National University of Computer & Emerging Sciences

CS 3001 - COMPUTER NETWORKS

Lecture 22

Chapter 5

7th November, 2023

Nauman Moazzam Hayat

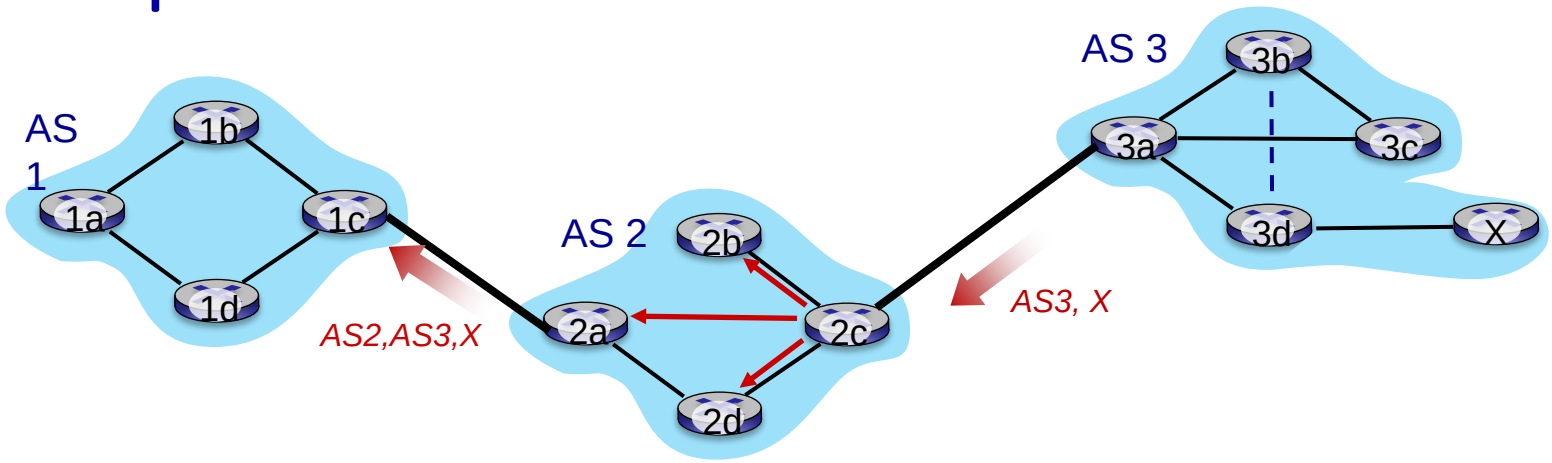
nauman.moazzam@lhr.nu.edu.pk

Office Hours: 02:30 pm till 06:00 pm (Every Tuesday & Thursday)

Path attributes and BGP routes

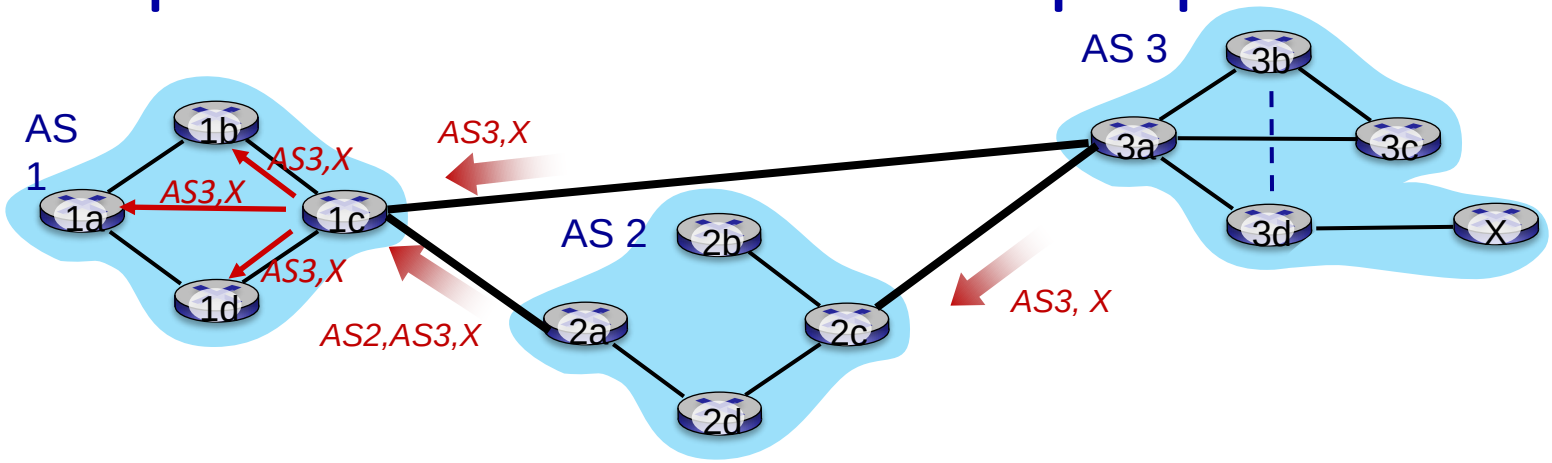
- BGP advertised route: prefix + attributes
  - prefix: destination being advertised
  - two important attributes:
    - AS-PATH: list of ASes through which prefix advertisement has passed
    - NEXT-HOP: indicates specific internal-AS router to next-hop AS
- policy-based routing:
  - gateway receiving 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

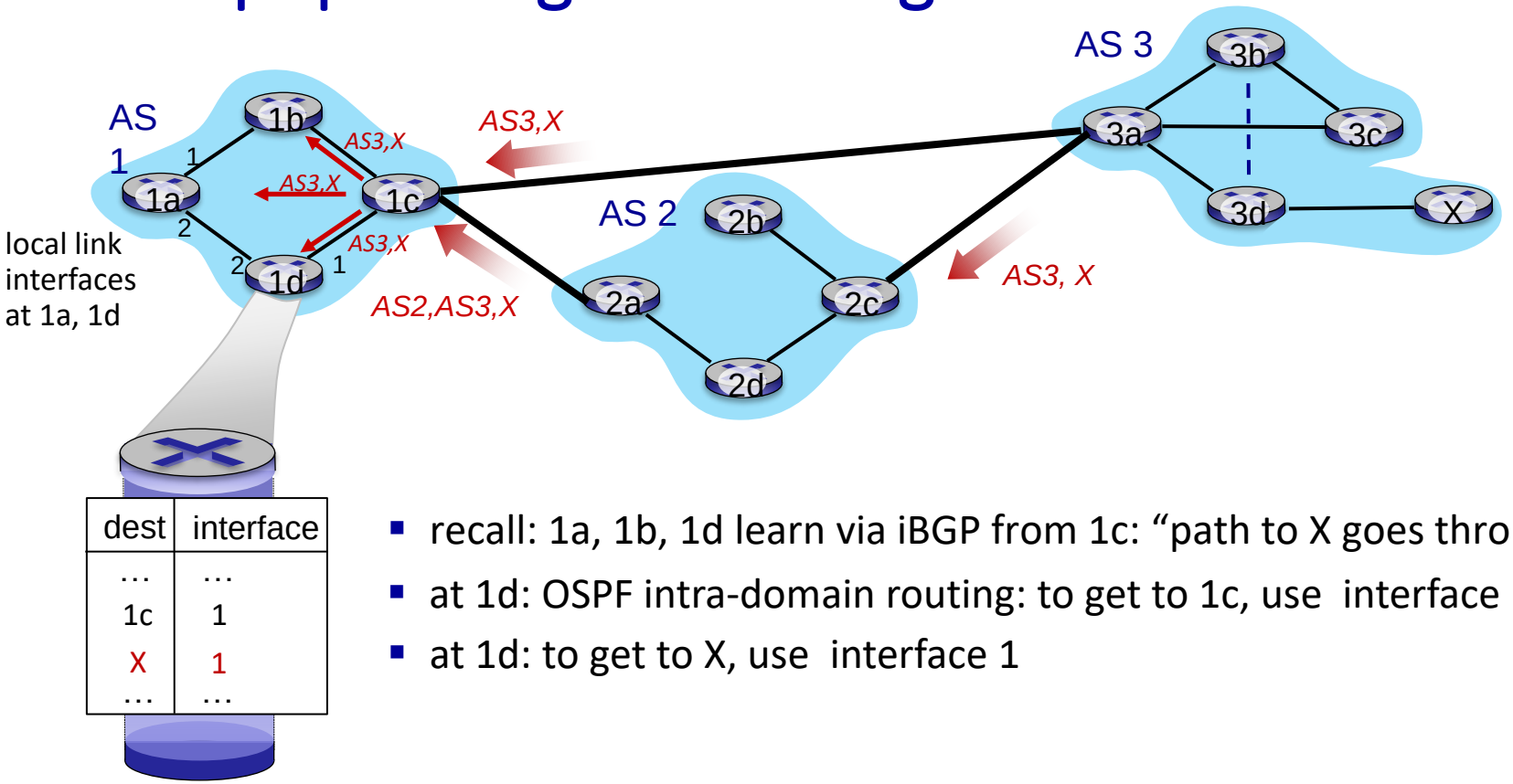
BGP path advertisement: multiple paths



gateway router may learn about multiple paths to destination:

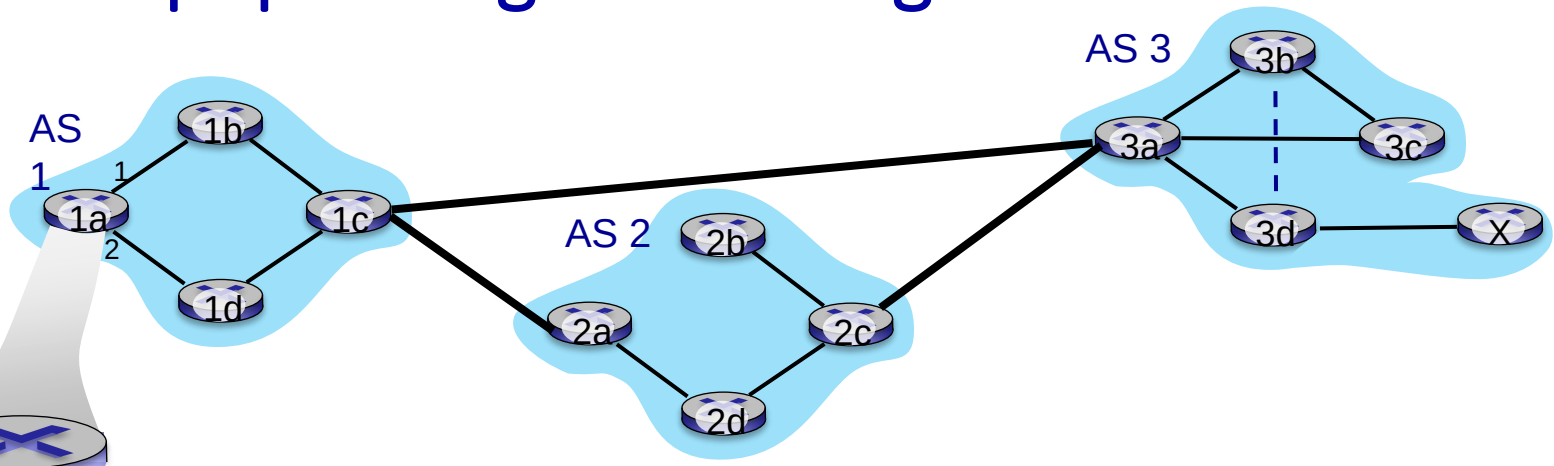
- AS1 gateway router 1c learns path *AS2,AS3,X* from 2a
- AS1 gateway router 1c learns path *AS3,X* from 3a
- based on *policy,* AS1 gateway router 1c chooses path *AS3,X* and advertises path within AS1 via iBGP

BGP: 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



dest	interface
…	…
1c	2
X	2
…	…

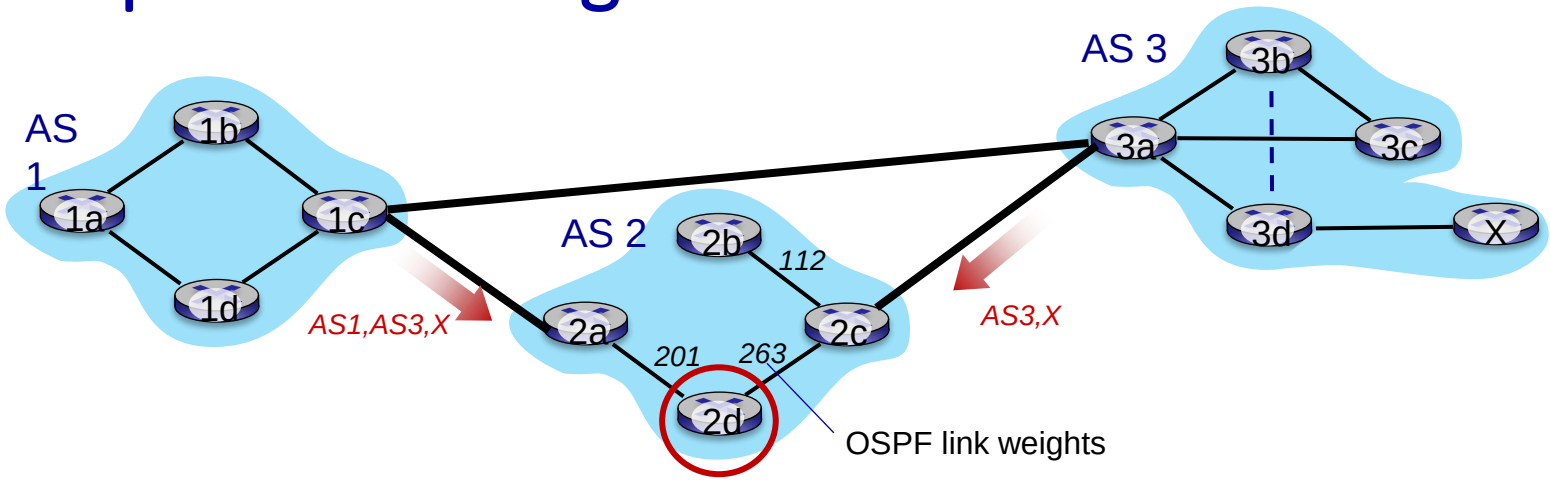
- recall: 1a, 1b, 1d learn via iBGP from 1c: “path to X goes through 1c”
- at 1d: OSPF intra-domain routing: to get to 1c, use interface 1
- at 1d: to get to X, use interface 1
- at 1a: OSPF intra-domain routing: to get to 1c, use interface 2
- at 1a: to get to X, use interface 2

How does entry get in forwarding table?

Summary

1. Router becomes aware of prefix
   - via BGP route advertisements from other routers
2. Determine router output port for prefix
   - Use BGP route selection to find best inter-AS route
   - Use OSPF (typically) to find best intra-AS route leading to best inter-AS route
   - Router identifies router port for that best route
3. Enter prefix-port entry in forwarding table

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!

Hot-Potato vs Cold-Potato Routing

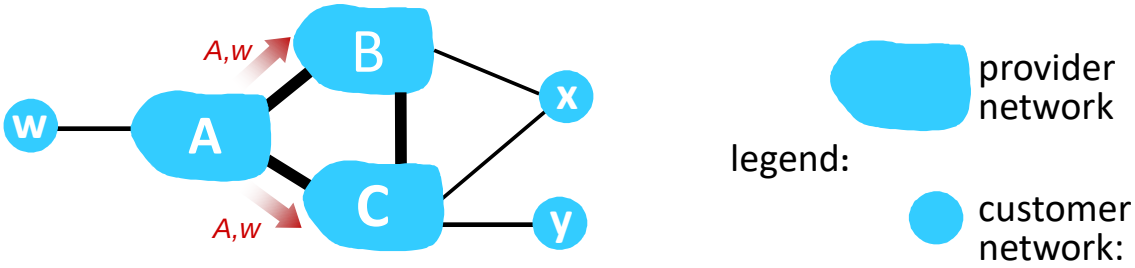
- ***Hot-potato*** *routing is the practice of passing traffic off to another AS as quickly as possible (thus using their network for wide-area transit.)*
  - *normal behavior of most peering agreements. It has the effect that the network receiving the data bears the cost of carrying it between cities. When the traffic ratio (traffic in both directions between peers) is reasonably even, this is considered fair.*
- ***Cold-potato*** *routing is the opposite: where the source AS holds onto the packet until it is as near to the destination as possible.*
  - *This is more expensive to do, but keeps the traffic under the network administrator's control for longer, allowing operators of well-provisioned networks to offer a higher quality of service to their customers. It can also be preferred when connecting to content providers.*

Example

- *Consider the case of two ISPs, A & B, who both have global networks. Additionally, they have peering agreements in both Europe and in Asia, which allows them to exchange data packets destined for the other's network at either location.*
- *Suppose a European customer of ISP A wants to transmit a data packet to an Asian customer of ISP B. ISP A will receive the packet in Europe and has to decide where to send the packet next.*
- *The first option is to hand off the packet to ISP B in Europe, and let ISP B carry the packet to Asia to be delivered to its destination. This is hot-potato routing, since ISP A hands off the packet at the earliest opportunity.*
- *The second option is for ISP A to carry the packet to Asia on its own internal network, and hand off to ISP B in Asia. This is called cold-potato, since ISP A keeps the packet in its internal network for as long as possible.*

Source: Wikipedia

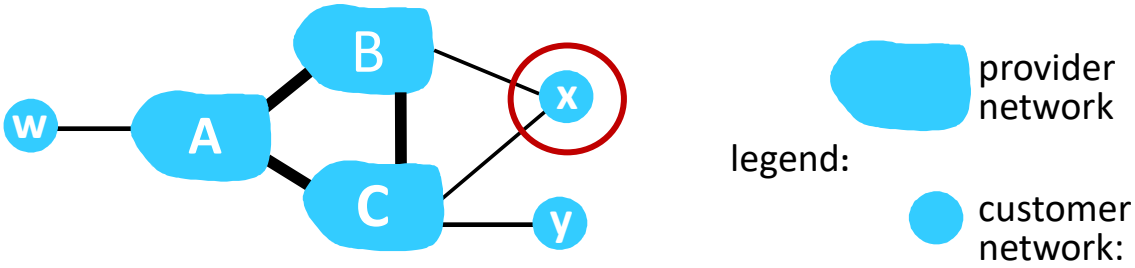
BGP: achieving policy via advertisements



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 stub network (a network where traffic either originates or terminates): which is dual homed, i.e. 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 route selection

- router may learn about more than one route to destination AS, selects route based on:
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
  3. closest NEXT-HOP router: hot potato routing
  4. additional criteria

Why different Intra-, Inter-AS 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

scale:

- hierarchical routing saves table size, reduced update traffic

performance:

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

VLSM Example

- VLSM Example below was solved

Given Classless IP Address Block 150.100.80.0/22

Four subnets required with the following pool sizes:

i) 300

ii) 200

iii) 100

iv) 120

Midterm II