Lesson 06-06: OSPF and BGP

CS 356 Computer Networks

Mikyung Han

mhan@cs.utexas.edu

Internet Reference Model

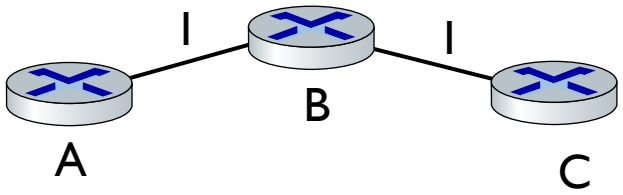
Example Protocols		Responsible for
FTP, HTTP, SMTP	Application	application specific needs
TCP, UDP	Transport	process to process data transfer
IP	Network	host to host data transfer across different network
Ethernet, WiFi	Link	data transfer between physically adjacent nodes
802.3 PHY	Physical	bit-by-bit or symbol-by-symbol delivery

2

Outline

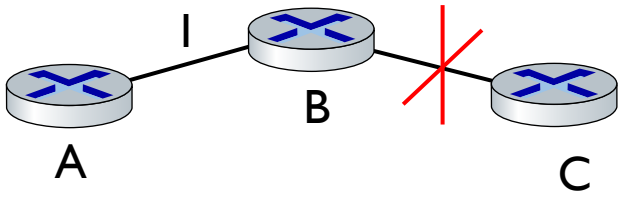
1. Distance Vector Recap

Count to Infinity Problem



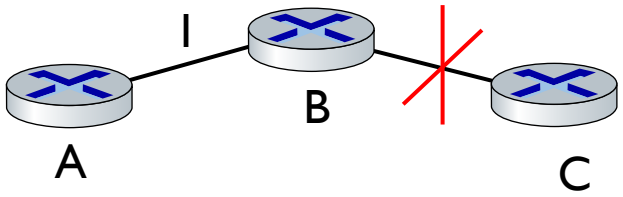
- Bellman-Ford algo after it converges
  - B knows it can get to C at a cost of 1
  - A knows it can get to C via B at cost of 2

Count to Infinity Problem
What happens link B-C fails?



- Upon link B-C failure
  - B knows it can't get to C so removes C from B's table
  - A advertises to B that it can get to C at cost 2
  - Since B can get to A at cost 1, B updates a route to C via A at cost 3
  - A then updates the cost to C to 4
  - B then updates the cost to C to 5
  - And so on

How to solve Count to Infinity problem? Via Split Horizon + Poison reverse



- Do not advertise the route back to B if it is via B
  - A according to split horizon rule does NOT advertises to B that it can get to C at cost 2 via B
  - B knows it can't get to C ~~so removes C from B's table~~ advertises B's cost to C is infinity by Poison reverse rule
  - A will eventually learn C's cost from B is infinity so A correctly updates its cost to C as infinity as well.

Outline

1. Distance Vector Recap
2. Motivation
3. OSPF

What are we missing?

our routing study thus far:
idealized

- all routers identical
- network "flat"

… not true in practice

scale: billions of destinations:

- can't store all destinations in routing tables!
- exchanging link-state or DV information would swamp links!

administrative autonomy:

- Internet: a network of networks
- each network admin may want to control routing in its own network

AS: Internet approach to scalable routing

aggregate routers into regions known as "autonomous systems" (AS) (a.k.a. "domains")

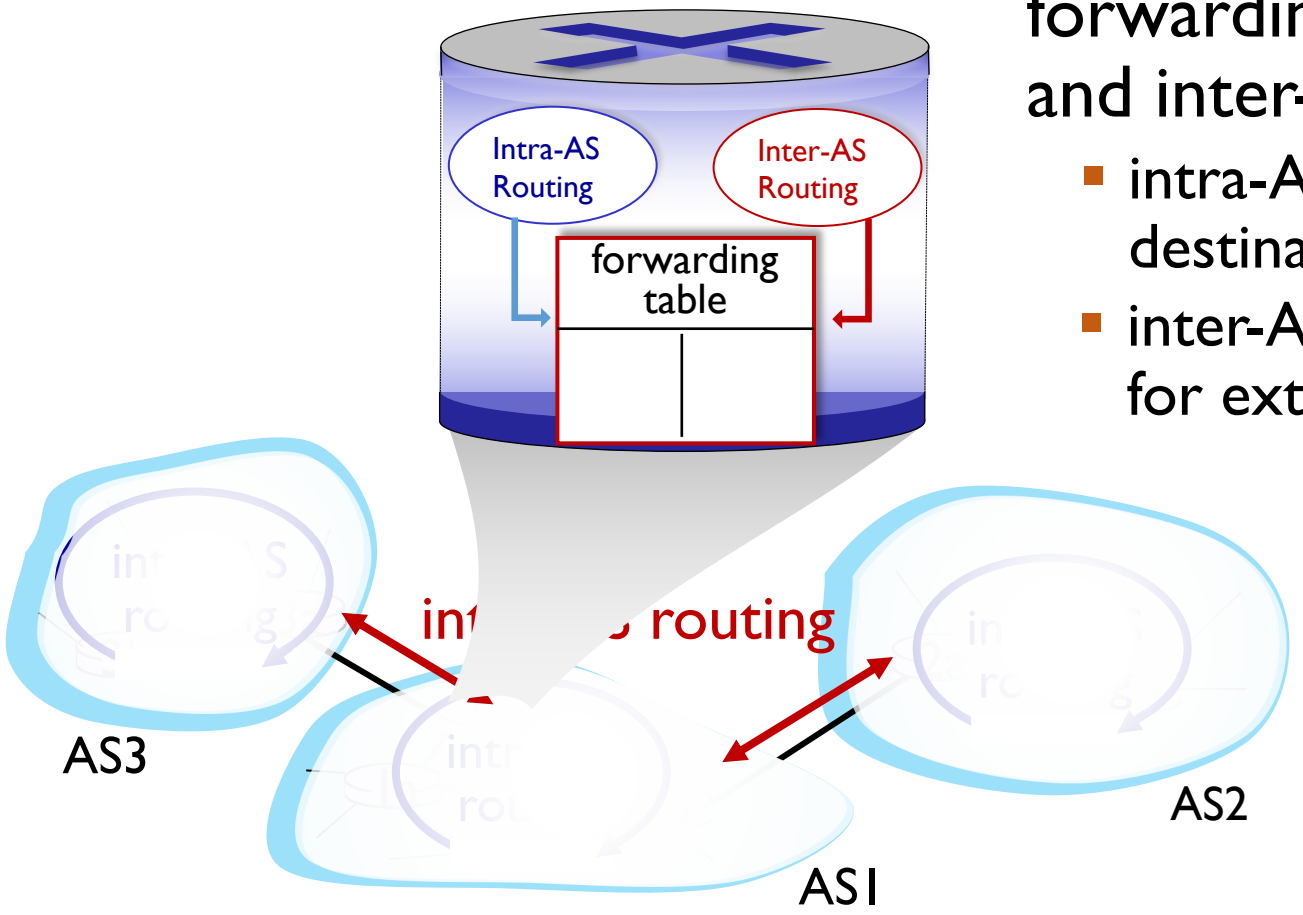
intra-AS (aka "intra-domain"): routing among routers within same AS ("network")

- all routers in AS must run same intra-domain protocol
- routers in different AS can run different intra-domain routing protocols
- gateway router: at "edge" of its own AS, has link(s) to router(s) in other AS'es

inter-AS (aka "inter-domain"): routing among AS'es

- gateways perform inter-domain routing (as well as intra-domain routing)

Interconnected ASes



forwarding table configured by intra- and inter-AS routing algorithms

- intra-AS routing determine entries for destinations within AS
- inter-AS & intra-AS determine entries for external destinations

Outline

1. Distance Vector Recap
2. Motivation
3. OSPF – Intra AS routing protocol

OSPF (Open Shortest Path First) routing

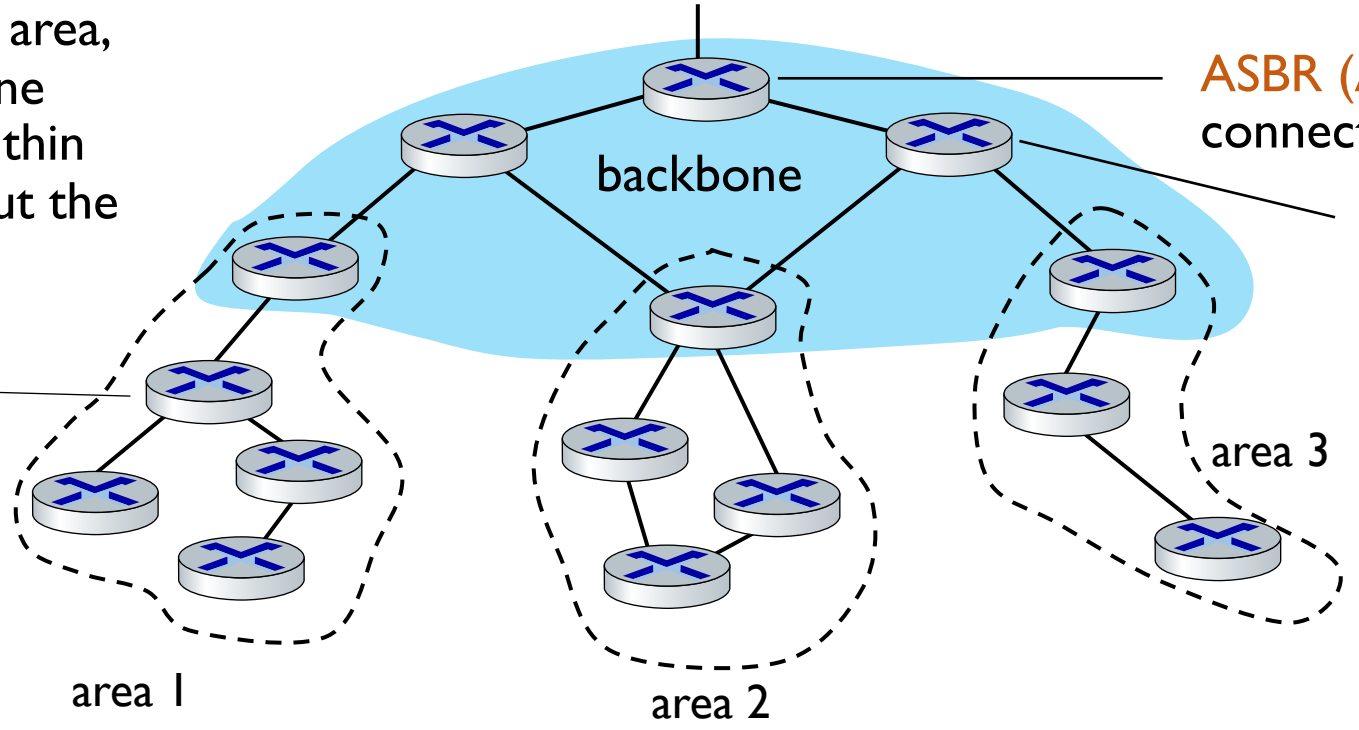
- "open": publicly available
- classic link-state
  - each router floods OSPF link-state advertisements (directly over IP rather than using TCP/UDP) to all other routers in entire AS
  - multiple link costs metrics possible: bandwidth, delay
  - each router has full topology, uses Dijkstra's algorithm to compute forwarding table
- security: all OSPF messages authenticated (to prevent malicious intrusion)

Isn’t flooding not scalable for bigger AS’s?

Hierarchical OSPF to solve scalability

ABR (Area Border router):
- "summarize" distances to destinations in own area, advertise in backbone
- Also lets routers within one area know about the other area

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



ASBR (AS Border Router):
connects to other ASes

Backbone router:
runs OSPF limited to backbone

Outline

1. Distance Vector Recap
2. Motivation
3. OSPF
4. BGP – Inter AS routing protocol

Why different Intra-, Inter-AS routing ?

scale: reducing forwarding table size, routing update traffic

- hierarchical routing: limiting the scope of full topological information
- BGP routing to CIDRized destination networks (summarized routes)

policy:

- inter-AS: admin wants control over how its traffic routed, who routes through its network
- intra-AS: single admin, so policy less of an issue

performance:

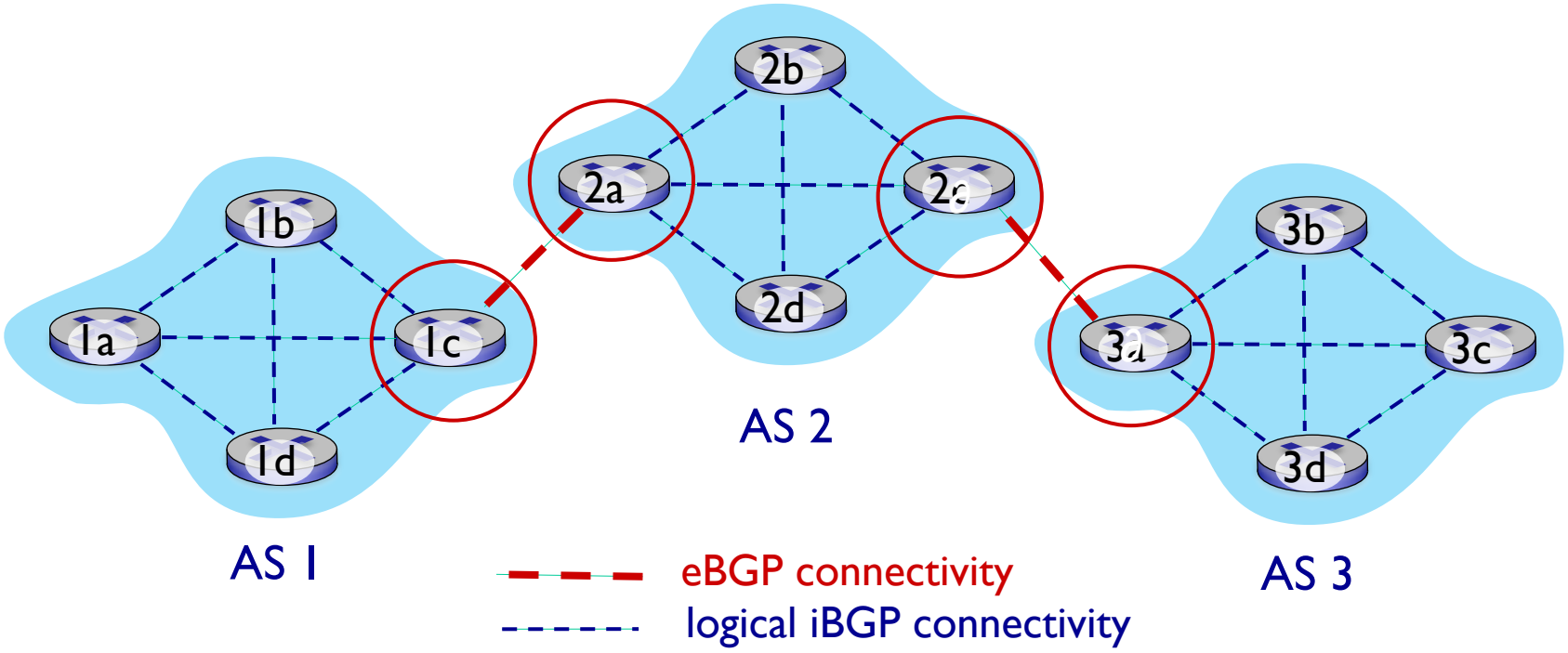
- intra-AS: can focus on performance
- inter-AS: policy dominates over performance

Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): the de facto inter-domain routing protocol
- allows subnet to advertise its existence, and destinations it can reach to rest of Internet

"I am here, and here is who I can reach and how"

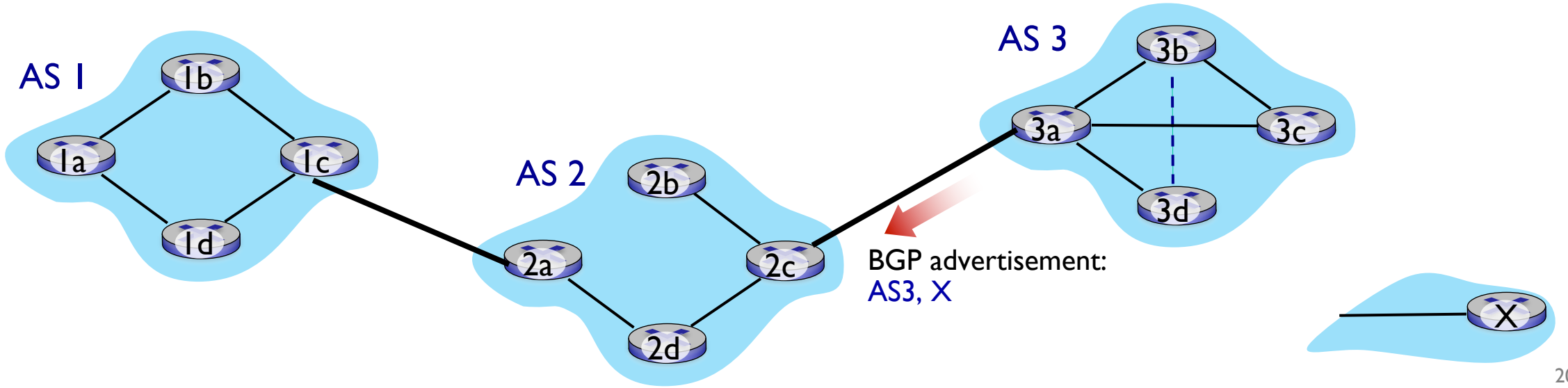
eBGP, iBGP connections



gateway routers run both eBGP and iBGP protocols

BGP basics

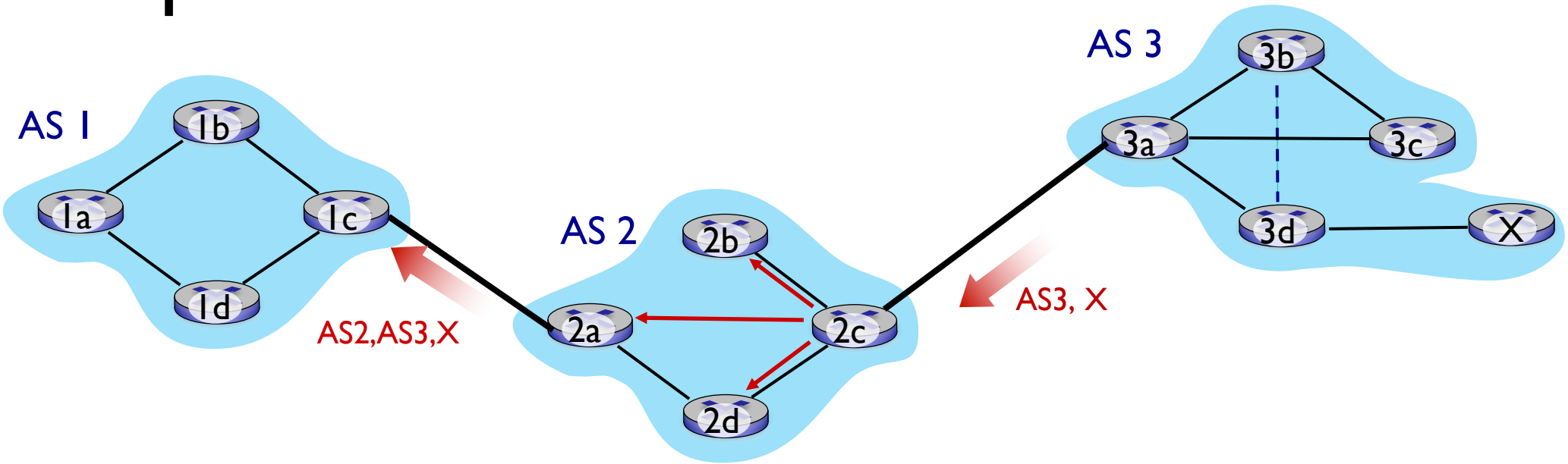
- BGP session: two BGP routers ("peers, speakers") exchange BGP messages over semi-permanent TCP connection:
  - advertising paths to different destination network prefixes (e.g., to a destination /16 network)
  - BGP is a "path vector" protocol
- when AS3 gateway 3a advertises path AS3,X to AS2 gateway 2c:
  - AS3 promises to AS2 it will forward datagrams towards X



Policy rules over performance in BGP

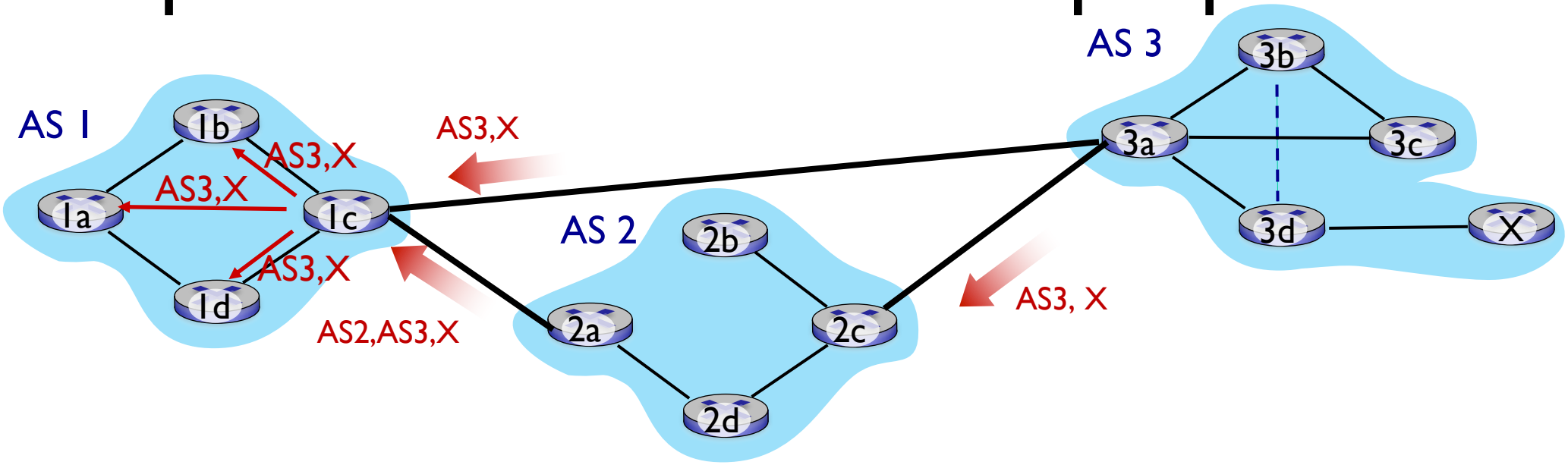
- policy-based routing:
  - router receiving route advertisement to destination X uses policy to accept/reject a path (e.g., never route through AS W, or country Y).
  - router uses policy to decide whether to advertise a path to neighboring AS Z (does router want to route traffic forwarded from Z destined to X?)

BGP path advertisement



- AS2 router 2c receives path advertisement AS3,X (via eBGP) from AS3 router 3a
- based on AS2 policy, AS2 router 2c accepts path AS3,X, propagates (via iBGP) to all AS2 routers
- based on AS2 policy, AS2 router 2a advertises (via eBGP) path AS2, AS3, X to AS1 router 1c

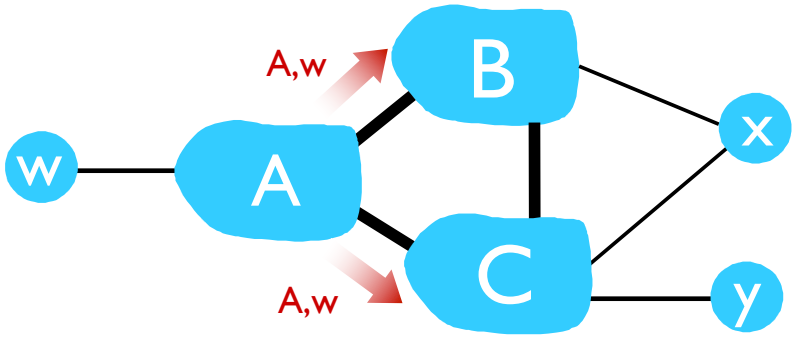
BGP path advertisement: multiple paths

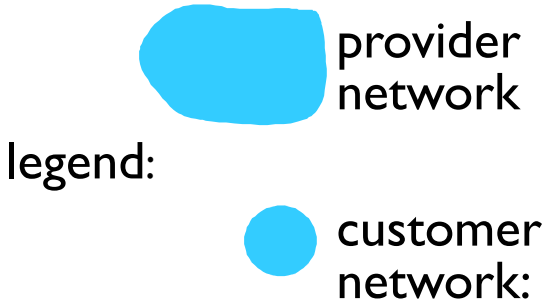


gateway routers may learn about multiple paths to destination:

- 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

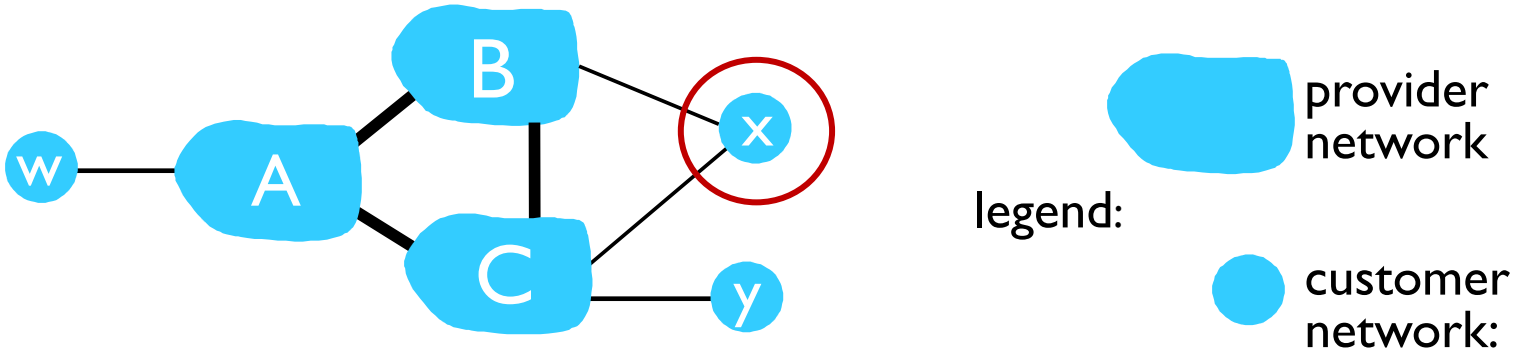




ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

- 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

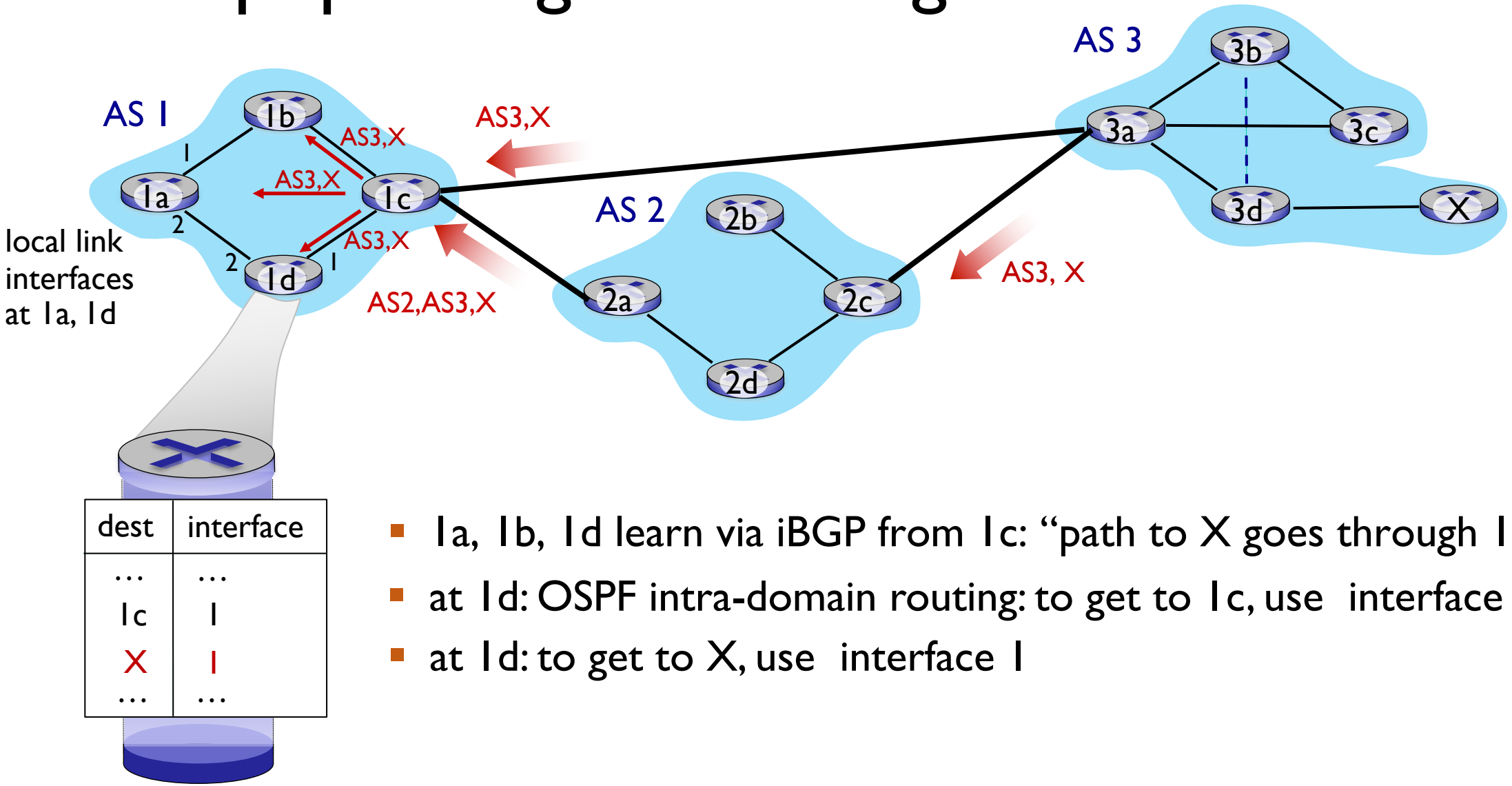
BGP: achieving policy via advertisements (more)



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)

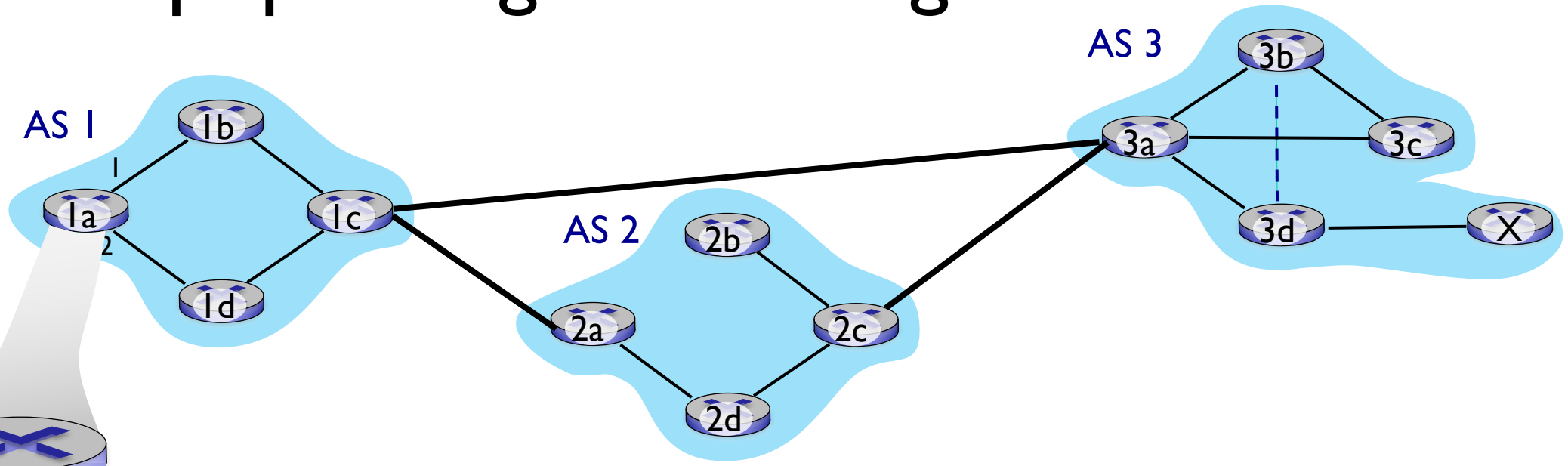
- A,B,C are provider networks
- x,w,y are customer (of provider networks)
- x is dual-homed: attached to two networks
- policy to enforce: x does not want to route from B to C via x
  - .. so x will not advertise to B a route to C

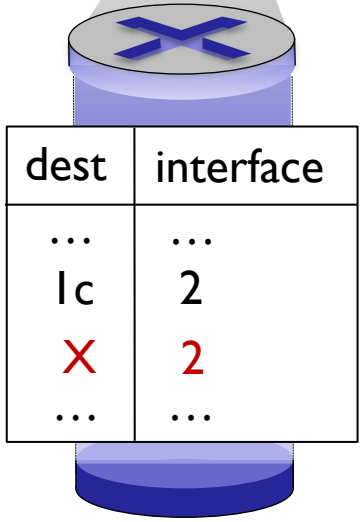
BGP: populating forwarding tables



- 1a, 1b, 1d learn via iBGP from 1c: "path to X goes through 1c"
- at 1d: OSPF intra-domain routing: to get to 1c, use interface 1
- at 1d: to get to X, use interface 1

BGP: populating forwarding tables

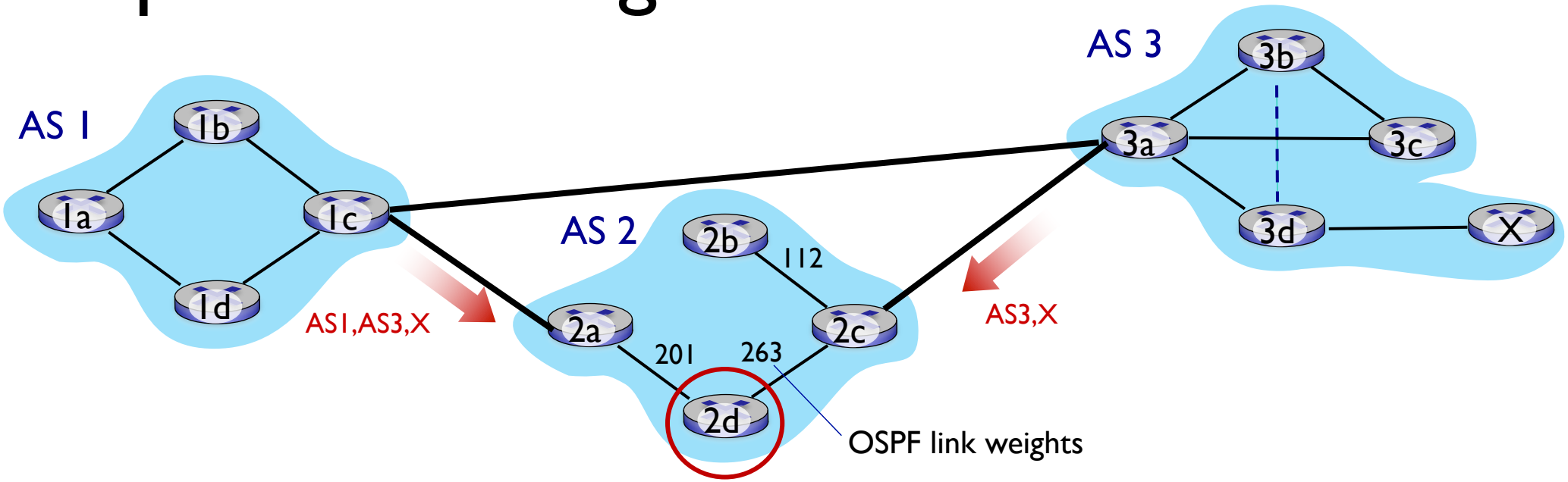




dest	interface
…	…
1c	2
X	2
…	…

- 1a, 1b, 1d learn via iBGP from 1c: "path to X goes through 1c"
- at 1d: OSPF intra-domain routing: to get to 1c, use interface 1
- at 1d: to get to X, use interface 1
- at 1a: OSPF intra-domain routing: to get to 1c, use interface 2
- at 1a: to get to X, use interface 2

Hot potato routing



- 2d learns (via iBGP) it can route to X via 2a or 2c
- hot potato routing: choose local gateway that has least intra-domain cost (e.g., 2d chooses 2a, even though more AS hops to X): don't worry about inter-domain cost!

Outline

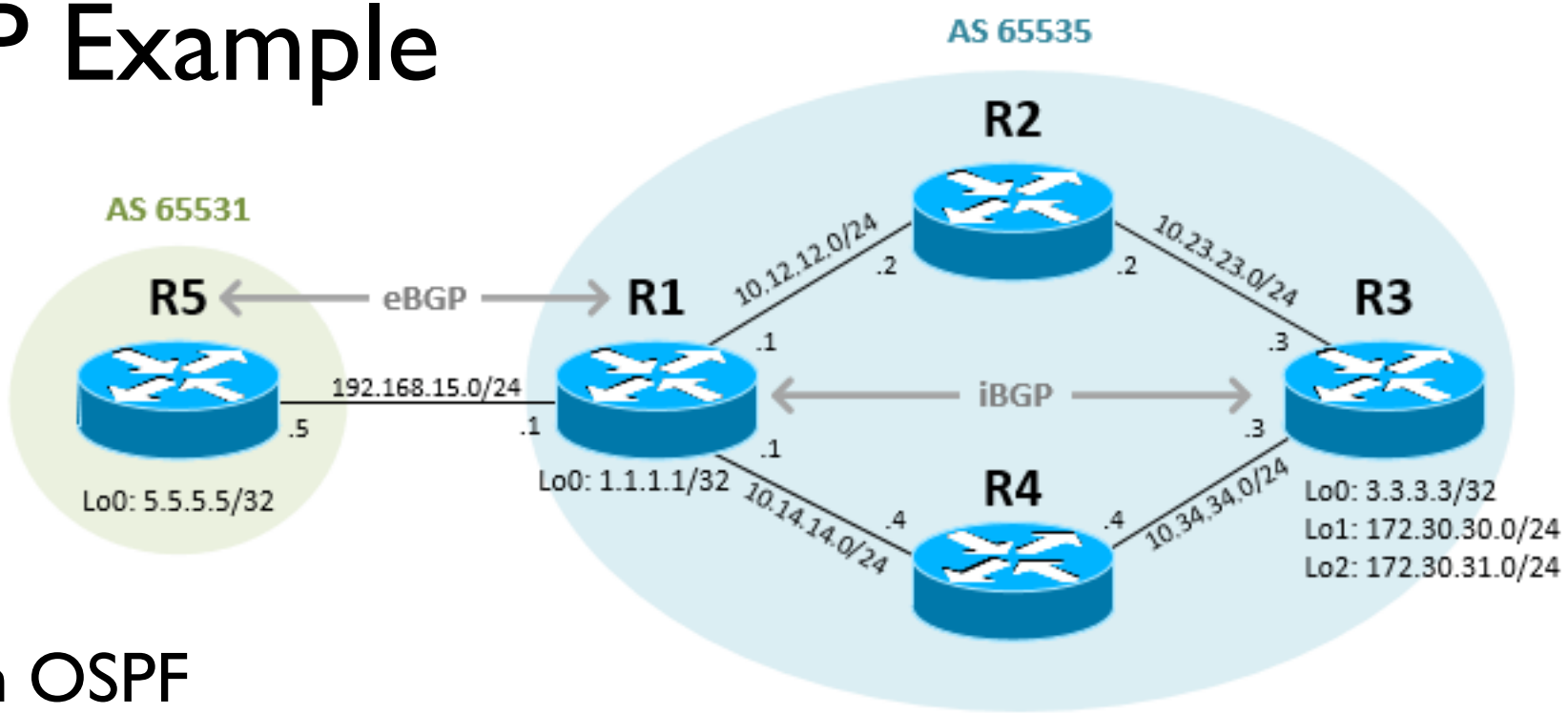
1. Distance Vector Recap
2. Motivation
3. OSPF
4. BGP
5. OSPF vs BGP

OSPF is for intra-AS routing whereas BGP is for inter-AS routing

OSPF vs BGP Comparison

	OSPF	BGP
Gateway protocol	Interior gateway protocol for intra-AS communication	Exterior gateway protocol for Inter-AS communication
Convergence	Fast	Slow
Design	Hierarchical network possible	Meshed
Scale	Smaller scale network	Large scale
Function	Fastest route is preferred over shortest (Dijkstra algo)	Policy dependent
Protocol	IP	TCP

OSPF and BGP Example

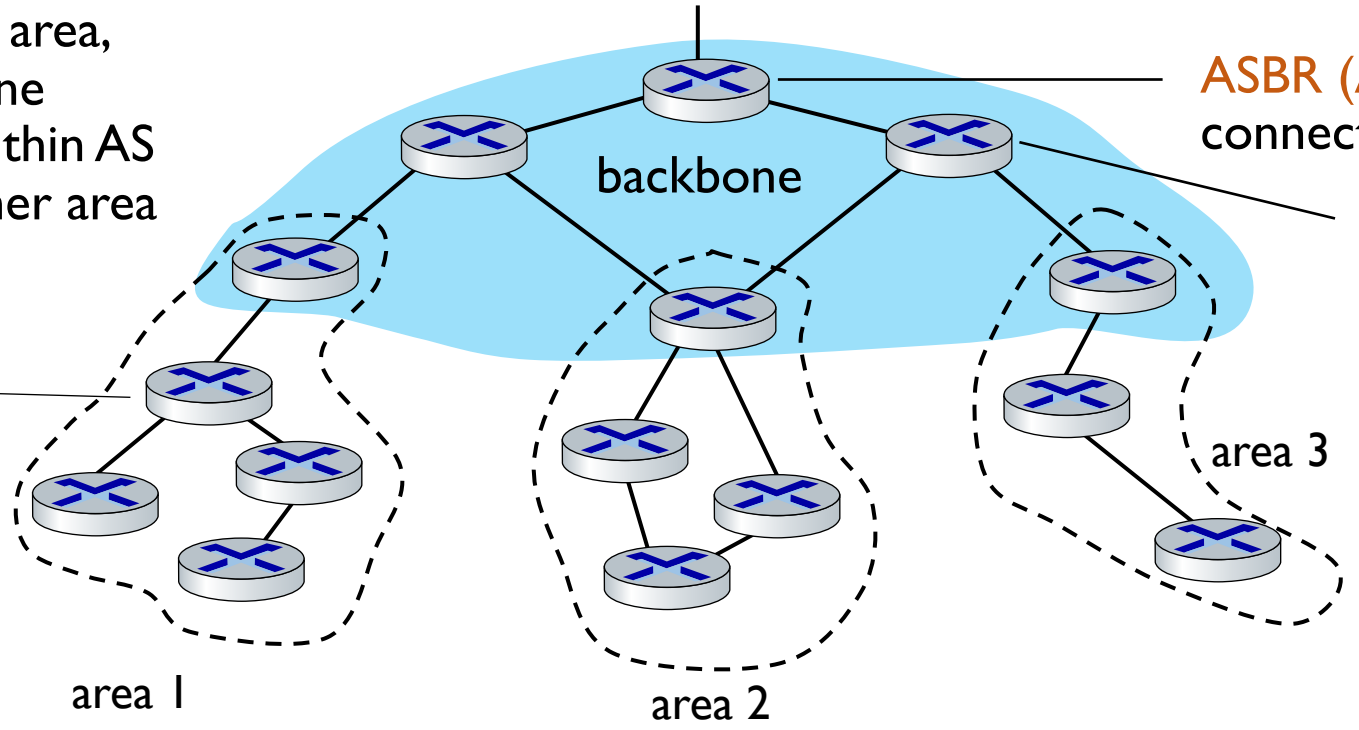


- All area routers run OSPF
- A few selected area routers and all border routers run BGP

- R1 and R5 are
  - ABR (Area Border Router) or ASBR (AS Border Router)

- T/F link-state advertisements are flooded across multiple areas/backbone
- T/F each node has detailed topology for its own area but just next hop for outside
- T/F ASBR runs only BGP not IGP (such as OSPF)

ABR (Area Border router):
- "summarize" distances to destinations in own area, advertise in backbone
- Also lets routers within AS know about the other area



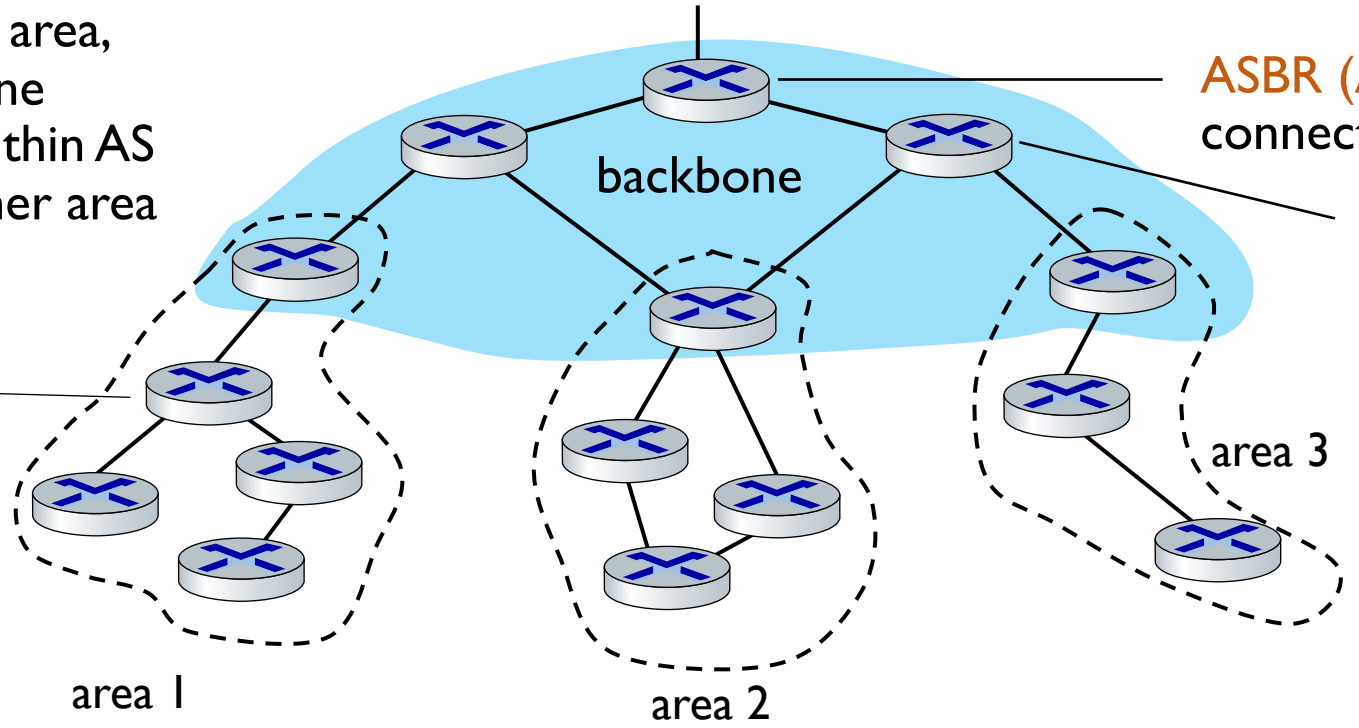
ASBR (AS Border Router): connects to other ASes

Backbone router: runs OSPF limited to backbone

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

- Link-state advertisements are NOT flooded across multiple areas/backbone
- Each node has detailed topology for its own area but just next hop for outside
- ASBR must run both BGP as well as IGP (such as OSPF)

ABR (Area Border router):
- "summarize" distances to destinations in own area, advertise in backbone
- Also lets routers within AS know about the other area



ASBR (AS Border Router): connects to other ASes

Backbone router: runs OSPF limited to backbone

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

Kahoot!