COMP 3234B
Computer and Communication Networks

2nd semester 2023-2024
Network Layer (IV)

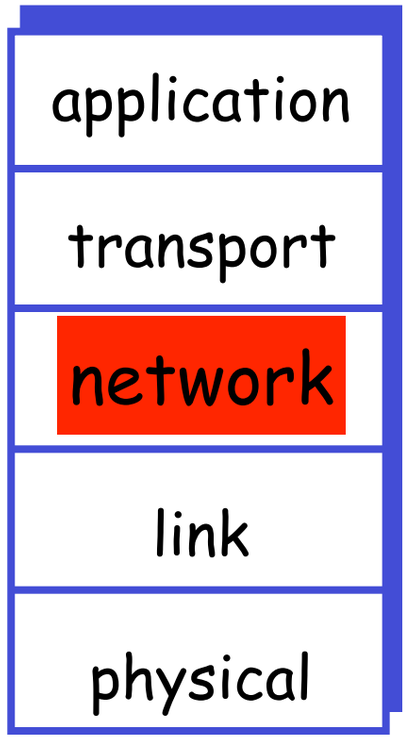
Prof. C Wu

Department of Computer Science
The University of Hong Kong

Roadmap

Network layer

- Principles behind network-layer services (ILO1)
  - forwarding vs. routing
  - network service models
- Router (ILO1)
- IP (ILO2,5)
  - DHCP
  - NAT
- ICMP (ILO2)
- Routing algorithms (ILO3)
- Routing in the Internet (ILO2,3)



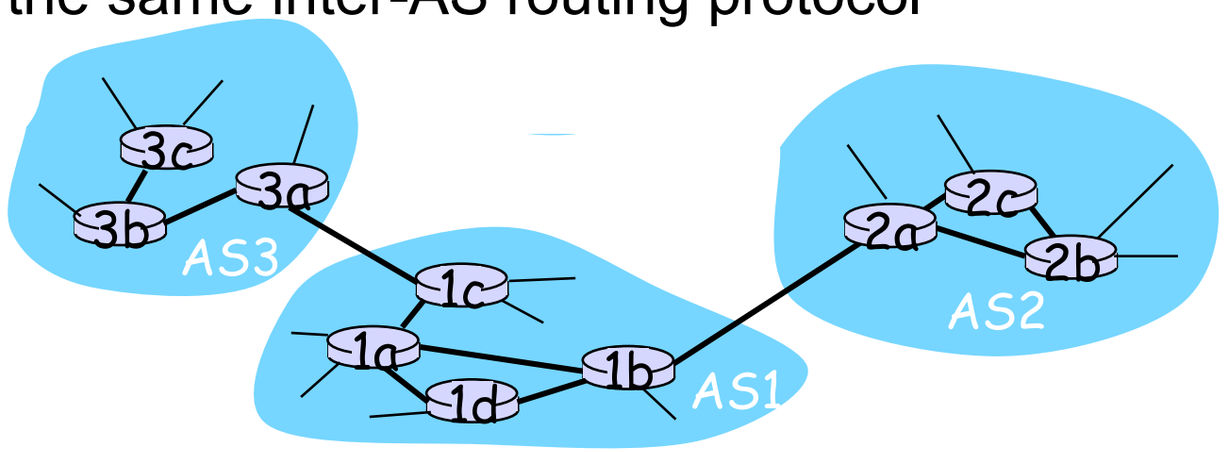
Routing in the Internet

- Hierarchy in the Internet
  - routers are aggregated into regions (or domains) — "autonomous systems" (*AS*)

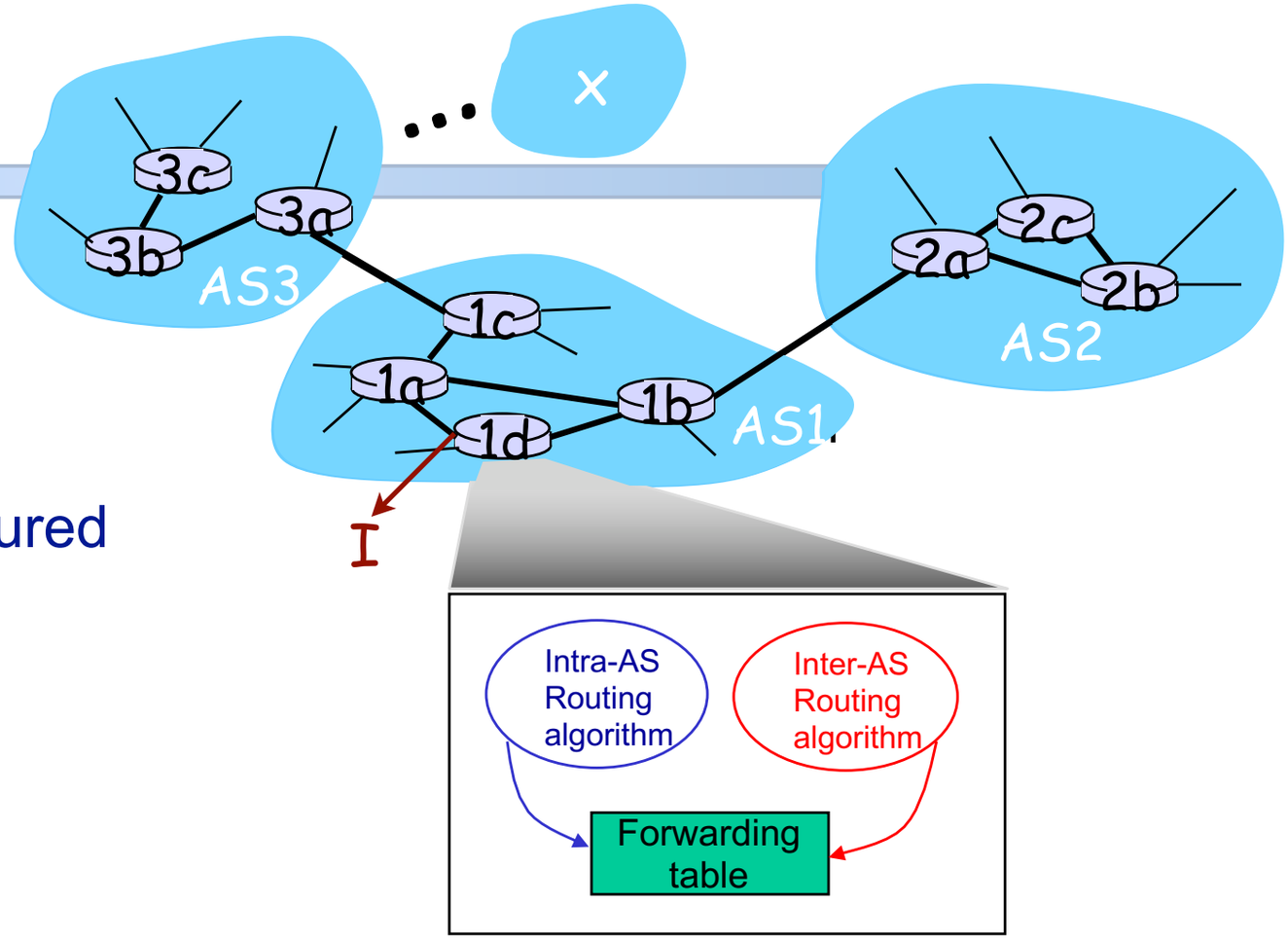
    each ISP's network consists of one or more ASes
  - each AS controls routing within its own network — administrative autonomy
  - gateway router: router with direct link(s) to router(s) in other ASes
- Hierarchical routing
  - *intra-AS routing*: routers in the same AS run the same routing protocol to decide paths in the AS

    routers in different ASes can run different intra-AS routing protocols
  - *inter-AS routing*: all ASes run to obtain reachability information for destinations in different ASes; gateways perform inter-domain routing (as well as intra-domain routing)

    all ASes run the same inter-AS routing protocol



Hierarchical routing



- Forwarding table at each router configured by both intra- and inter-AS routing protocols
  - intra-AS sets entries for internal destinations
  - inter-AS & intra-AS set entries for external destinations

Example: setting forwarding table entry in router 1d for external destinations in subnet x

- inter-AS protocol obtains that subnet x reachable via AS3 (gateway router 1c) for AS1
- inter-AS protocol propagates reachability info to all internal routers in AS1
- router 1d determines with intra-AS routing protocol that its interface I is on the least-cost path to 1c.
- router 1d installs forwarding table entry (x,I)

Internet routing protocols

- Intra-AS routing protocols (also known as Interior Gateway Protocols)
  - RIP: Routing Information Protocol (distance vector algorithm, largely obsoleted)
  - OSPF: Open Shortest Path First (link state algorithm)
  - Others:

    (E)IGRP: (Enhanced) Interior Gateway Routing Protocol (Cisco proprietary for decades before standardization in 2016, distance vector protocol)

    IS-IS: Intermediate System to Intermediate System (closely related to OSPF)
- Inter-AS routing protocol
  - BGP

OSPF

- Open Shortest Path First
  - widely used in Internet

    “open”: protocol specification publicly available
  - Link State algorithm

    each router has complete topology map of the AS

    runs Dijkstra’s algorithm to decide shortest-path tree to all subnets

    link cost configured by network administrator
  - OSPF advertisement broadcast to all routers in the AS

    when one neighboring link’s state changes (e.g., cost)

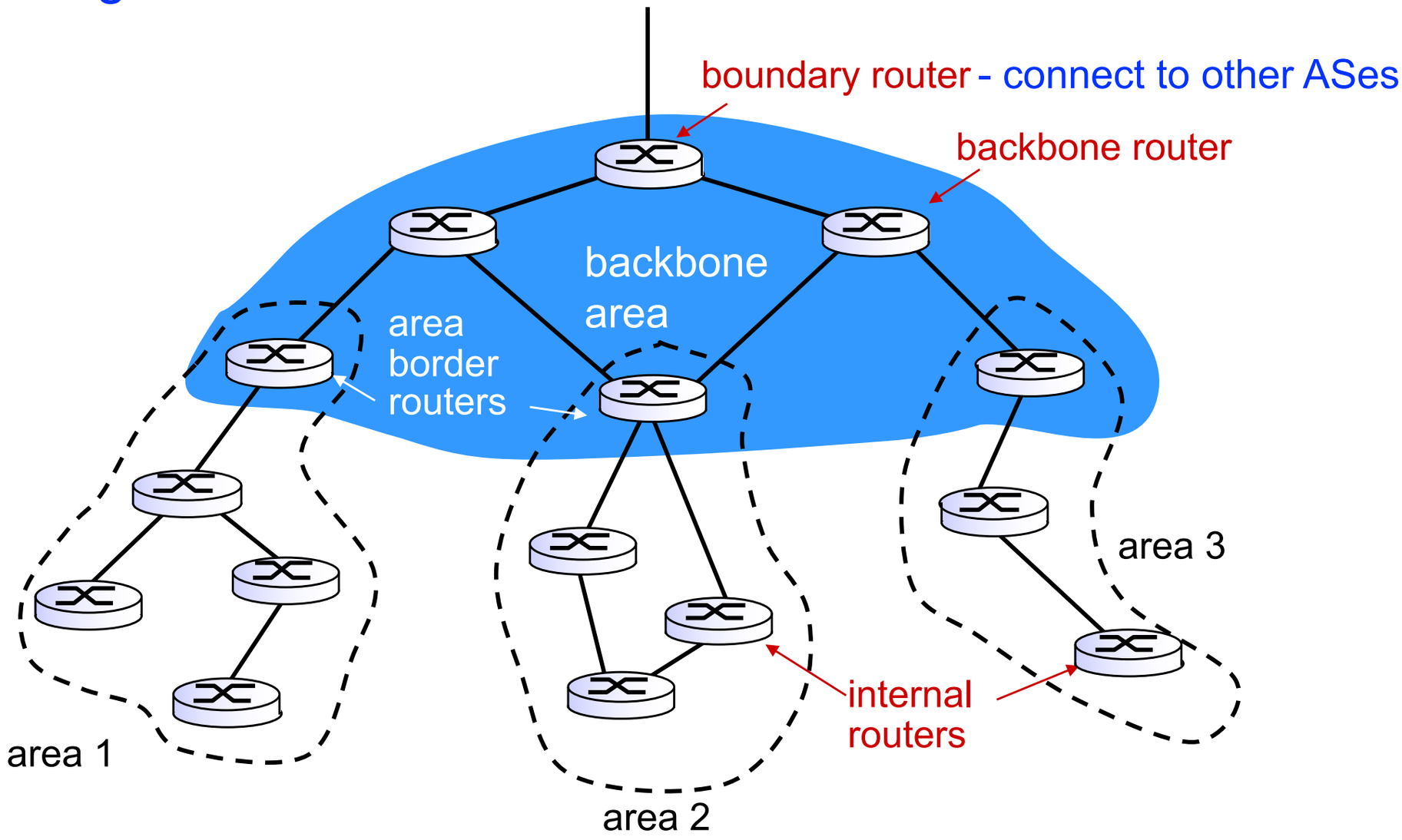
    or periodically even if no changes (at least once every 30 minutes)

    carried in OSPF messages over IP

OSPF (cont'd)

- Hierarchical OSPF in a large AS
  - two-level hierarchy: local areas, backbone area
  - link-state advertisements only broadcast within an area

    each nodes has detailed area topology
  - area border routers: forward datagrams to other areas' border routers through the backbone area



BGP

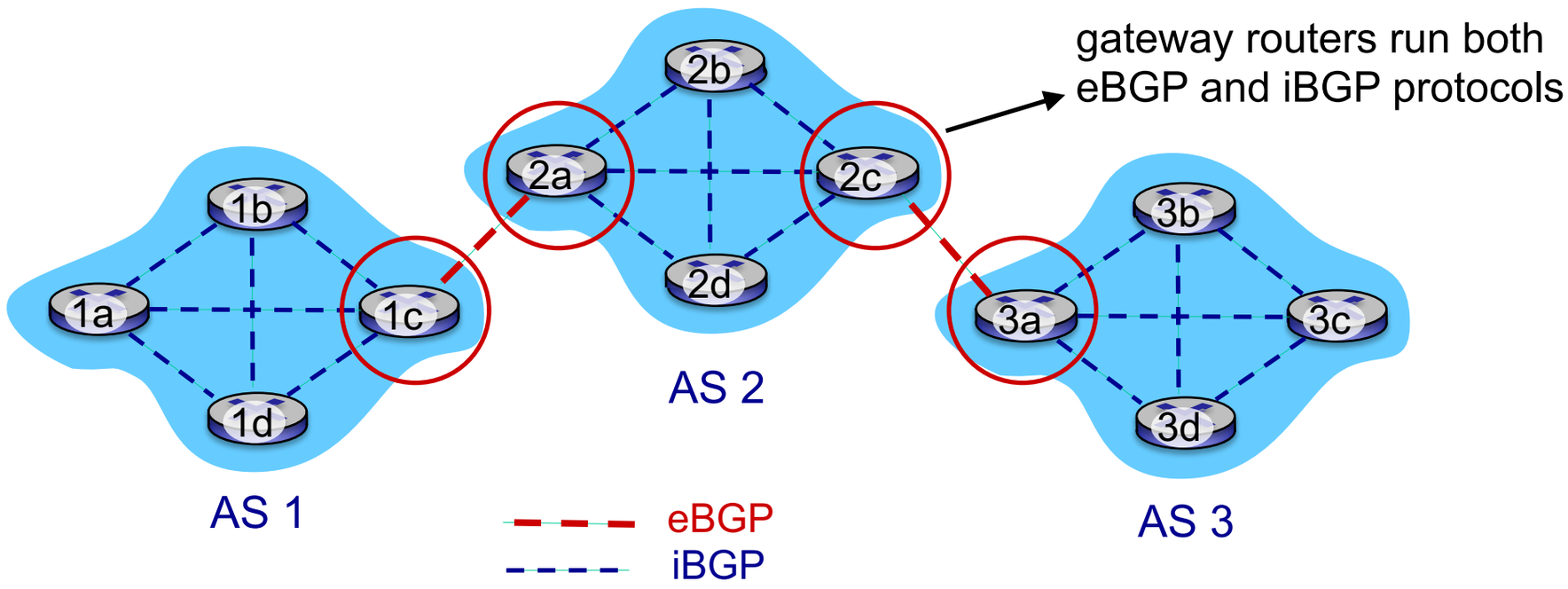
- Border Gateway Protocol
  - BGP is the de facto inter-AS routing protocol in today's Internet
  - BGP provides each AS a means to:

    (eBGP session) obtain subnet reachability information from neighboring ASes

    (iBGP session) propagate reachability information to all AS-internal routers

    determine "good" routes to subnets based on reachability information and AS policy

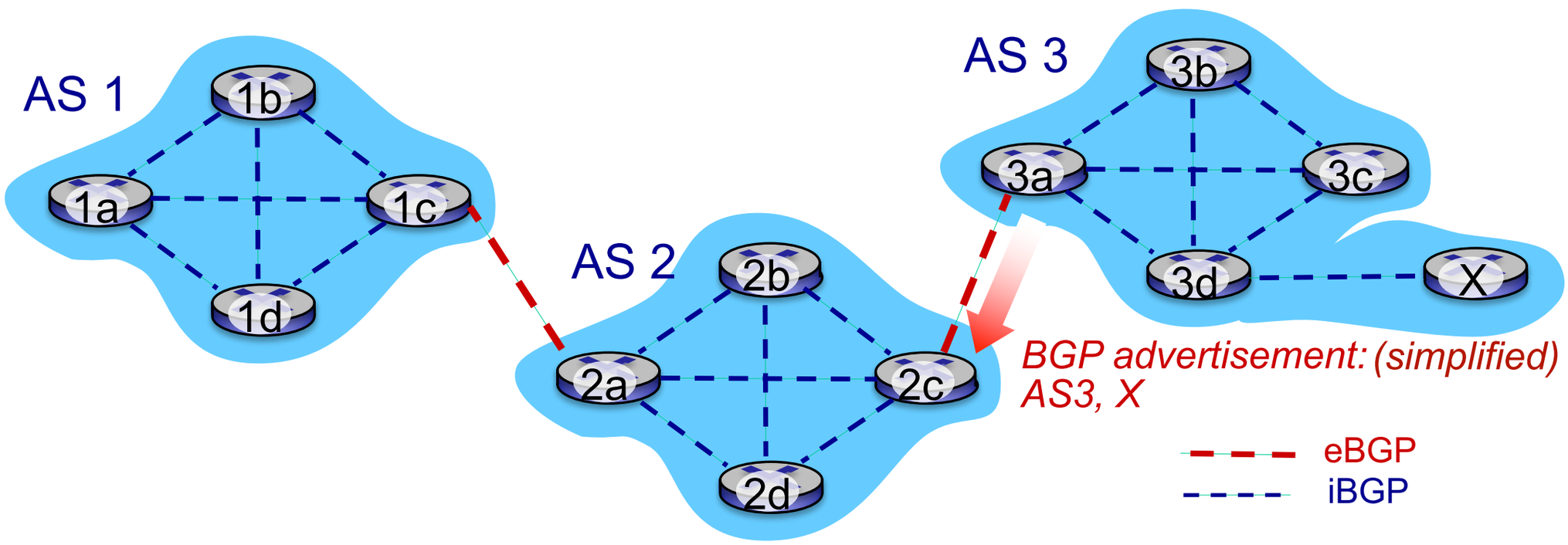
    allow subnet to advertise its existence to the rest of Internet



BGP session

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

For example, when AS3’s gateway router 3a advertises path AS3,X to AS2’s gateway router 2c, AS3 promises to AS2 that it will forward datagrams towards X

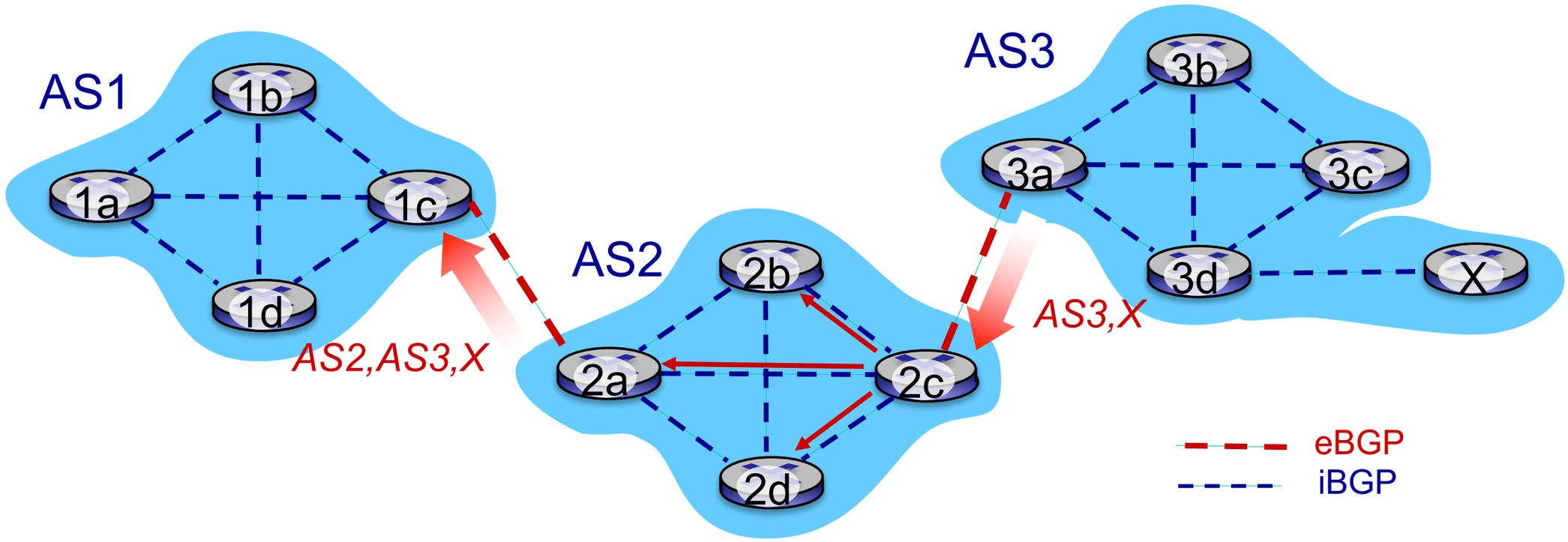


BGP path advertisement

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

Based on AS2's policy, AS2's router 2c accepts path AS3,X, propagates it (via iBGP) to all AS2's routers

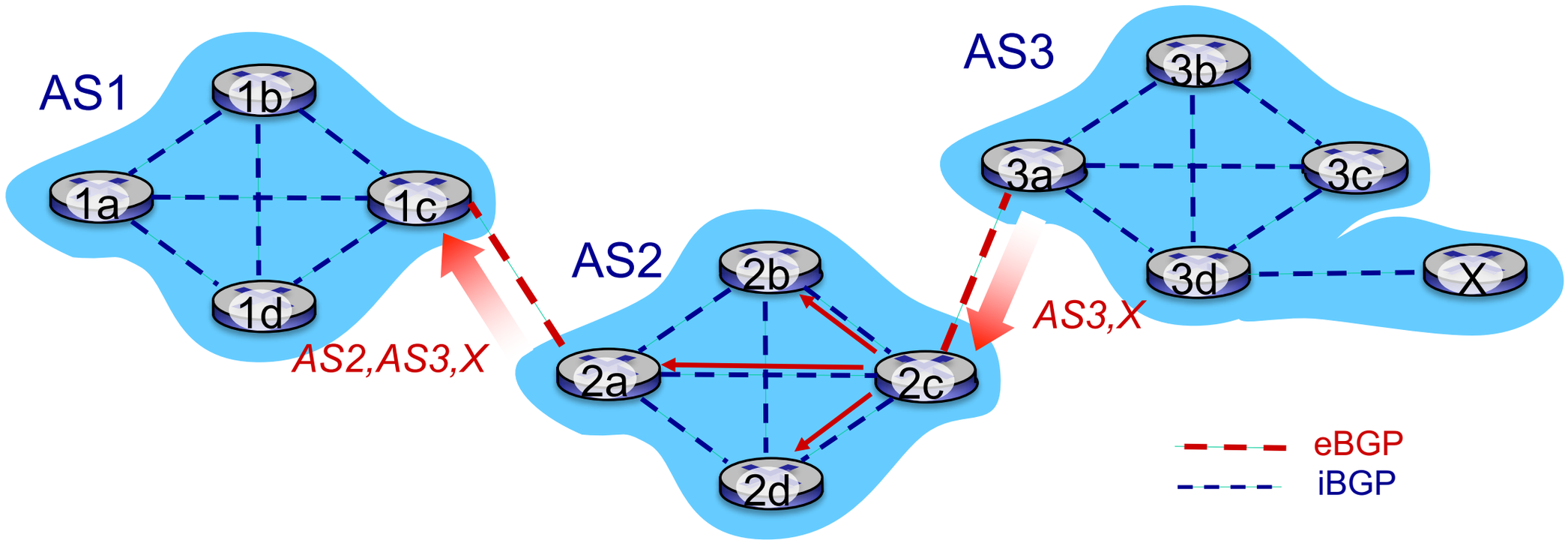
Based on AS2's policy, AS2's router 2a advertises (via eBGP) path AS2, AS3, X to AS1's router 1c



BGP path attributes

- BGP advertises destination network prefix (*x*) with BGP attributes, e.g.:
  - AS-PATH: list of ASes through which the advertisement has passed
  - NEXT-HOP: IP address of the interface of the router leading into next-hop AS
  - example BGP *route* that each router in AS1 knows to reach prefix x:

    IP address of leftmost interface of router 2a; AS2 AS3; x



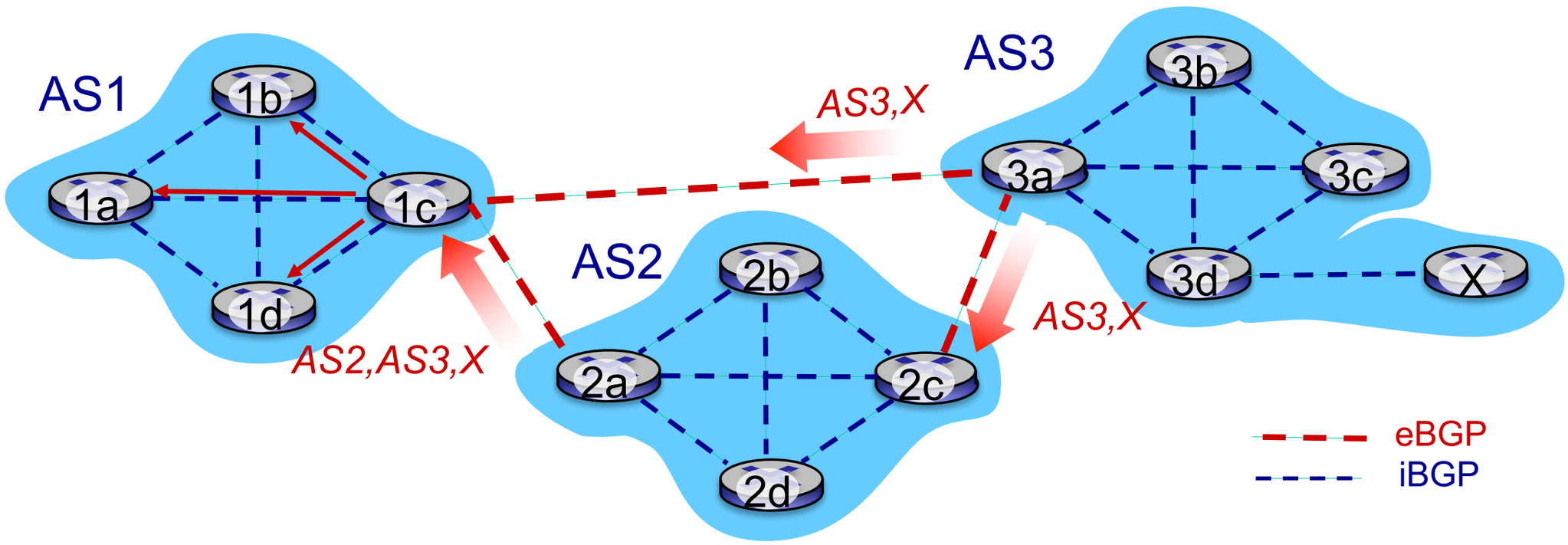
BGP path advertisement (cont'd)

Gateway router may learn about multiple paths to destination prefix x:

— AS1's gateway router 1c learns path AS2,AS3,X from 2a

— AS1's gateway router 1c learns path AS3,X from 3a

— Based on AS1's policy, AS1's gateway router 1c chooses path AS3,X, and advertises this path within AS1 via iBGP

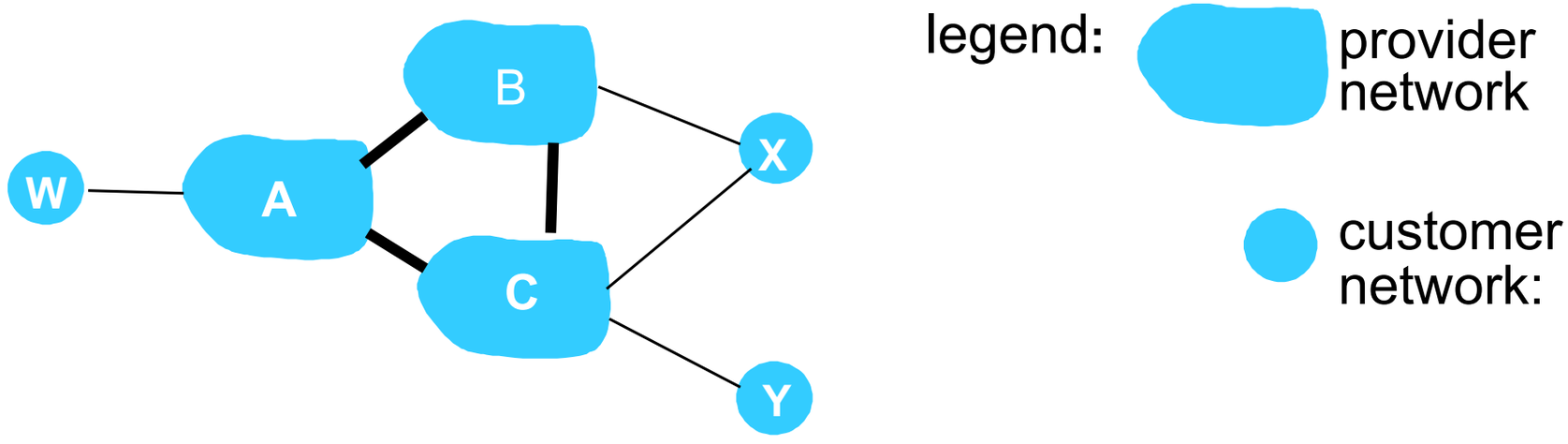


Policy-based routing:

gateway router receiving route advertisements uses import policy to accept/decline path (e.g., never route through AS Y);

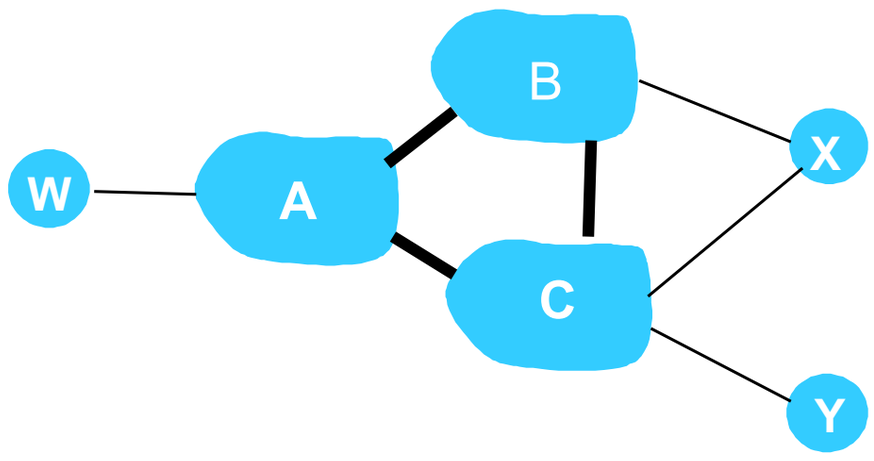
AS policy also determines whether to advertise path to other neighboring ASes

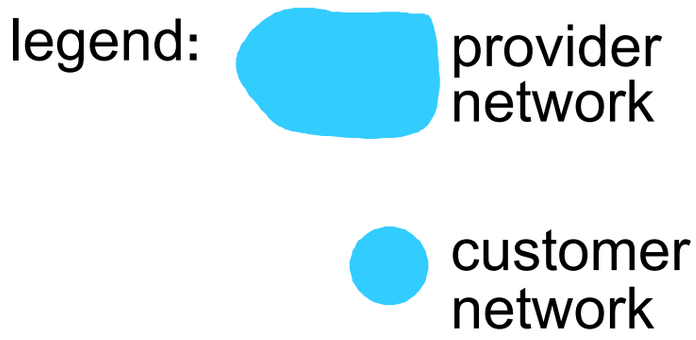
Achieving policy via advertisements



- A,B,C are provider networks (e.g., global transit ISPs' ASes)
- X, Y, W are customer networks of provider networks (e.g., access ISPs' ASes who pay the global transit ISPs for routing traffic to/from themselves)
  - X is dual-homed: attached to two provider networks

Achieving policy via advertisements (cont'd)



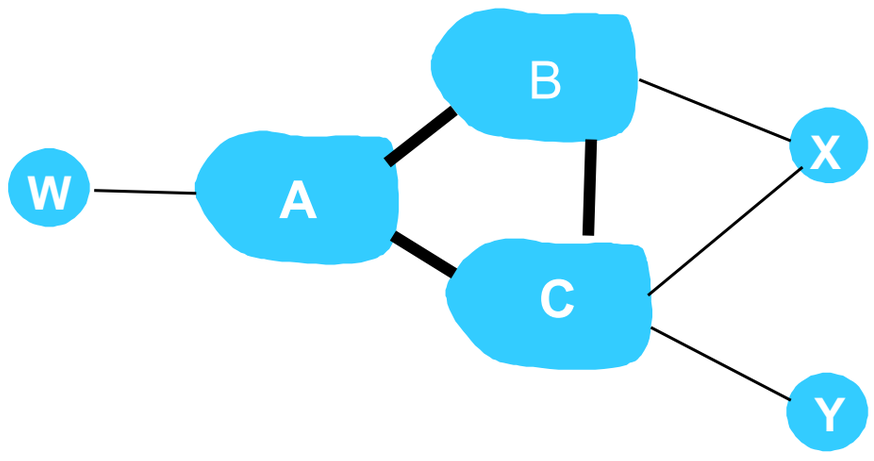


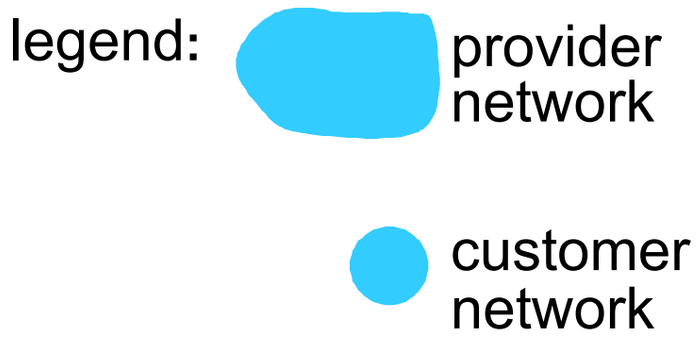
Suppose a global transit ISP only wants to route traffic to/from its customer networks and does not want to carry transit traffic between other ISPs

- A advertises path A,W to B and C
- B chooses not to advertise B,A,W to C:

  B gets no "revenue" for routing C,B,A,W, since none of C, A, W are B's customers
- C does not learn about C,B,A,W path
- C will route C,A,W (not using B) to get to W

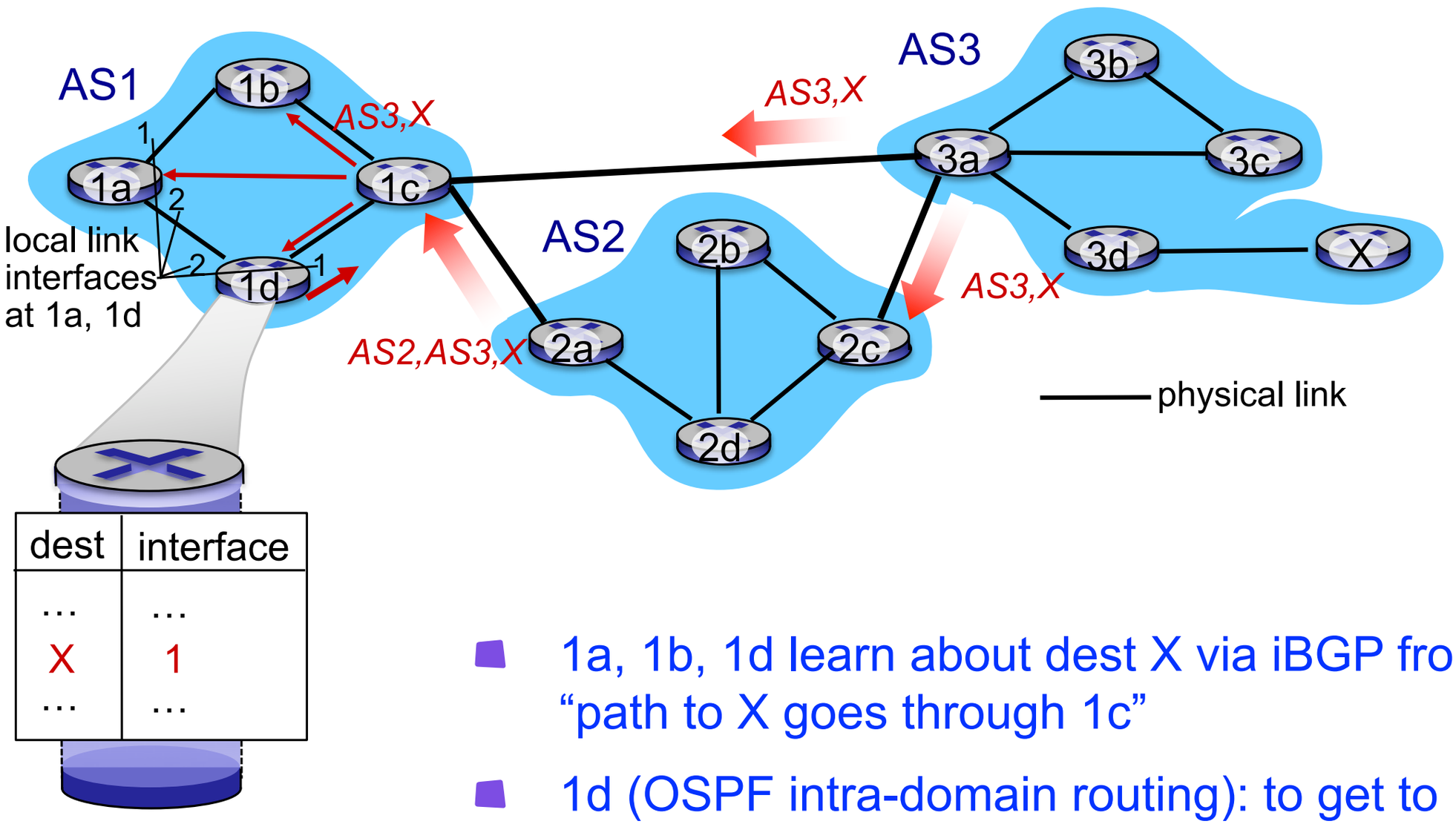
Achieving policy via advertisements (cont'd)





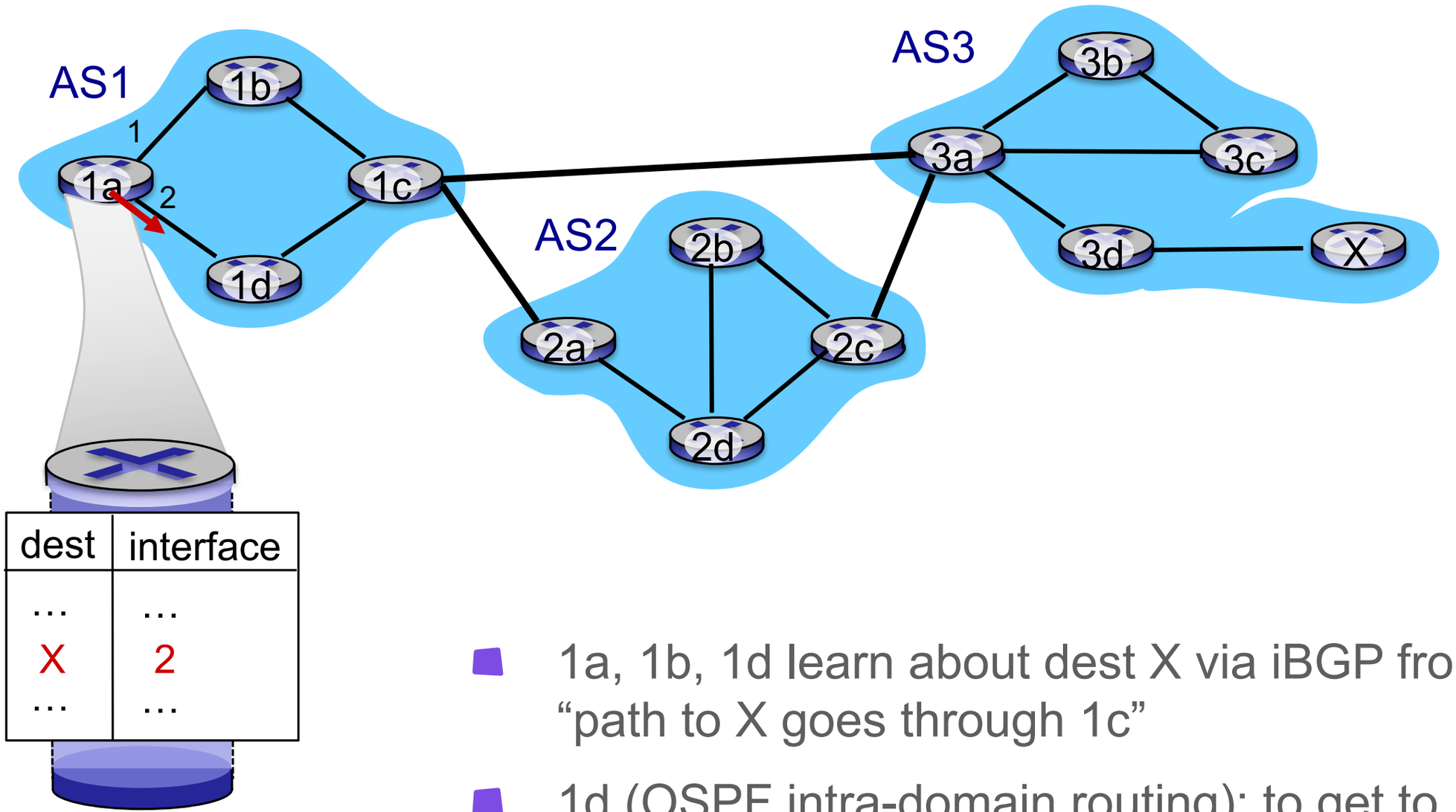
If X does not want to route B,X,C, then X will not advertise to B a route to C

Forwarding table entries



- 1a, 1b, 1d learn about dest X via iBGP from 1c: “path to X goes through 1c”
- 1d (OSPF intra-domain routing): to get to 1c, forward over outgoing local interface 1

Forwarding table entries (cont'd)



- 1a, 1b, 1d learn about dest X via iBGP from 1c: "path to X goes through 1c"
- 1d (OSPF intra-domain routing): to get to 1c, forward over outgoing local interface 1
- 1a (OSPF intra-domain routing): to get to 1c, forward over outgoing local interface 2

- Required reading:
  - Computer Networking: A Top-Down Approach (8th Edition)
    Ch 5.3, 5.4

- Acknowledgement:
  - Some materials are extracted from the slides created by Prof. Jim F. Kurose and Prof. Keith W. Ross for the textbook.