Lesson 06-06:
OSPF and BGP

CS 326E Elements of Networking

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. Network layer functionality Recap

Two functionalities in Network layer: data plane vs control plane

- Data plane performs forwarding according to forwarding table
- Control plane performs routing according to routing algorithms

Refer to textbook for routing algorithms – Dijsktra, Bellman-ford

Outline

1. Network layer functionality Recap
2. Motivation
3. OSPF

What are we missing?

Routing algorithms: idealized

- all routers identical
- network is "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

For scalability and autonomy, per-region routing is needed

- Such region is defined by Autonomous System (AS)
- AS is a set of routers that belong to the same region, also governed by the same management entity
- AS is sometimes also referred as "domain"

Now routing is simplified between intra-AS routing and inter-AS routing

Intra-AS routing defines routing within the same AS

- intra-AS routing runs routing algorithm among routers within same AS only
- 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 AS has link(s) to router(s) in other AS'es

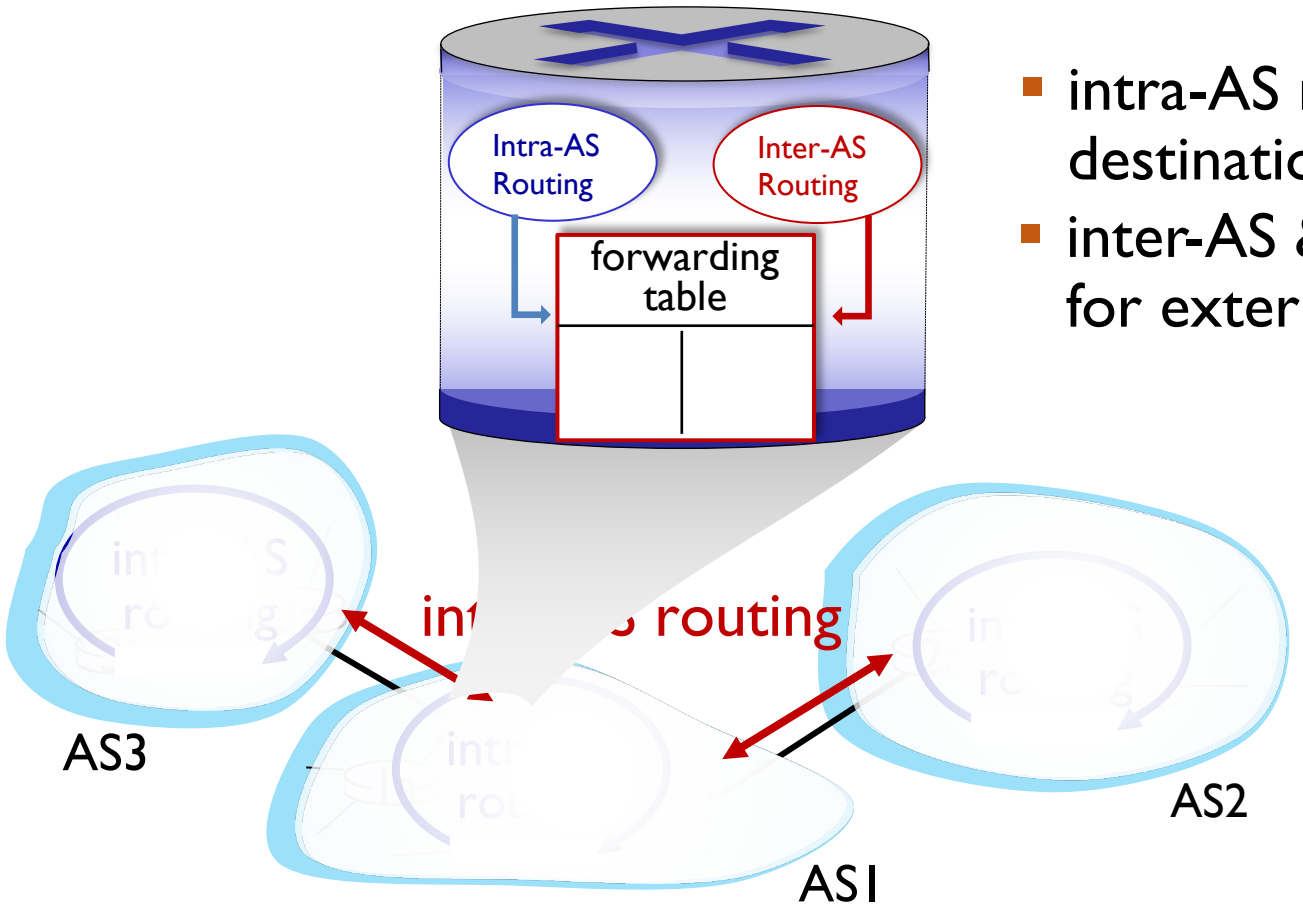
Inter-AS routing defines routing among different AS'es

- All outside traffic goes to gateway routers
- Gateway routers perform inter-domain routing
- Policy rules over efficiency

Gateway routers perform both intra- and inter-AS routing

How many inter-AS routing protocols are used in practice?

Forwarding table is configured by both intra- and inter-AS routing algorithms



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

How many routers in one AS?

- Small AS (enterprise or local ISP): 5-50
- Medium AS (Regional ISP, Cloud providers: 50-500
- Large AS (Tier 1 ISP, Global Cloud providers): 1000+
  - AT&T, Lumen, Cogent
  - Google, AWS, Microsoft

Does UT Austin have its own AS?

Yes! AS15307

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 link-state advertisements (directly over IP rather than using TCP/UDP) to all other routers in entire AS
  - each router has full topology
  - Uses Dijkstra's algorithm to compute forwarding table
  - multiple link costs metrics possible: bandwidth, delay
- security: all OSPF messages authenticated (to prevent malicious intrusion)
  - Authenticity: message is sent by the claimed sender
  - Integrity: message is not altered

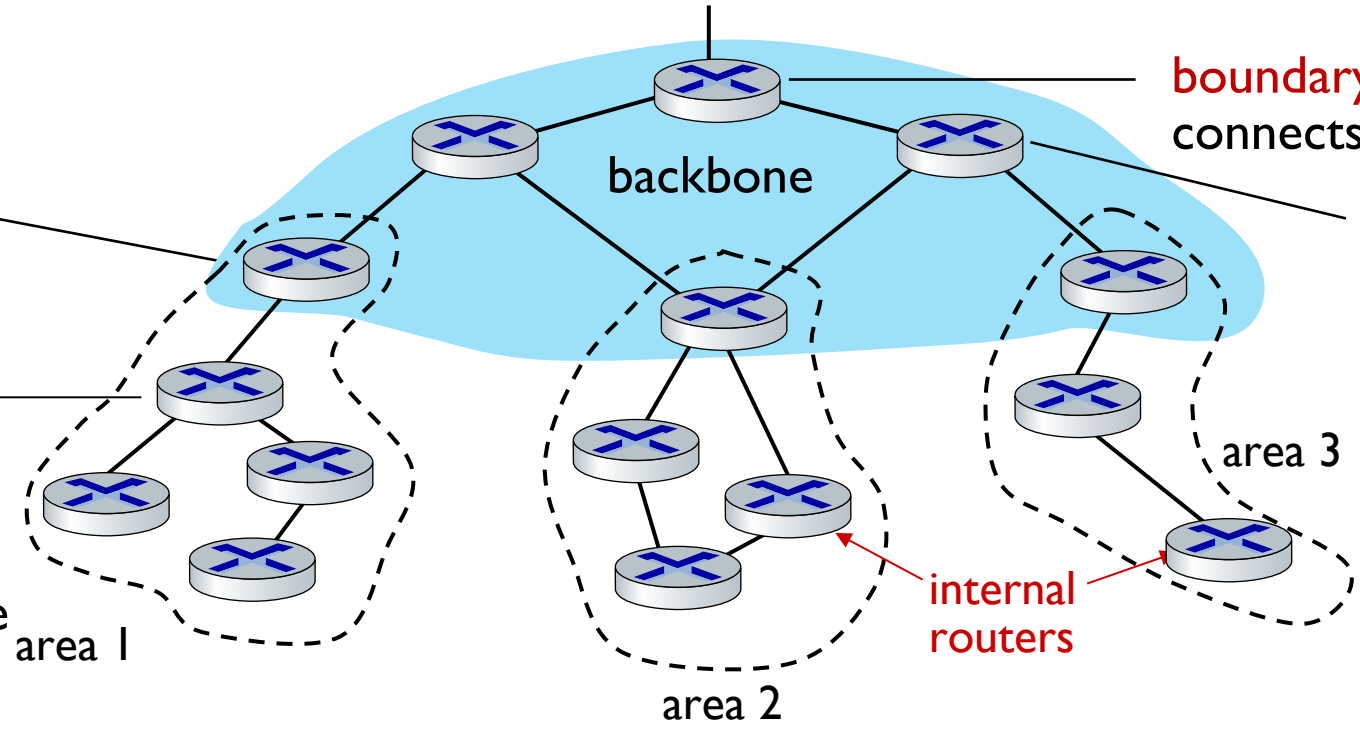
Isn't flooding not scalable for bigger AS's?

Hierarchical OSPF to solve scalability

- two-level hierarchy: local area, backbone.
  - link-state advertisements flooded only in area, or backbone
  - each node has detailed area topology
    knows only the next hop (direction) for out of area destinations

area border routers (ABR):
"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(ASBR):
connects to other ASes

backbone router:
runs OSPF limited to backbone

Hierarchical routing limits the scope of full topological information

This reduces forwarding table size and routing update traffic thus provides better scalability!

Outline

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

Why different Intra-, Inter-AS routing?

There are two factors that determines routing

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

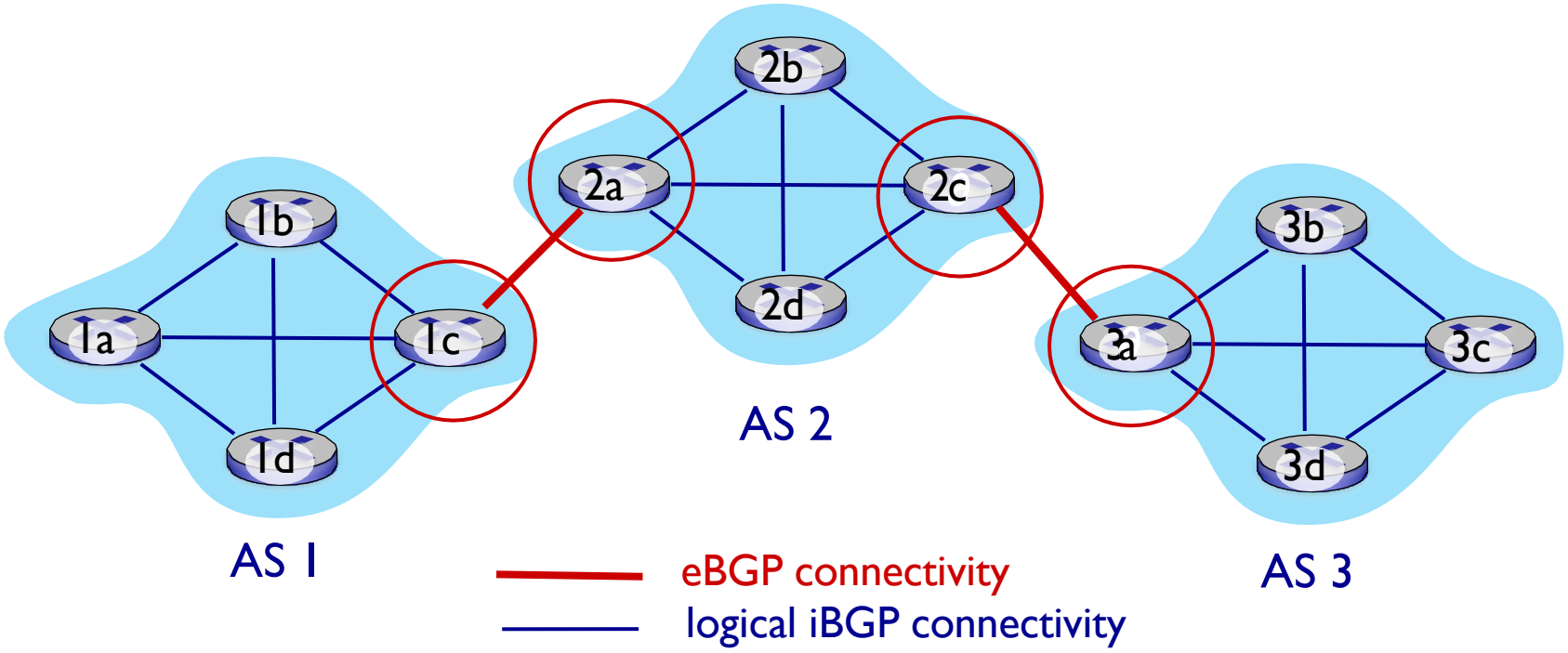
"I am here, and here is who I can reach and how"

BGP is designed for scalability

- External BGP (eBGP) sessions are run among border AS routers
  - Obtain destination network reachability info from neighboring ASes
  - Summarized routes are used

- Internal BGP(iBGP) sessions are run among selected AS-internal routers
  - propagate reachability info to all AS-internal routers (iBGP) about destinations outside the AS

eBGP, iBGP connections



gateway routers run both eBGP and iBGP protocols

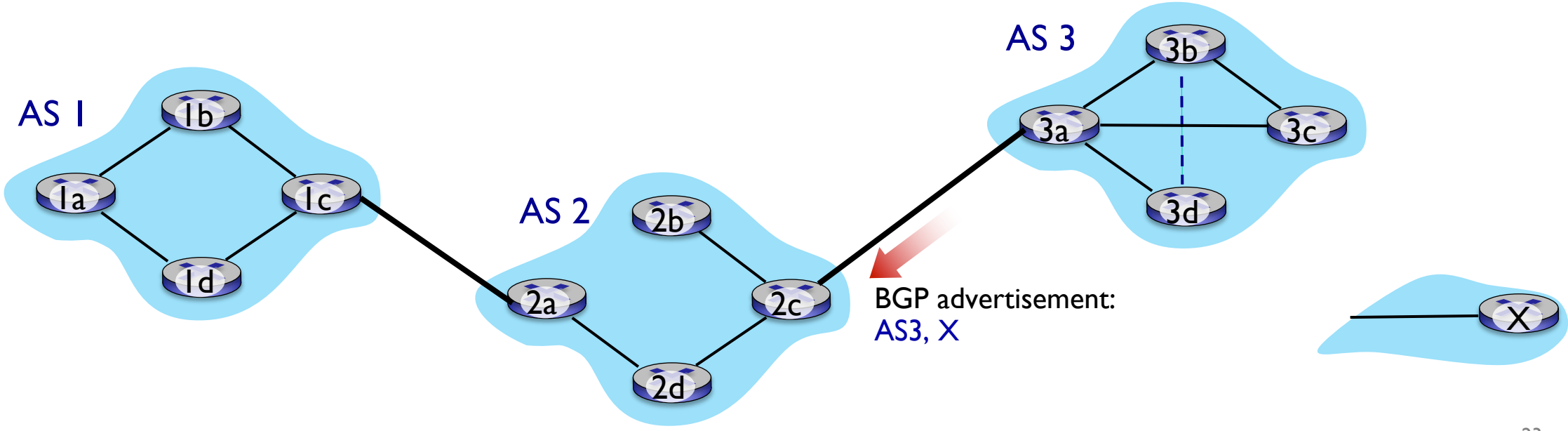
iBGP vs intra-AS routing

- iBGP propagates external routes within AS, handles external AS routes
- intra-AS routing handles local routes within AS

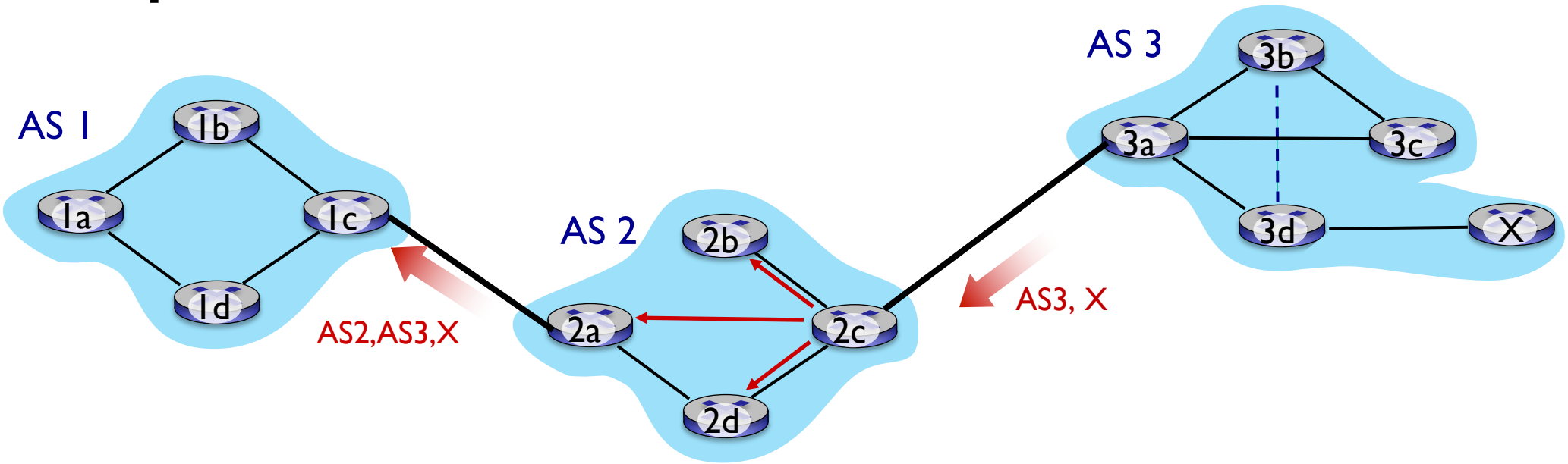
BGP is a "path vector" protocol

- BGP advertises paths to different destination network prefixes
- BGP session: two BGP routers ("peers") exchange BGP advertisements over semi-permanent TCP connection:

- when AS3 gateway 3a advertises path AS3,X to AS2 gateway 2c:
  - AS3 promises to AS2 it will forward datagrams towards 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

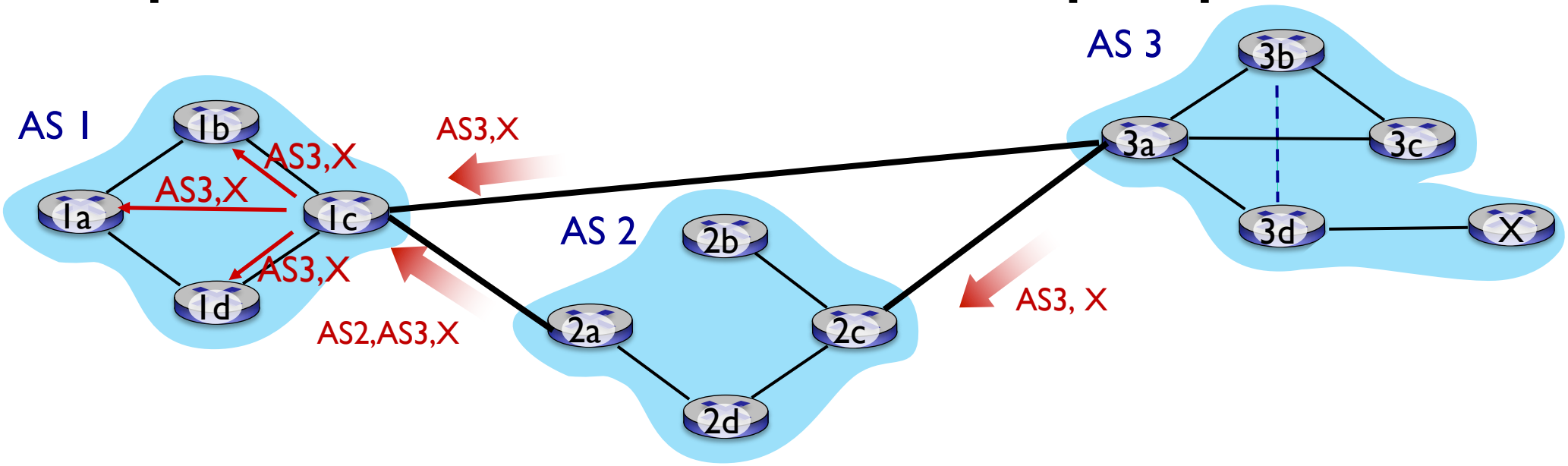
BGP Path attributes

- BGP advertised route consists of prefix + attributes
  - prefix: destination being advertised
    - Why just prefix?
  - two important attributes:
    - AS-PATH: list of ASes through which prefix advertisement has passed
    - NEXT-HOP: indicates specific internal-AS router to next-hop AS

Policy rules over performance in BGP

- 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 (Do I really want to route traffic forwarded from Z destined to X?)

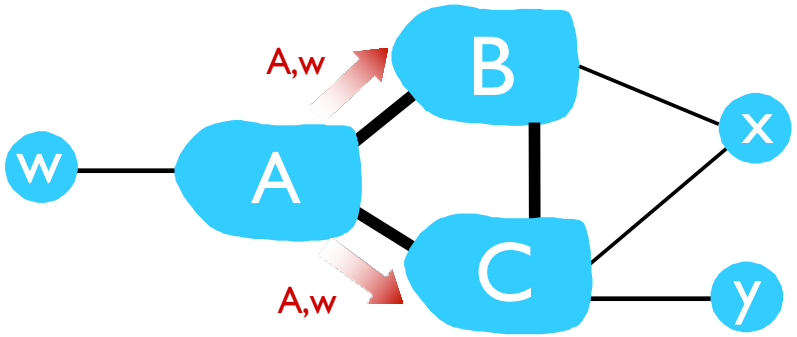
BGP path advertisement: multiple paths

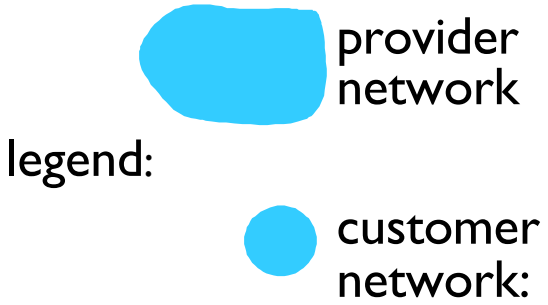


gateway router may learn about multiple paths to destination:

- 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

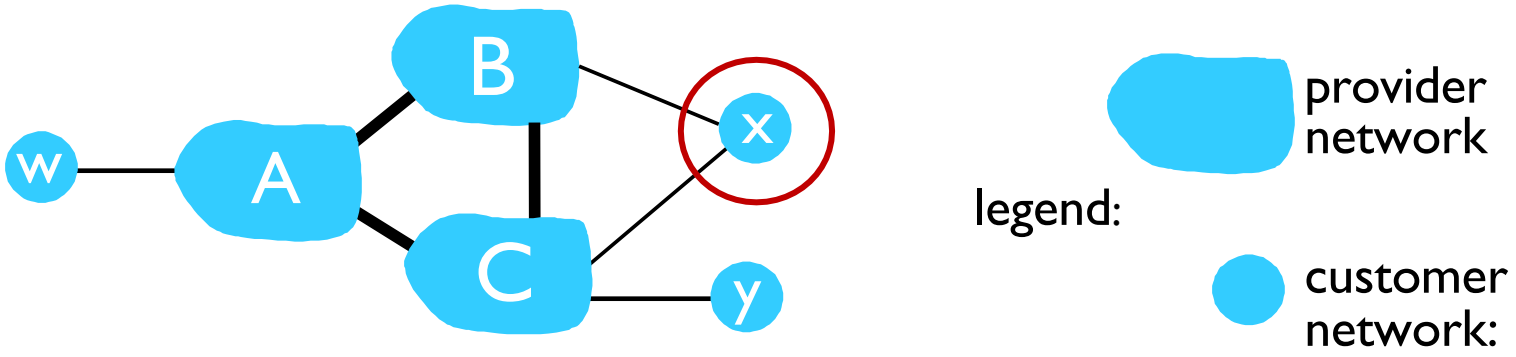




ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

- 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

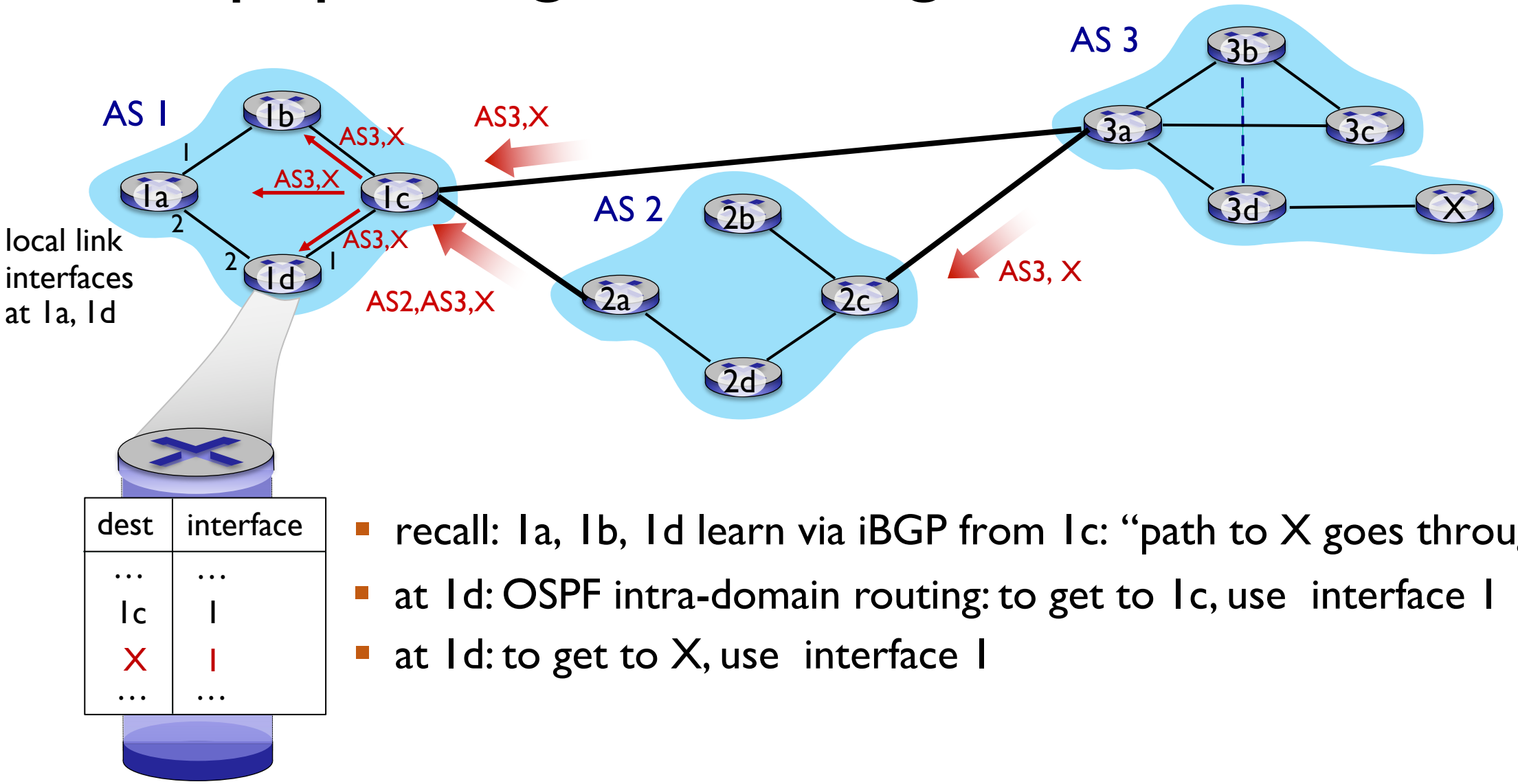
BGP: achieving policy via advertisements (more)



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)

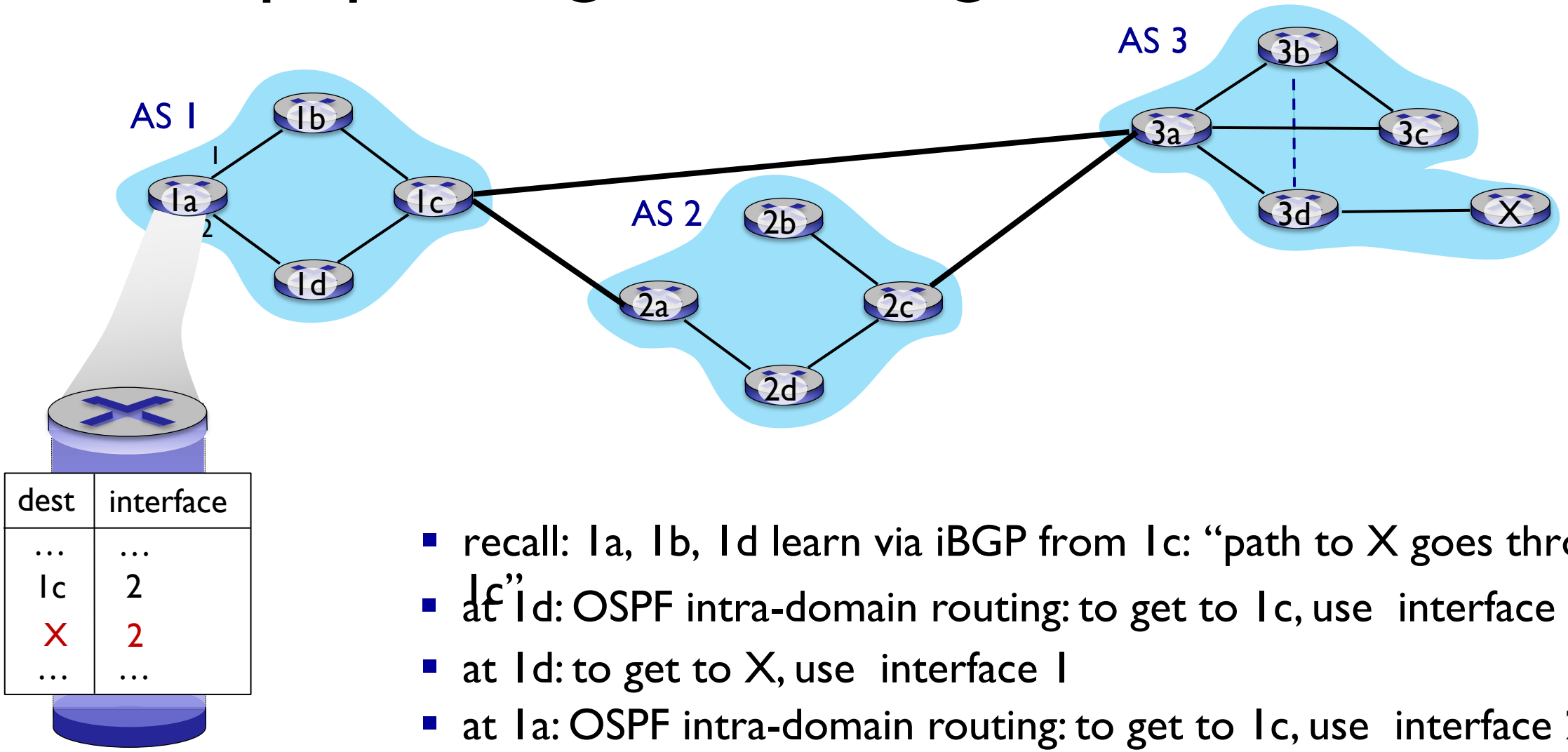
- A,B,C are provider networks
- x,w,y are customer (of provider networks)
- x is dual-homed: attached to two networks
- policy to enforce: x does not want to route from B to C via x
  - .. so x will not advertise to B a route to C

BGP: populating forwarding tables



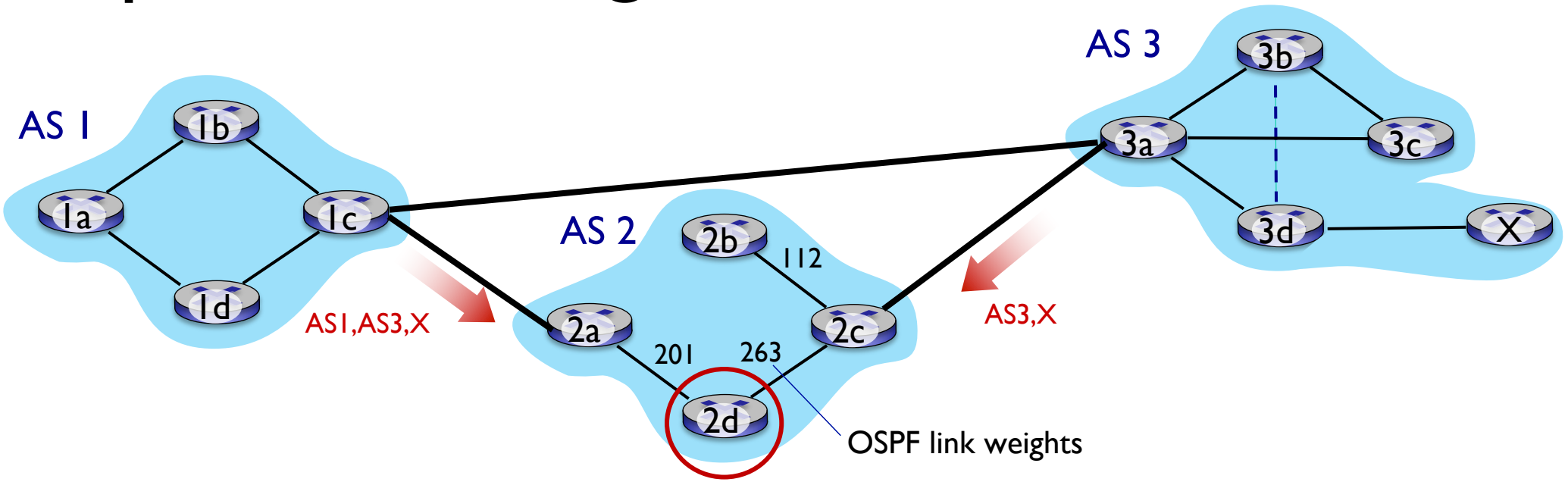
- recall: 1a, 1b, 1d learn via iBGP from 1c: "path to X goes through 1c"
- at 1d: OSPF intra-domain routing: to get to 1c, use interface 1
- at 1d: to get to X, use interface 1

BGP: populating forwarding tables



- recall: 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

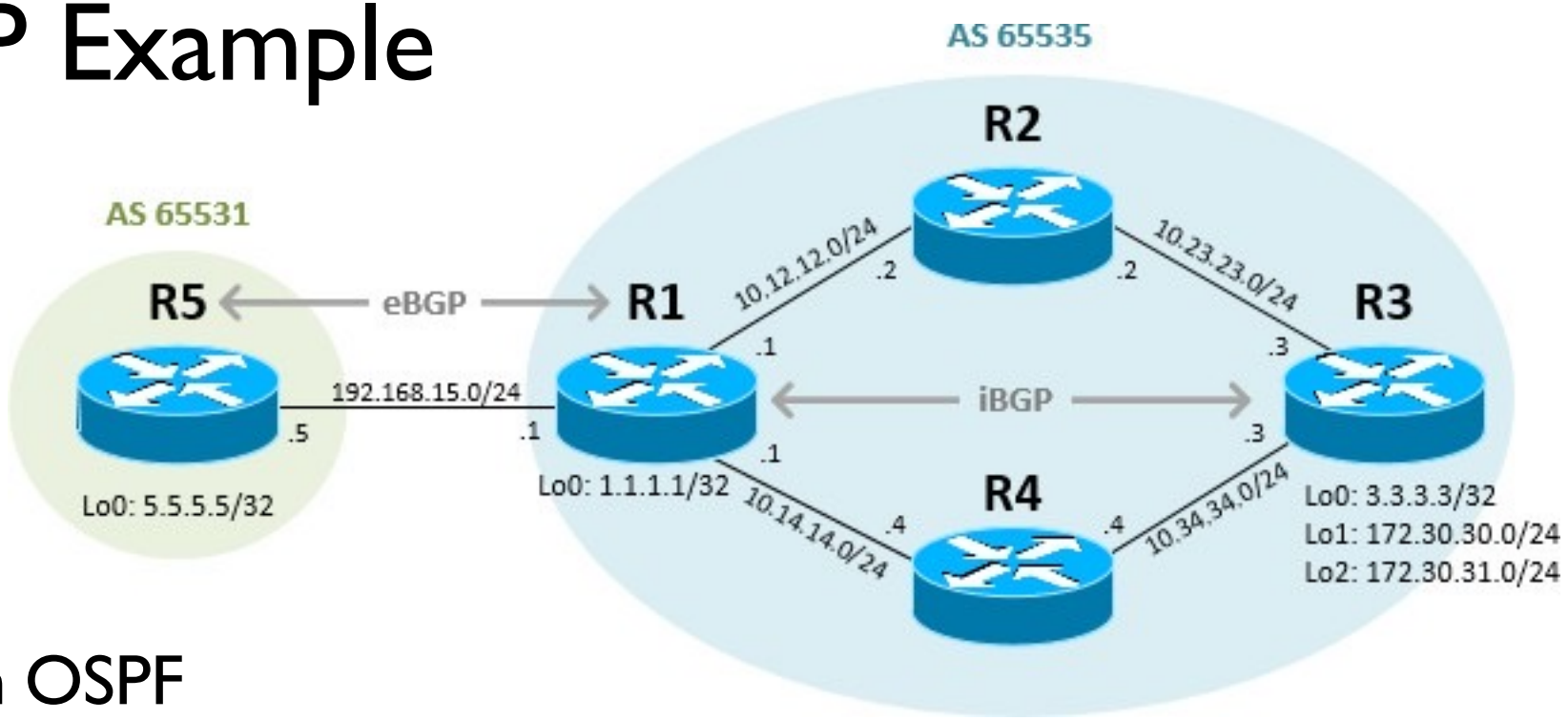
1. Distance Vector Recap
2. Motivation
3. OSPF
4. BGP
5. OSPF vs BGP

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

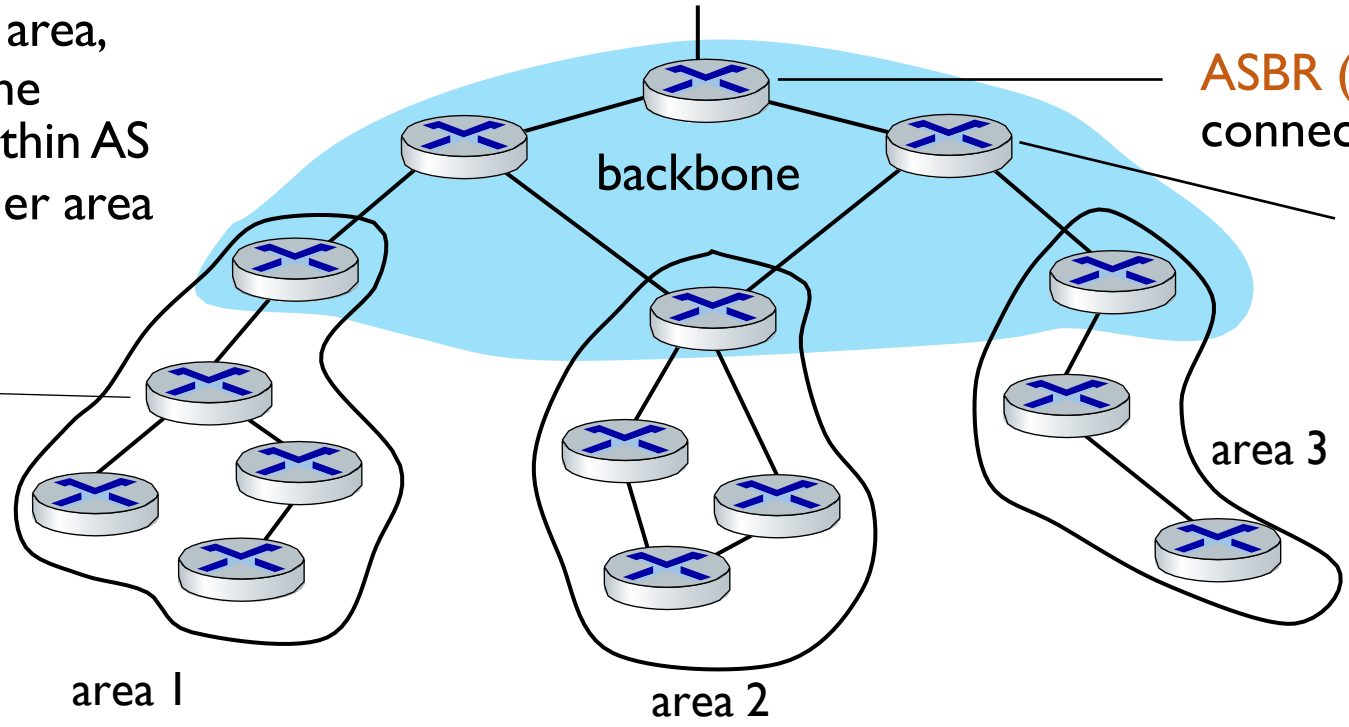


- All area routers run OSPF
- A few selected area routers and all border routers run BGP

- 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



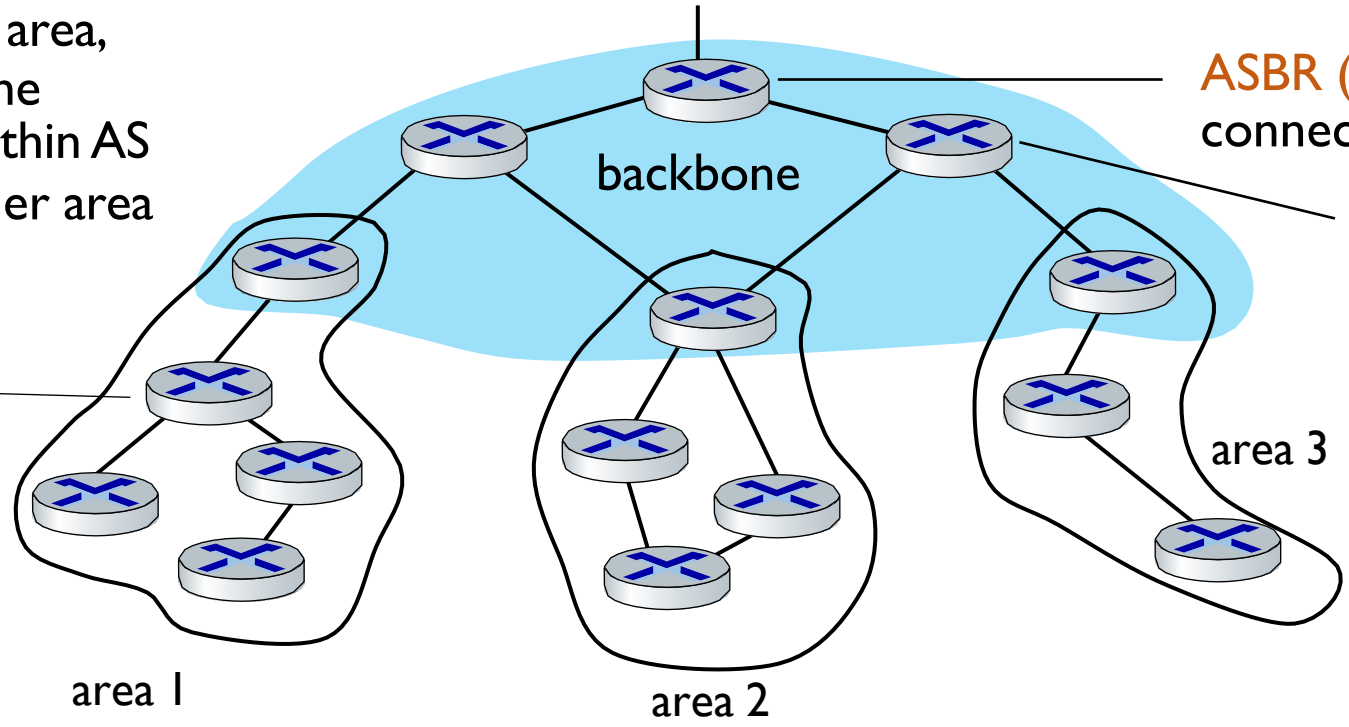
ASBR (AS Border Router): connects to other ASes

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

In-class Exercise!