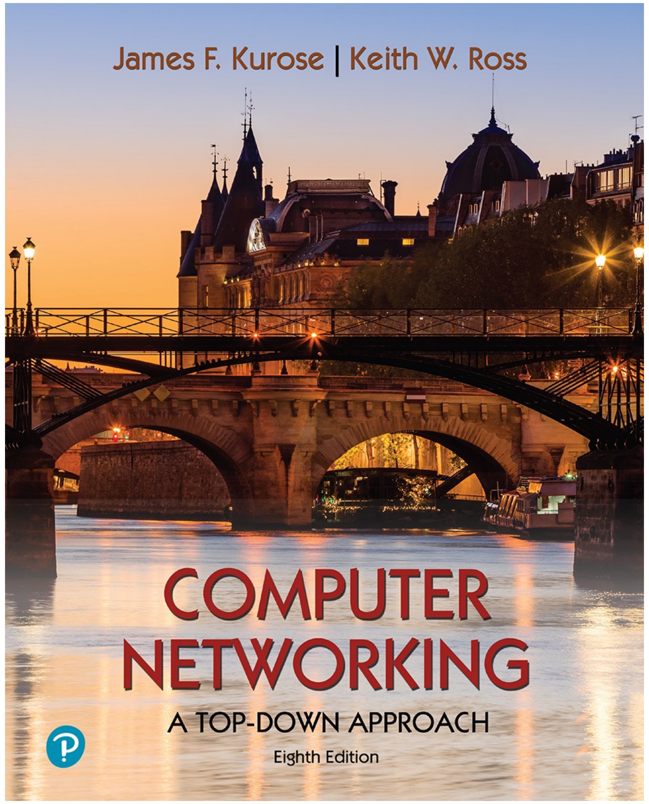
Chapter 5
Network Layer: Control Plane

Yaxiong Xie

Department of Computer Science and Engineering
University at Buffalo, SUNY

Adapted from the slides of the book's authors



Computer Networking: A Top-Down Approach
8^{th} edition
Jim Kurose, Keith Ross
Pearson, 2020

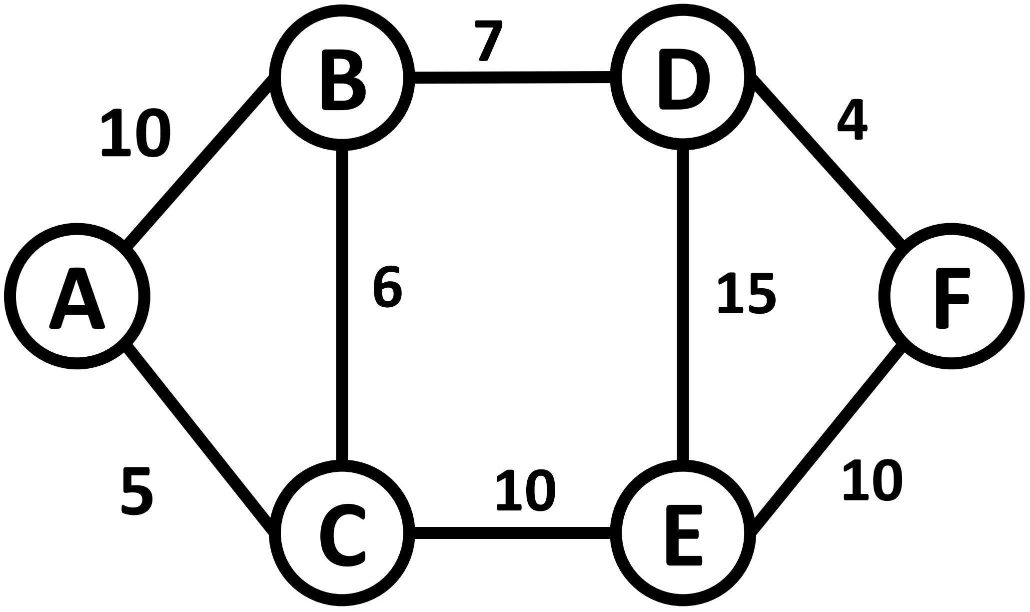
Network layer: “control plane” roadmap

- introduction
- routing protocols
- Scalable Routing Algorithm
  - intra-AS routing: OSPF
  - inter-AS routing: BGP
- SDN control plane
- Internet Control Message Protocol

- network management, configuration
  - SNMP
  - NETCONF/YANG

Routing protocols we have discussed

Problem: Given a network topology, how do we compute good paths from one node to another?



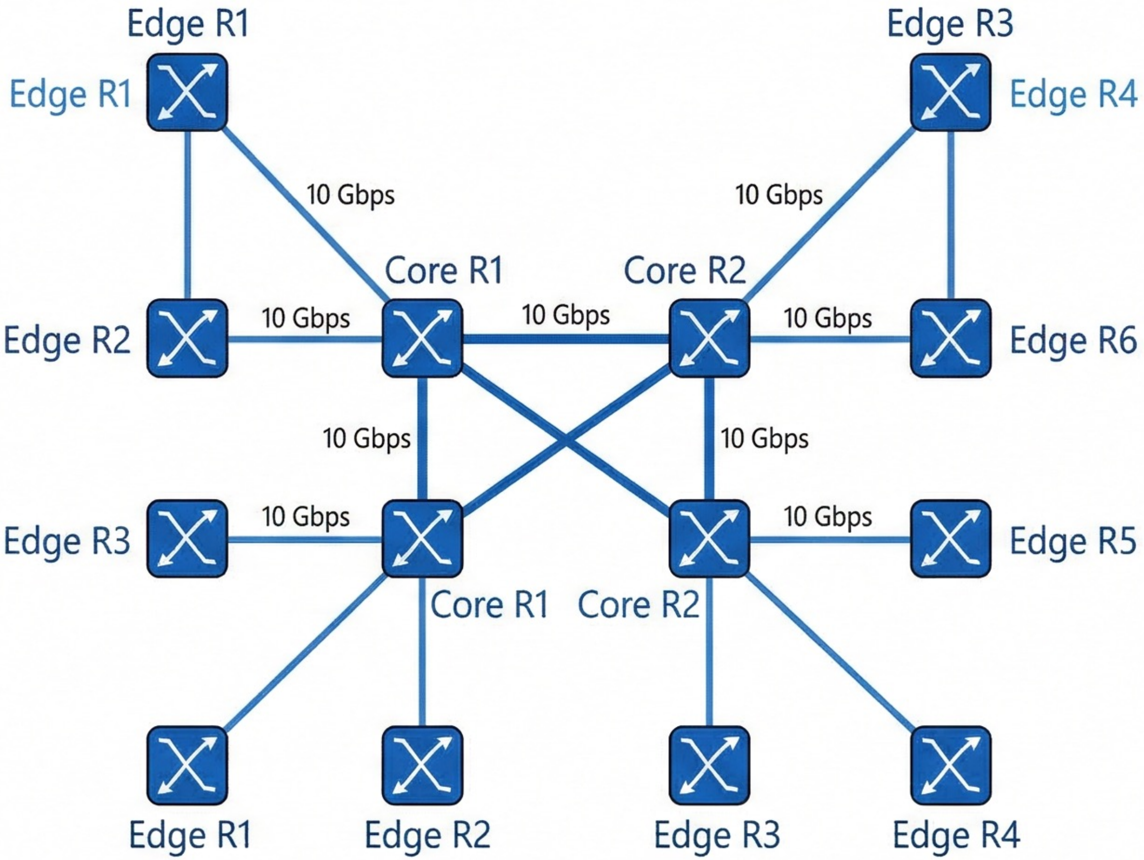
Network topology

Assumptions

- Aligned objectives
  - All routers aim to deliver packets to the destination as efficiently as possible.
- Willingness to cooperate
  - Routers are willing to share routing information, and sometimes even full topology information.
- A common metric
  - shortest path, least cost, or lowest delay.
- A single administrative domain
  - The network is managed by one organization, or at least follows one common set of rules
- Basic mutual trust among node
  - All routers aim to deliver packets to the destination as efficiently as possible
- Manageable scale

Works inside Small Managed Network

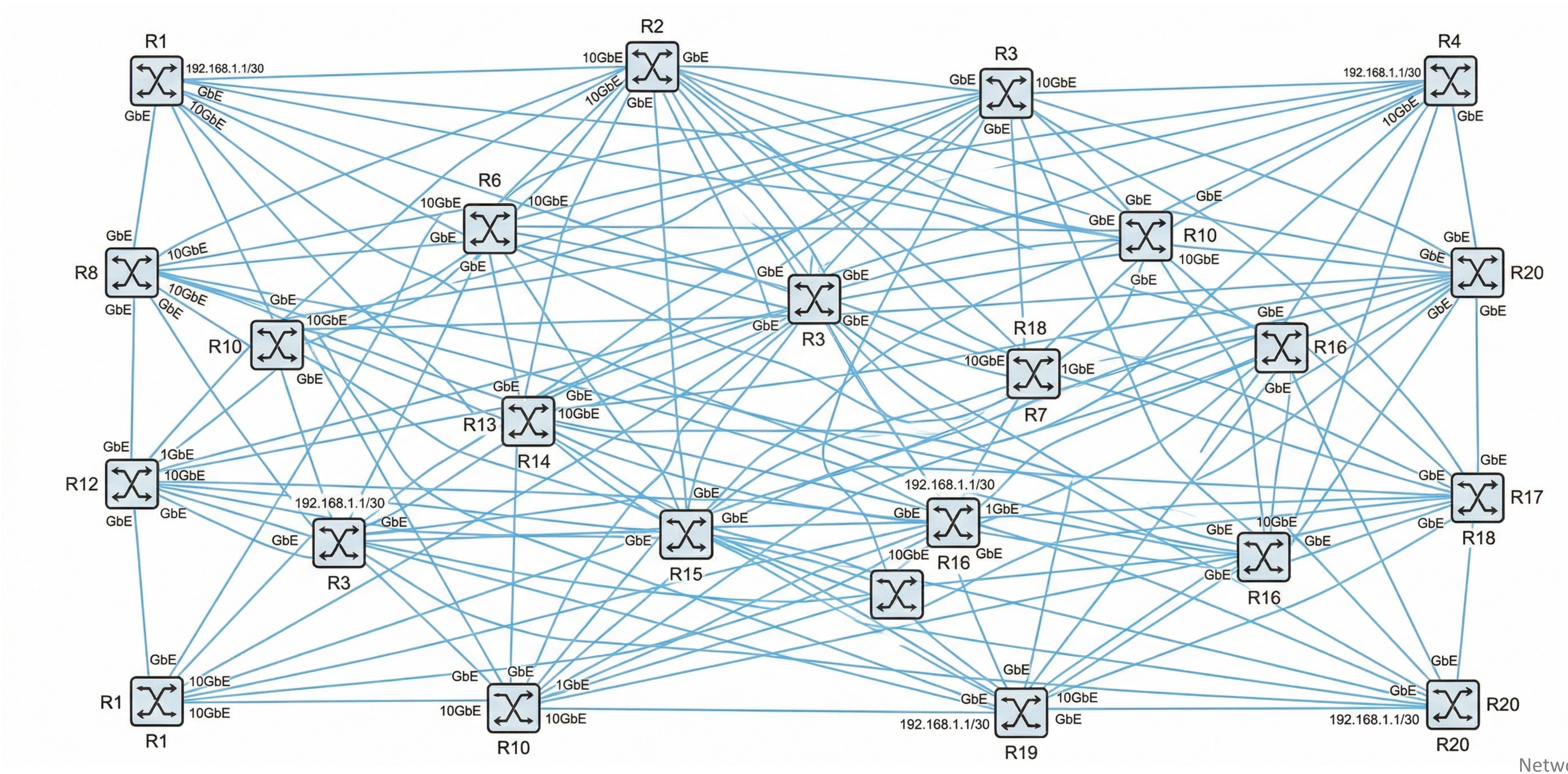
UB Campus Network



SYSTEM ASSUMPTIONS

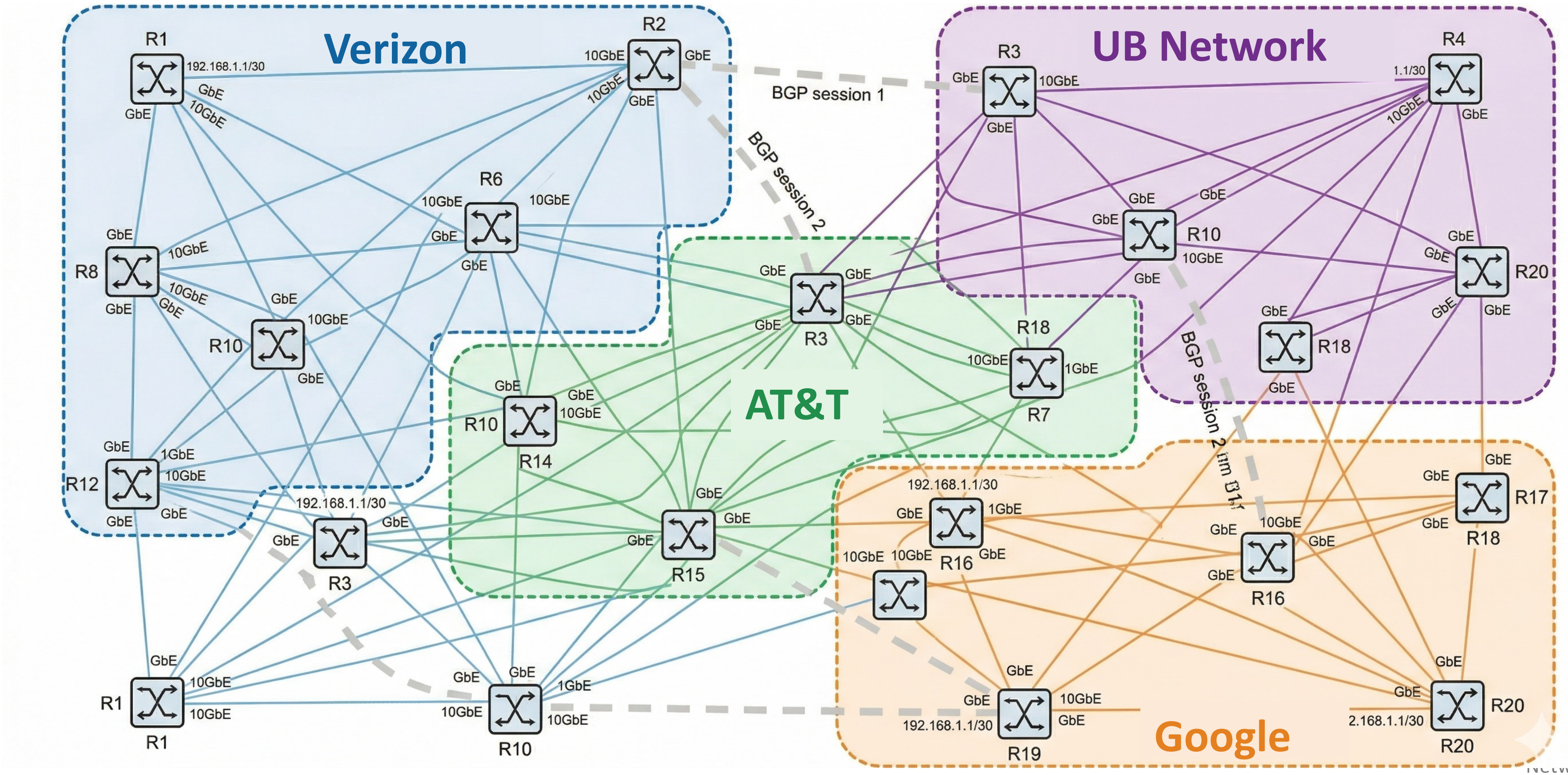
Why the Internet Breaks Those Assumptions

Scale



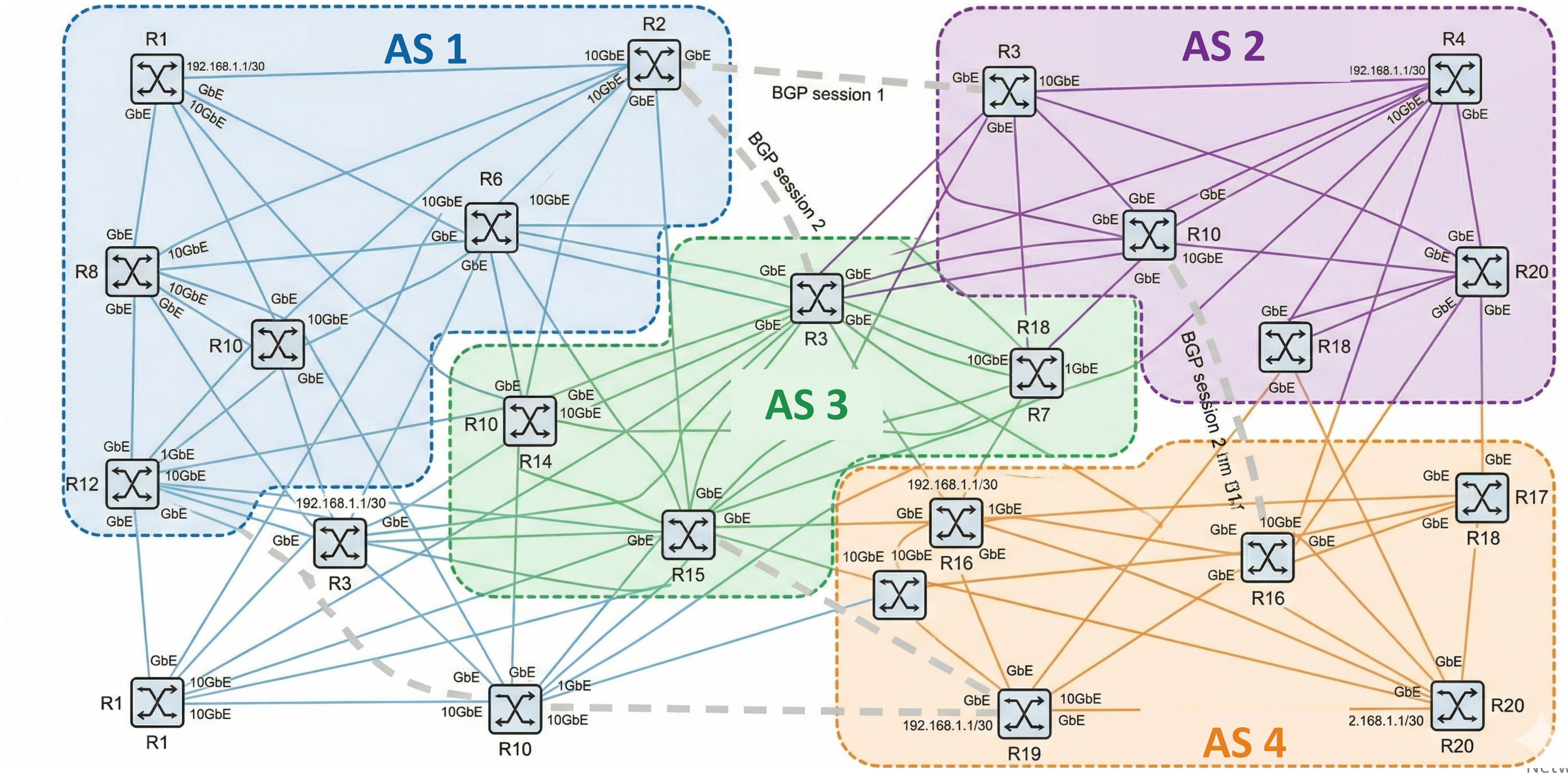
Why the Internet Breaks Those Assumptions

Scale **Independently operated** **Business Policy** **Privacy** **Trust**

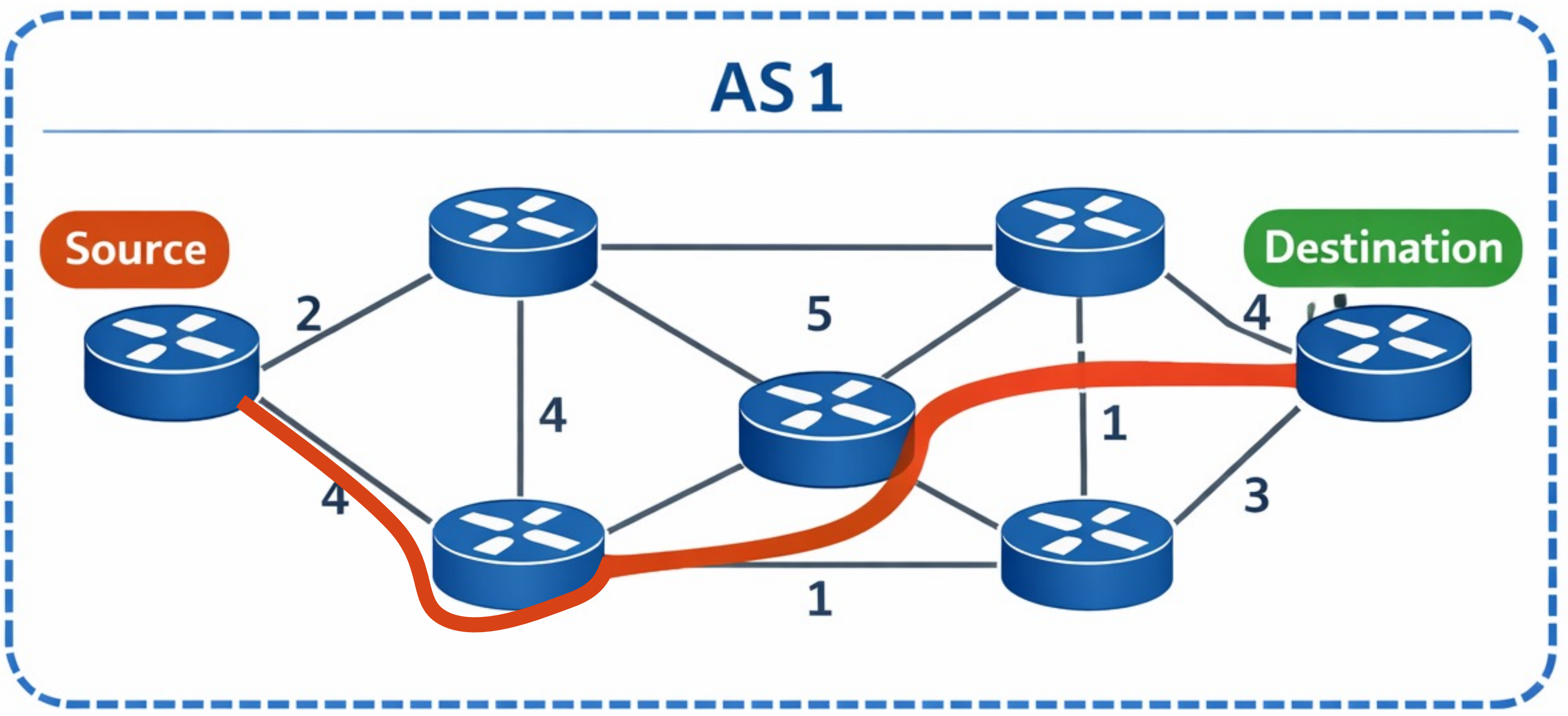


Internet Is a Network of Networks

Autonomous System: AS is a network under a single administrative control

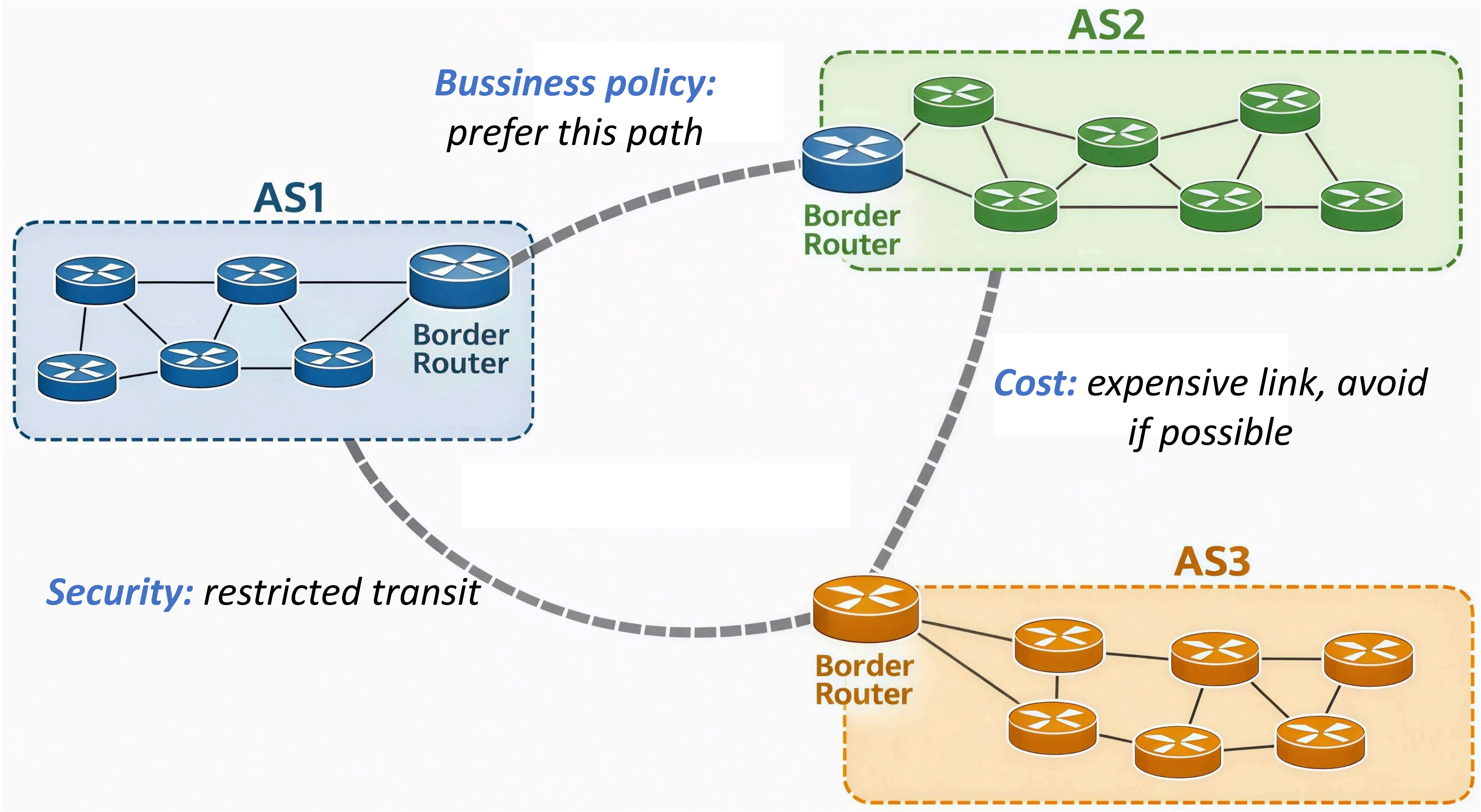


Intra-AS Routing Algorithm



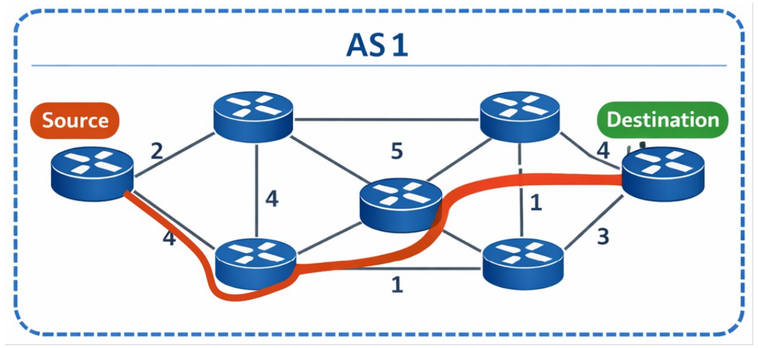
Goal: find the shortest path to destination

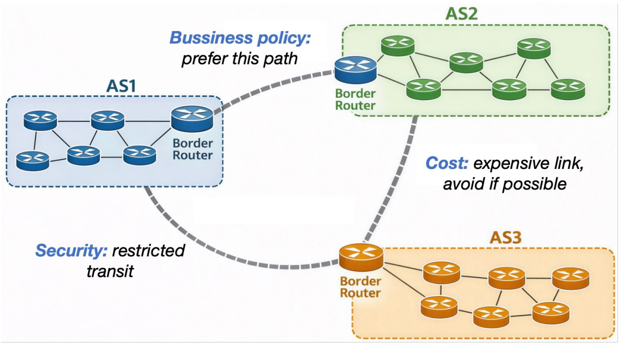
Inter-AS Routing Algorithm



Goal: balance policy, cost, and aggreements -- not just shortest path

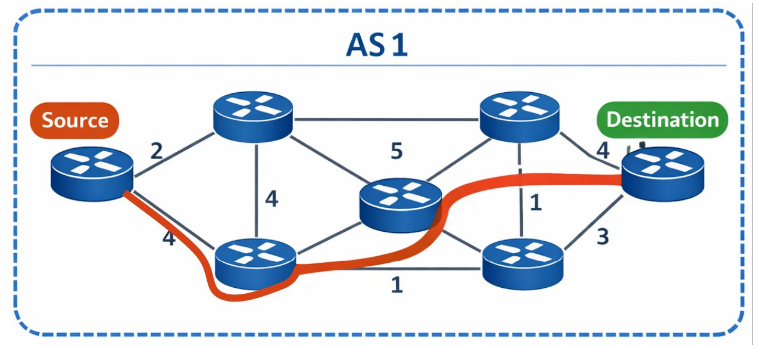
Two-level of Routing

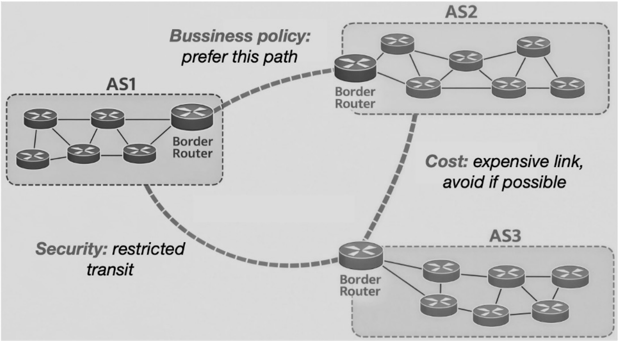




	Intra-AS Routing	**Inter-AS Routing**
Scope	within one AS	across ASes
Goal	efficiency / shortest path	policy + reachability + scalability
Example Protocol	OSPF, EIGRP, RIP	BGP

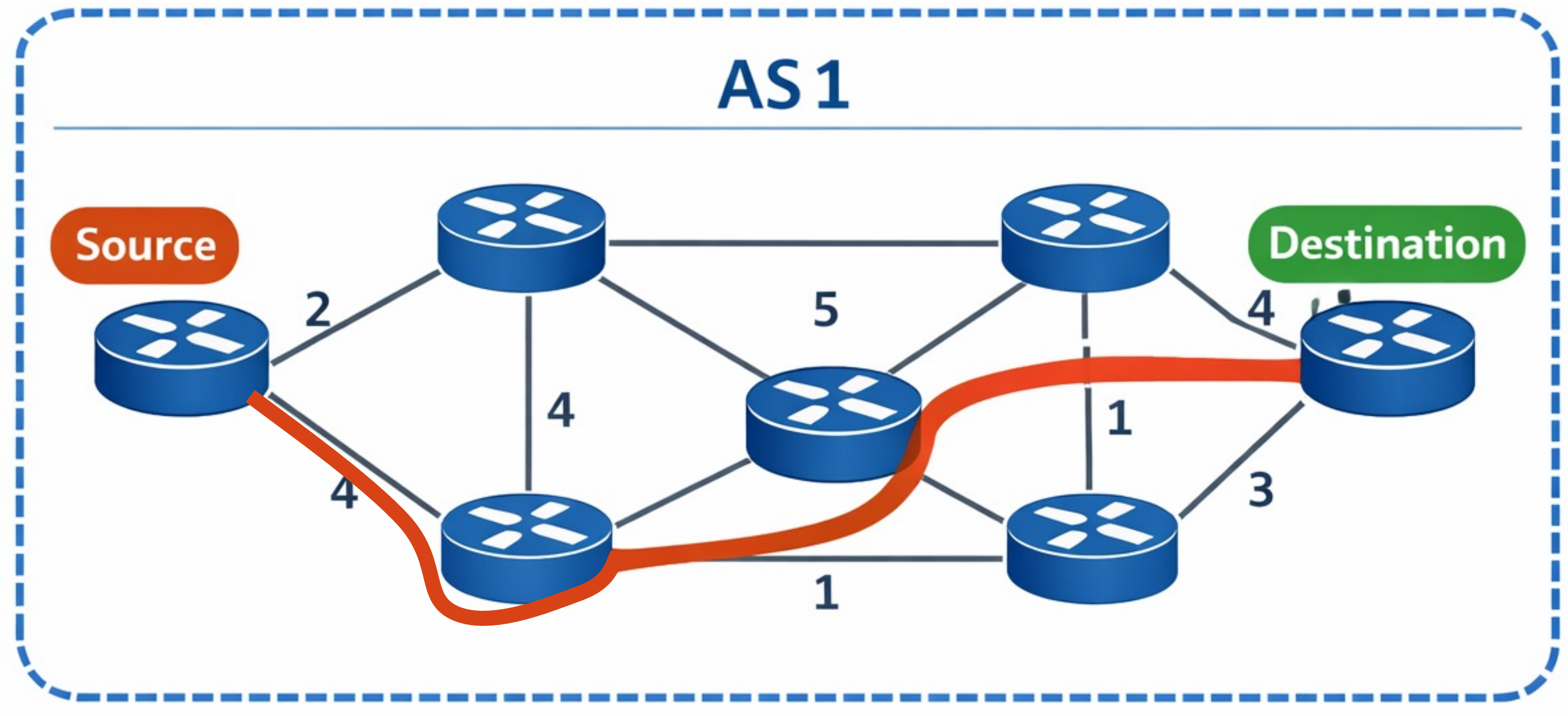
Two-level of Routing





	Intra-AS Routing	Inter-AS Routing
Scope	within one AS	across ASes
Goal	efficiency / shortest path	policy + reachability + scalability
Example Protocol	OSPF, EIGRP, RIP	BGP

Intra-AS Routing Algorithm



Goal: find the shortest path to destination

Link state

Distance Vector

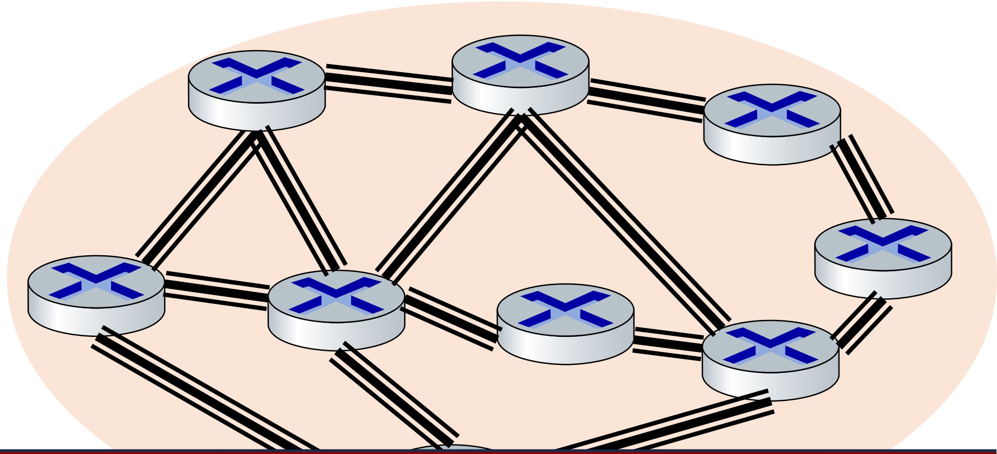
Intra-AS routing: routing within an AS

most common intra-AS routing protocols:

- RIP: Routing Information Protocol [RFC 1723]
  - classic DV: DVs exchanged every 30 secs
  - no longer widely used
- EIGRP: Enhanced Interior Gateway Routing Protocol
  - DV based
  - formerly Cisco-proprietary for decades (became open in 2013 [RFC 7868])
- OSPF: Open Shortest Path First [RFC 2328]
  - link-state routing
  - IS-IS protocol (ISO standard, not RFC standard) essentially same as OSPF

OSPF (Open Shortest Path First) routing

- "open": publicly available
- classic link-state
  - each router floods OSPF link-state advertisements (directly over IP rather than using TCP/UDP) to all other routers in entire AS
  - multiple link costs metrics possible: bandwidth, delay



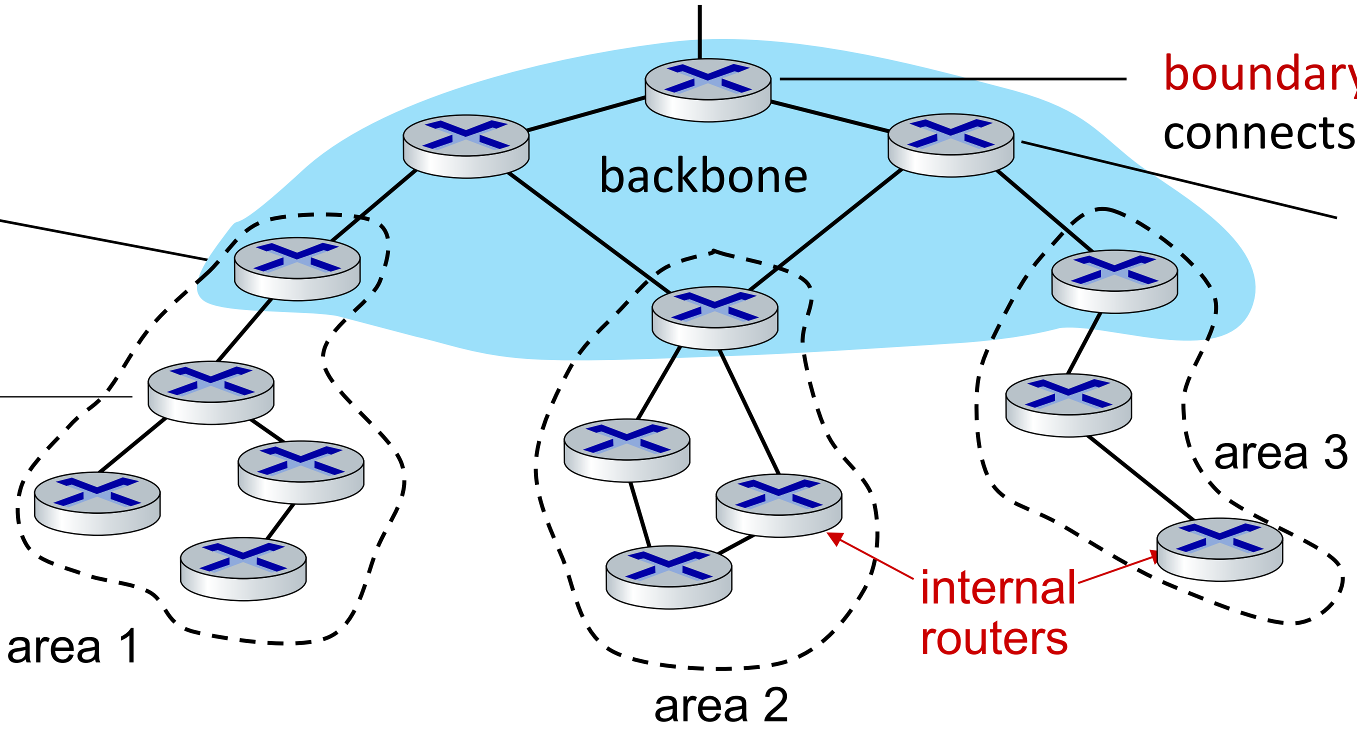
What if the AS itself becomes extremely large?

Hierarchical OSPF

- two-level hierarchy: local area, backbone.
  - link-state advertisements flooded only in area, or backbone
  - each node has detailed area topology; only knows direction to reach other destinations

area border routers:
"summarize" distances to destinations in own area, advertise in backbone

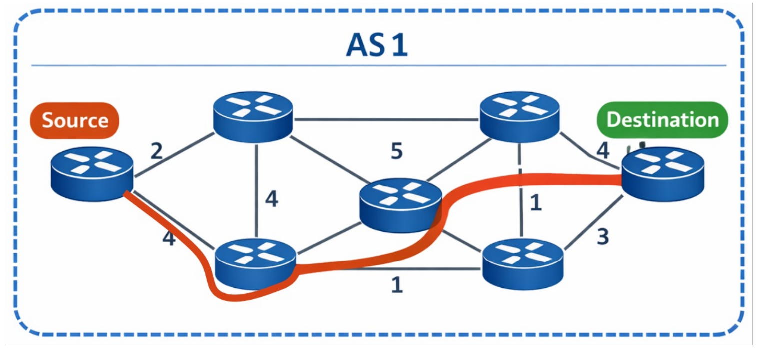
local routers:
- flood LS in area only
- compute routing within area
- forward packets to outside via area border router

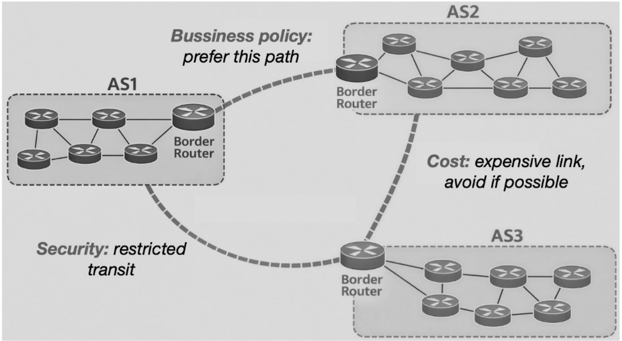


boundary router:
connects to other ASes

backbone router:
runs OSPF limited to backbone

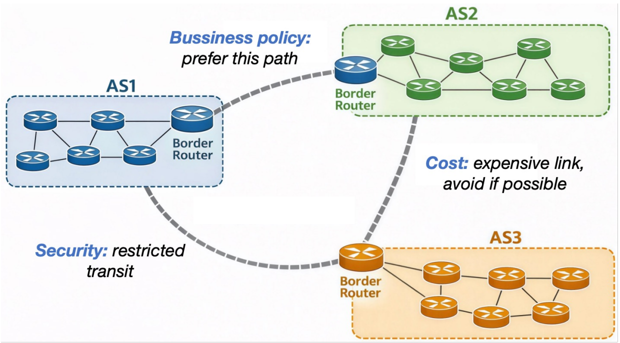
Two-level of Routing





	Intra-AS Routing	Inter-AS Routing
Scope	within one AS	across ASes
Goal	efficiency / shortest path	policy + reachability + scalability
Example Protocol	OSPF, EIGRP, RIP	BGP

Two-level of Routing

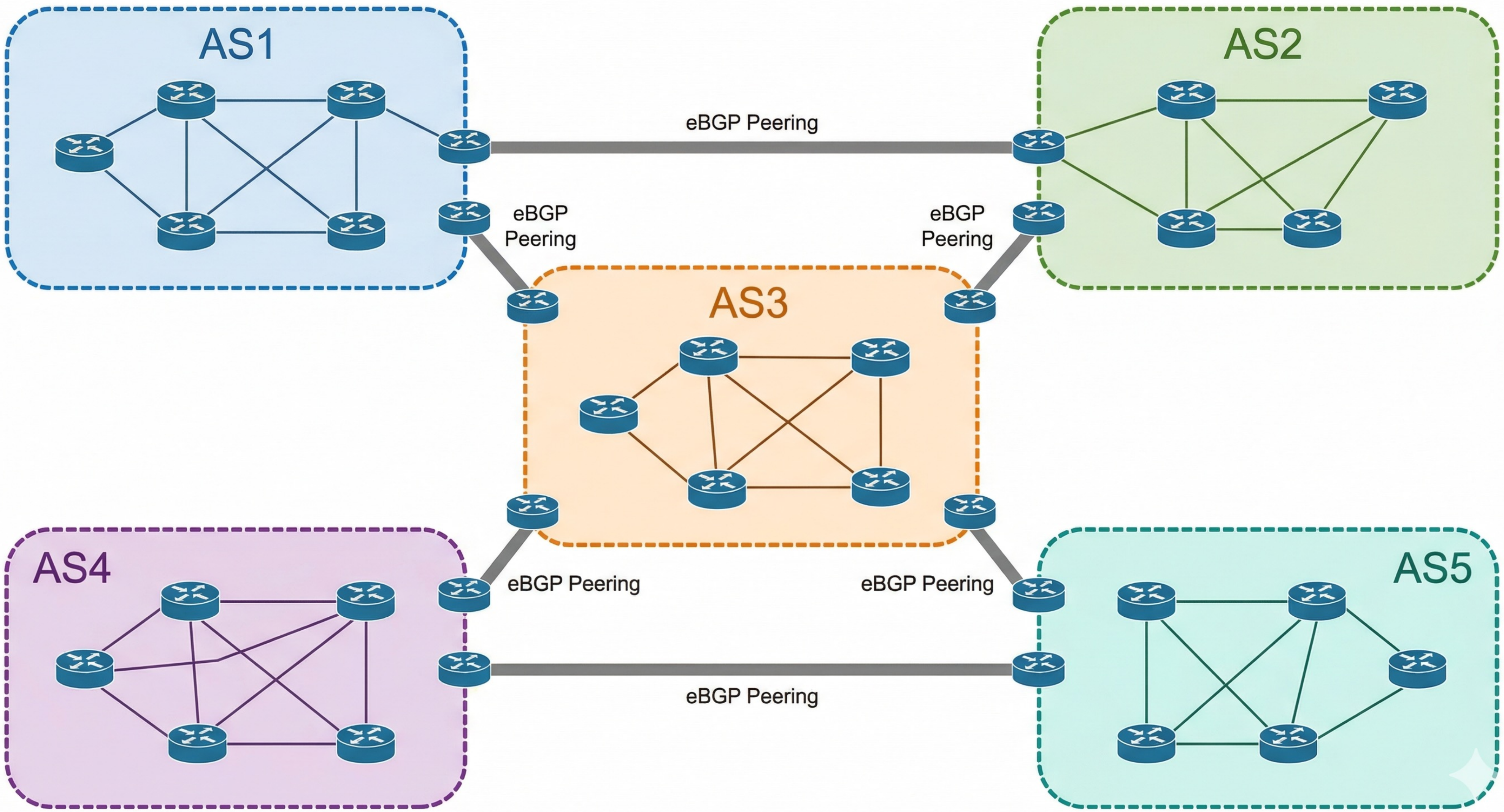


	Intra-AS Routing	Inter-AS Routing
Scope	within one AS	across ASes
Goal	efficiency / shortest path	policy + reachability + scalability
Example Protocol	OSPF, EIGRP, RIP	BGP

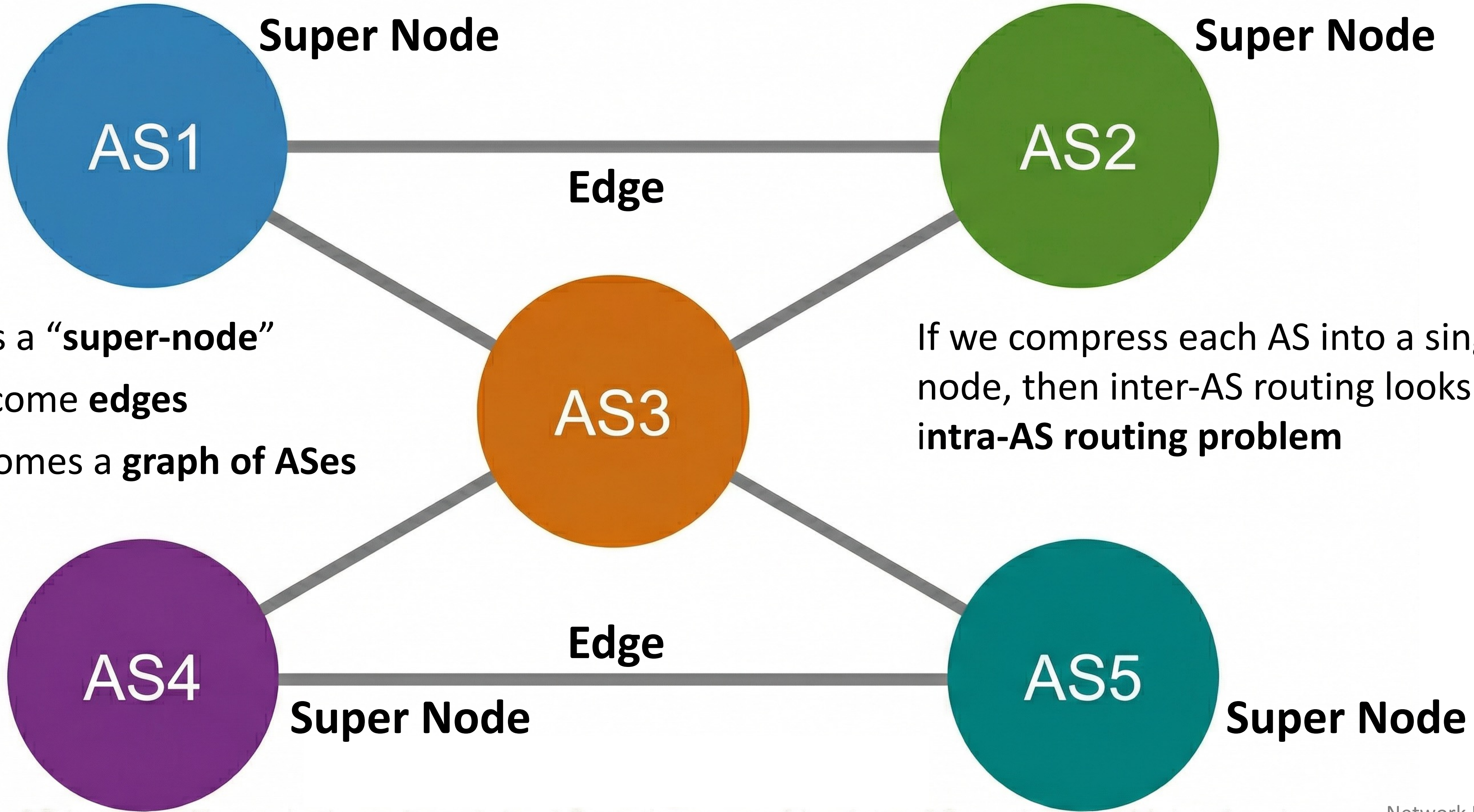
Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto inter-domain routing protocol
  - "glue that holds the Internet together"
- allows subnet to advertise its existence, and the destinations it can reach, to rest of Internet: *"I am here, here is who I can reach, and how"*

BGP at the High-Leve



A Tempting Abstraction: AS as Node

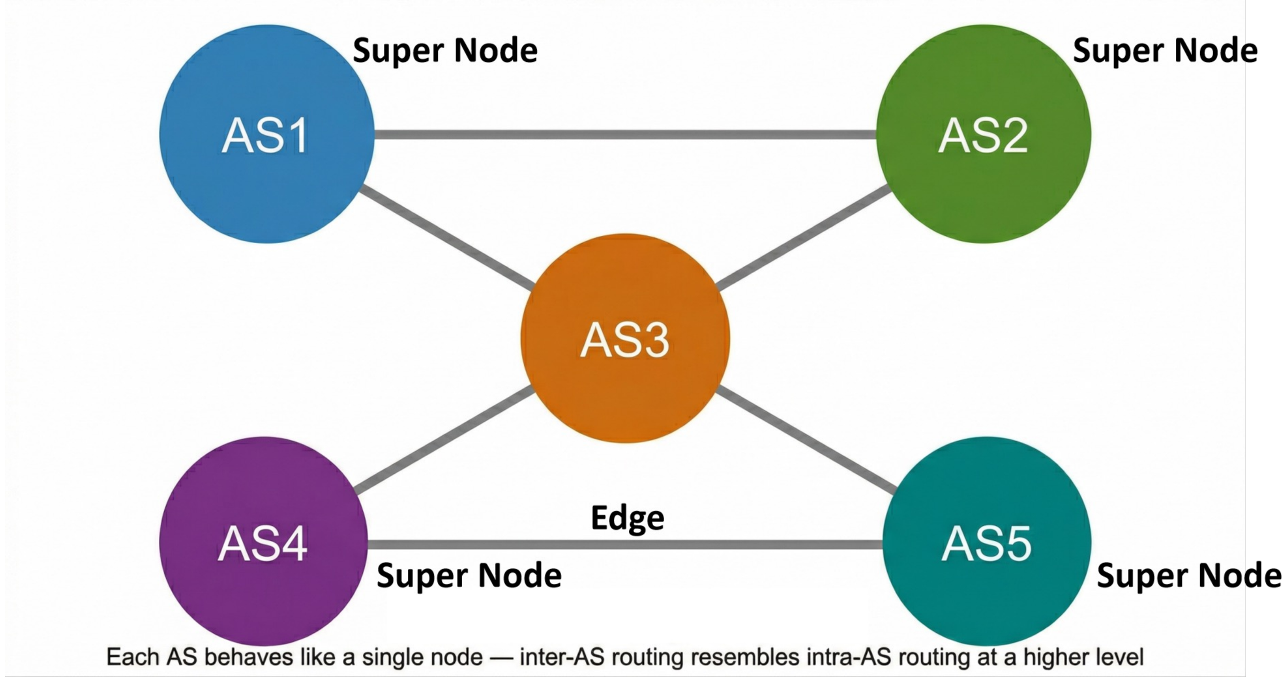


- Each AS becomes a “**super-node**”
- Inter-AS links become **edges**
- The Internet becomes a **graph of ASes**

If we compress each AS into a single node, then inter-AS routing looks like **intra-AS routing problem**

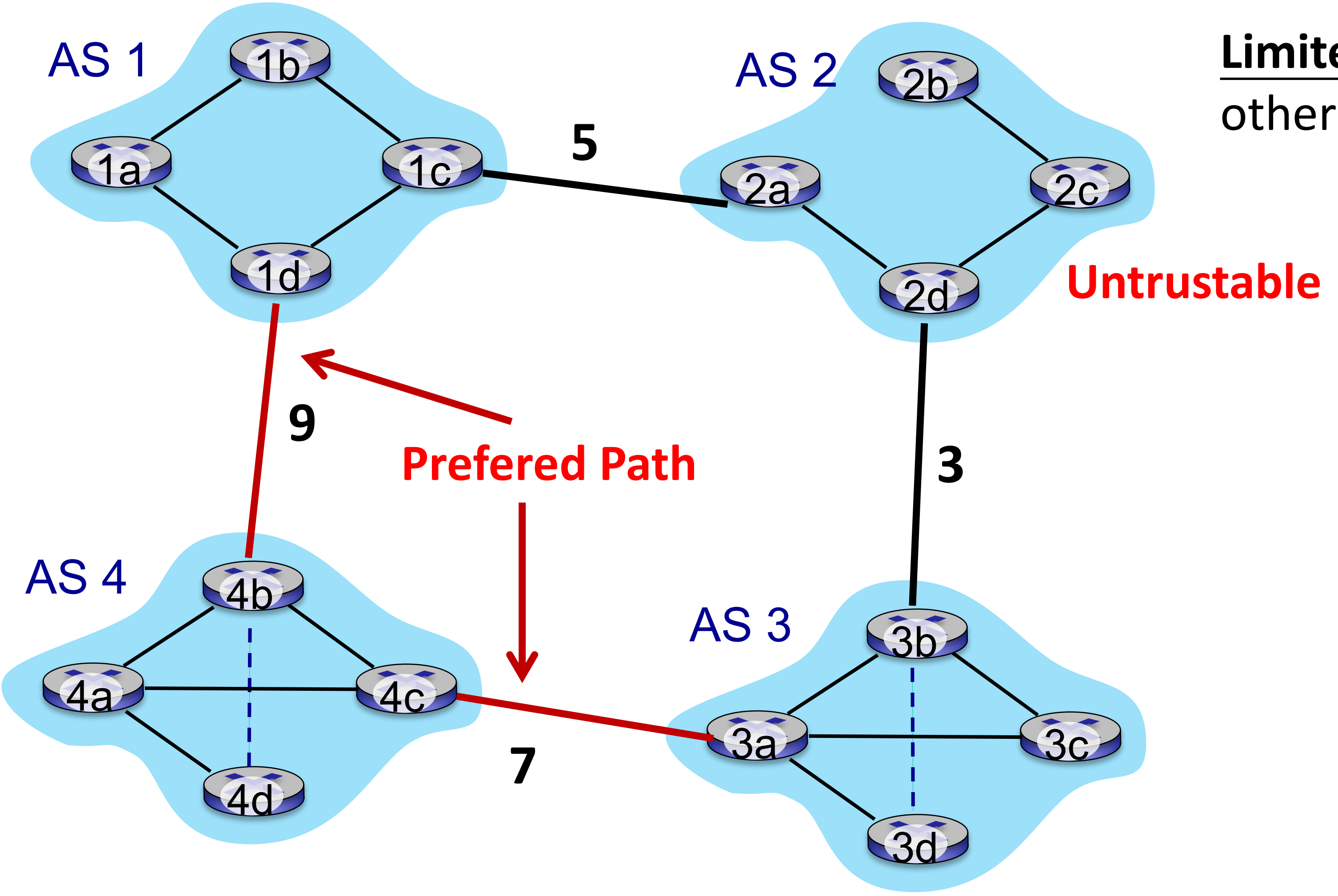
Each AS behaves like a single node — inter-AS routing resembles intra-AS routing at a higher level

Inter-AS Routing Looks Like a “Vector” Protocol



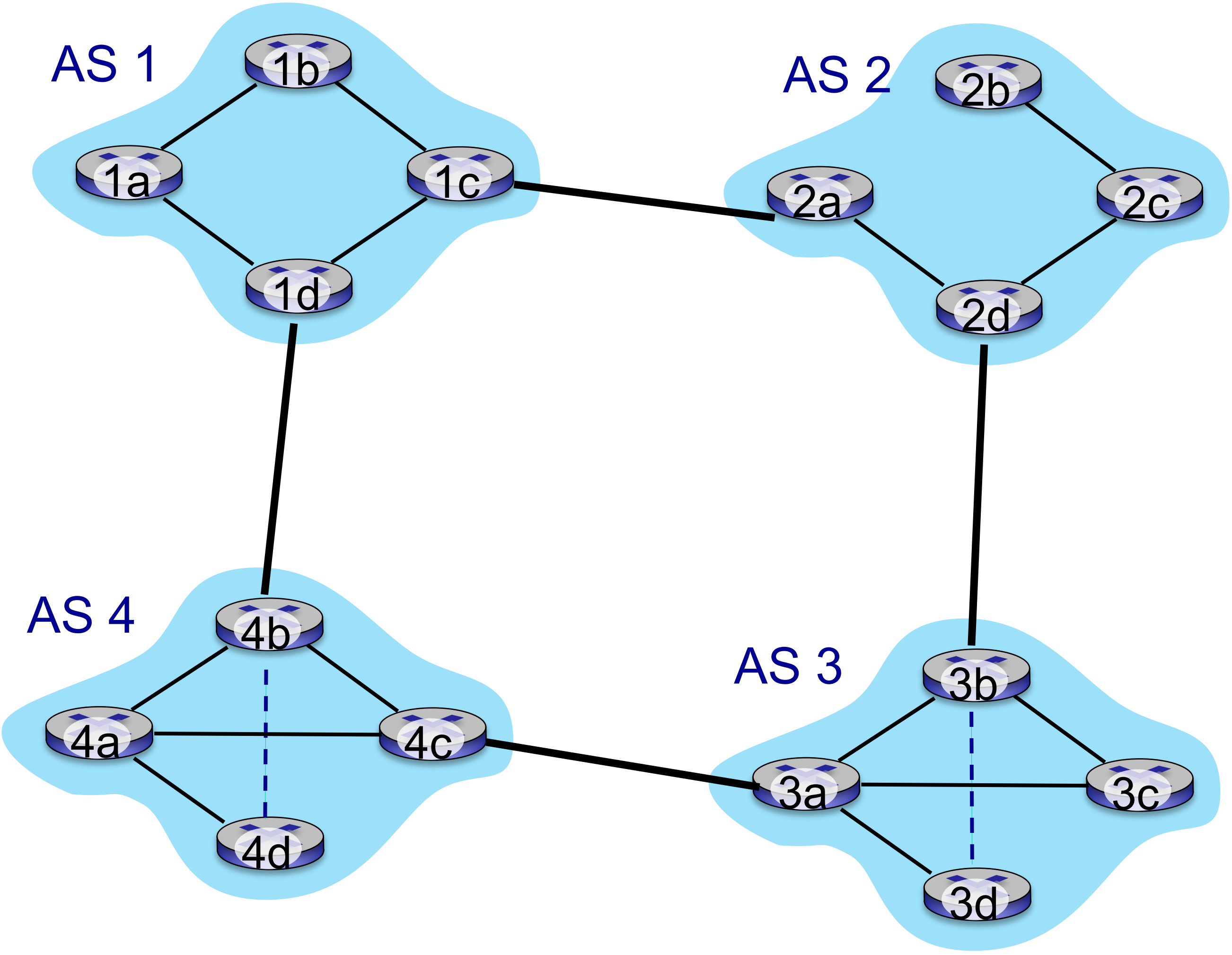
- No AS has a complete view of the whole Internet
- Routing information is exchanged neighbor by neighbor
- Reachability is propagated incrementally
- In that sense, BGP still feels “vector-like”

Inter-AS Routing Is Not Just About Efficiency



Limited trust: they may not trust each other fully

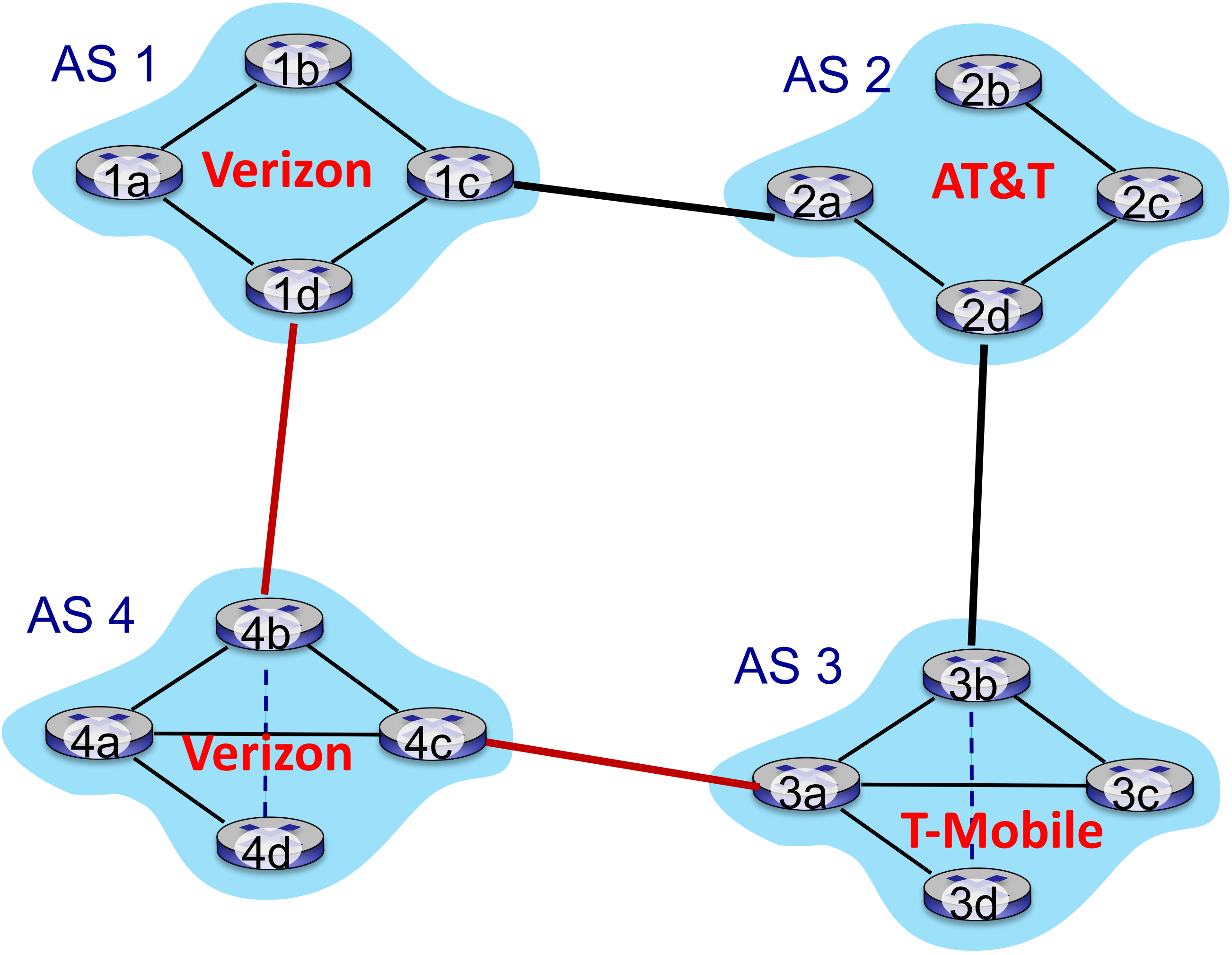
Inter-AS Routing Is Not Just About Efficiency



Limited trust: they may not trust each other fully

Hide internal topology: they may not want to reveal internal topology

Inter-AS Routing Is Not Just About Efficiency

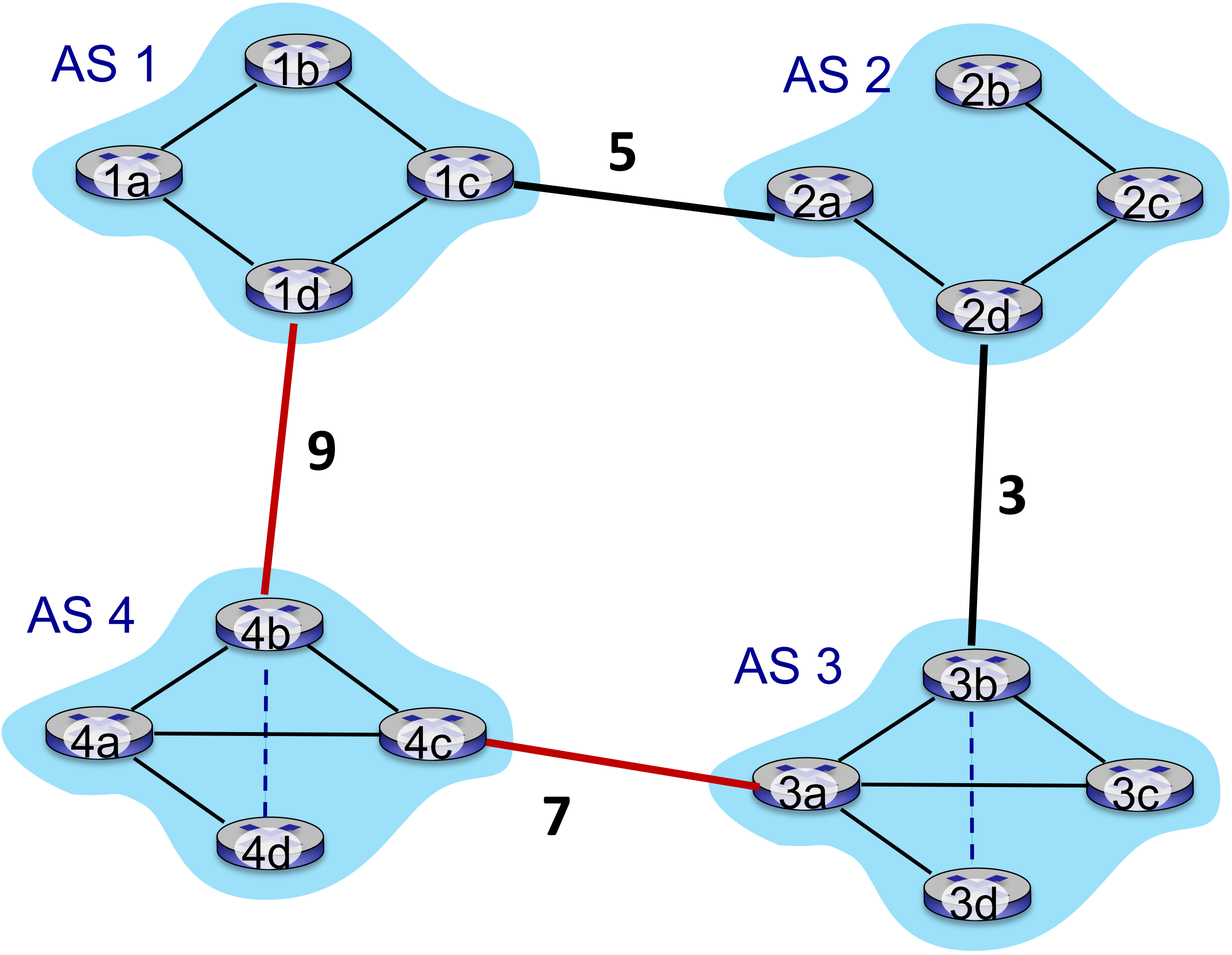


Limited trust: they may not trust each other fully

Hide internal topology: they may not want to reveal internal topology

Business relationships: competitors

Inter-AS Routing Is Not Just About Efficiency



Limited trust: they may not trust each other fully

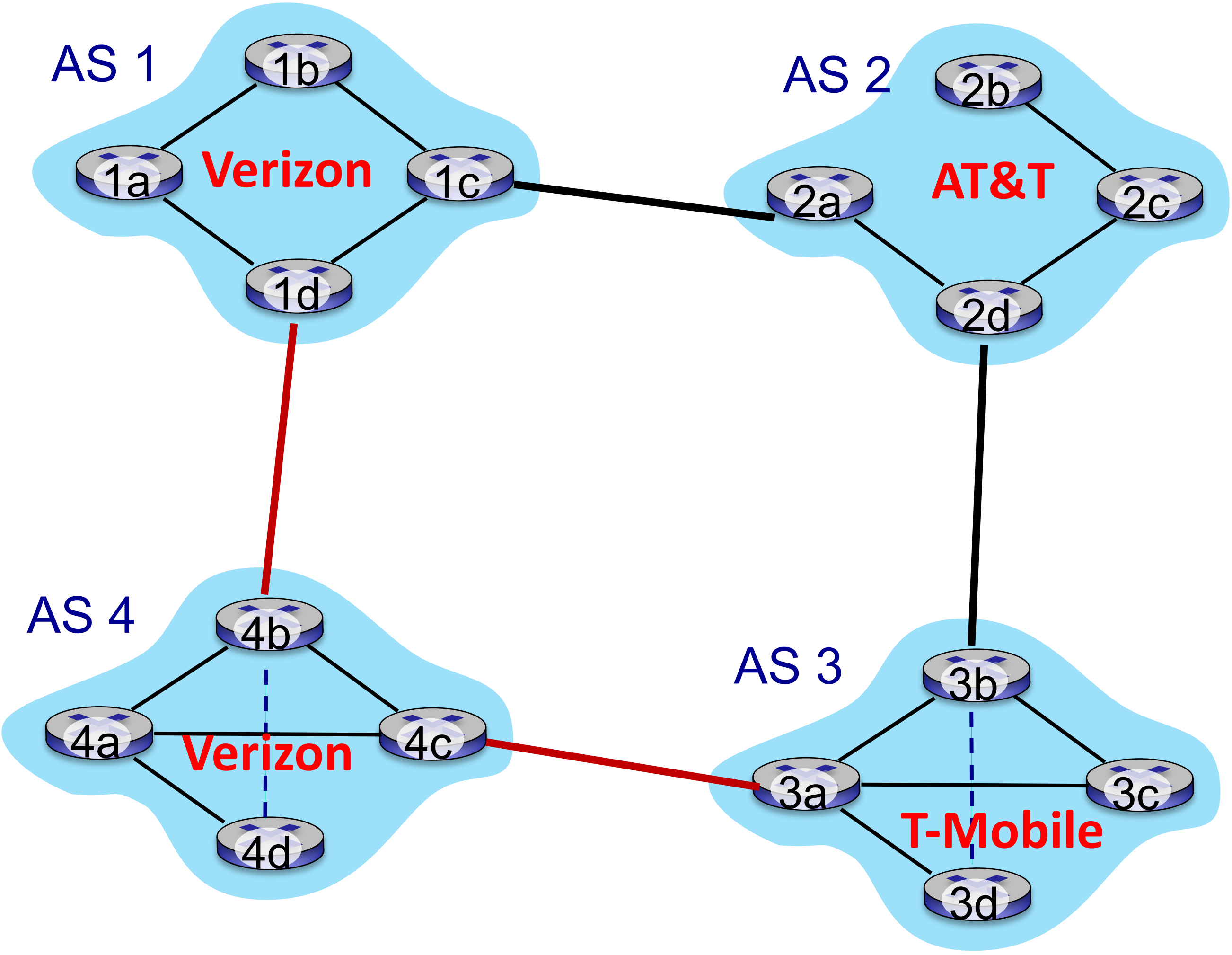
Hide internal topology: they may not want to reveal internal topology

Business relationships: competitors

Policy over shortest path: they may prefer one route for policy reasons, not for shortest distance

Once we move across AS boundaries, routing is no longer just a technical optimization problem. It becomes a problem involving economics, policy, and autonomy.

Distance Alone Does Not Capture What Matters



Distance: a number tells you "***how far***"

Identidy of AS: the identity of the ASes on the path matters

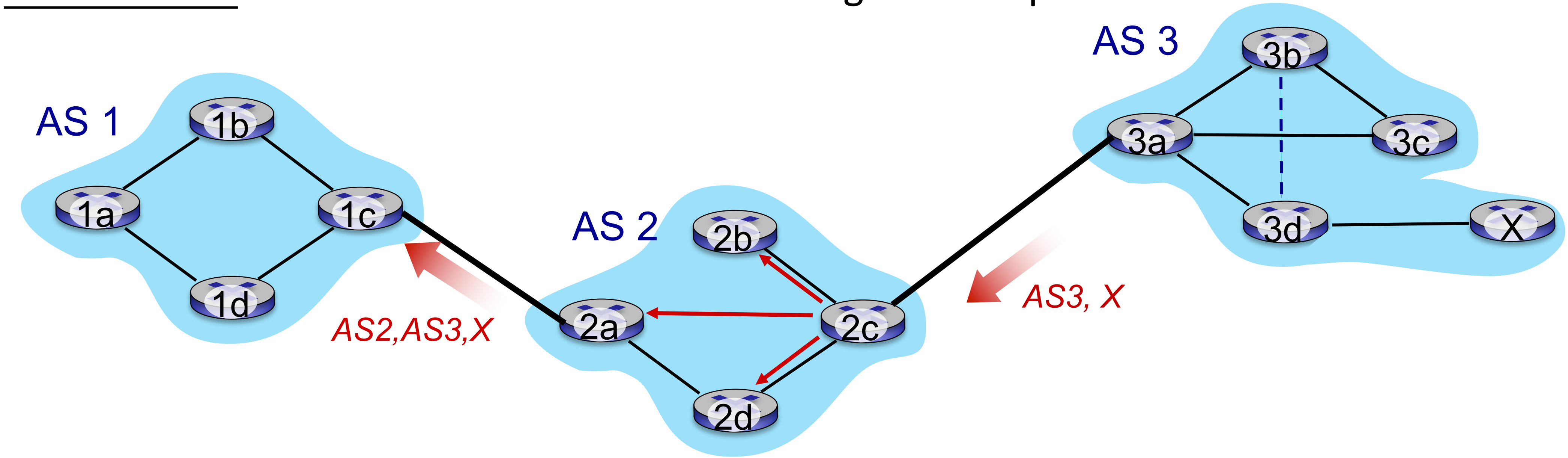
- Policy
- Security
- Business

Path Vector: Advertise the Path, Not Just the Distance

- BGP session: two BGP routers ("peers") exchange BGP messages over semi-permanent TCP connection:
  - advertising *paths* to different destination network prefixes (BGP is a "path vector" protocol)

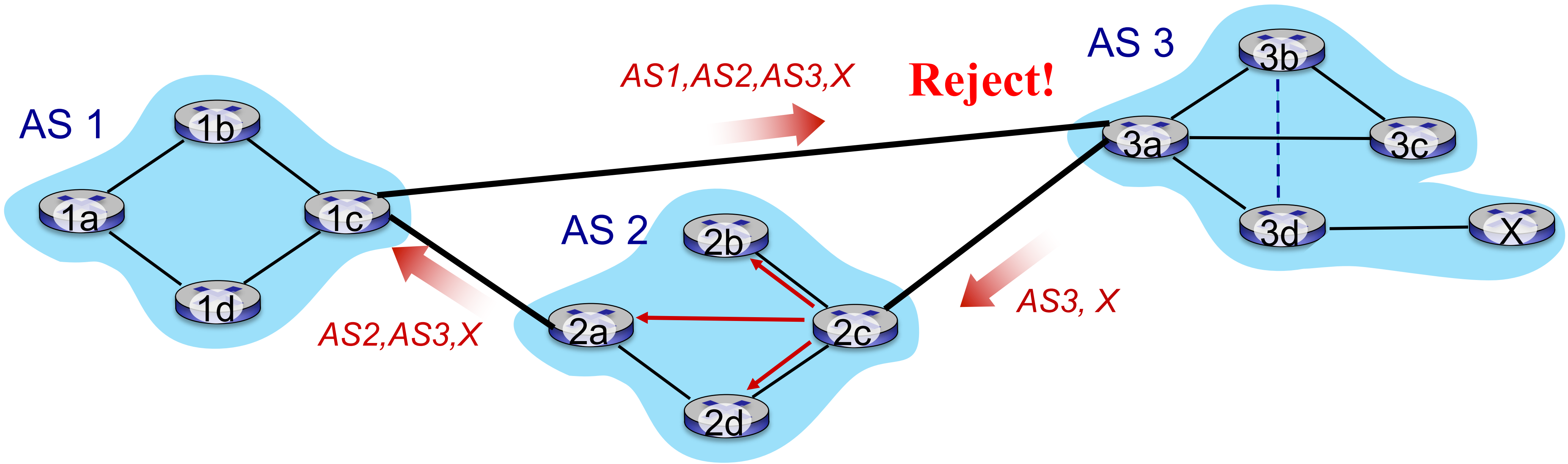
Distance Vector: "I can reach this destination with cost X"

Path Vector: "I can reach this destination through this AS path"



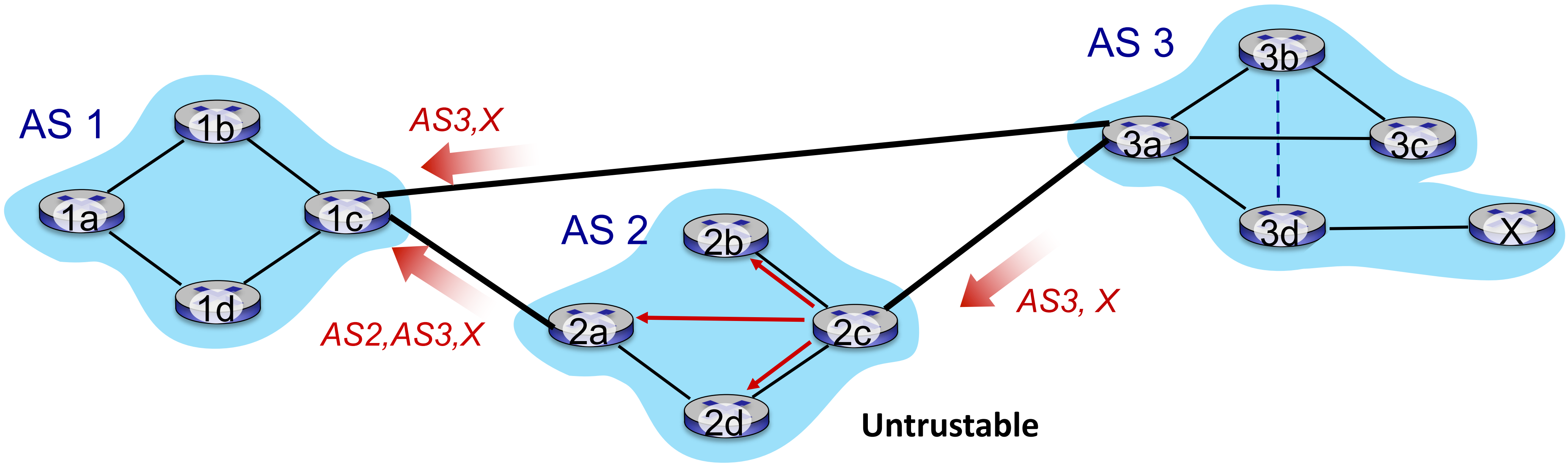
Why path vector is necessary: loop prevention

- For pure distance vector, loop is hard to detect "count-to-infinity"
- In BGP, Loop detection becomes much more explicit since a route carries the ASes it has traversed
  - AS rejects a path if itself appears in the path

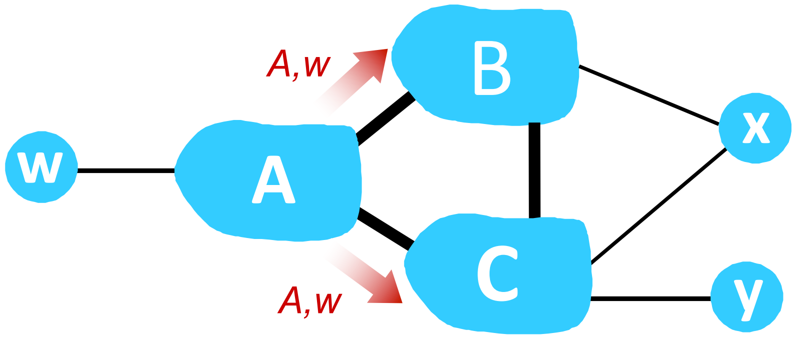


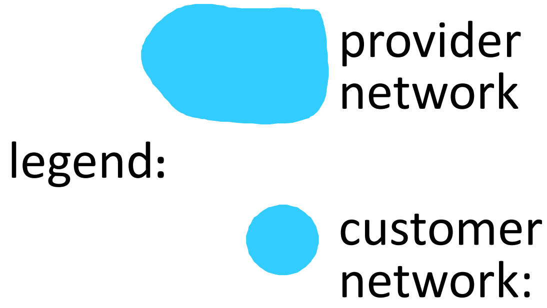
Why path vector is necessary: policy-based routing

- AS1 gateway router 1c learns path *AS2,AS3,X* from 2a
- AS1 gateway router 1c learns path *AS3,X* from 3a
- based on *policy,* AS1 gateway router 1c chooses path *AS3,X* and advertises path within AS1 via iBGP



BGP: achieving policy via advertisements

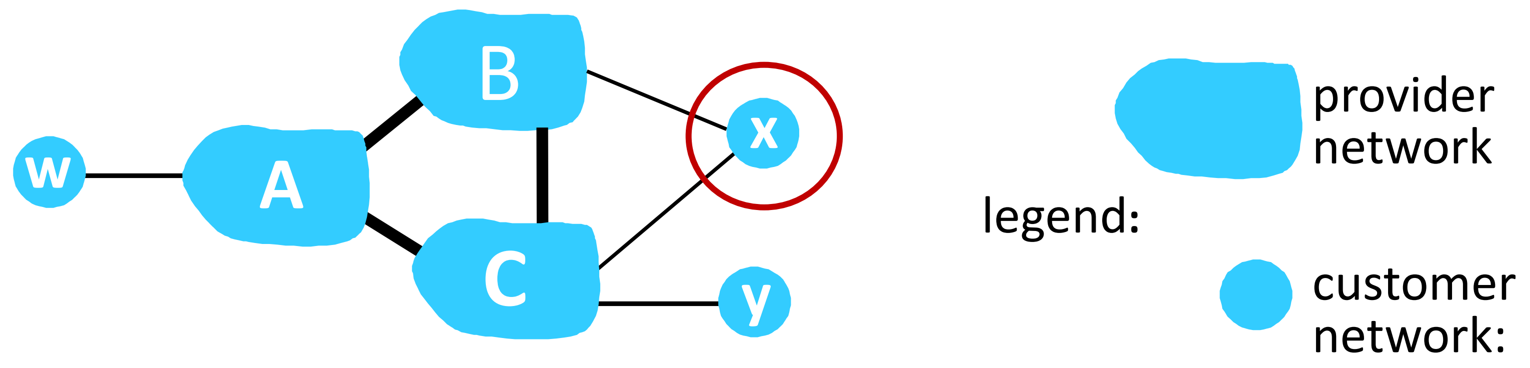




ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

- 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

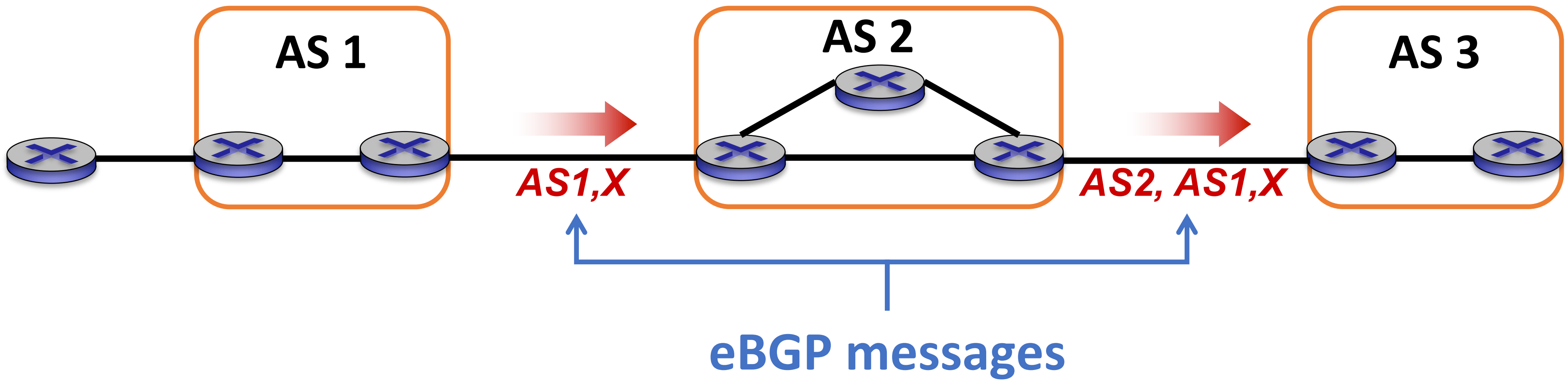
BGP: achieving policy via advertisements (more)

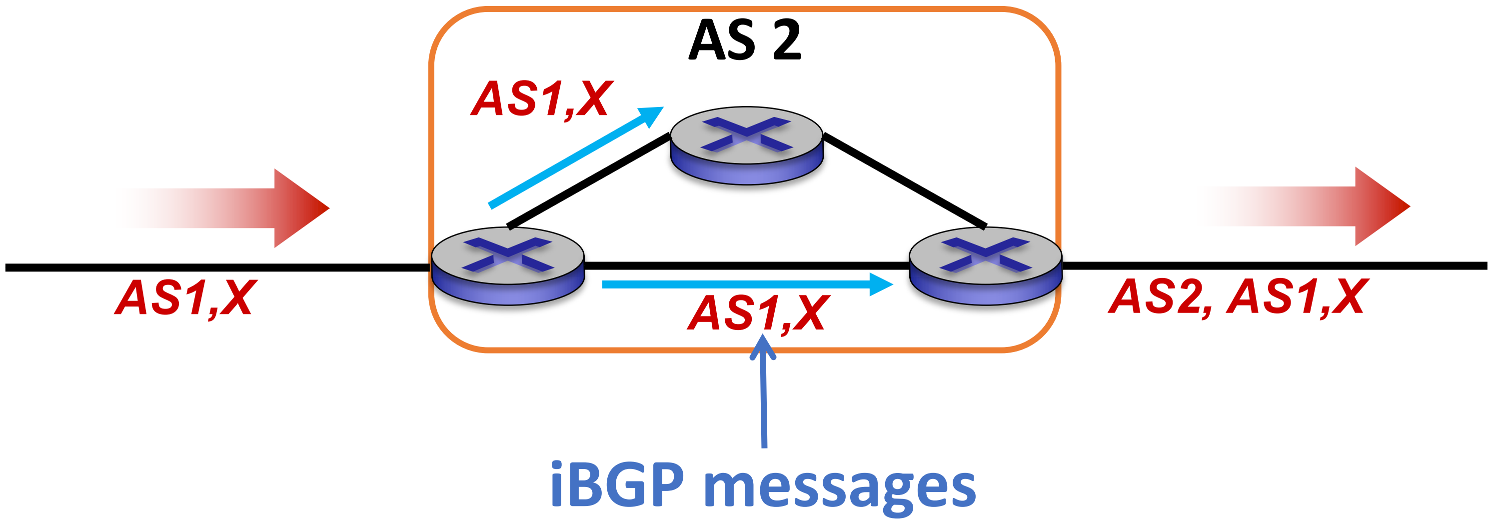


ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

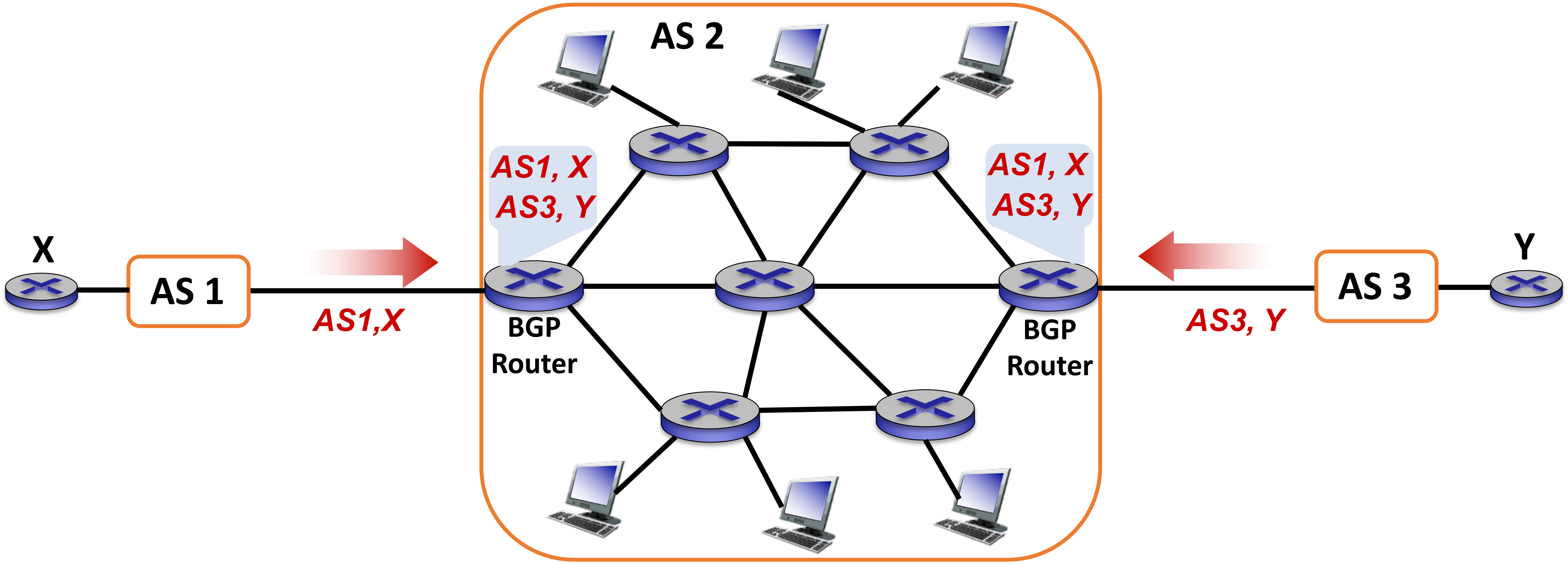
- A,B,C are provider networks
- x,w,y are customer (of provider networks)
- x is dual-homed: attached to two networks
- *policy to enforce:* x does not want to route from B to C via x
  - .. so x will not advertise to B a route to C

iBGP (internal) and eBGP (external)

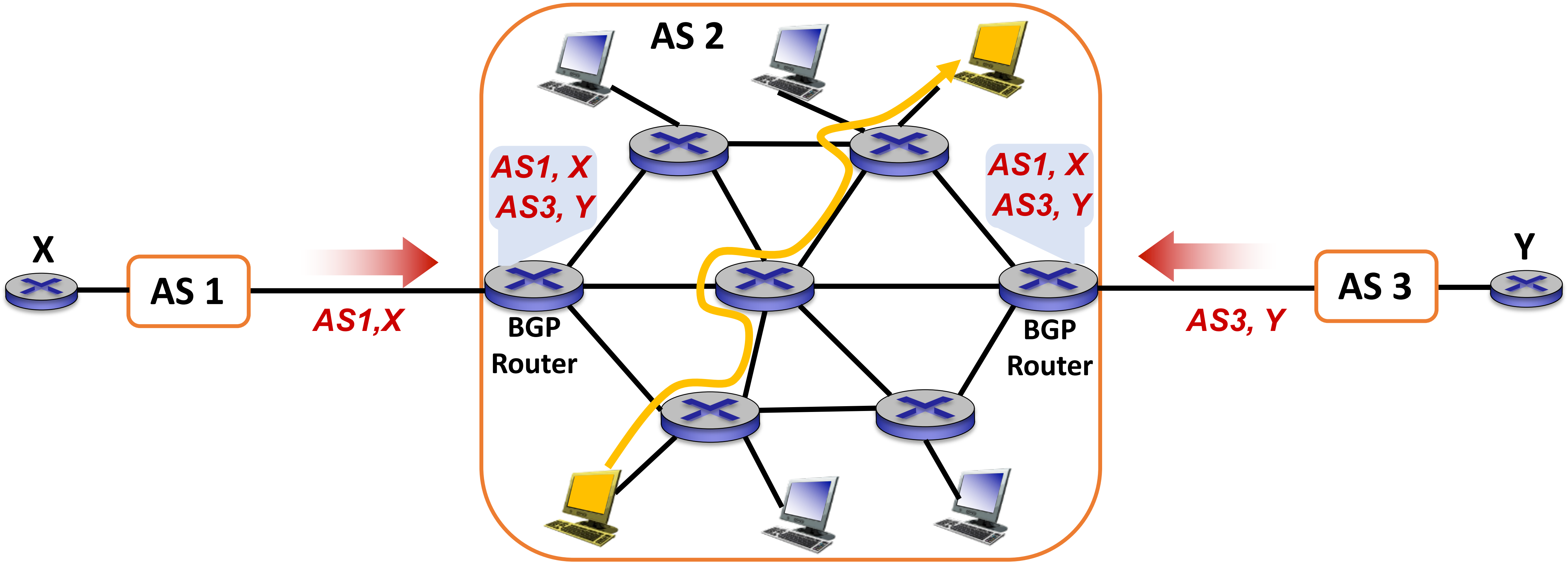




iBGP + eBGP + Intra-AS routing

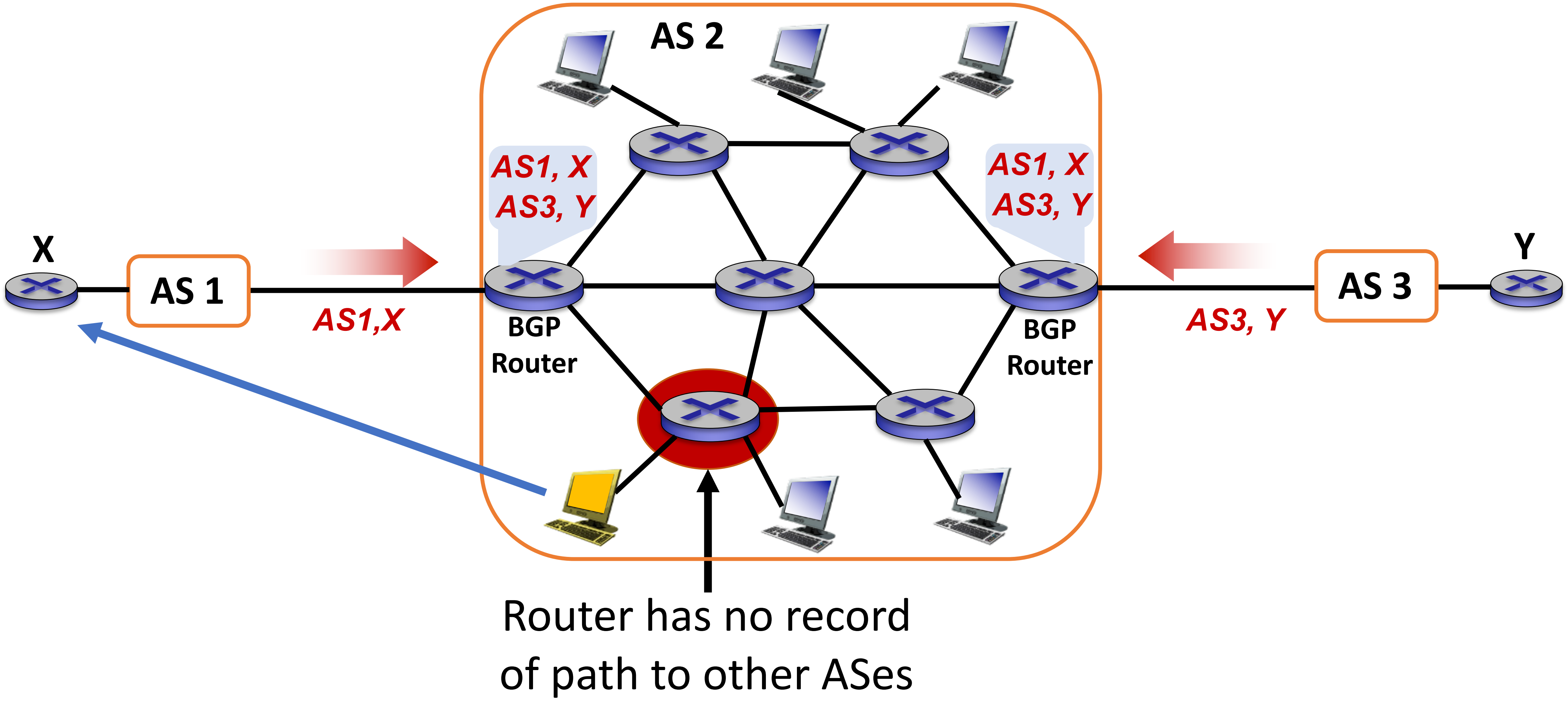


iBGP + eBGP + Intra-AS routing

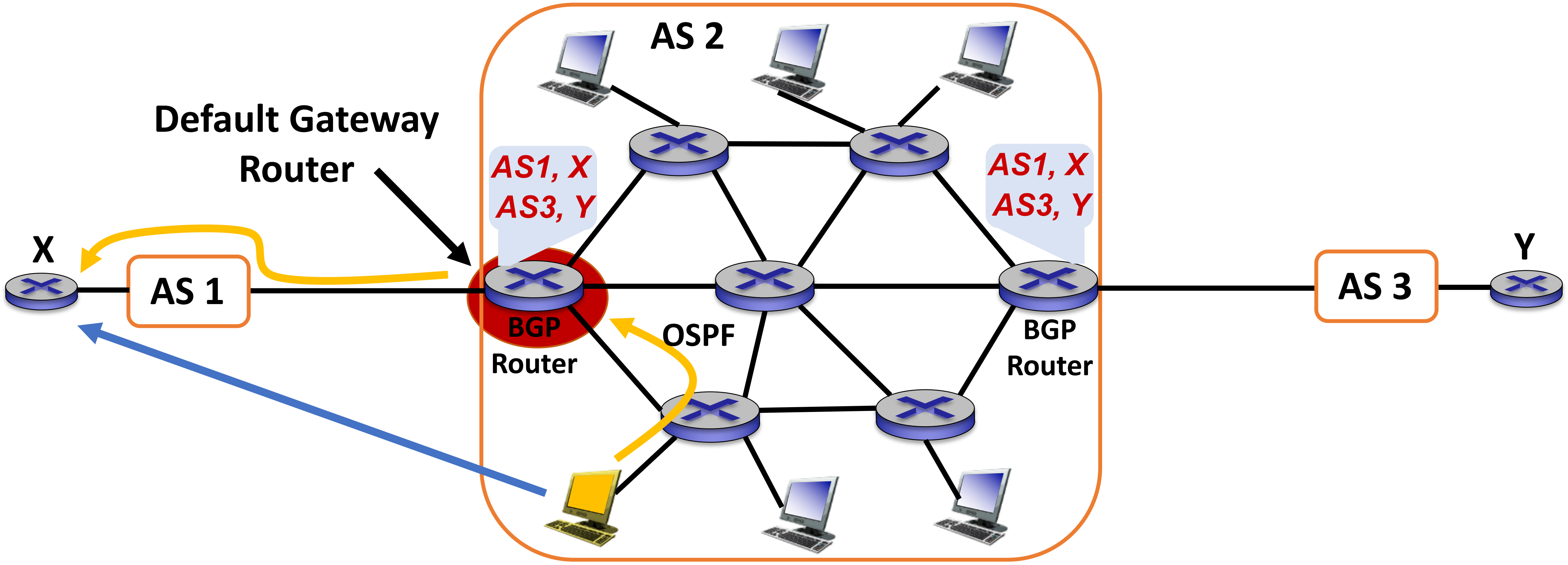


Intra-AS routing: OSPF

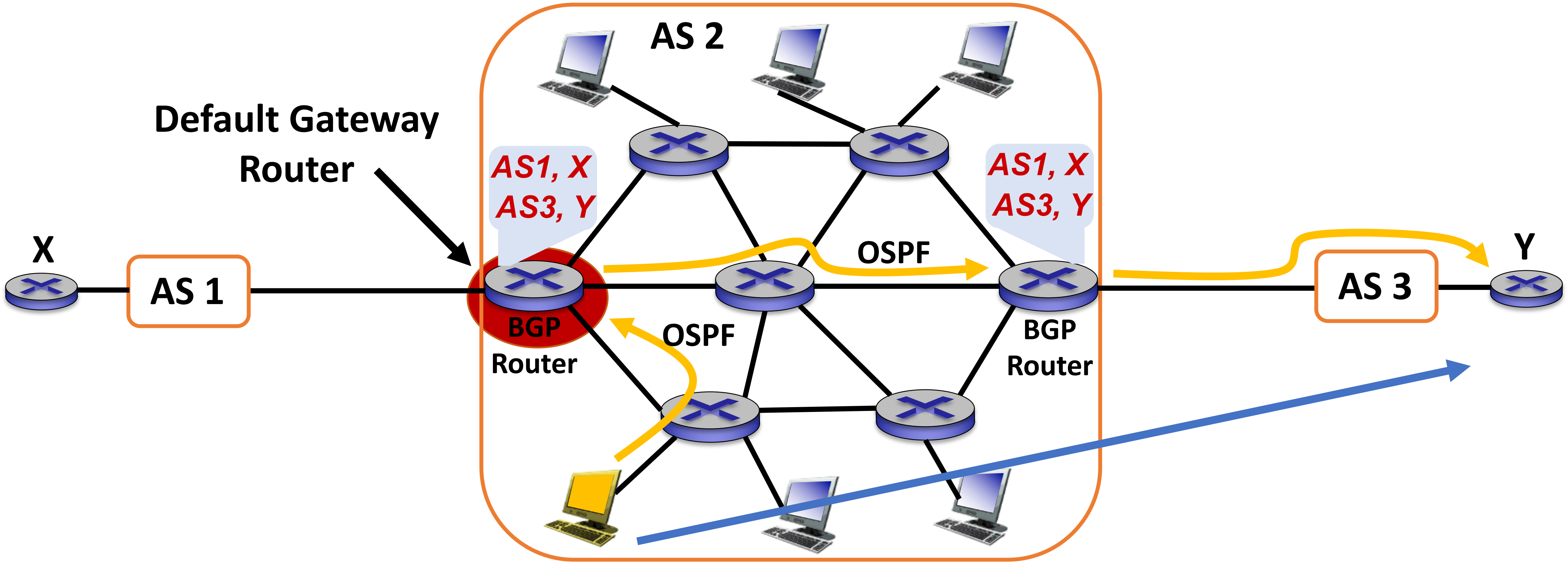
iBGP + eBGP + Intra-AS routing



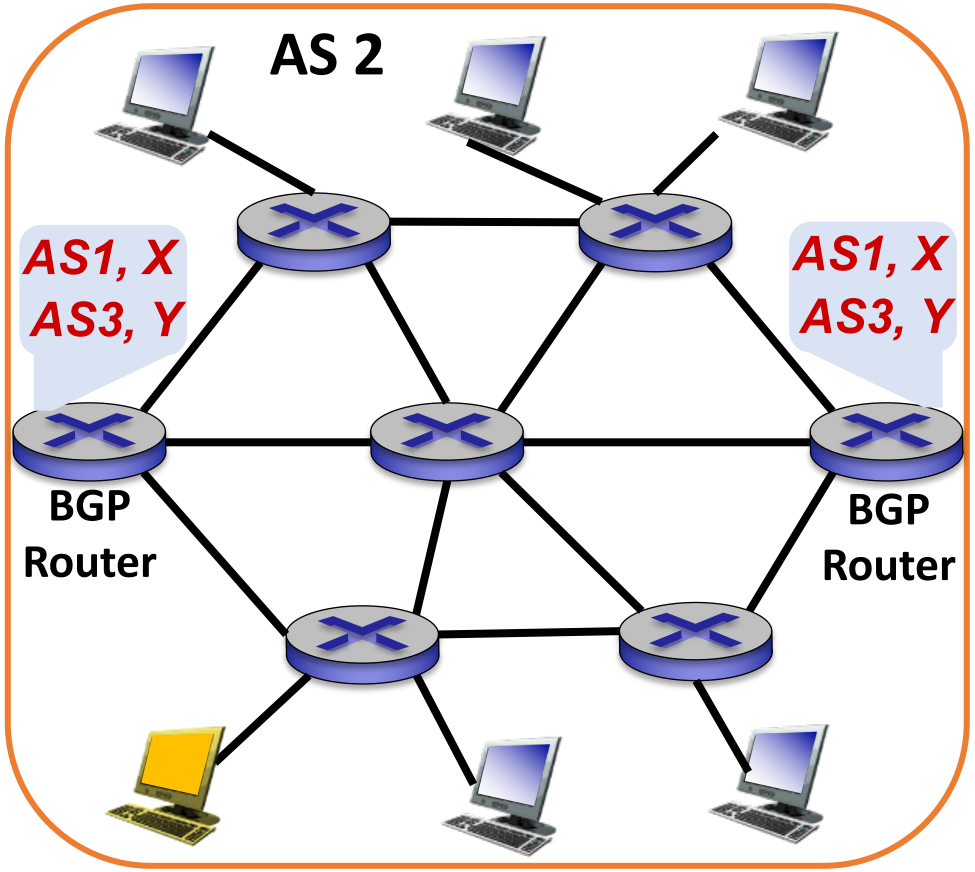
iBGP + eBGP + Intra-AS routing

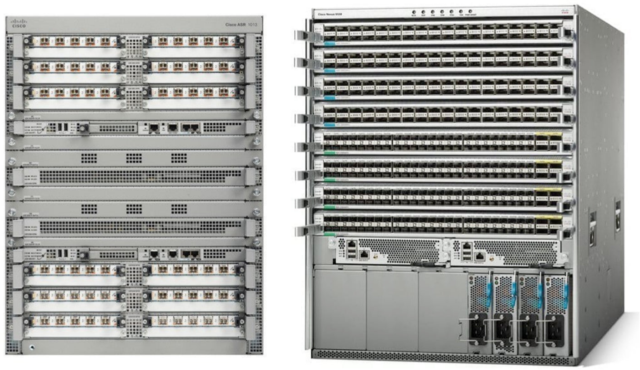


iBGP + eBGP + Intra-AS routing



Normal AS router VS BGP router





CISCO Routers

- ❑ Carry Full Route Table
- ❑ The global **IPv4 routing table** has surpassed **900,000 prefixes**
- ❑ **IPv6 table** has exceeded **200,000 prefixes**
- ❑ Special hardware

Normal Routers

- ❑ Maintain hundreds of prefixes within a AS
- ❑ A normal CPU can handle this
- ❑ Cheap