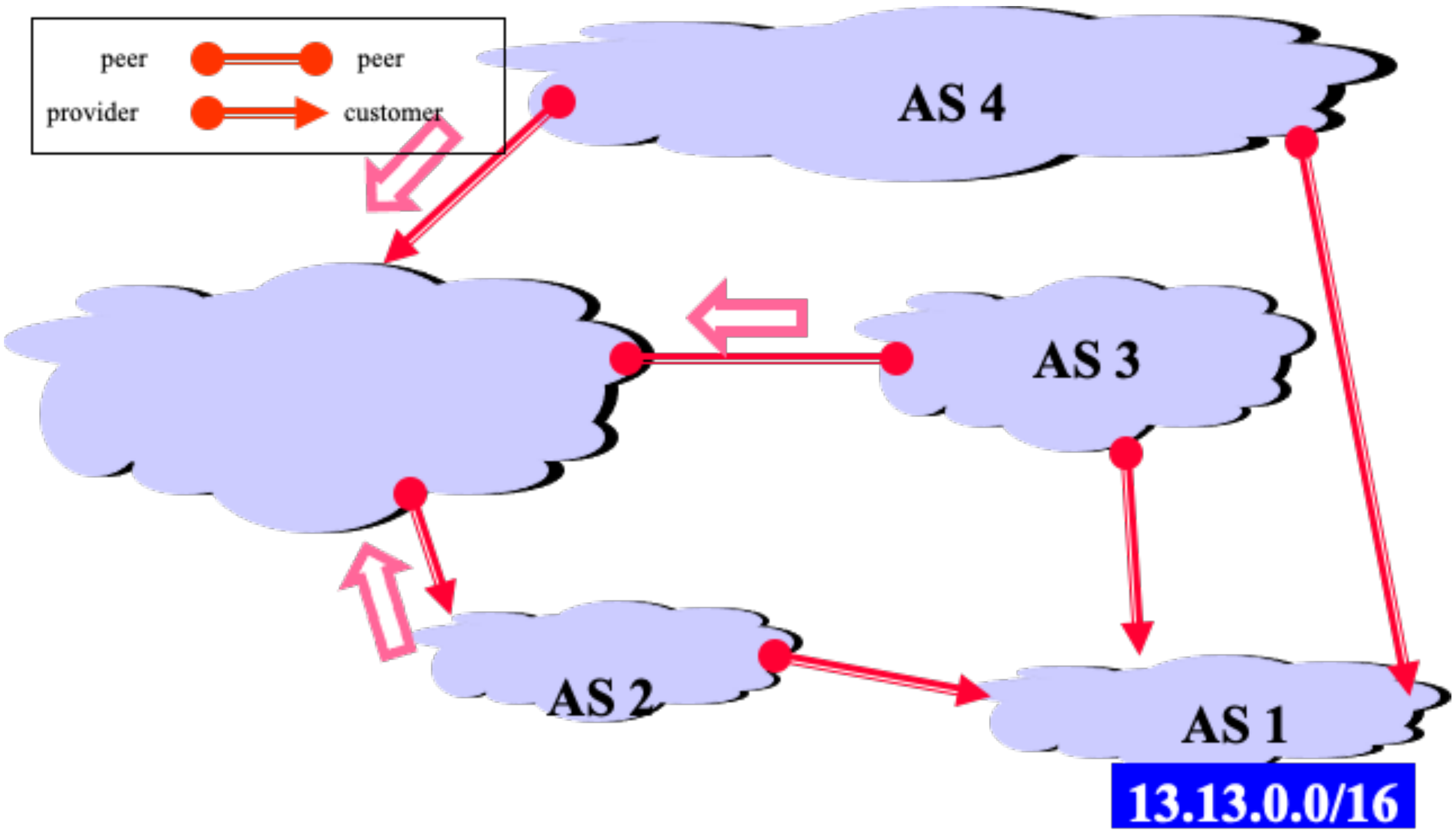
BGP Policy Example



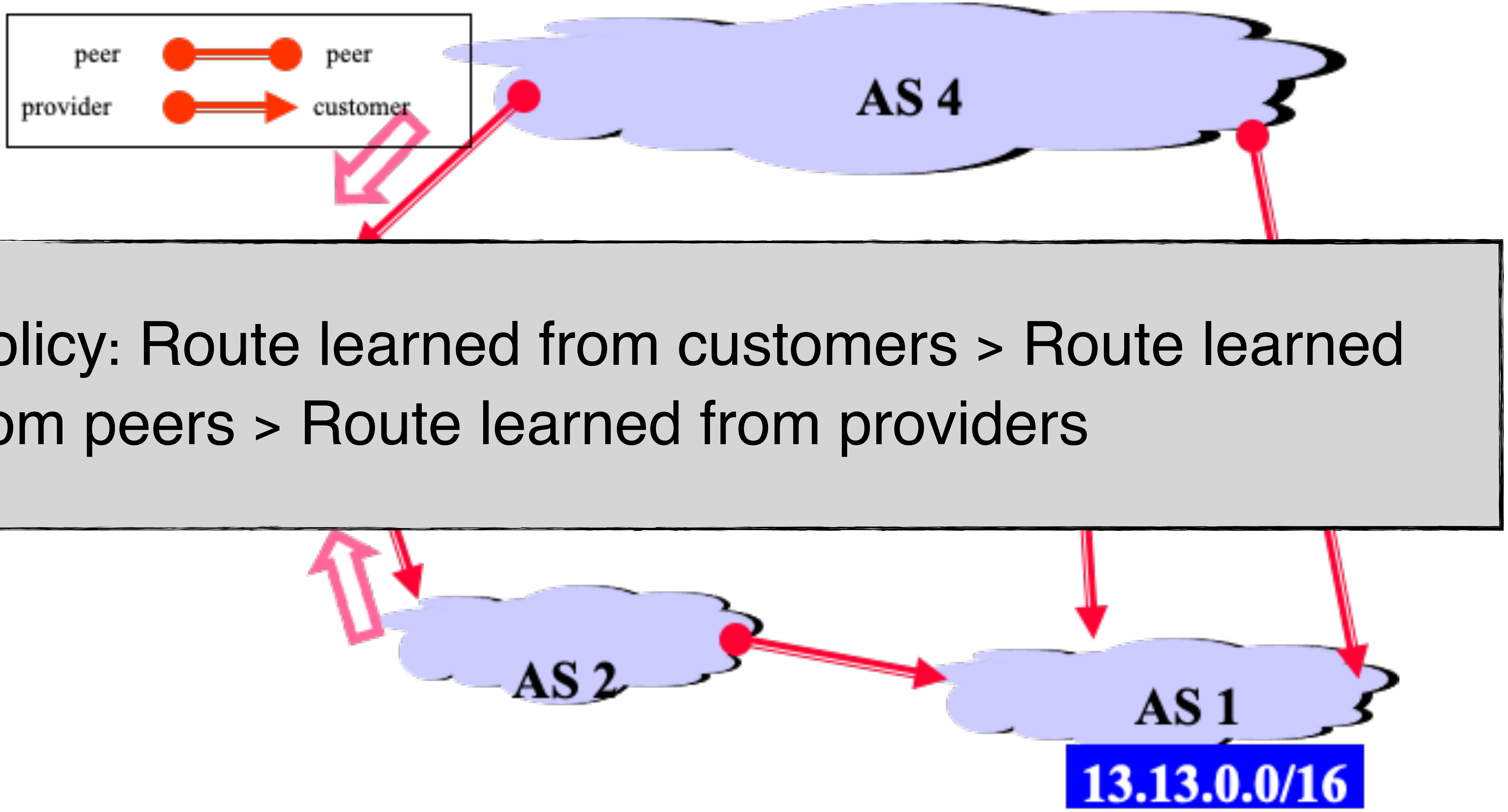
Peering and Customer-Provider

- Peering relationship
  - Peers provide transit to each other
  - Peering relationships are free with no cost
- Customer-Provider relationship
  - Customers use providers to reach the rest of the Internet
  - Customers pay providers for this

Import Policy: Prefer Customer Routing

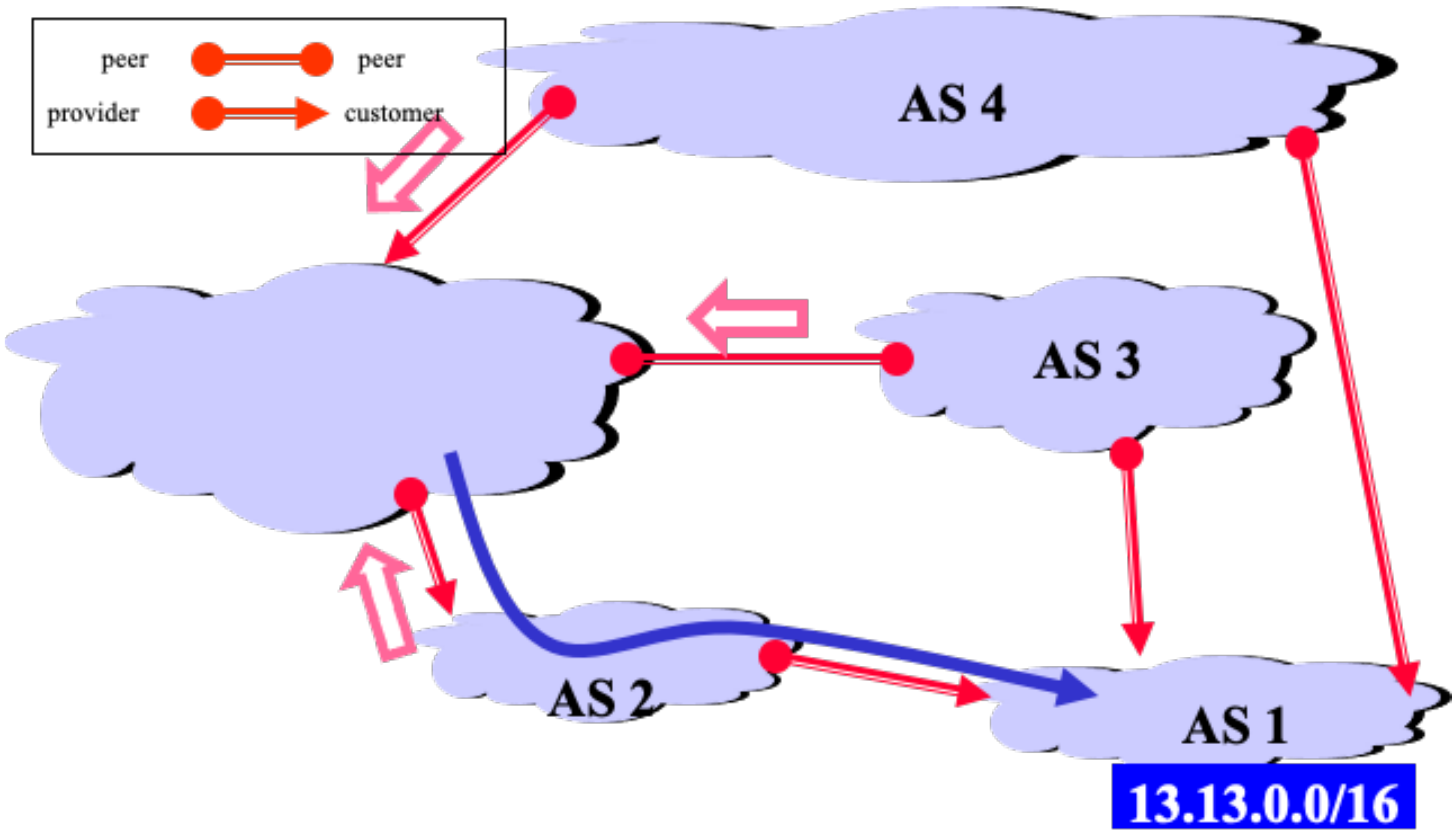


Import Policy: Prefer Customer Routing



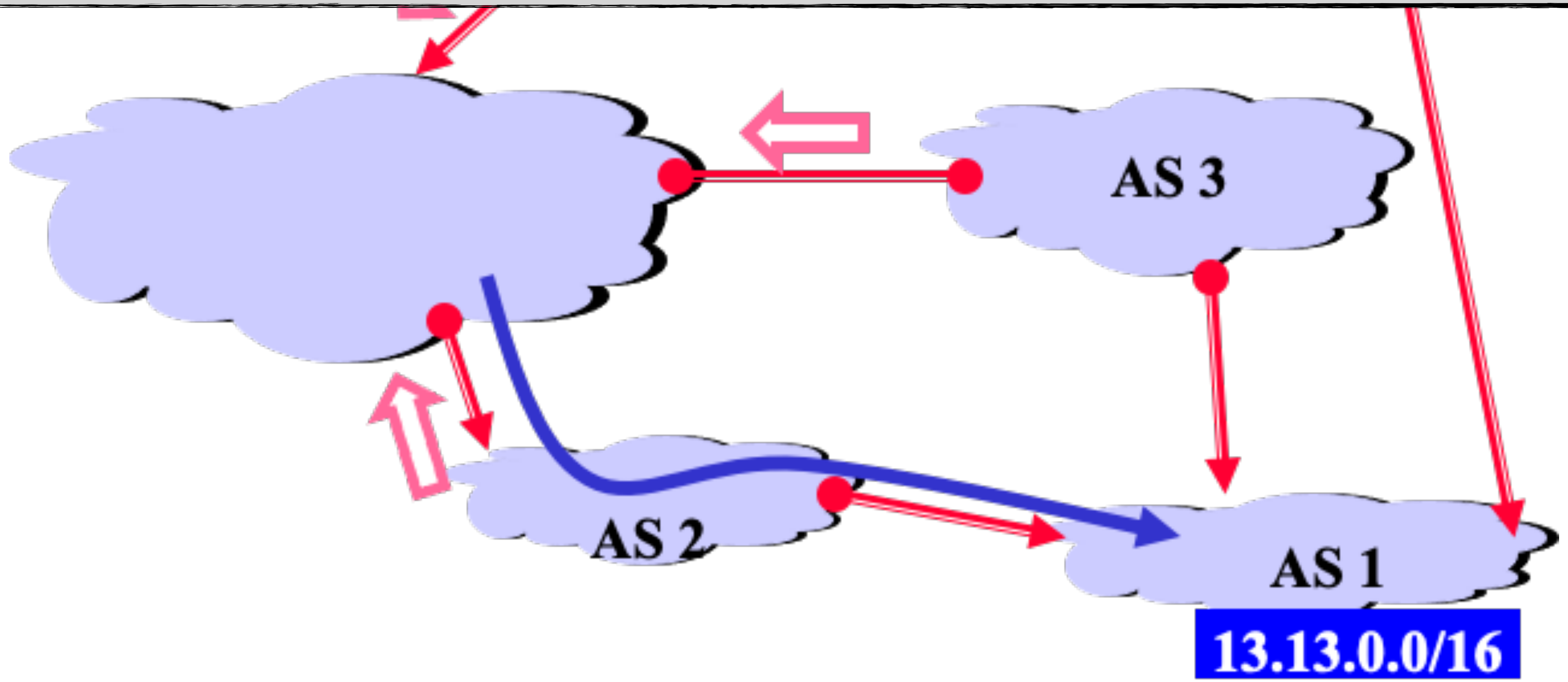
Policy: Route learned from customers > Route learned from peers > Route learned from providers

Import Policy: Prefer Customer Routing

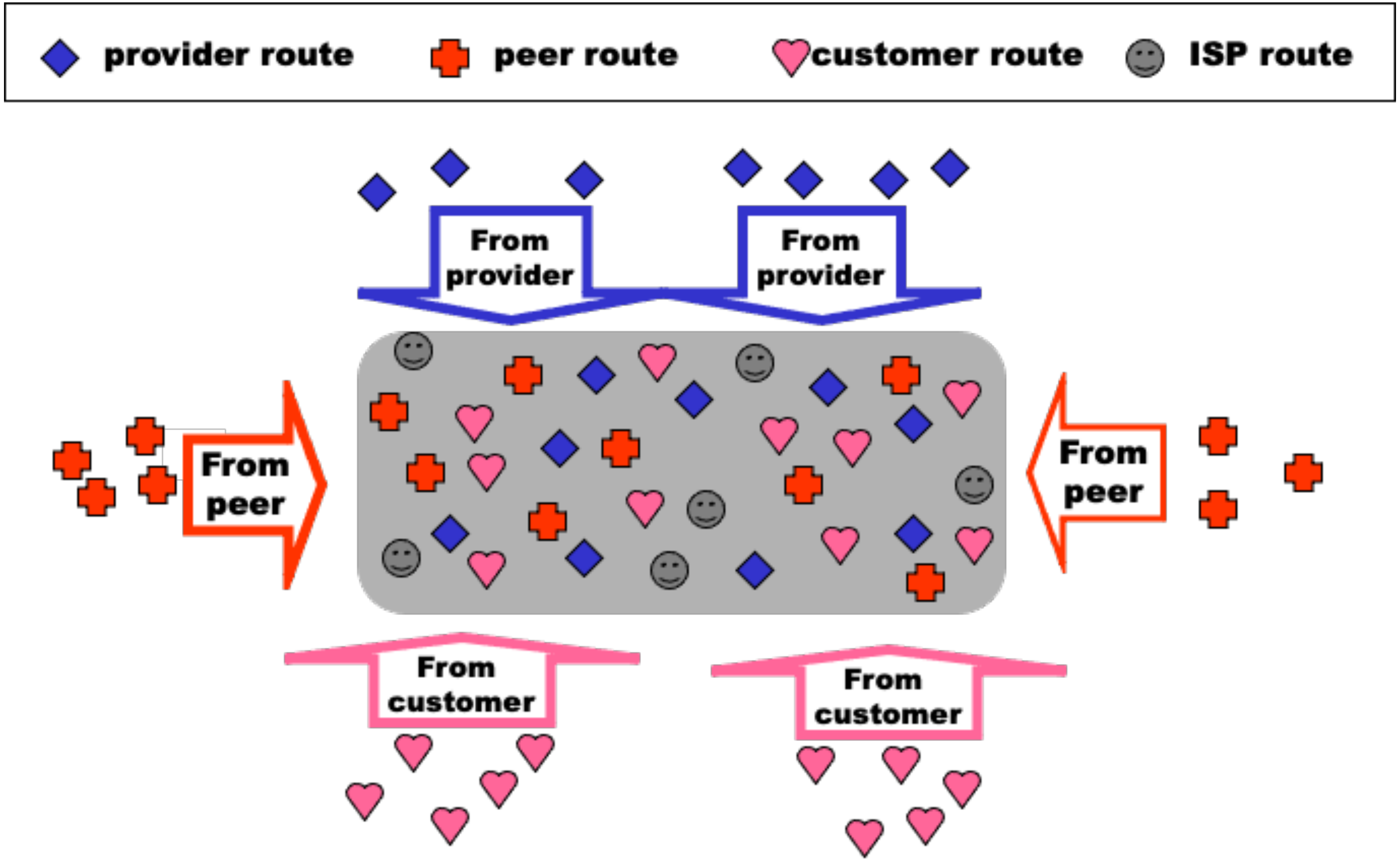


Import Policy: Prefer Customer Routing

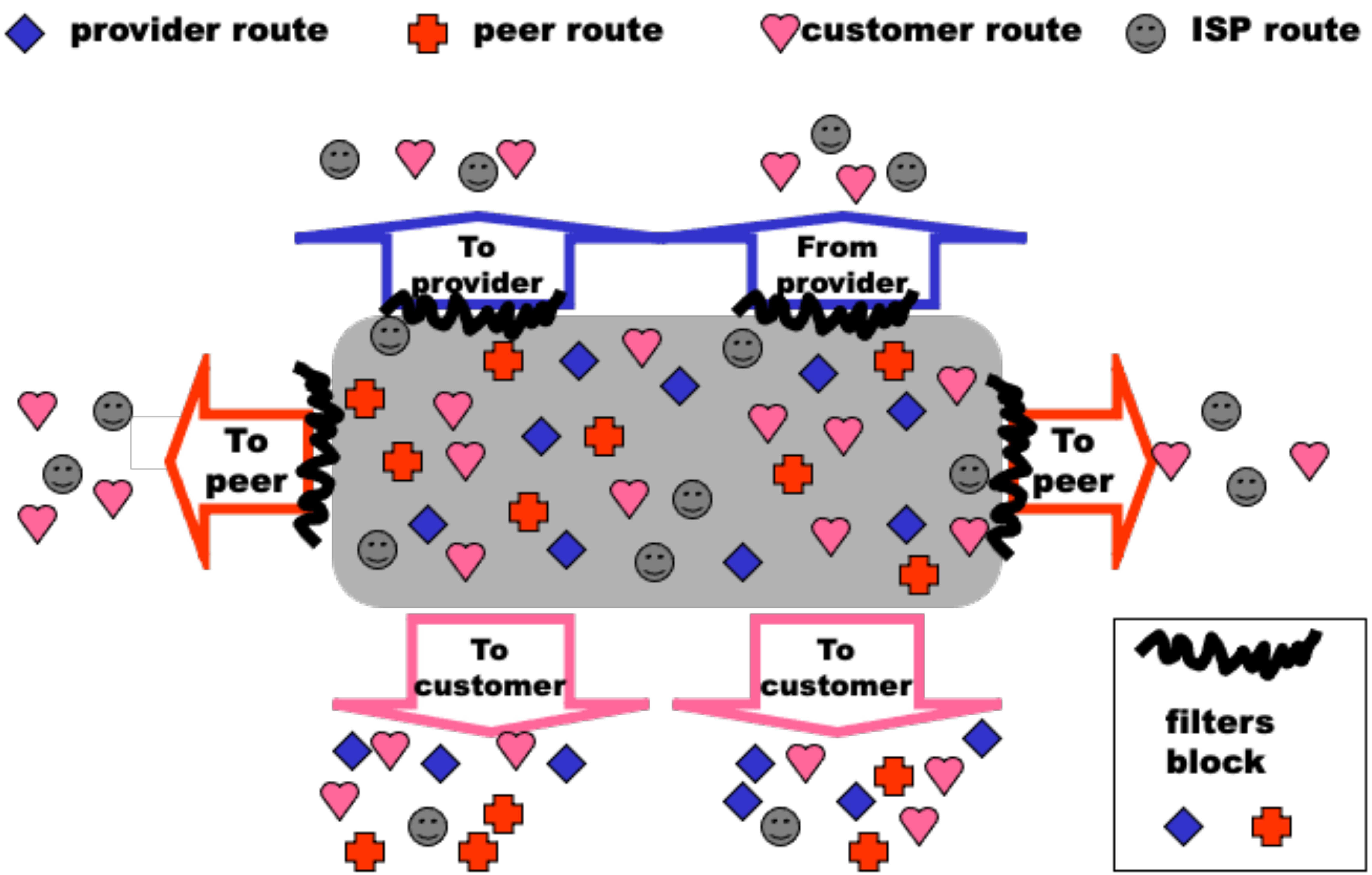
Set appropriate "local pref" to reflect preferences: higher local preference values are preferred



Import Routes



Export Routes



BGP Export Policies

Route learned From ↓	Customer	Provider	Peer
	Advertise to →		
Customer	✓	✓	✓
Provider	✓	✗	✗
Peer	✓	✗	✗

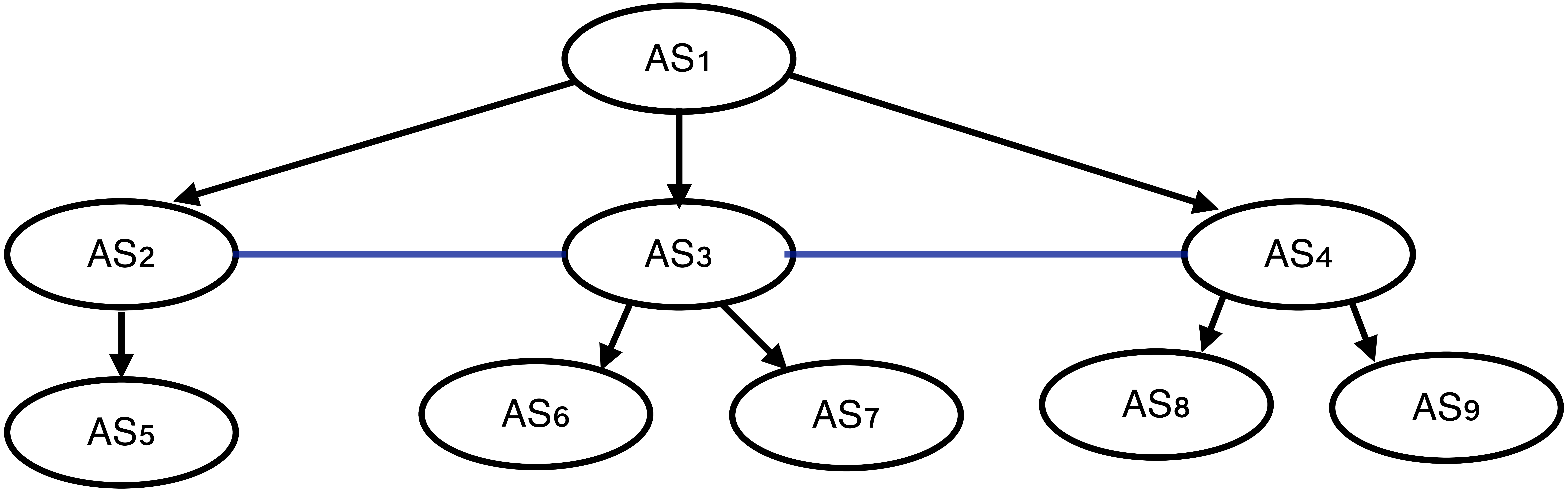
A BGP Example

Consider a network with 9 ASes. They have the following relationships:

- AS1 is the provider for AS2, AS3, and AS4
- AS2 is the provider for AS5
- AS2 and AS3 are peers; AS3 and AS4 are peers
- AS3 is the provider for AS6 and AS7
- AS4 is the provider for AS8 and AS9

A BGP Example

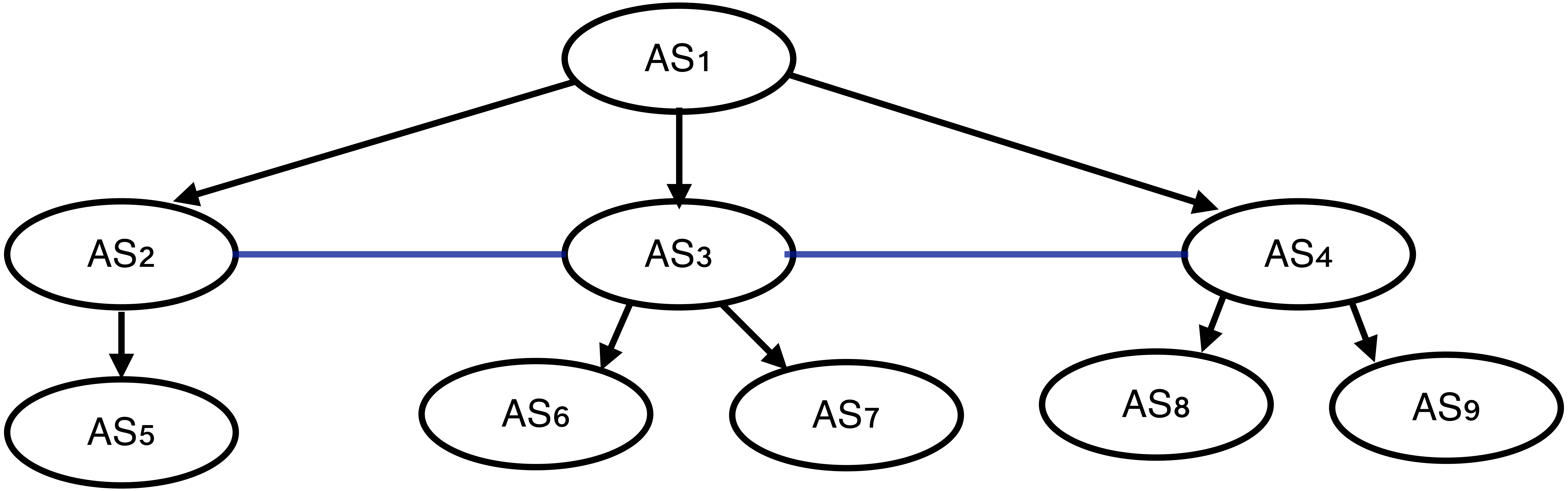
What is the AS path used for AS8-> AS7?



A BGP Example: Case 1

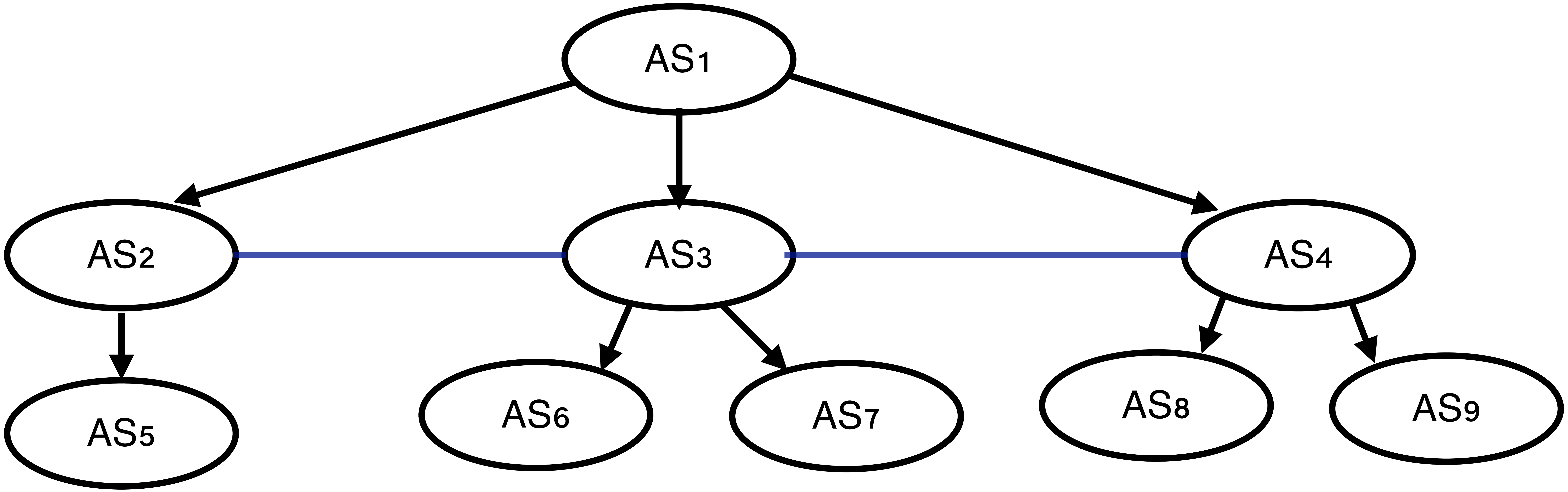
What is the AS path used for AS8-> AS7?

- AS8 -> AS4 -> AS3 -> AS7



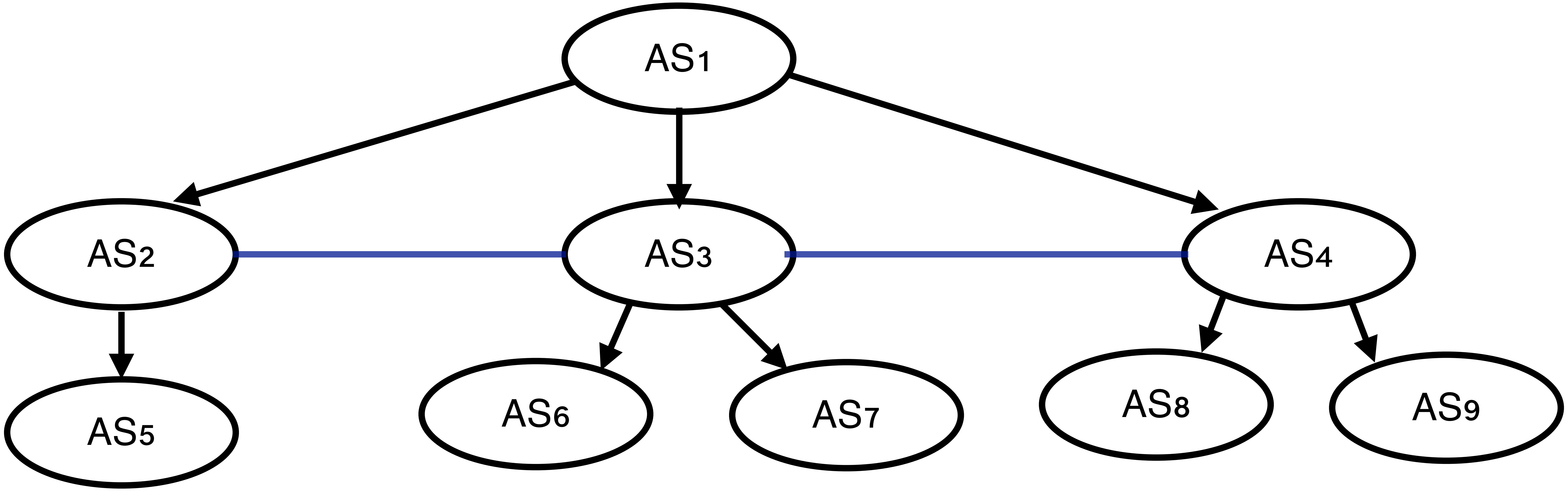
A BGP Example: Case 2

Is (AS5, AS2, AS3, AS4, AS8) a valid path to go from a host in AS5 to a host in AS8?



A BGP Example: Case 2

Is (AS5, AS2, AS3, AS4, AS8) a valid path to go from a host in AS5 to a host in AS8? => No!



BGP in Reality

AS 7007 incident

From Wikiped

The AS 7007
sometimes ir
Probably bec
and had the
problems tha
of these facto
Analysis

How Pakistan knocked YouTube offline (and how to make sure it never happens again)

Suspicious event hijacks Amazon traffic for 2 hours, steals cryptocurrency

Almost 1,300 addresses for Amazon Route 53 rerouted for two hours.

DAN GOODIN - 4/24/2018, 2:00 PM

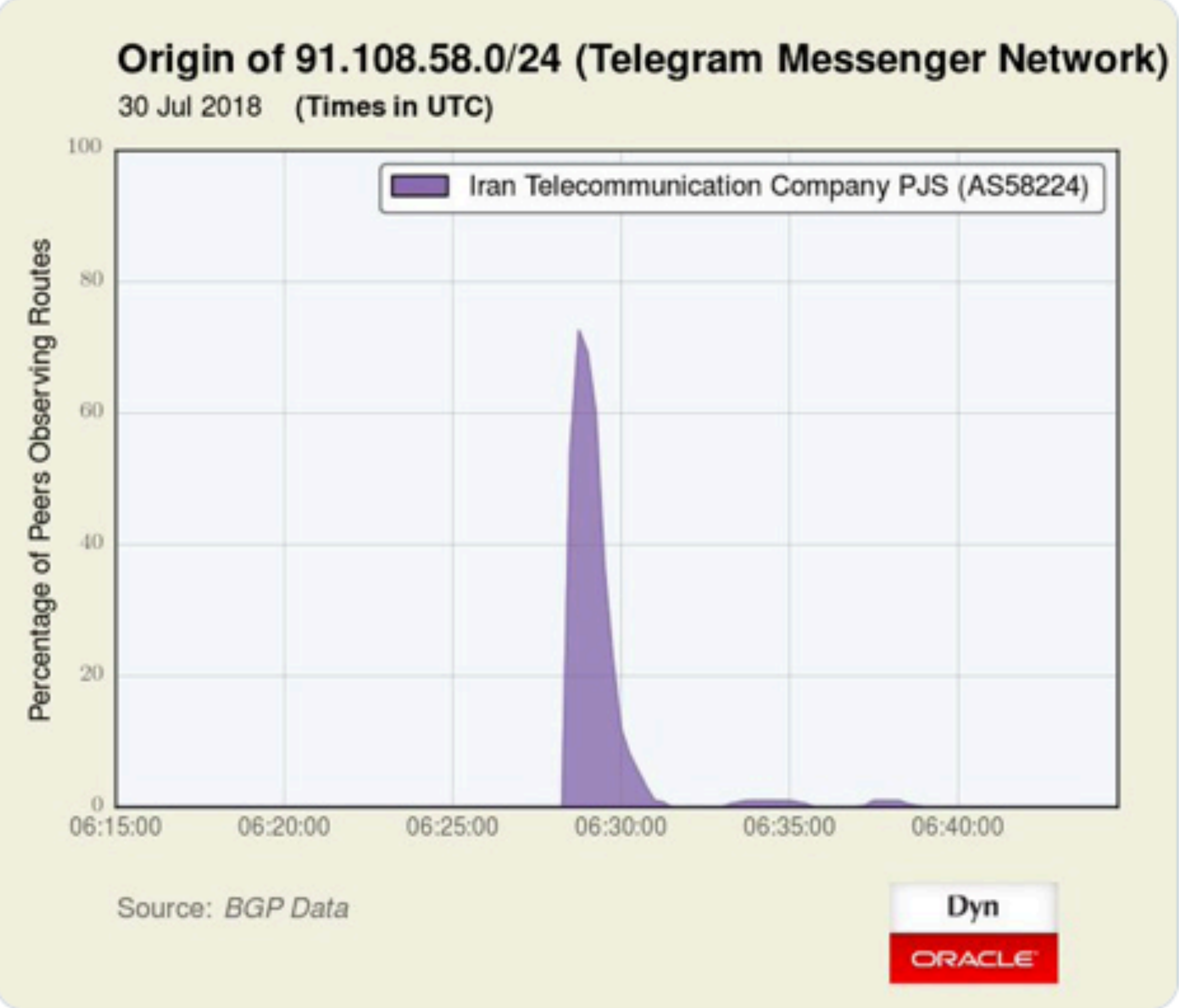
amazon.com

125

Amazon lost control of a small number of its cloud services IP addresses for two hours on Tuesday morning when hackers exploited a known Internet-protocol weakness that let them to redirect traffic to rogue destinations. By subverting Amazon's domain-resolution service, the attackers masqueraded as cryptocurrency website MyEtherWallet.com and stole about $150,00 in digital coins from unwitting end users. They may have targeted other Amazon customers as well.

InternetIntelligence
@InternetIntel

At 06:28 UTC earlier today (30-Jul), an Iranian state telecom network briefly leaked over 100 prefixes. Most were Iranian networks, but the leak also included 10 prefixes of popular messaging app @telegram (8 were more-specifics).



BGP in Reality

Summary

- Today
  - Inter-domain Routing

- Next lecture
  - NAT, IPv6, and Multicast