Computer Networks

Inter-AS Routing
Border Gateway Protocol (BGP)

Amitangshu Pal

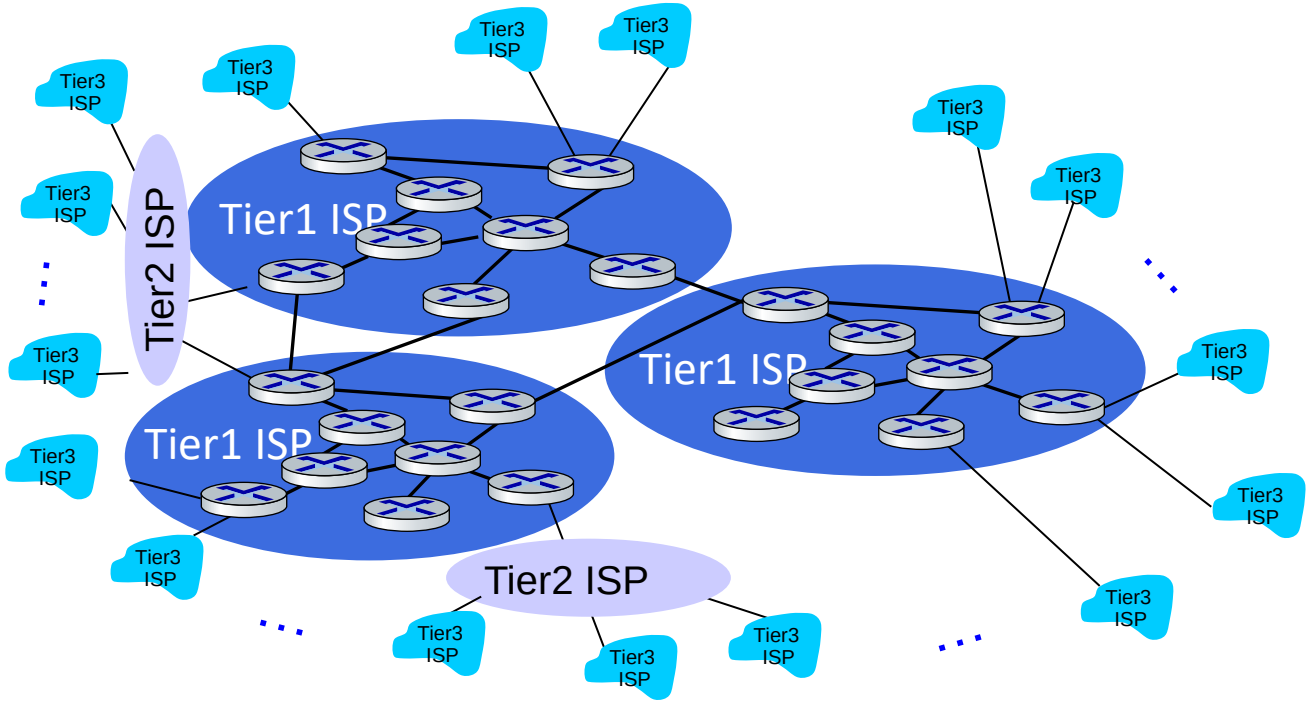
Computer Science and Engineering

IIT Kanpur

Inter-AS and Intra-AS Routing

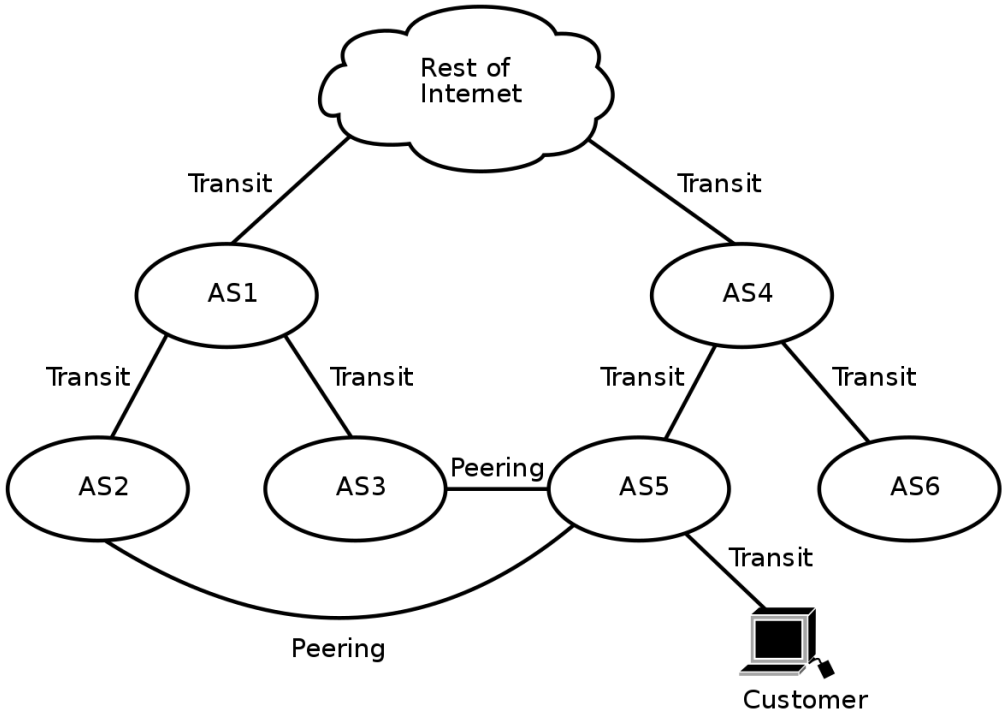
- AS: controlled by a single administrative entity, i.e. an ISP

Src:https://commons.wikimedia.org/wiki/File:India_and_Neighbouring_Countries_Map_(official_borders).png



Inter-AS Routing: BGP

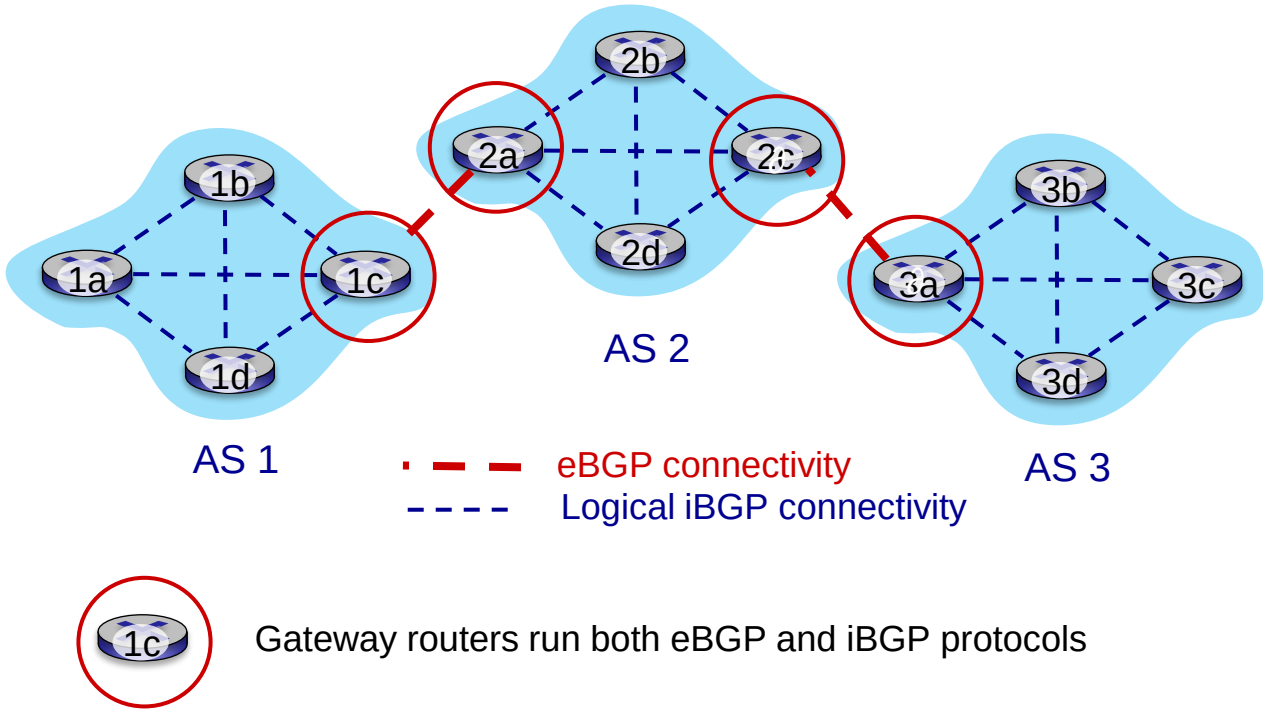
- BGP (Border Gateway Protocol): the de facto inter-domain routing protocol
  - "Glue that holds the Internet together"
- Allows subnet to advertise its existence, and the destinations it can reach, to rest of Internet
- Relations among the ASes depend on multiple factors → competitions, politics, economy etc.
  - Transmit: One AS pays another one for access to the Internet
  - Peer: Two ASes exchange traffic between their users freely, and for mutual benefit



Src: https://fr.wikipedia.org/wiki/Peering#/media/Fichier:AS-interconnection.svg

Inter-AS Routing: BGP

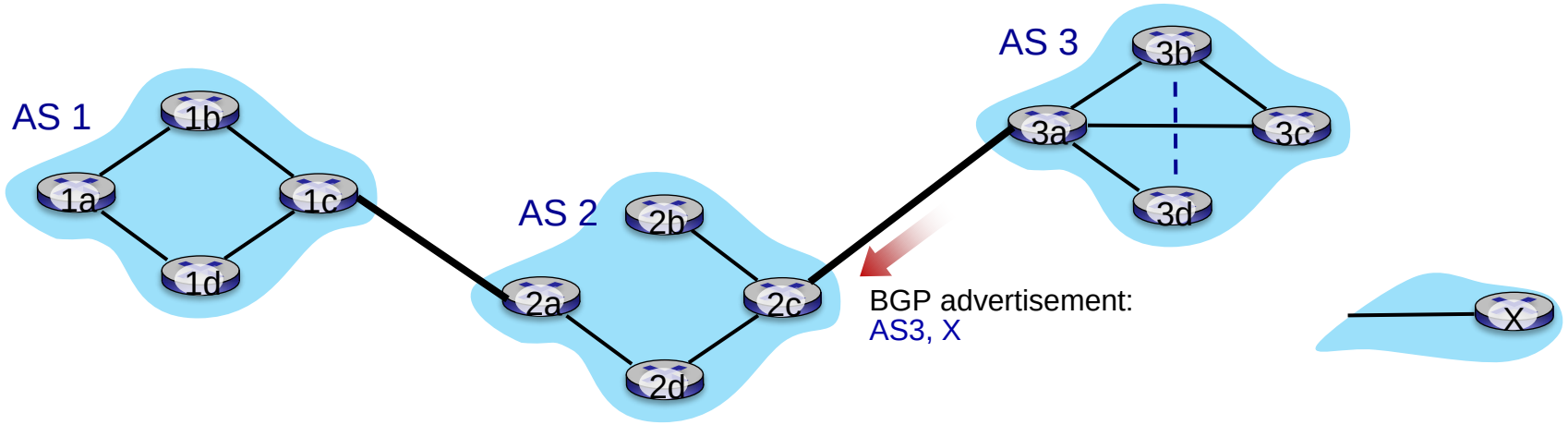
- BGP provides each AS a means to:
  - eBGP: obtain subnet reachability information from neighboring ASes
  - iBGP: propagate reachability information to all AS-internal routers.
  - Determine “good” routes to other networks based on reachability information and policy



BGP and Policy Based Path Advertisement

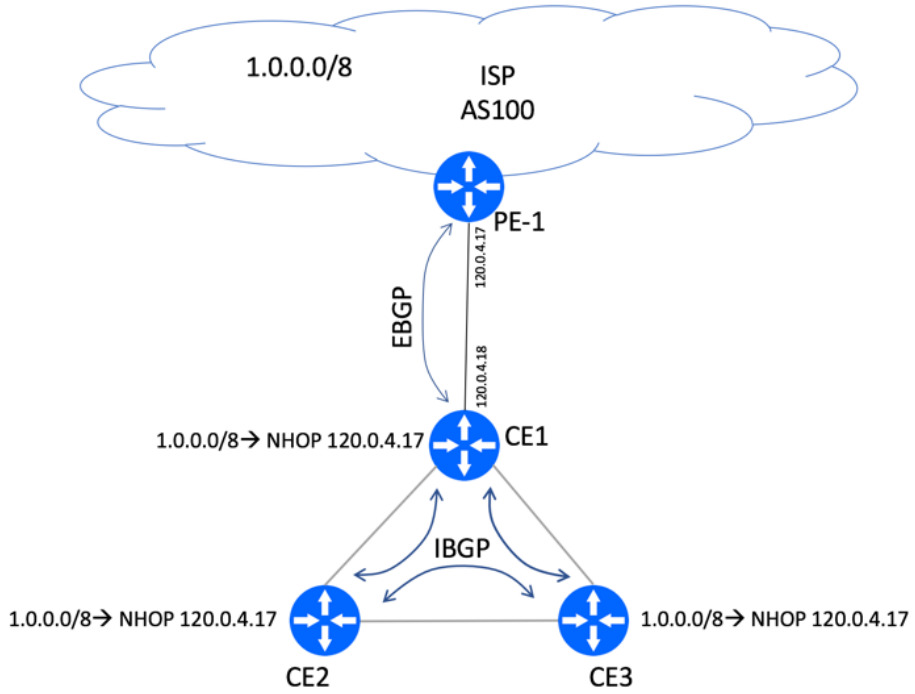
BGP Path Advertisement

- BGP session: two BGP routers (“peers”) exchange BGP messages over semi-permanent TCP connection (port 179):
  - Advertising paths to different destination network prefixes (BGP is a “path vector” protocol)
  - BGP sessions do not correspond to physical links
- 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

- Route advertisement contain an IP prefix + attributes
  - Prefix: destination being advertised
  - Two important attributes:
    - AS-PATH: list of ASes through which prefix advertisement has passed (AS50, AS76)
    - NEXT-HOP: indicates the router interface that begins the AS-PATH
    - The intra-AS routing protocol will determine the least-cost paths to all subnets attached to the routers in the AS, including the subnet for the link between CE1 and PE-1



Src: http://bgphelp.com/2017/03/05/bgp-next-hop-self-explained/

Key Points About BGP

1. What BGP Does:

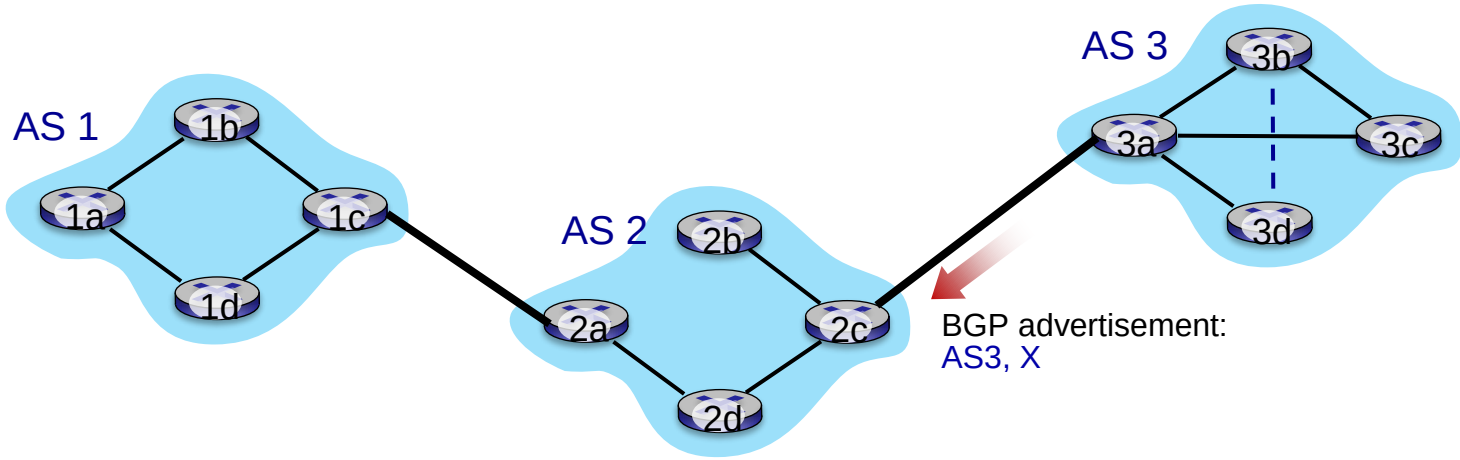
   Advertising Routes: Think of each network or Autonomous System (AS) as a city that tells other cities which roads (routes) it has. BGP helps these networks share this information so they can find the best paths for data to travel.
   Choosing Paths: When a network knows different ways (routes) to send data to the same place, BGP helps it choose the best one based on things like distance, speed, or traffic conditions.

2. Types of Relationships in BGP:

   Transit: This is like one city paying a larger city to use its roads. In internet terms, a smaller network might pay a larger network to connect to the rest of the internet.
   Peering: This is like two cities agreeing to let each other's cars use their roads for free. Two networks exchange data directly without having to pay each other because it benefits both.

BGP Path Advertisement

- Policy-based advertisement:
  - Import policy: Gateway receiving route advertisement uses import policy to accept/decline path (e.g., never route through AS Y).
  - Export policy: AS policy also determines whether to advertise path to other other neighboring ASes



Let's simplify BGP path advertisement and its policy-based controls. This revolves around how networks choose which routes to use and share with others. Think of it like deciding which friends to tell about a secret path or deciding whether to accept advice on a shortcut from someone.
BGP Path Advertisement Explained

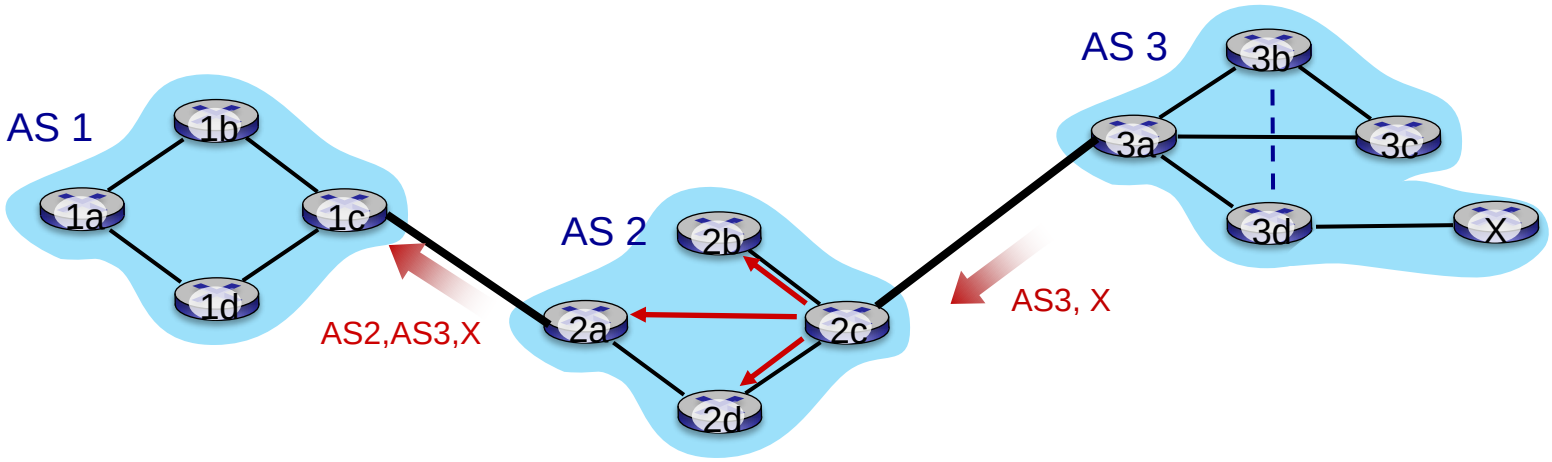
BGP (Border Gateway Protocol) uses policies to control which routes are advertised to other networks and which routes are accepted from them. These policies are set by network administrators based on various business, security, or technical considerations.
Two Main Types of Policies

Import Policy:
What It Is: This is a set of rules a network uses to decide whether to accept or reject the route information coming from another network.

Example: Suppose there's a rule like "never route through AS Y." This means if a route advertisement comes in and involves going through AS Y, the network will reject this route based on its import policy. This might be because AS Y is considered unreliable, too slow, or too expensive.

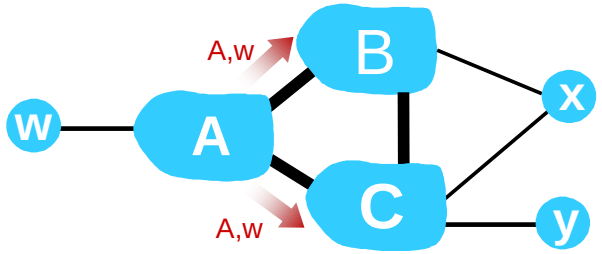
Export Policy:
What It Is: This determines whether a network will share its own route information with neighboring networks.
Example: A network might decide not to advertise certain routes to specific neighbors if it wants to manage its traffic flow better, keep certain routes less congested, or for reasons related to business agreements or costs.

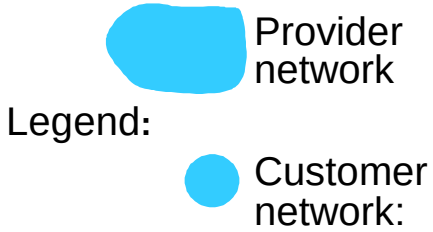
BGP Path Advertisement



- 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
  → import policy
- Based on AS2 policy, AS2 router 2a advertises (via eBGP) path AS2, AS3, X to AS1 router 1c
  → export policy

BGP: Achieving Policy via Advertisements



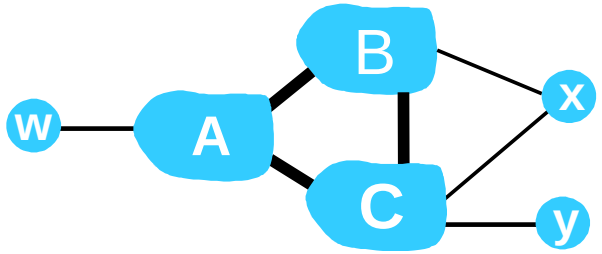


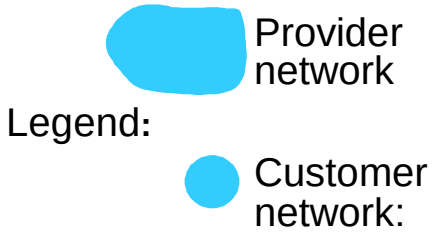
- A,B,C are provider networks
- x,w,y are customer (of provider networks)
- x is dual-homed: attached to two networks

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





- A,B,C are provider networks
- x,w,y are customer (of provider networks)
- x is dual-homed: attached to two networks

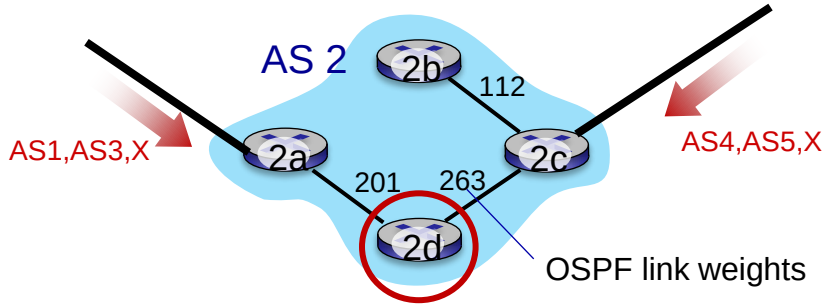
ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

- 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 and Policy Based Routing

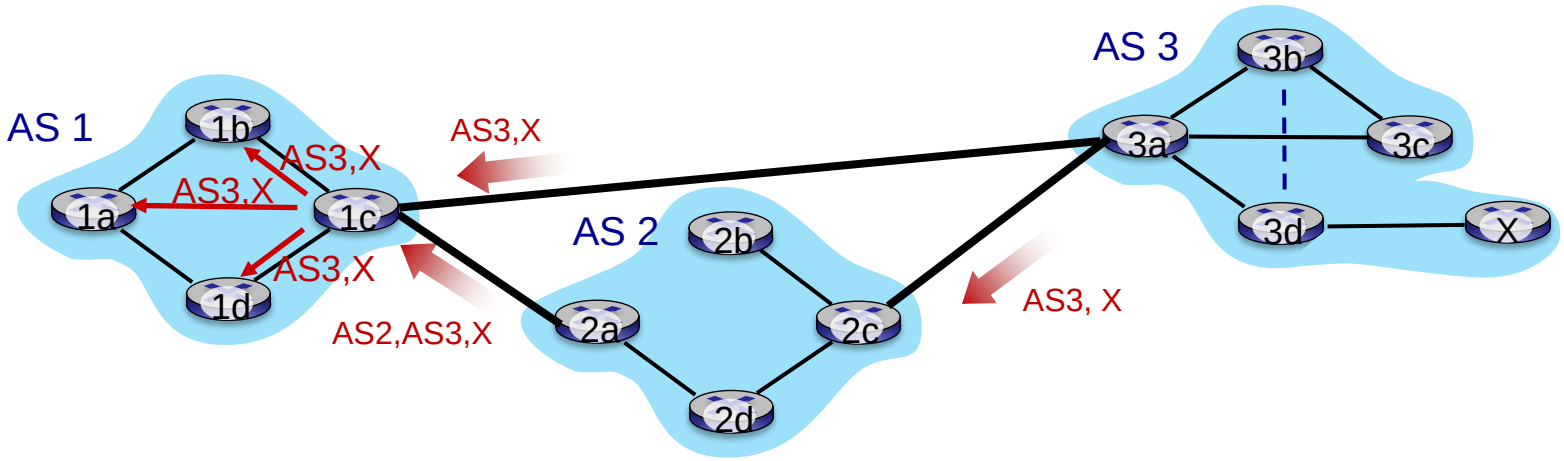
BGP Route Selection

- Router may learn about more than one route to destination AS, selects route based on:
  1. Local preference value attribute: policy decision
     - Set by the import policy
  2. Shortest AS-PATH
     - Included in the route advertisement
  3. Closest NEXT-HOP router: Hot potato routing
     - Based on intra-domain routing protocol (i.e. OSPF)
  4. Additional criteria (like smallest NEXT-HOP router id etc.)



Local Preference Value Attribute (Policy Decision): This attribute represents the preference for routes received from different neighbors. It's a policy decision made by the network administrator. Routes with higher local preference values are preferred. This value is set by the import policy configured on the router.
Shortest AS-PATH: The AS-PATH attribute contains the sequence of AS numbers through which the route has traversed. BGP prefers routes with the shortest AS-PATH, as shorter paths are usually more efficient and have fewer potential points of failure.
Closest NEXT-HOP Router (Hot Potato Routing): BGP routers select the closest next-hop router based on the IGP (Interior Gateway Protocol) metric within their own AS. This is known as hot potato routing, where routers try to hand off traffic to the next-hop router as quickly as possible to minimize transit delays within their own network. Usually, the closest next-hop router is determined by an intra-domain routing protocol such as OSPF (Open Shortest Path First) or IS-IS (Intermediate System to Intermediate System).
Additional Criteria: There may be other criteria used for route selection, such as the smallest NEXT-HOP router ID. This can be a tie-breaker when multiple routes have the same values for the above criteria.

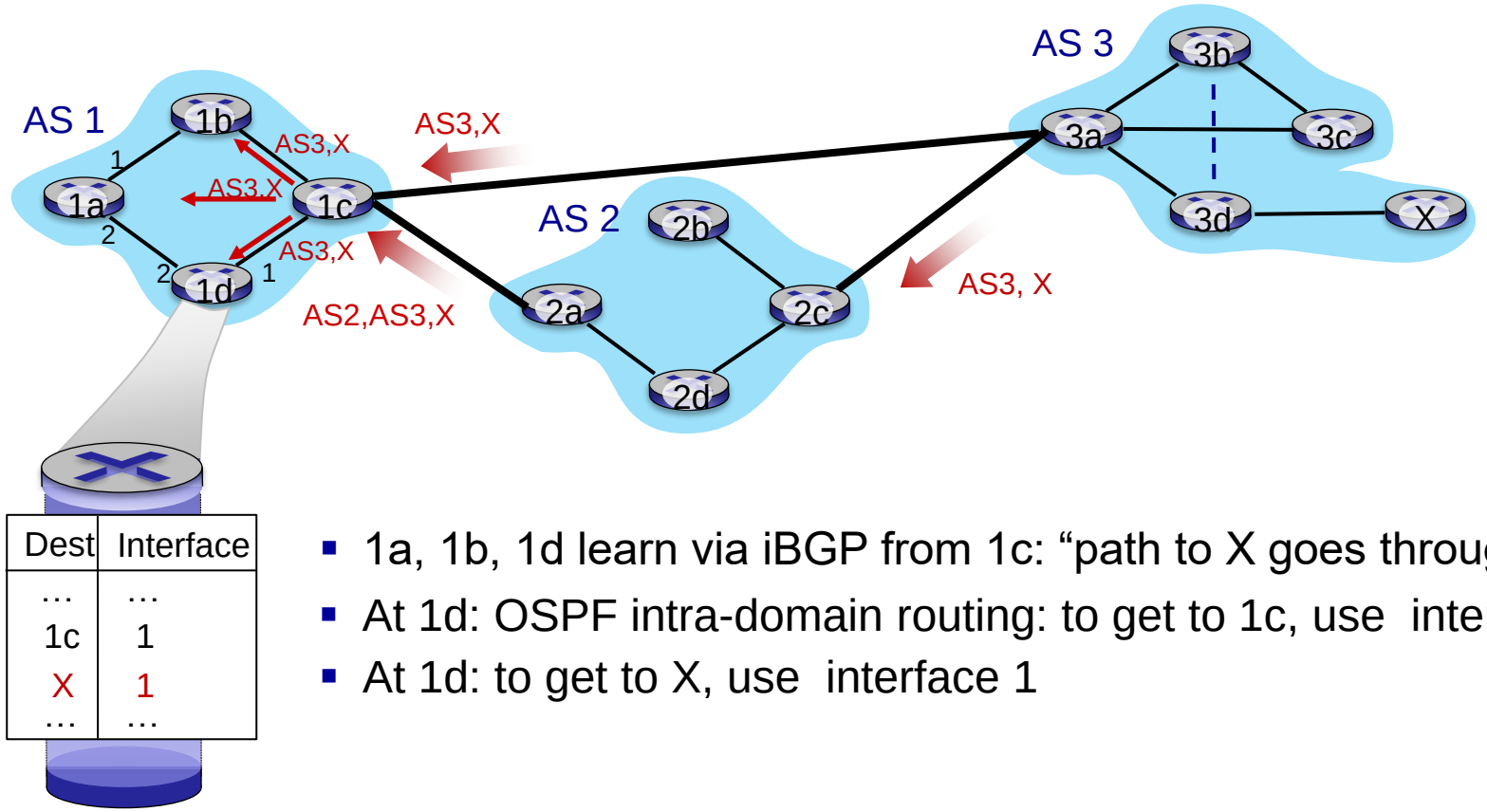
BGP Route Selection



Gateway router may learn about multiple paths to destination:

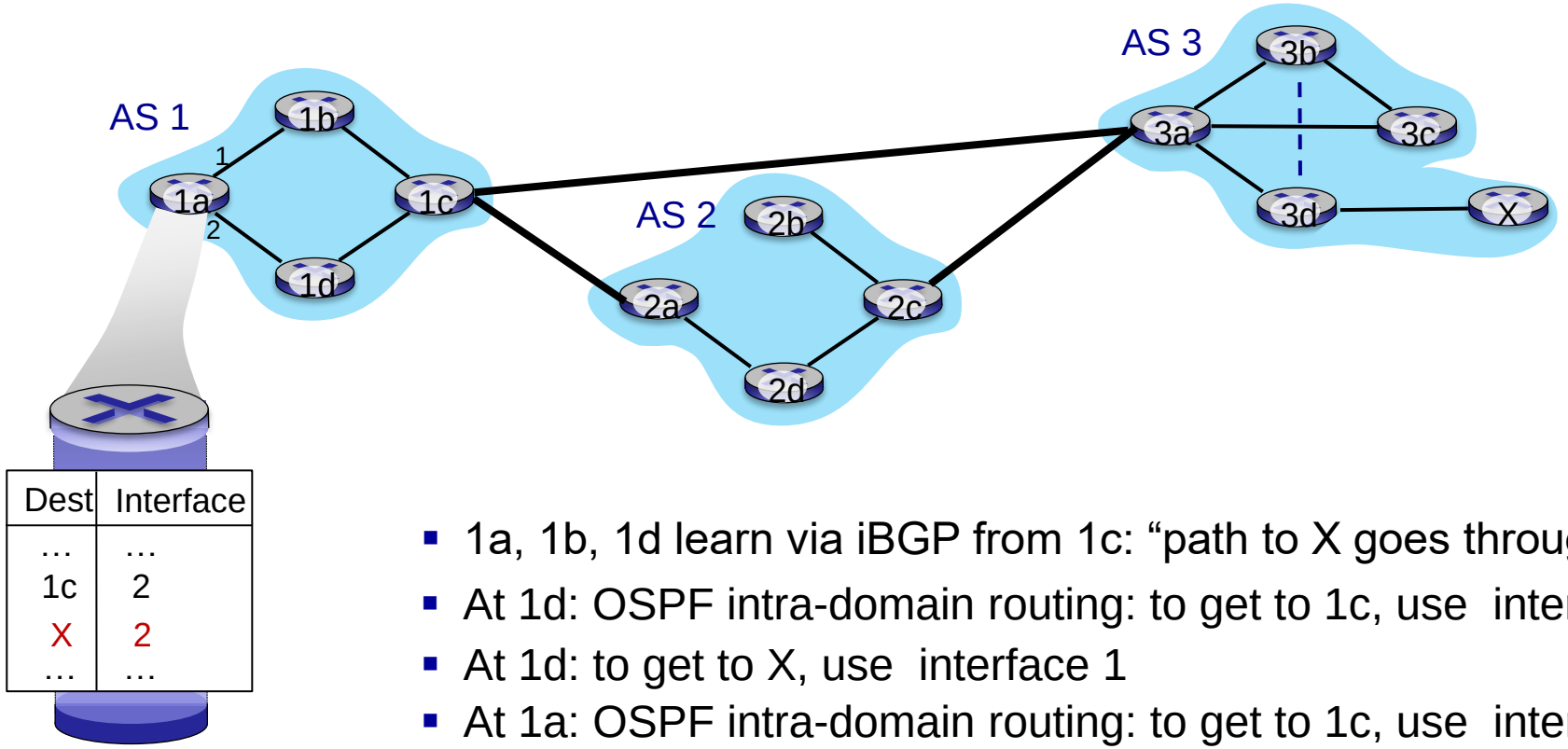
- AS1 gateway router 1c learns path AS2,AS3,X from 2a
- AS1 gateway router 1c learns path AS3,X from 3a
- Based on policy, AS1 gateway router 1c chooses path AS3,X and advertises path within AS1 via iBGP

BGP Route Selection



- 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 Route Selection

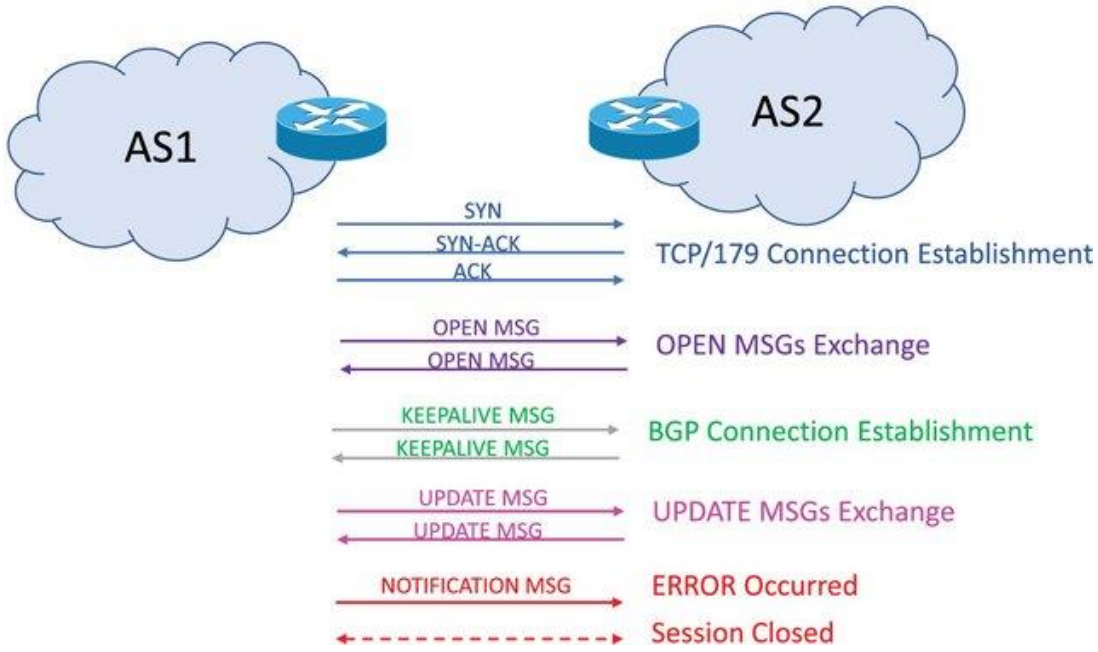


- 1a, 1b, 1d learn via iBGP from 1c: “path to X goes through 1c”
- At 1d: OSPF intra-domain routing: to get to 1c, use interface 1
- At 1d: to get to X, use interface 1
- At 1a: OSPF intra-domain routing: to get to 1c, use interface 2
- At 1a: to get to X, use interface 2

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



Src: https://onlinelibrary.wiley.com/doi/10.1002/dac.5266?af=R

Inter-AS vs Intra-AS Routing

- Inter-AS: admin wants control over how its traffic routed, who routes through its network
  - Policy dominates over performance

- Intra-AS: single admin, so policy less of an issue → can focus on performance