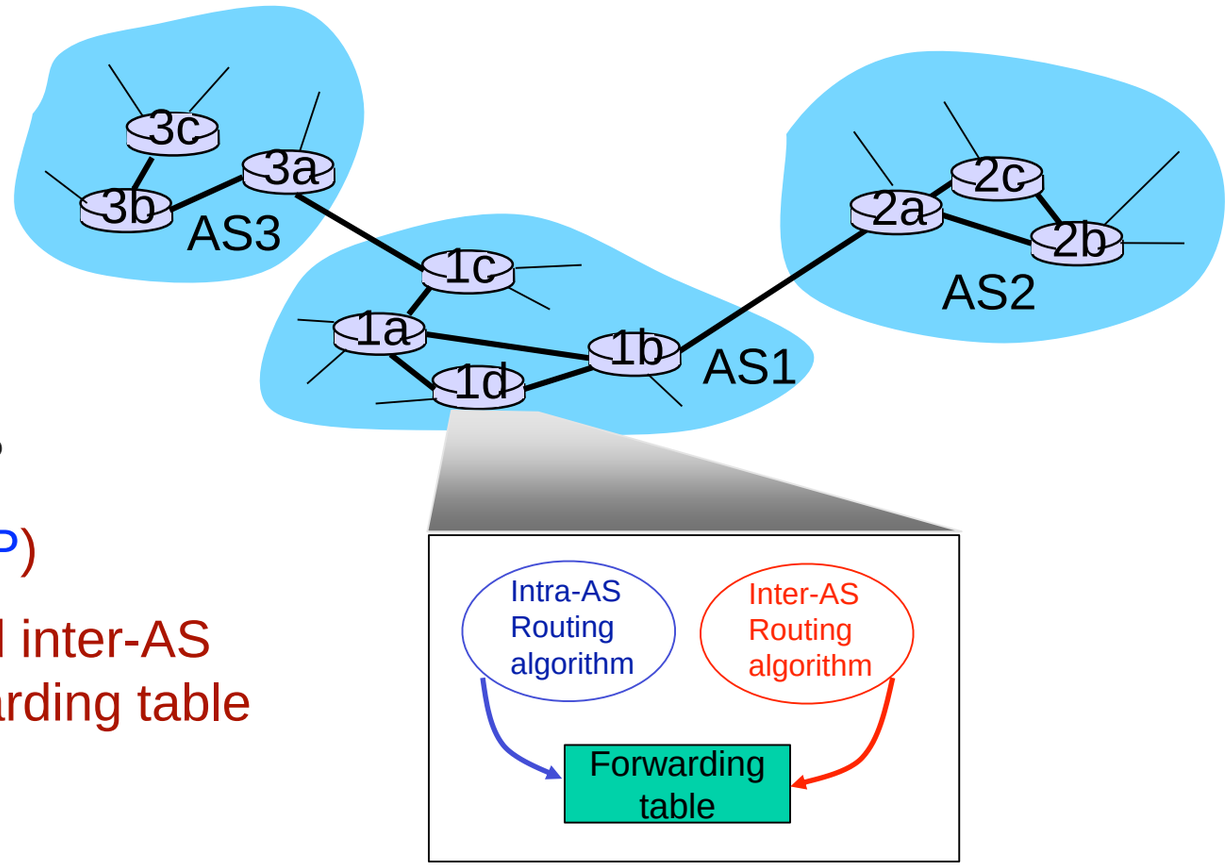
CS 330: Network Applications & Protocols

Network Layer

Galin Zhelezov
Department of Physical Sciences
York College of Pennsylvania

Hierarchical Routing (overview)

- **Routing protocols within a single network (an autonomous system) are determined by some network administrator**
  - Provide routing information for hosts *within the network* (i.e. Intra-AS routing)
  - For example, a company or ISP would use some routing protocol for their network
- **Routers within an AS also need a way to route packets to hosts on other networks**
  - Route packets to gateway router that will direct packets to the next network (AS)
    - All packets destined for another network are routed to the gateway router
  - What if there are multiple gateways? To which gateway should packets get routed?
    - Handled by inter-AS routing protocols (BGP)
    - Routers get information from both intra and inter-AS routing protocols and maintain info in forwarding table



Making routing scalable

Our routing study thus far - idealized

- all routers identical
- network "flat"

… *not* true in practice

scale: **with billions of destinations:**

- **can't store all destinations in routing tables!**
- **routing table exchange would swamp links!**

administrative autonomy

- **internet = network of networks**
- **each network admin may want to control routing in its own network**

Internet approach to scalable routing

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

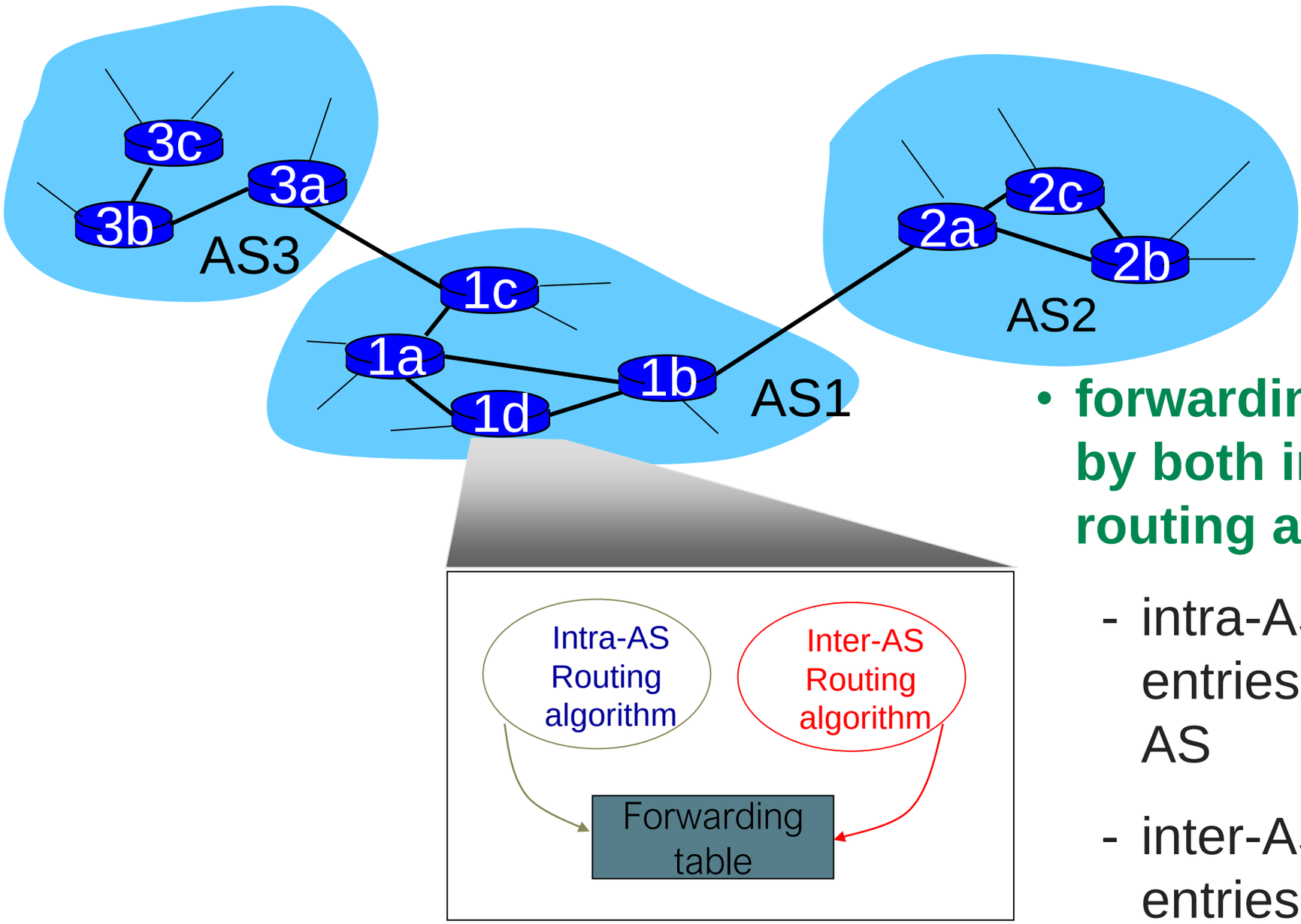
intra-AS routing

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

inter-AS routing

- **routing among AS'es**
- **gateways perform inter-domain routing (as well as intra-domain routing)**

Interconnected ASes



- **forwarding table configured by both intra- and inter-AS routing algorithm**
  - intra-AS routing determine entries for destinations within AS
  - inter-AS & intra-AS determine entries for external destinations

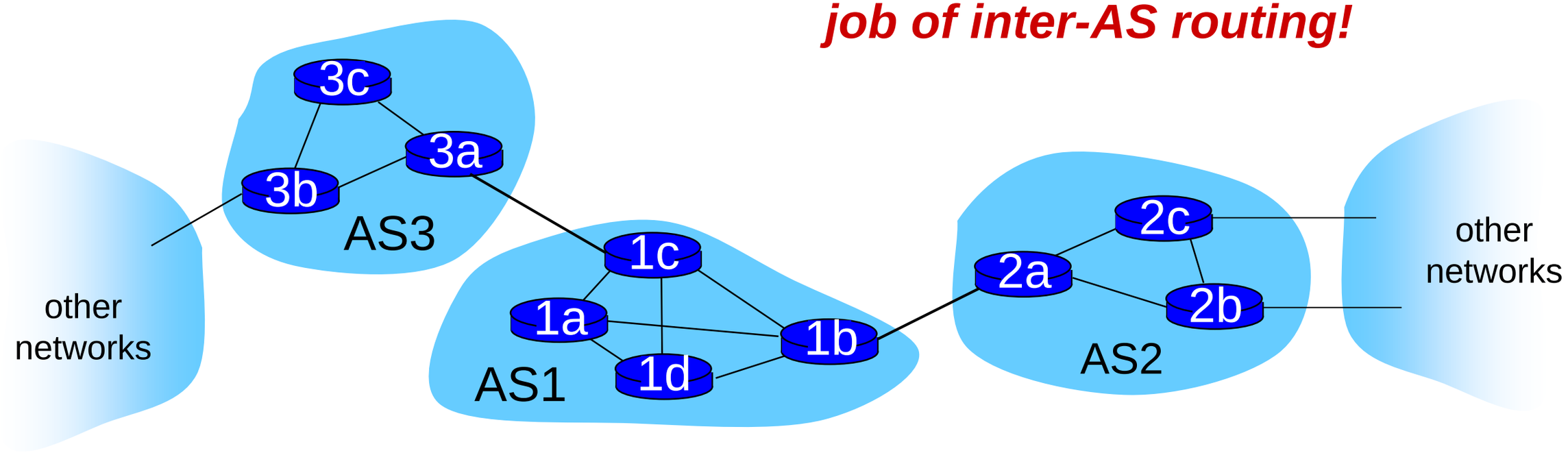
Inter-AS tasks

- **suppose router in AS1 receives datagram destined outside of AS1:**
  - router should forward packet to gateway router, but which one?

AS1 must:

1. **learn which dests are reachable through AS2, which through AS3**
2. **propagate this reachability info to all routers in AS1**

job of inter-AS routing!



Overview of Network Layer

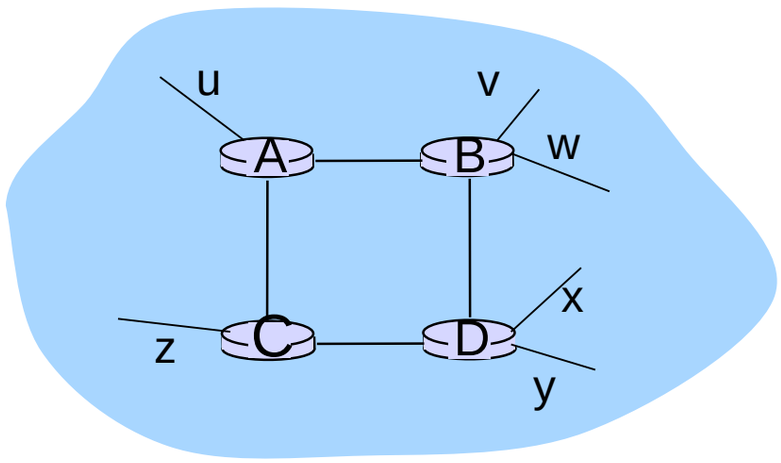
- **Virtual Circuit and Datagram Networks**
- **Router Architectures**
- **IP: Internet Protocol**
- **Routing algorithms**
- **Routing in the Internet**
  - RIP
  - OSPF
  - BGP
- **Broadcast and multicast routing**

Intra-AS Routing

- **also known as *interior gateway protocols (IGP)***
- **most common intra-AS routing protocols:**
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First (IS-IS protocol essentially same as OSPF)
  - IGRP: Interior Gateway Routing Protocol (Cisco proprietary for decades, until 2016)

RIP (Routing Information Protocol)

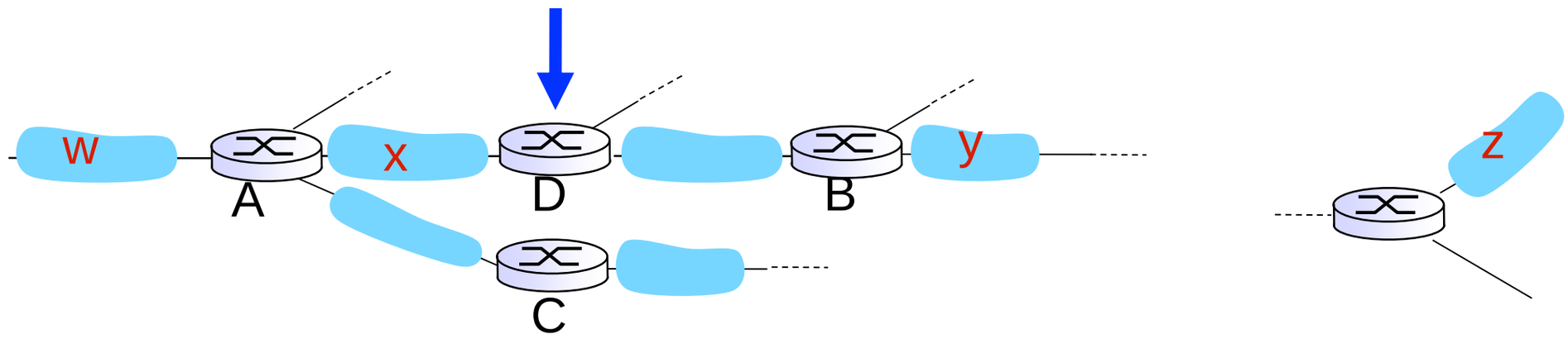
- **Included in BSD-UNIX distribution in 1982**
- **A distance-vector algorithm**
  - Distance metric is number of hops
    - Maximum hops is 15, each link costs 1 hop
  - DVs are exchanged with neighbors every 30 seconds in RIP Response Messages (aka RIP advertisements)



From router A to destination *subnets:*

Subnet	Hops
u	1
v	2
w	2
x	3
y	3
z	2

RIP Example



Routing Table in Router D

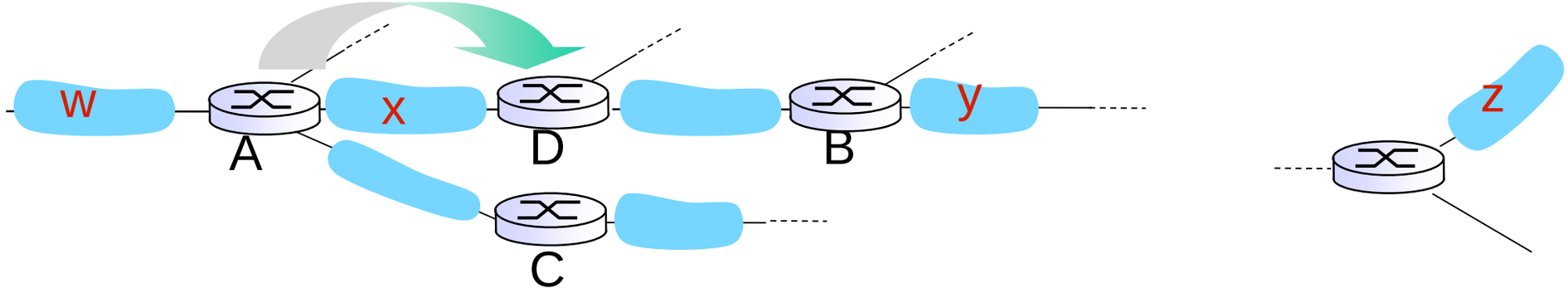
Destination Subnet	Next Router	# Hops to dest
w	A	2
y	B	2
z	B	7
x	--	1
..	..	..

Currently takes 7 hopes to get to subnet z through router B

RIP Example

A-to-D Advertisement

dest	next	hops
w	-	1
x	-	1
z	C	4
.	.	.



Routing Table in Router D

Destination Subnet	Next Router	# Hops to dest
w	A	2
y	B	2
z	B A	7 5
x	--	1
..	..	..

RIP: Link Failure & Recovery

- **If no advertisement is heard after 180 seconds**
  - Neighbor/link is declared dead
  - Routes via neighbor are invalidated
  - New advertisements sent to neighbors
  - Neighbors, in turn, send out new advertisements (if tables changed)
  - Link failure info quickly propagates to the entire network

Overview of Network Layer

- **Virtual Circuit and Datagram Networks**
- **Router Architectures**
- **IP: Internet Protocol**
- **Routing algorithms**
- **Routing in the Internet**
  - RIP
  - OSPF
  - BGP
- **Broadcast and multicast routing**

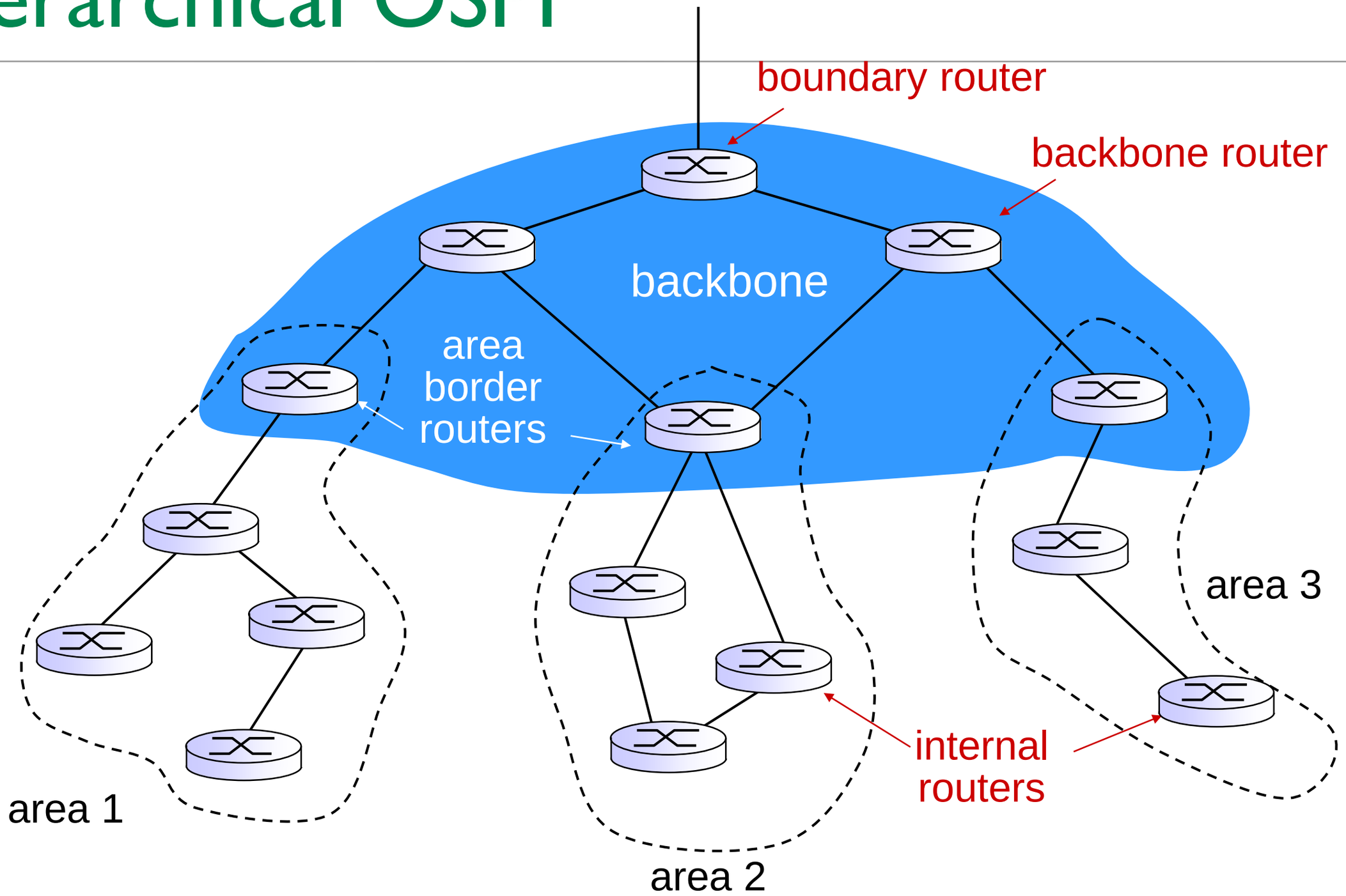
OSPF (Open Shortest Path First)

- **Uses link-state algorithm**
  - Network topology map stored at each node
  - Route computation using Dijkstra's algorithm
- **OSPF advertisement carries one entry per neighbor**
- **Advertisements flooded to entire AS (not just neighbors)**
  - Carried in OSPF messages directly over IP (rather than TCP or UDP)
  - IP header Protocol field is set to 89
- **Designed as a successor to RIP and therefore has some more advanced features**

OSPF Advanced Features

- **Security - all OSPF messages can be authenticated**
  - Prevent intruders from messing with router forwarding tables
- **Multiple same-cost paths allowed (only one path allowed in RIP)**
- **Supports multiple cost metrics per link for different TOS**
- **Integrated unicast and multicast routing support**
- **Multicast OSPF (MOSPF) uses same topology data base as OSPF**
- **Support for structuring large domains hierarchically**

Hierarchical OSPF



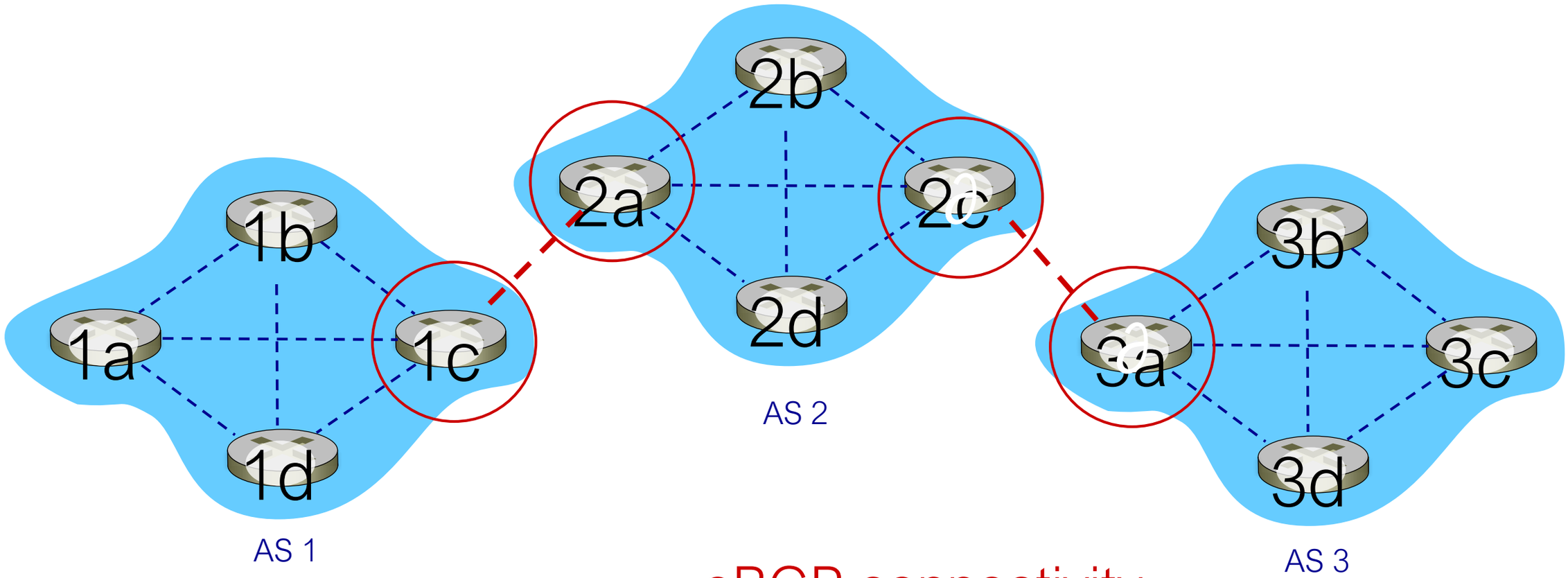
Overview of Network Layer

- **Virtual Circuit and Datagram Networks**
- **Router Architectures**
- **IP: Internet Protocol**
- **Routing algorithms**
- **Routing in the Internet**
  - RIP
  - OSPF
  - BGP
- **Broadcast and multicast routing**

Internet Inter-AS Routing: BGP

- **BGP (Border Gateway Protocol) is *the* de facto inter-domain routing protocol**
  - The “glue that holds the Internet together”
- **BGP provides each AS a means to:**
  - eBGP (external BGP) - obtain subnet reachability information from neighboring ASs
  - iBGP (internal BGP) - propagate reachability information to all AS-internal routers
  - Determine “good” routes to other networks based on reachability information and policy
- **Allows subnet to advertise its existence to rest of Internet**

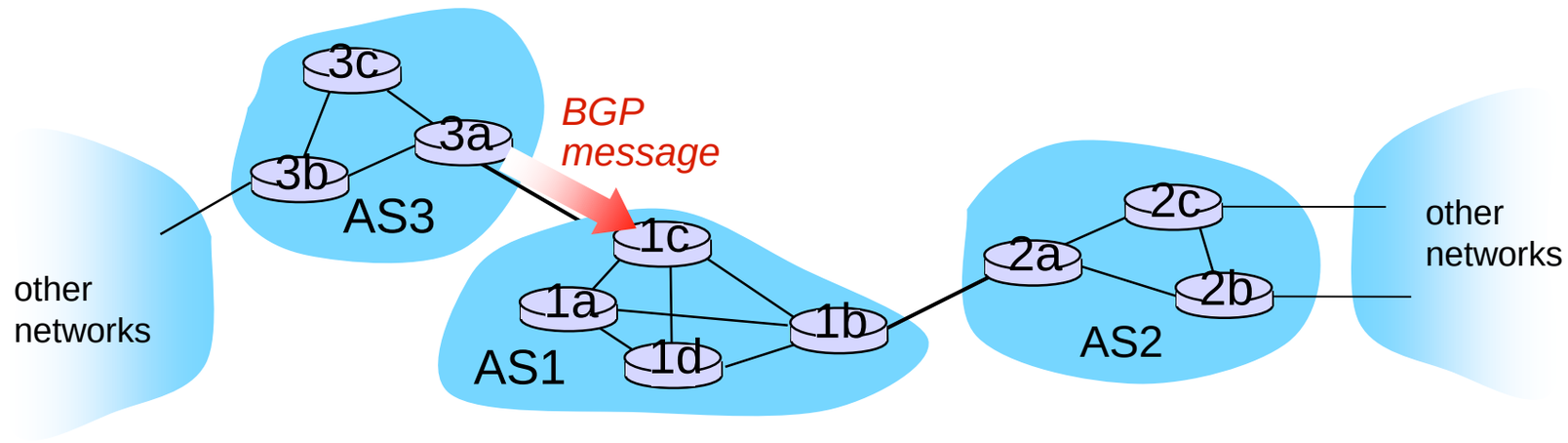
eBGP, iBGP connections



BGP Basics

- **BGP session - two BGP routers ("peers") exchange BGP messages**
  - Advertising paths to different destination network prefixes ("path vector" protocol)
  - Exchanged over semi-permanent TCP connections

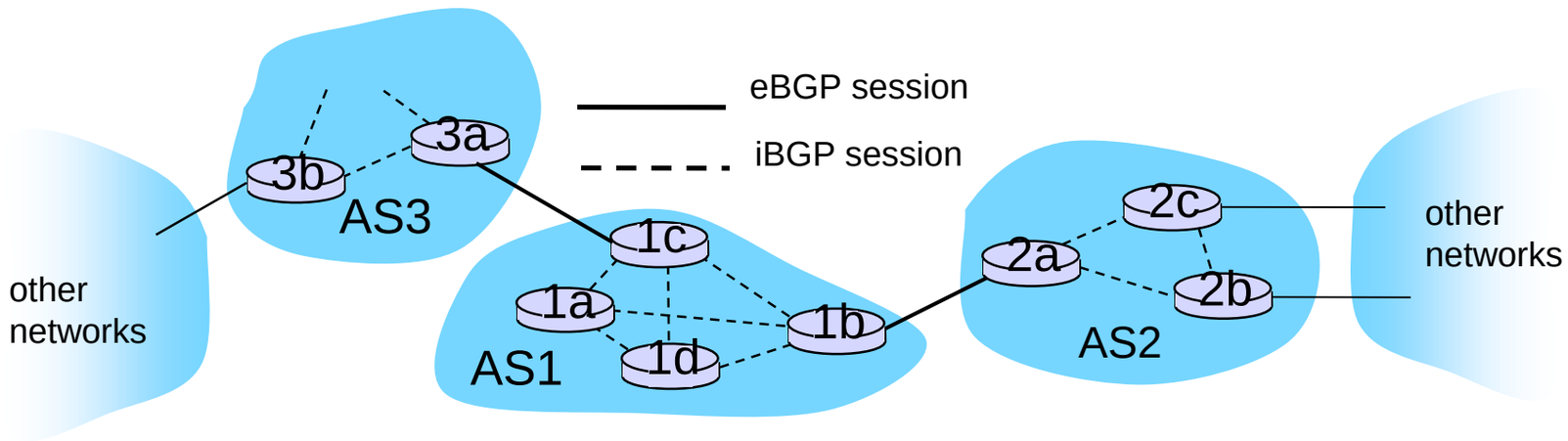
- **When AS3 advertises a prefix to AS1**
  - AS3 promises it will forward datagrams towards that prefix
  - AS3 can aggregate prefixes in its advertisement



BGP Basics: Distributing Path Information

- **Using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.**
  - 1c can then use iBGP do distribute new prefix info to all routers in AS1
  - 1b can then re-advertise new reachability info to AS2 over 1b-to-2a eBGP session

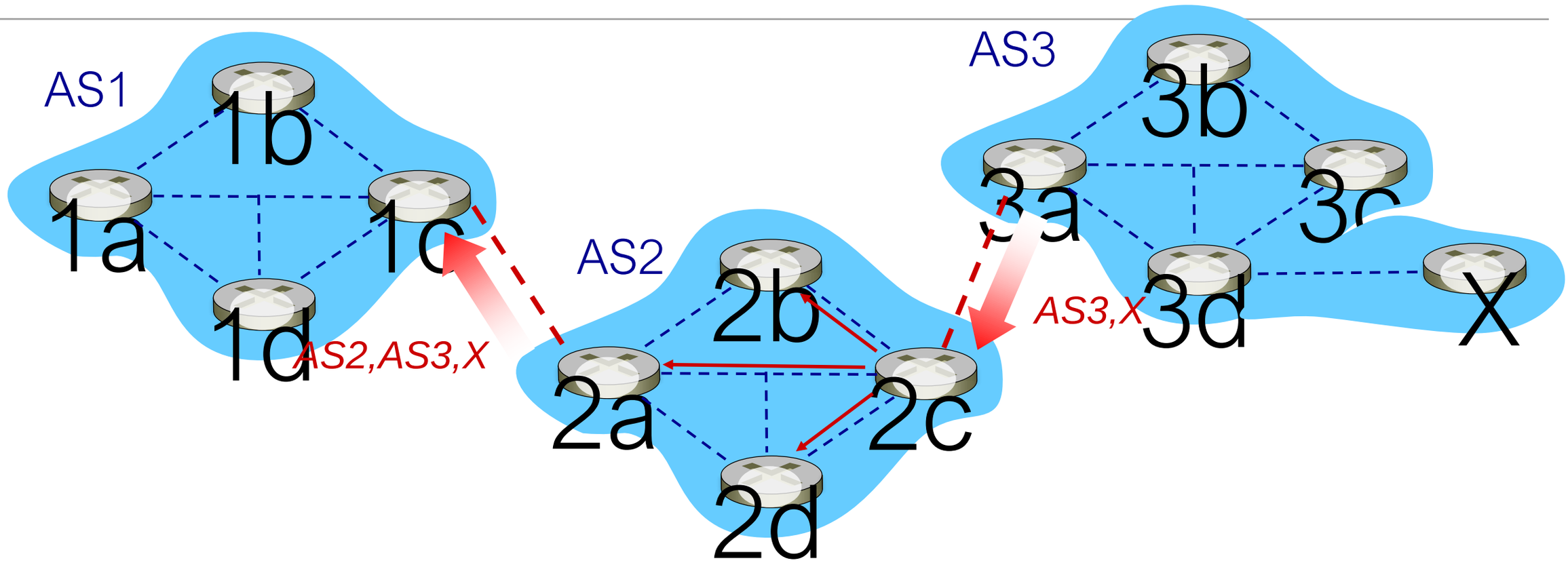
- **When router learns of new prefix, it creates entry for prefix in its forwarding table**



Path attributes and BGP routes

- **advertised prefix includes BGP attributes**
  - prefix + attributes = "route"
- **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-based routing:***
  - gateway receiving route advertisement uses *import policy* to accept/decline path (e.g., never route through AS Y).
  - AS policy also determines whether to *advertise* path to other other neighboring ASes

BGP path advertisement

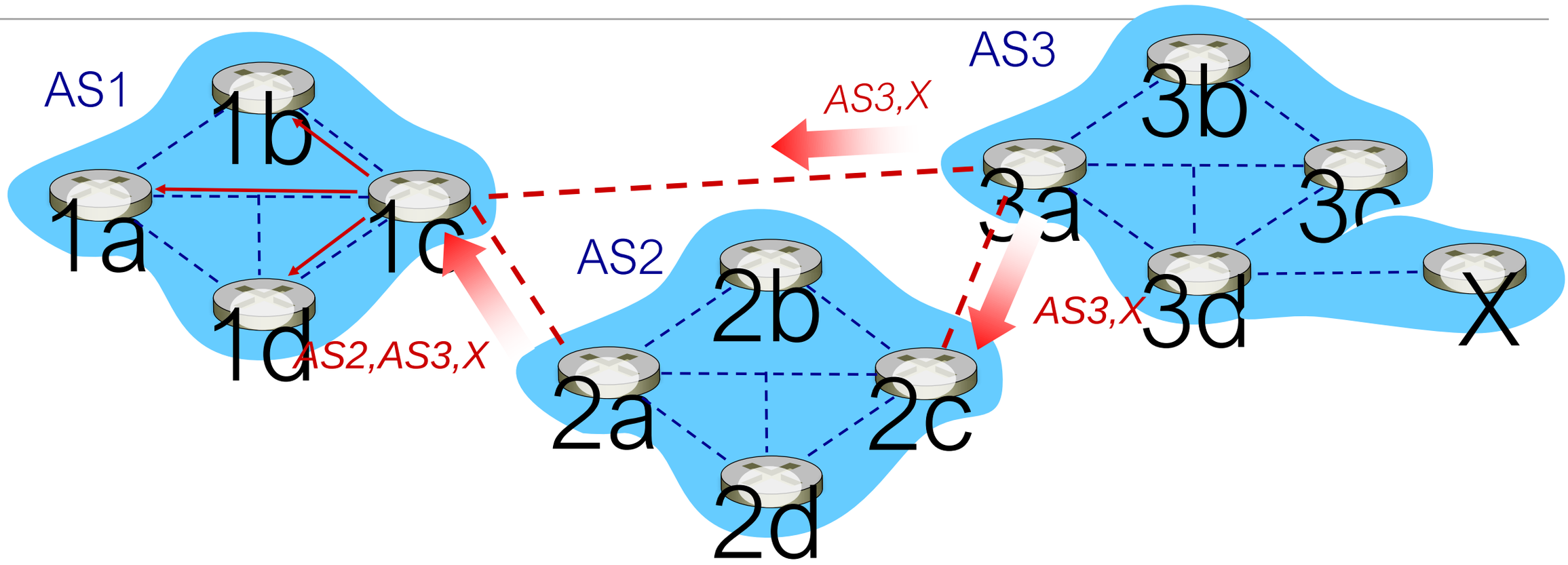


- AS2 router 2c receives path advertisement AS3,X (via eBGP) from AS3 router 3a

- **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



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 messages

- **BGP messages exchanged between peers over TCP connection**
- **BGP messages:**
  - OPEN: opens TCP connection to remote BGP peer and authenticates sending BGP peer
  - UPDATE: advertises new path (or withdraws old)
  - KEEPALIVE: keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - NOTIFICATION: reports errors in previous msg; also used to close connection

Why Different Intra-AS, Inter-AS routing?

- **Policy**
  - Inter-AS - admin wants control over how its traffic is routed, who routes through its networks
  - Intra-AS - single admin, so no policy decisions needed
- **Scale**
  - Hierarchical routing saves table size, reduced update traffic
- **Performance**
  - Intra-AS - can focus on performance
  - Inter-AS - policy may dominate over performance