Link Layer: Scheduling
Network Layer: BGP

CS204: Advanced Computer Networks

Nov 1, 2023

Adapted from Jiasi's Slides for Spring 2022

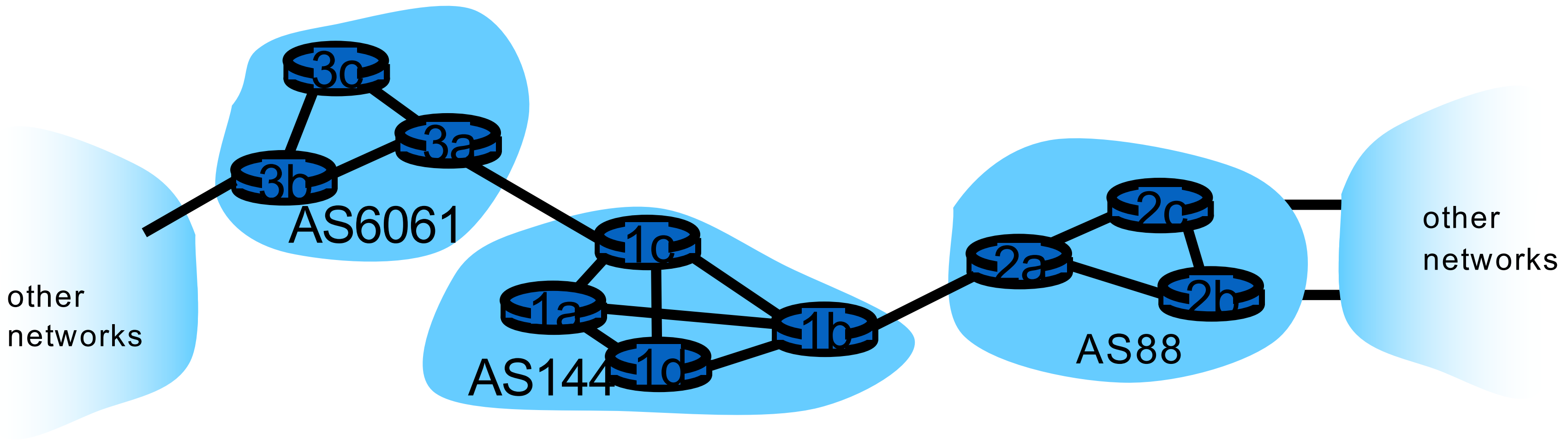
Overview

- Inter-AS routing
  - BGP
  - Example
  - Policy decisions
- Stable BGP routing
- Economics of AS relationships

Q: How to “glue together” a network of networks?

Autonomous Systems

- Autonomous system (AS)
  - Unit of routing policy
  - ~50k ASes in use
  - E.g., UCR has AS#6061, AT&T has AS#144, …



Examples

- Suppose you want to map the Internet…

IP address	AS#
173.246.82.76/30	40191
14.163.30.128/30	45899
14.163.30.144/30	45899
197.178.128.0/17	33771
199.241.113.0/24	30047
24.72.183.0/24	47887
24.72.167.0/24	47887
62.24.127.0/24	12455
103.237.33.0/24	58558
103.254.169.0/24	59149
142.234.125.0/24	15003
194.135.25.0/24	2118
208.81.73.0/24	27621
66.193.109.0/24	4323
110.78.161.0/24	23456
131.60.200.0/23	391
23.122.48.0/23	7018
177.154.78.0/23	28349
69.32.246.0/23	27365
177.8.252.0/23	262707
197.42.86.0/23	8452

AS#	AS#	Relationship
4323	12122	p2p
23151	19406	unknown
1103	21345	unknown
12714	20562	p2p
9298	4808	c2p
3267	196695	unknown
24626	43265	p2c
37432	33763	unknown
3741	50683	unknown
11030	45896	unknown
54345	33660	c2p
28310	52768	p2c
209	62507	p2c
47541	12637	unknown
23624	4713	c2p
59467	42632	c2p
9505	7529	unknown
20505	25180	p2p
62480	33659	c2p
56730	41678	unknown
29636	34407	unknown

IP Address

- UCR IP ranges
  - 138.23.0.0/16 AS6061
  - 192.31.146.0/24
  - 192.31.148.0/24
  - 192.35.223.0/24

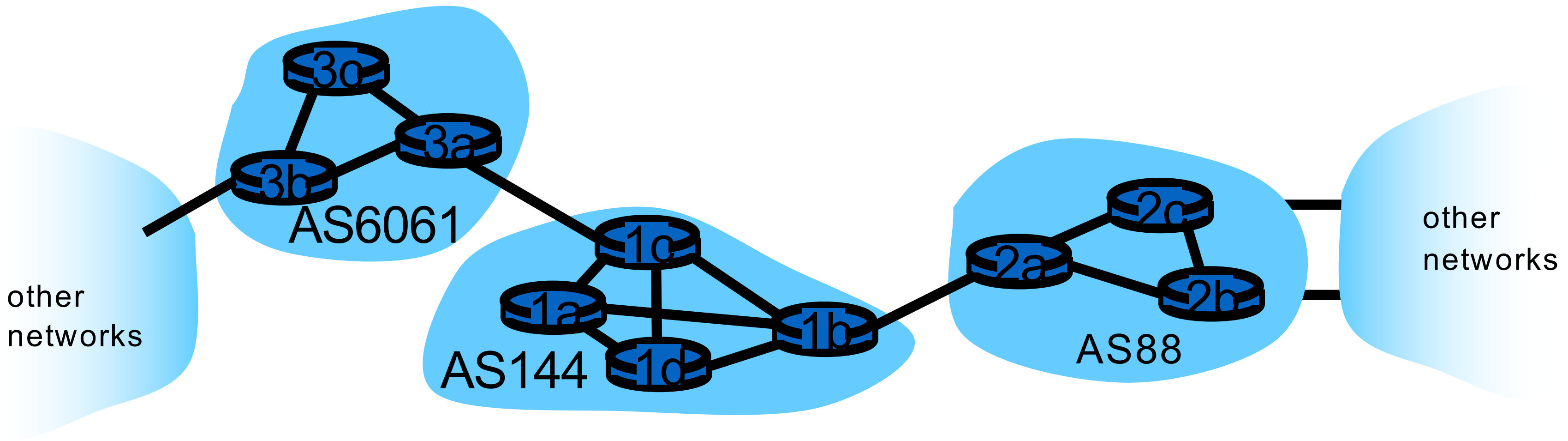
Q: What is the difference between IP address blocks and AS#?

Review of Routing

- Inter-AS routing
  - BGP
- Intra-AS routing
  - RIP
  - OSPF

Link-state? (e.g., Dijkstra)

Distance vector? (e.g., Bellman-Ford)



Why different Intra-, Inter-AS routing ?

policy:

- inter-AS: admin wants control over how its traffic routed, who routes through its net
- intra-AS: single admin, so no policy decisions needed

scale:

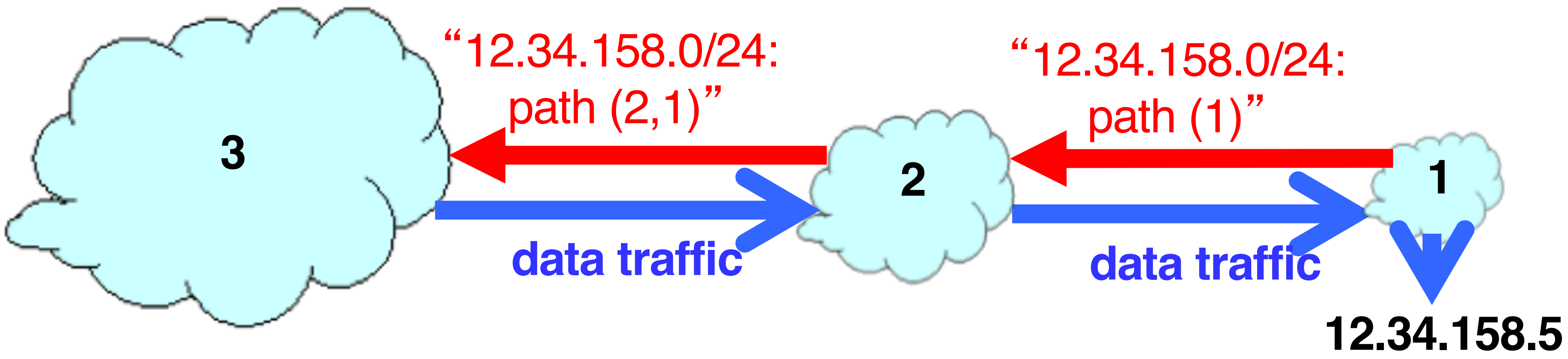
- hierarchical routing saves table size, reduced update traffic

performance:

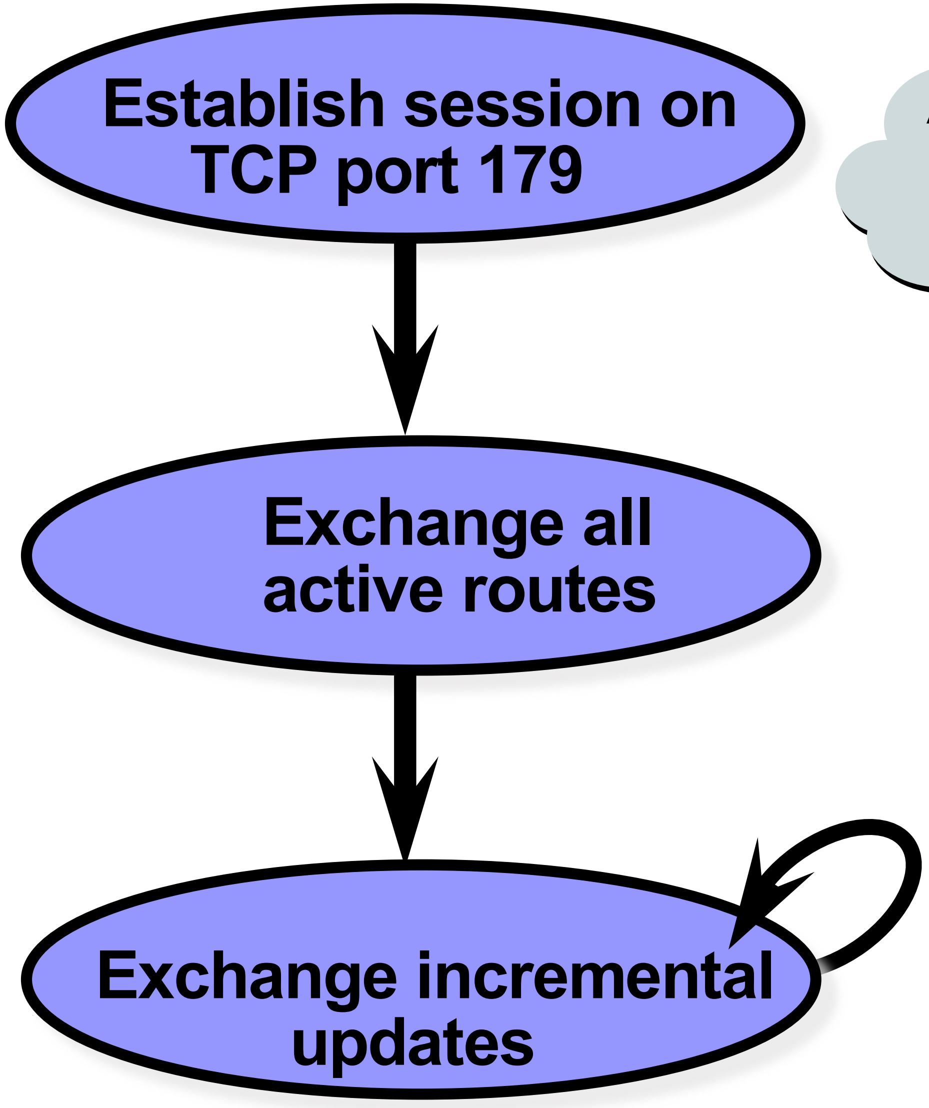
- intra-AS: can focus on performance
- inter-AS: policy may dominate over performance

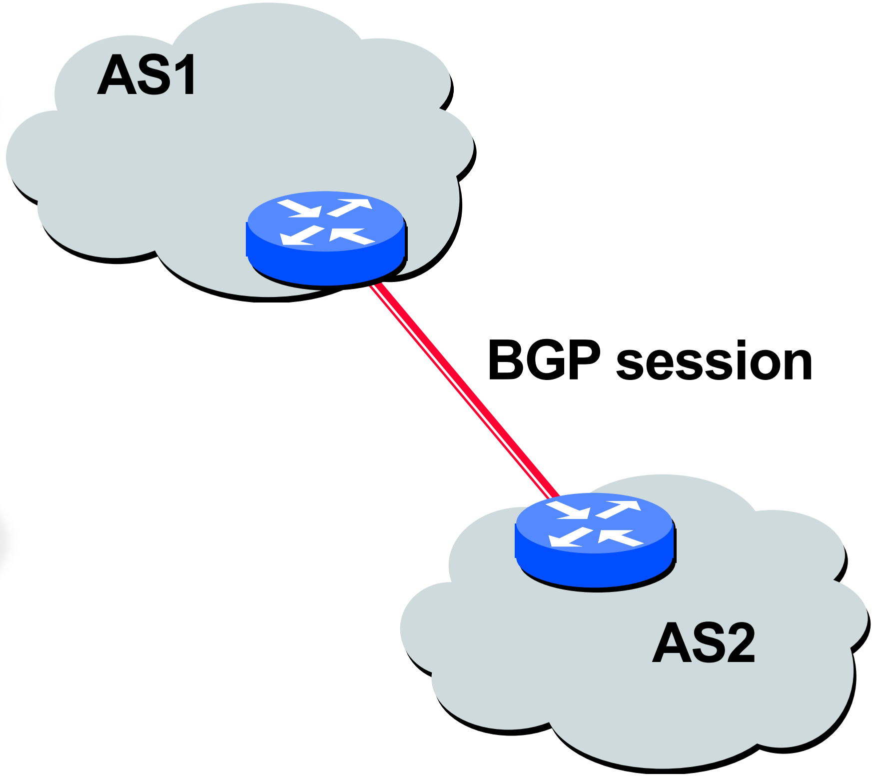
Border Gateway Protocol

- ASes exchange info about who they can reach
  - IP prefix: block of destination IP addresses
  - AS path: sequence of ASes along the path
- Policies configured by the AS's operator
  - Path selection: which of the paths to use?
  - Path export: which neighbors to tell?



BGP Session Operation





While connection is ALIVE exchange route UPDATE messages

Incremental Protocol

- A node learns multiple paths to destination
  - Stores all of the routes in a routing table
  - Applies policy to select a single active route
  - … and may advertise the route to its neighbors
- Incremental updates
  - Announcement
    - Upon selecting a new active route, add node id to path
    - … and (optionally) advertise to each neighbor
  - Withdrawal
    - If the active route is no longer available
    - … send a withdrawal message to the neighbors

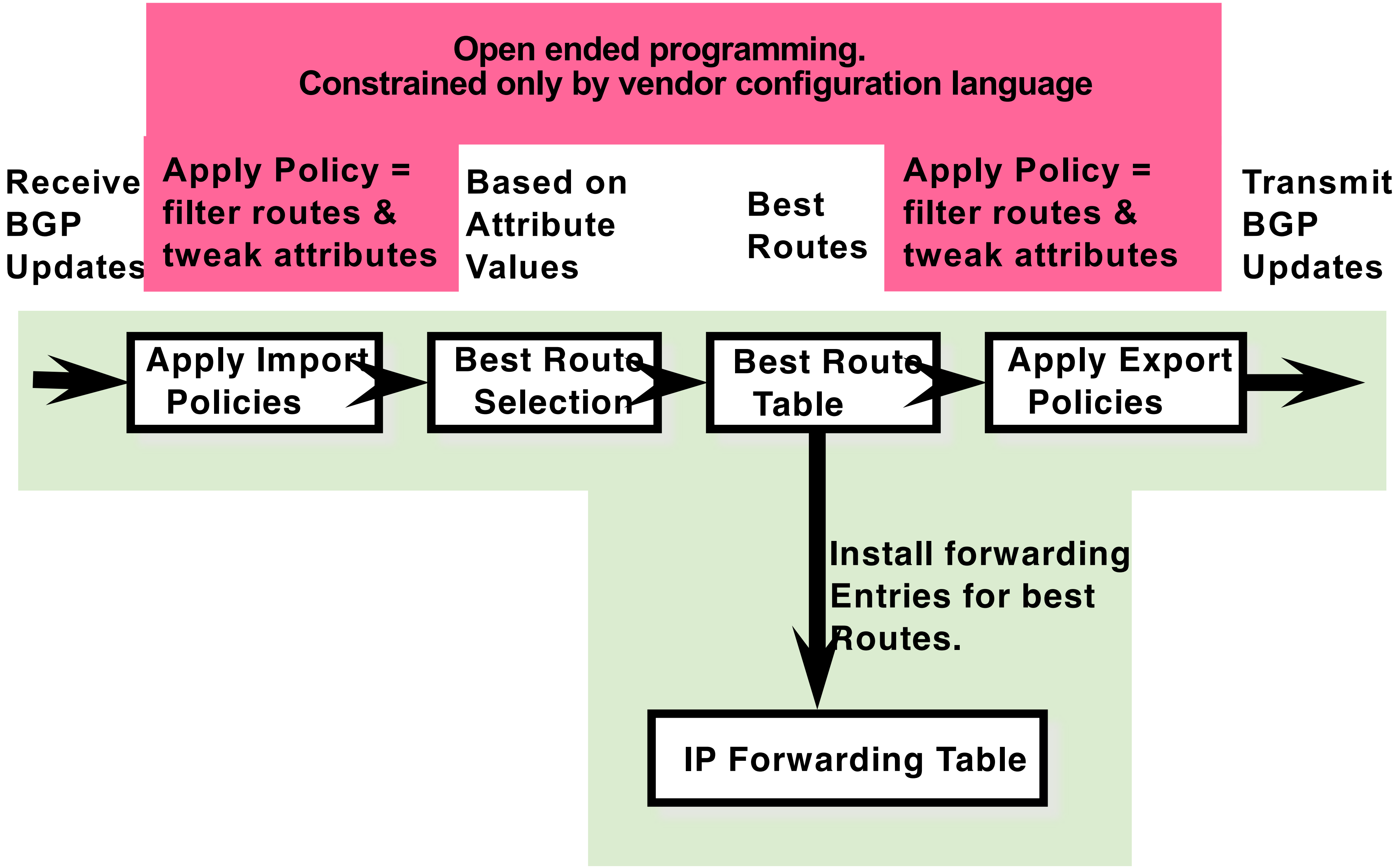
Routing Policies

- Economics
  - Enforce business relationships
  - Pick routes based on revenue and cost
  - Get traffic out of the network as early as possible
- Traffic engineering
  - Balance traffic over edge links
  - Select routes with good end-to-end performance
- Security and scalability
  - Filter routes that seem erroneous
  - Prevent the delivery of unwanted traffic
  - Limit the dissemination of small address blocks

Applying Policy to Routes

- Import policy
  - Filter unwanted routes from neighbor
    - E.g. prefix that your customer doesn't own
  - Manipulate attributes to influence path selection
    - E.g., assign local preference to favored routes
- Export policy
  - Filter routes you don't want to tell your neighbor
    - E.g., don't tell a peer a route learned from other peer
  - Manipulate attributes to control what they see
    - E.g., make a path look artificially longer than it is

BGP Policy: Influencing Decisions



Diving Deeper into an Example

How does entry get in forwarding table?

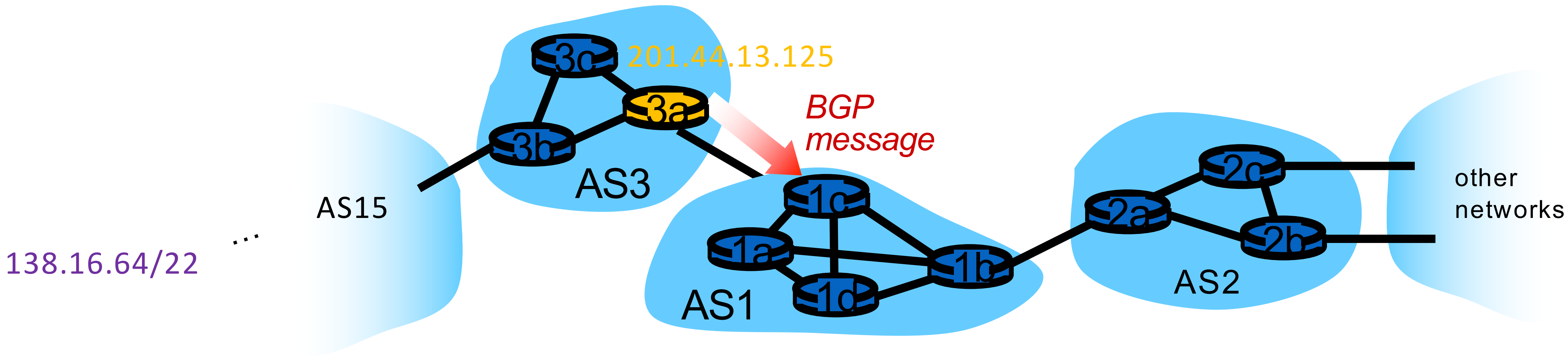
High-level overview

1. Router becomes aware of destination prefix
2. Filter the route based on policy ($?)
3. Router determines output port for prefix
4. Router enters prefix-port in forwarding table

Path attributes and BGP routes

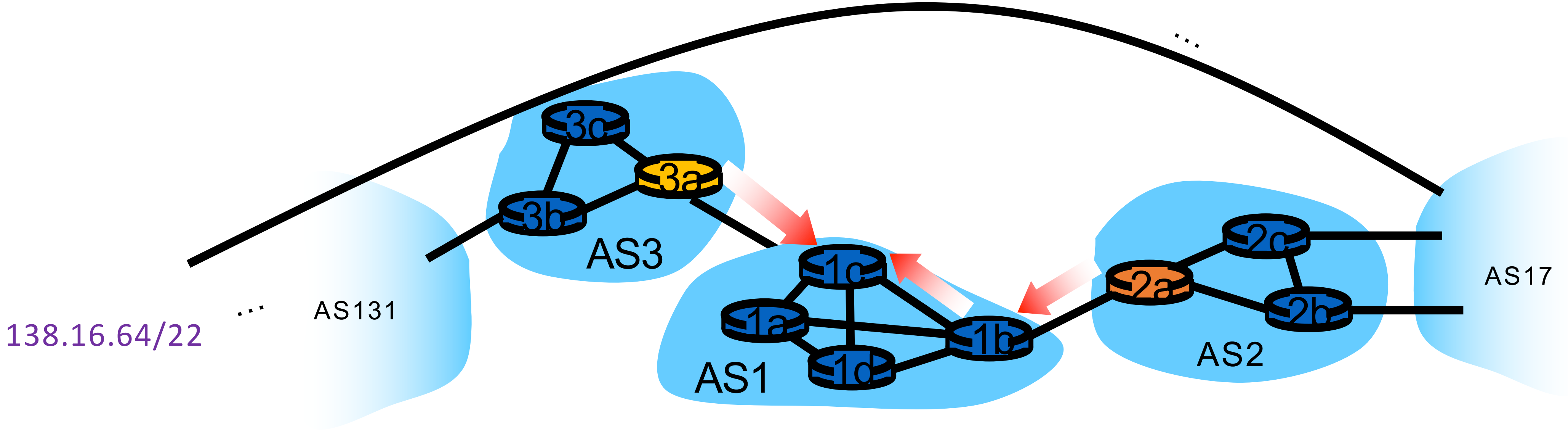
- advertised **destination prefix** includes BGP attributes
  - prefix + attributes = “route”
- two important attributes:
  - AS-PATH: contains ASs through which prefix advertisement has passed
  - NEXT-HOP: indicates specific internal-AS router to next-hop AS
- Example
  - Prefix: 138.16.64/22
  - AS-PATH: AS3 AS15 …
  - NEXT-HOP: 201.44.13.125

Router becomes aware of prefix



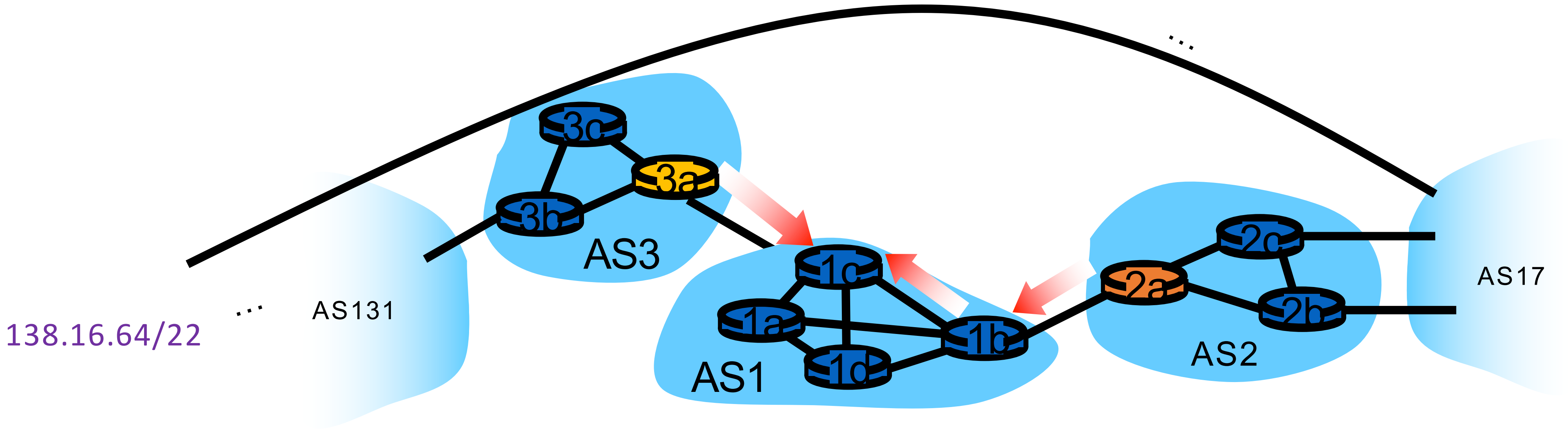
- BGP message contains "routes"
- "route" is a prefix and attributes: AS-PATH, NEXT-HOP,…
- Example
  - Prefix: 138.16.64/22
  - AS-PATH: AS3 AS15 …
  - NEXT-HOP: 201.44.13.125

Router may receive multiple routes



- Router may receive multiple routes for same prefix
- Which route to pick?
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
  3. closest NEXT-HOP router: hot potato routing (i.e., pass off to another AS as quickly as possible)
  4. additional criteria

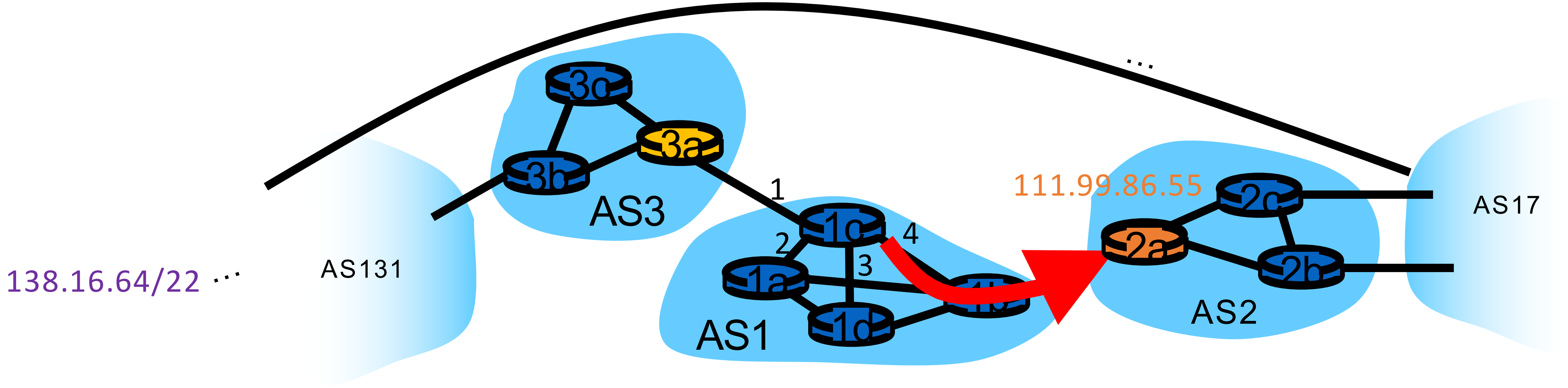
2. Shortest AS Path



- AS3 AS131 AS201 to 138.16.64/22 Hot potato route
- AS2 AS17 to 138.16.64/22 Shortest AS path

Next, use intra-domain routing to get to AS2

- Use selected route's NEXT-HOP attribute
  - NEXT-HOP = IP address of the router interface that begins the AS PATH
- Example:
  - AS-PATH: AS2 AS17 …; NEXT-HOP: 111.99.86.55
- Router 1c uses OSPF to find shortest internal path to 111.99.86.55
- Insert entry (138.16.64/22, 4) into router 1c's forwarding table



How does entry get in forwarding table?

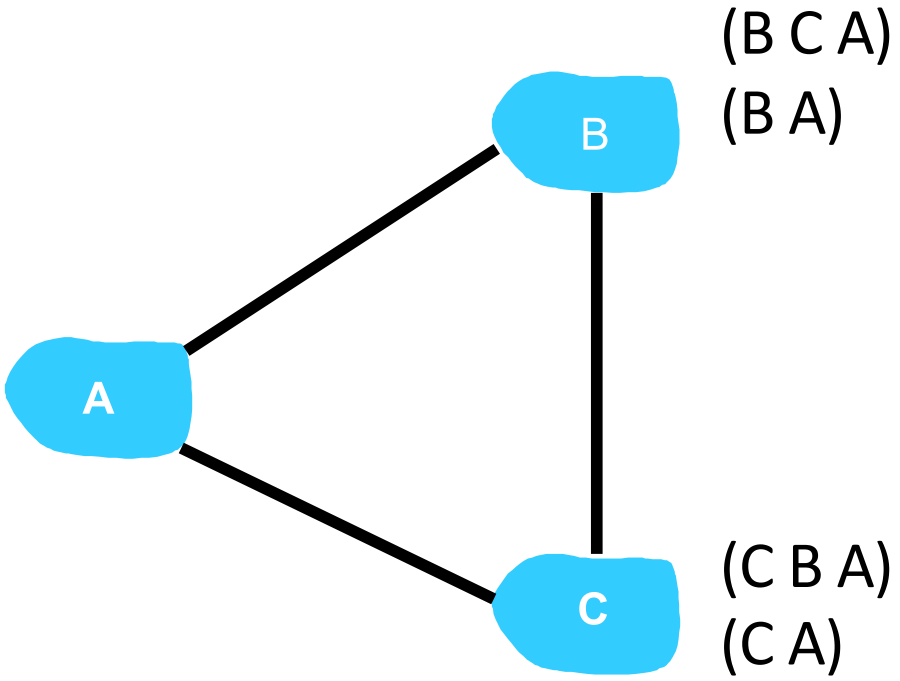
Summary

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

Is BGP Stable?

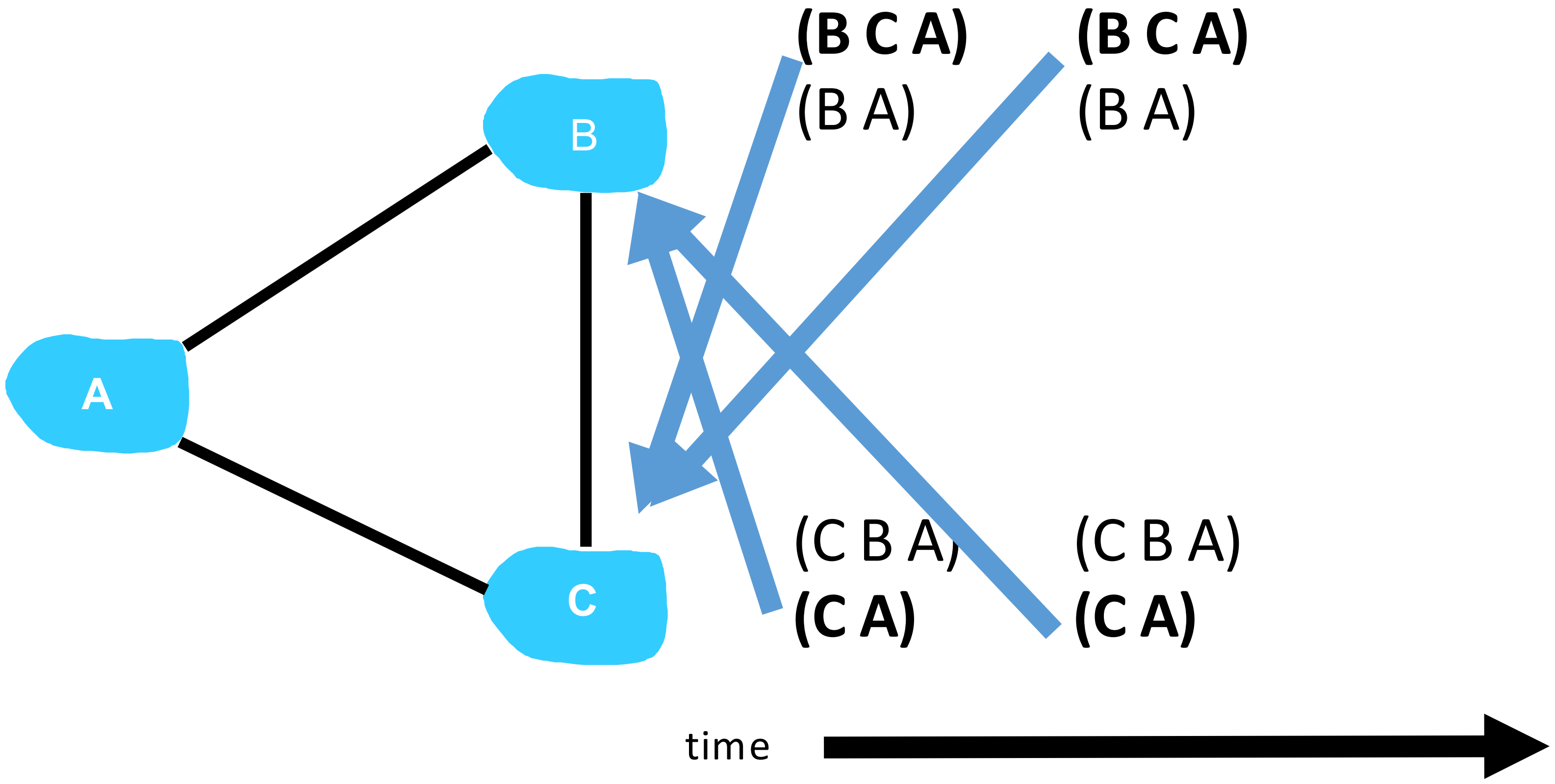
Toy Example

- Trying to get to destination A
- Routes listed in order of preference



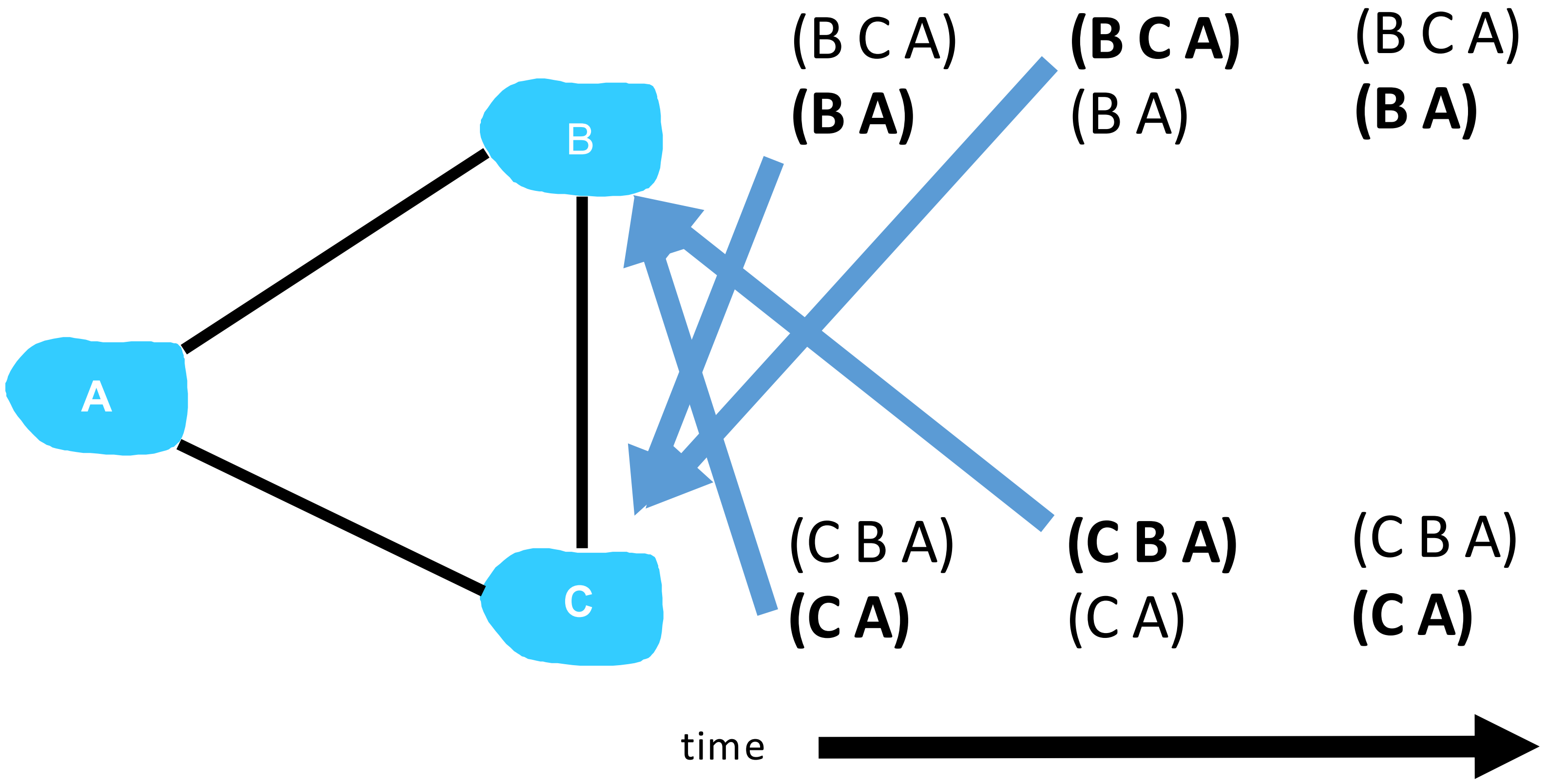
Stable BGP routing

- Suppose we start off with a certain initial configuration



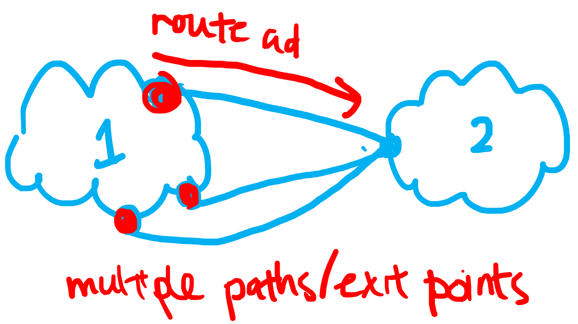
Unstable BGP routing

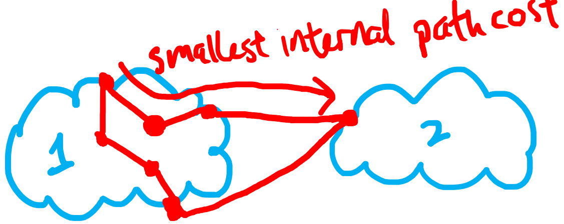
- Suppose we start off with the second choice options…



BGP Decision Process

- Highest local preference
  - *Set by import policies* upon receiving advertisement
- Shortest AS path
  - Included in the route advertisement
- Lowest origin type (e.g., eBGP preferred over iBGP)
  - Included in advertisement or reset by import policy
- Smallest multiple exit discriminator
  - When multiple paths (exit points) exist to same AS
  - Included in the advertisement or reset by import policy
- Smallest internal path cost to the next hop
  - Based on intradomain routing protocol (e.g., OSPF)
- Smallest next-hop router id
  - Final tie-break





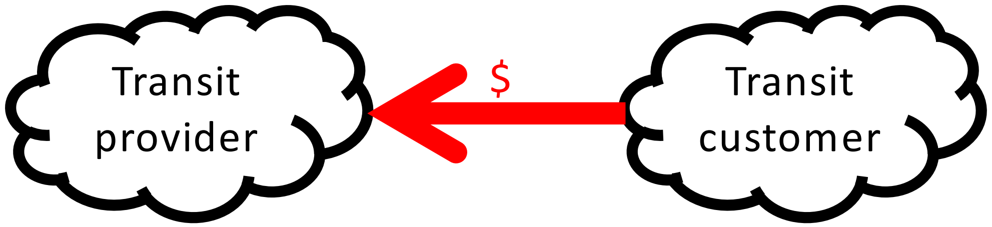
Economic Policies

Peering and Transit ISPs

- Peering

- Traffic flows are bi-directional
- ISPs jointly pay for equipment costs

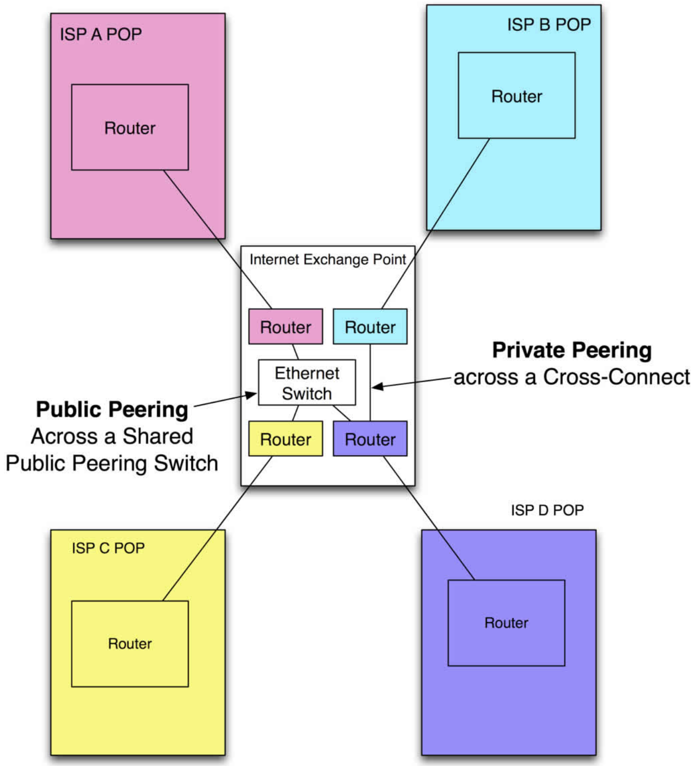
- Transit



- Traffic flows are bi-directional
- Arrow = Payment from customer to provider for upstream and downstream traffic

Internet Exchange Points (IXP)

- When two networks peer, it attracts other networks to peer there too
- Direct connection between ISPs still preferred
- Run as non-profits (Europe) or private business (USA)
  - Provide network equipment, switches, etc.
  - Monthly fee to join the IXP



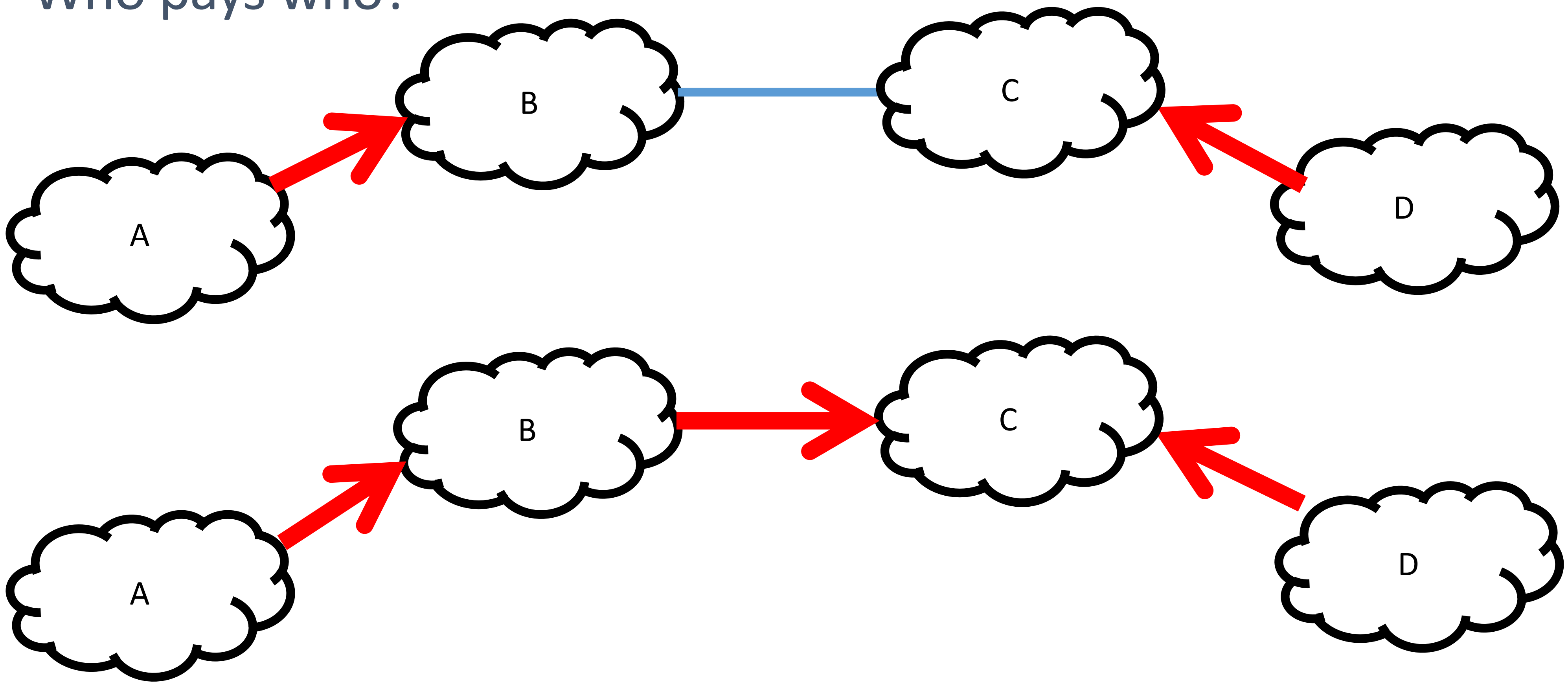
Pricing Contract

- Pricing contracts are typically not public information
- $/Mbps/month for all traffic to all destinations

- Variations
  - Paid peering
    - E.g. Netflix paid Comcast for direct peering
  - Backplane peering
    - Charge small ISPs for access to ISP's peers
  - Regional pricing
    - Pay to access different geographical regions

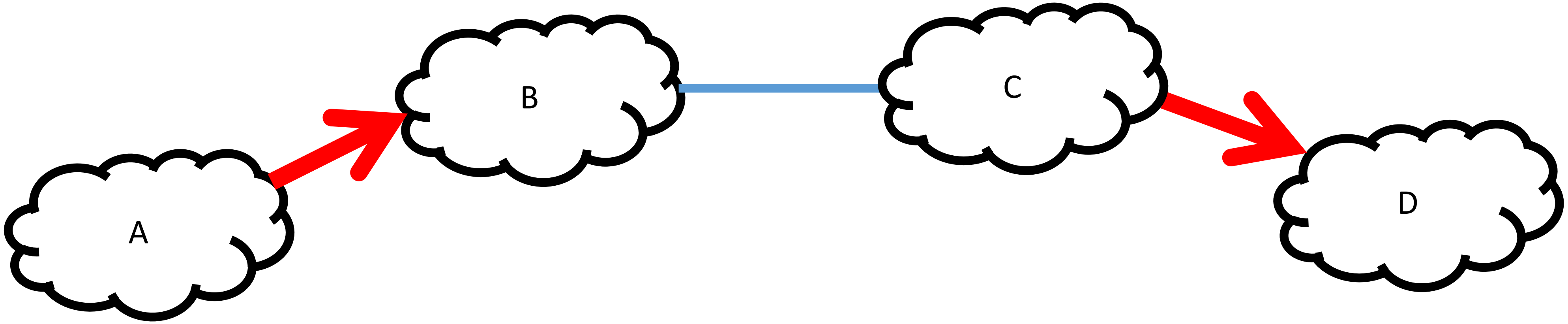
Examples

- Does traffic flow from A to D?
- Who pays who?



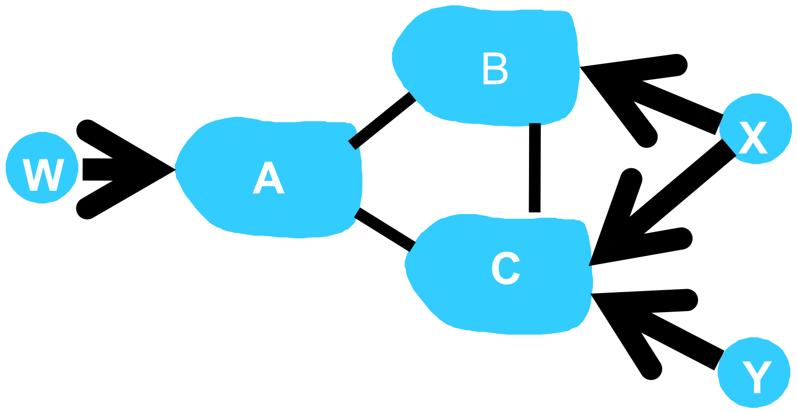
Examples (2)

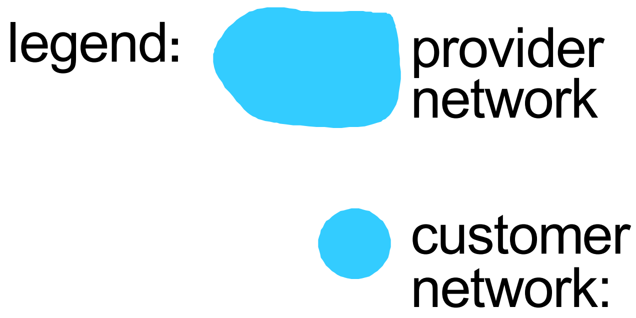
Peering connection only open to customers, not to other peers or networks it purchases transit from



Q: Will C announce B to D?

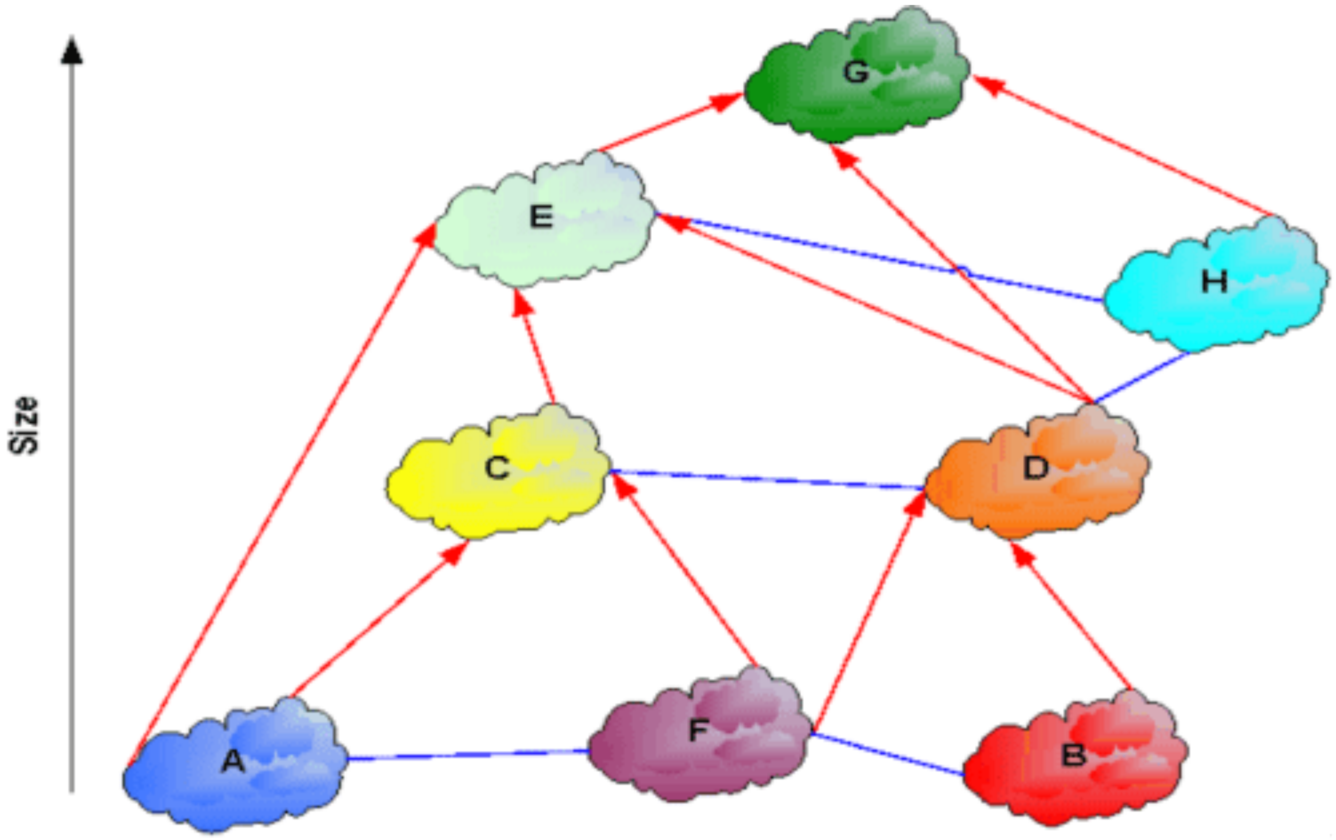
Examples (3)





- A advertises path AW to B
- B advertises path BAW to X
- Q: Should B advertise path BAW to C?
  - No way! B gets no "revenue" for routing CBAW since neither W nor C are B's customers
  - B wants to force C to route to w via A
  - B wants to route *only* to/from its customers!

Examples (4)



Who can network G see?

- All networks, because E, D, H buy transit from it.

Can A see B through F?

- No, because F doesn't advertise B to A

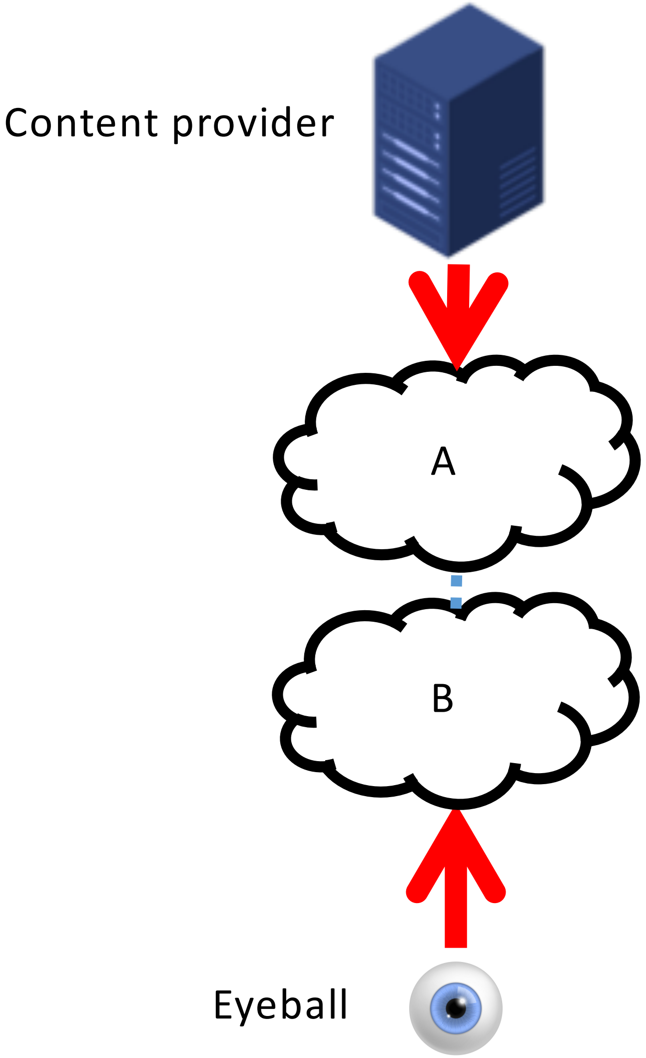
Can C see B through D or F?

- Can see B through its peer D
- Cannot see B through its transit customer F

Will traffic from C to H go through E or D or both?

- Through E
- Not through D

Content providers can also get involved



- **November 2010**: After Internet backbone provider Level 3 signs a deal with Netflix to distribute video, Comcast demands money from Level 3 for carrying traffic over the proverbial "last mile" to Comcast subscribers.
- **January 2011**: European ISPs Deutsche Telekom, Orange (formerly France Telecom), Telecom Italia, and Telefónica commission a report saying companies like Netflix and Google's YouTube service should give ISPs a lot more money.
- **August 2011**: Cogent, another Internet backbone provider that handles Netflix traffic, files a complaint in France against Orange, saying the ISP is providing inadequate connection speeds.
- **January 2013**: Free, a French ISP, is accused of slowing down YouTube traffic by failing to upgrade infrastructure (but is later cleared of intentionally degrading YouTube traffic by the French regulator). Free also temporarily blocks ads on YouTube and other video services by sending an update to its modems.
- **January 2013**: Orange and Google have a similar dispute, with Orange CEO Stephane Richard claiming victory. He says that Google is paying Orange to compensate the operator for mobile traffic sent from Google servers.
- **January 2013**: Time Warner refuses Netflix's offer of a free caching service that would provide better performance to Netflix users on Time Warner's network.
- **June 2013**: Cogent accuses Verizon of allowing "ports" between the two providers to fill up, degrading Netflix performance for Verizon customers.
- **July 2013**: The European Commission opens an antitrust probe into whether ISPs abused market positions in negotiations with content providers, and it searches the offices of Orange, Deutsche Telekom, and Telefónica. Separately, the French government demands details of interconnection agreements involving AT&T and Verizon.

Source: https://arstechnica.com/information-technology/2013/07/why-youtube-buffers-the-secret-deals-that-make-and-break-online-video/

Reading

- *Computer Networking: A Top-Down Approach*, Kurose & Ross
- Lixin Gao and Jennifer Rexford, "Stable Internet Routing Without Global Coordination," *IEEE Trans. Networking*, 2001.
- "How the 'Net works: an introduction to peering and transit", *ArsTechnica,* https://arstechnica.com/features/2008/09/peering-and-transit/.