EECS 489
Computer Networks

Winter 2025

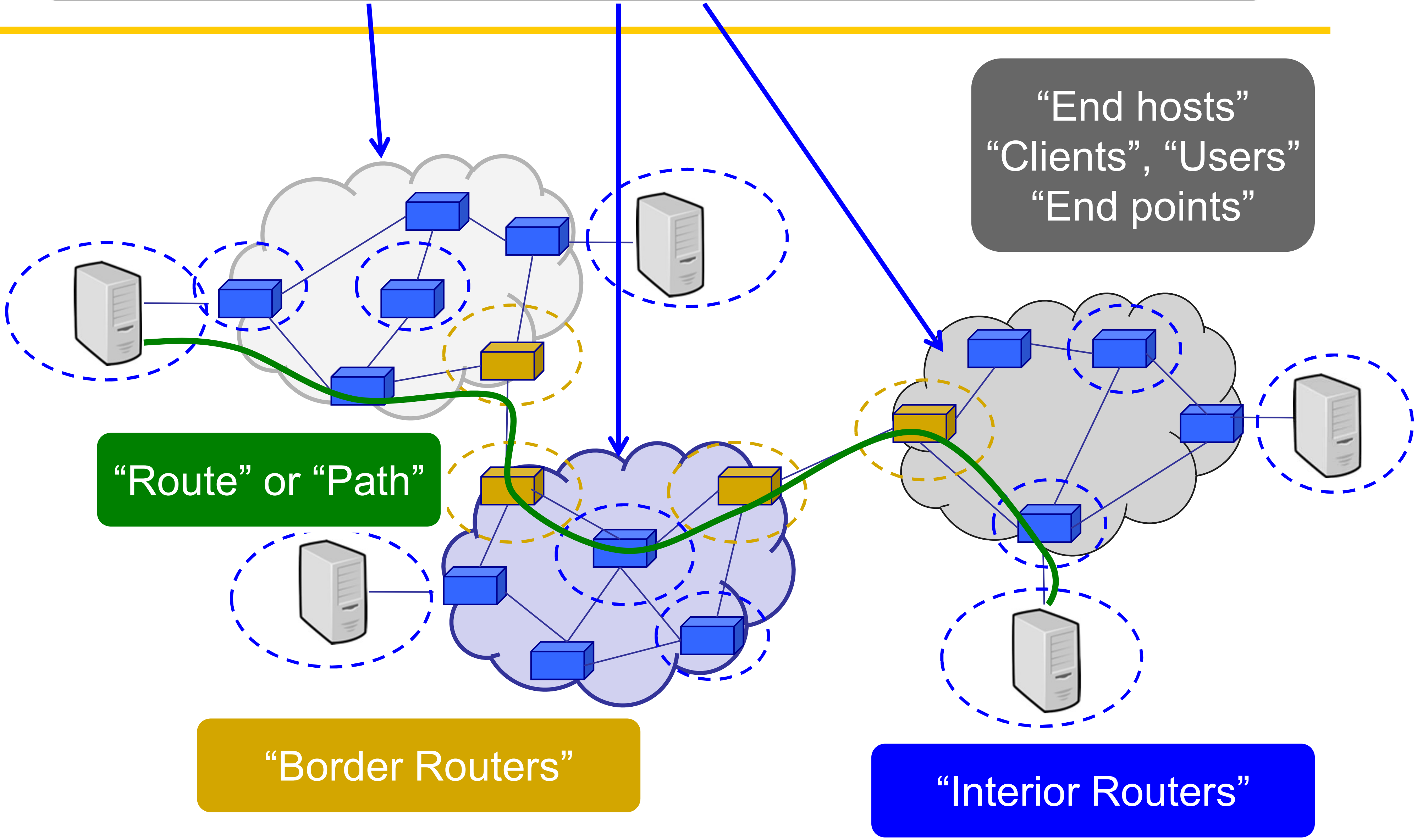
Mosharaf Chowdhury

Material with thanks to Aditya Akella, Sugih Jamin, Philip Levis, Sylvia Ratnasamy, Peter Steenkiste, and many other colleagues.

Agenda

- Inter-domain-routing

“Autonomous System (AS)” or “Domain”
Region of a network under a single administrative entity



Autonomous systems (AS)

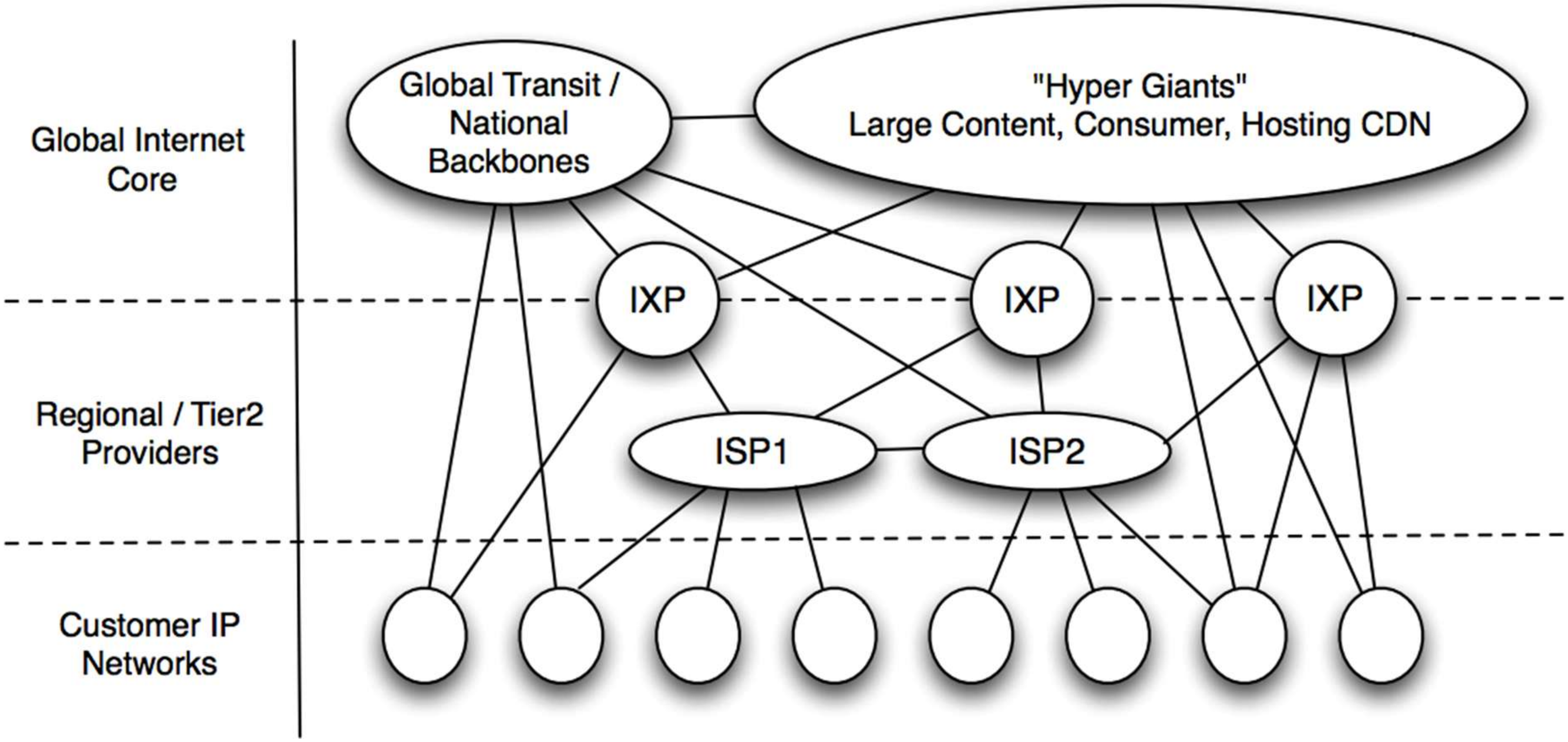
- An AS is a network under a single administrative control
  - Currently, ~77,000 IPv4 ASes & 34,750+ IPv6 ASes
    - Source: https://radar.cloudflare.com/routing
- ASes are sometimes called "domains"
- Each AS is assigned a unique identifier (ASN)
  - E.g., University of Michigan owns ASNs 177 to 180

AS-level Internet

- Used to be a large graph of ASes
  - In 2007, half of the Internet's total traffic came from ~2000 ASes

- It's consolidating since then
  - In 2009, the largest 150 ASes contributed to half the traffic

AS-level Internet



Internet Inter-Domain Traffic, SIGCOMM, 2010

AS-level Internet Today

- By 2019, half the Internet traffic came from only five hypergiants such as Google, Netflix, Meta, Akamai
  - Seven Years in the Life of Hypergiants’ Off-nets, SIGCOMM, 2021

“Intra-domain” routing: Within an AS

- Link-State (e.g., OSPF) and Distance-Vector (e.g., RIP)
- Primary focus
  - Finding least-cost paths
  - Fast convergence

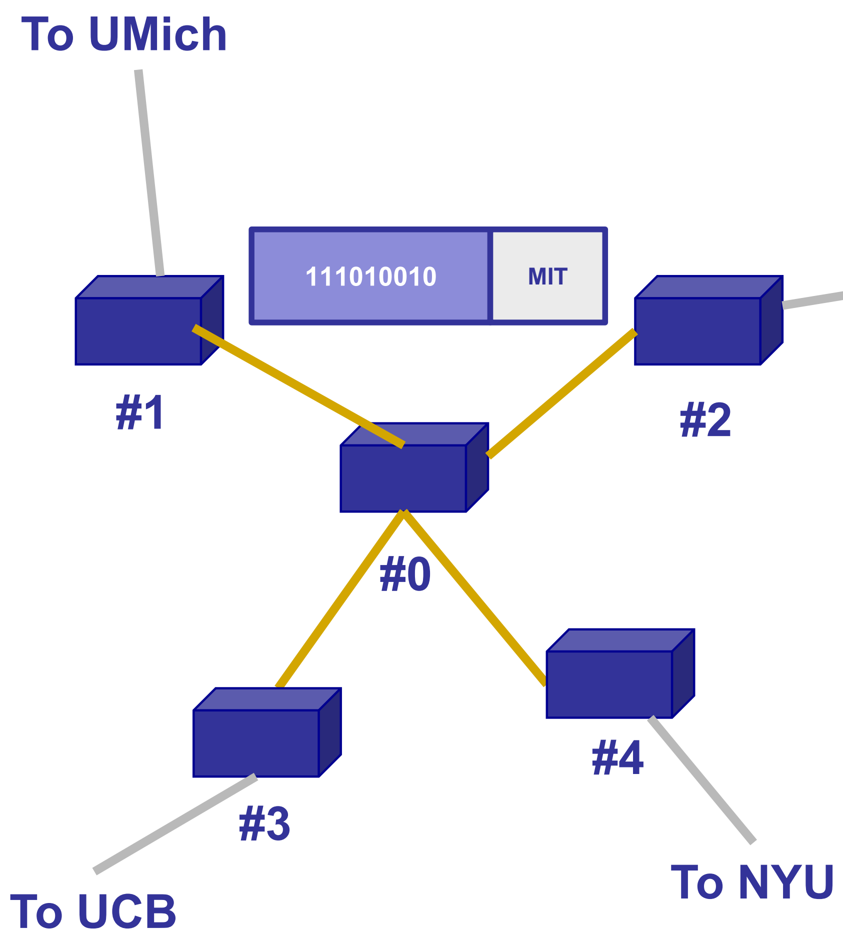
"Inter-domain" routing: Between ASes

- Two key challenges
  - Scaling
  - Administrative structure
    - Issues of autonomy, policy, privacy

Recall: Addressing (so far)

- Each host has a unique ID
- No particular structure to those IDs

Recall: Forwarding



Forwarding Table @ #0

Dst	Next Hop
UMICH	1
UCB	3
MIT	2
NYU	4

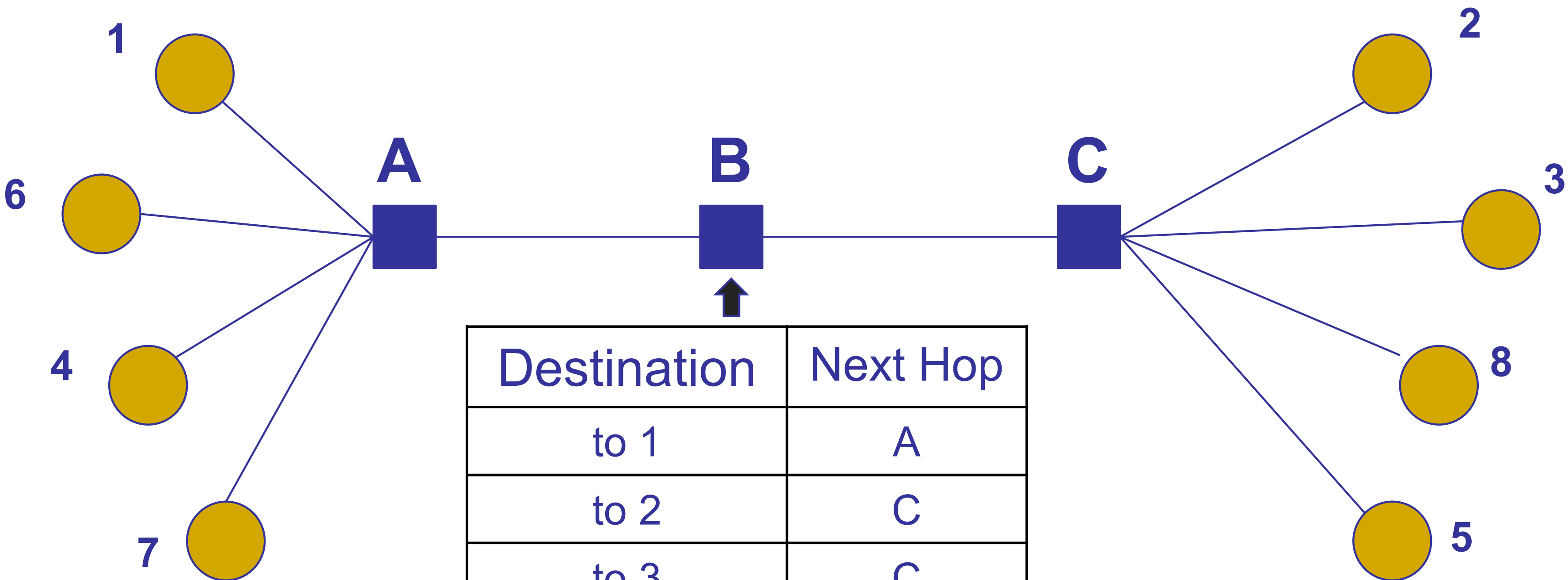
Two key challenges

- Scaling
- Administrative structure
  - Issues of autonomy, policy, privacy

Scaling

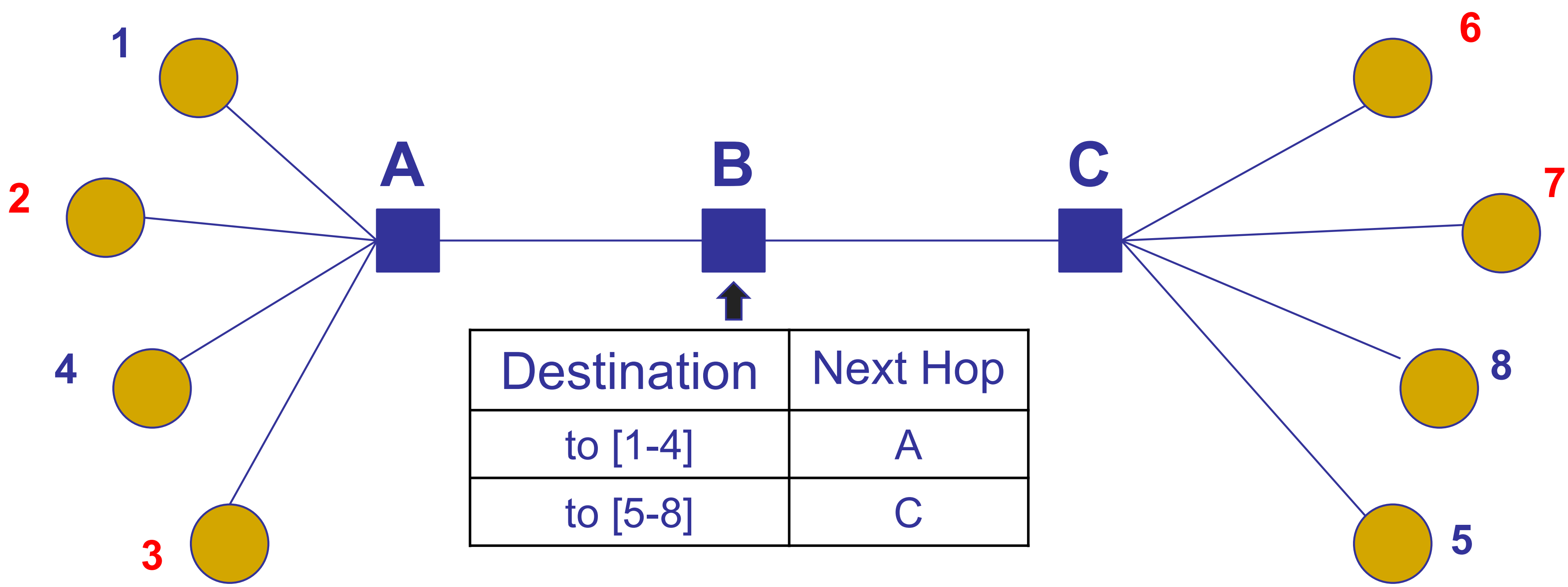
- A router must be able to reach any destination
  - Given packet's destination address, lookup next hop
- Naive: Have an entry for each destination
  - There would be over 10^8 entries!
  - AND routing updates per destination!
- How can we improve scalability?
  - We have already seen an example: longest-prefix matching

A smaller table at node B?



Destination	Next Hop
to 1	A
to 2	C
to 3	C
to 4	A
to 5	C
to 6	A
to 7	A
to 8	C

Re-number the end-systems?



Destination	Next Hop
to [1-4]	A
to [5-8]	C

- Careful address assignment → can *aggregate* multiple addresses into one range → scalability!
- Akin to reducing the number of destinations

Scaling

- A router must be able to reach any destination
- Naive: Have an entry for each destination
- Better: Have an entry for a range of addresses
  - Can’t do this if addresses are assigned randomly!
  - How addresses are allocated will matter!

- Host addressing is key to scaling

Two key challenges

- Scaling
- Administrative structure
  - Issues of autonomy, policy, privacy

Administrative structure shapes inter-domain routing

- ASes want freedom in picking routes
  - "My traffic can't be carried over my competitor's network"
  - "I don't want to carry A's traffic through my network"
  - Not expressible as Internet-wide "least cost"
- ASes want autonomy
  - Want to choose their own internal routing protocol
  - Want to choose their own policy
- ASes want privacy
  - Choice of network topology, routing policies, etc.

Choice of routing algorithm

- Link-state
  - No privacy – broadcasts all network information
  - Limited autonomy – needs agreement on metric, algo
- Distance-vector is a decent starting point
  - Per-destination updates give some control
  - BUT wasn't designed to implement policy
  - AND is vulnerable to loops

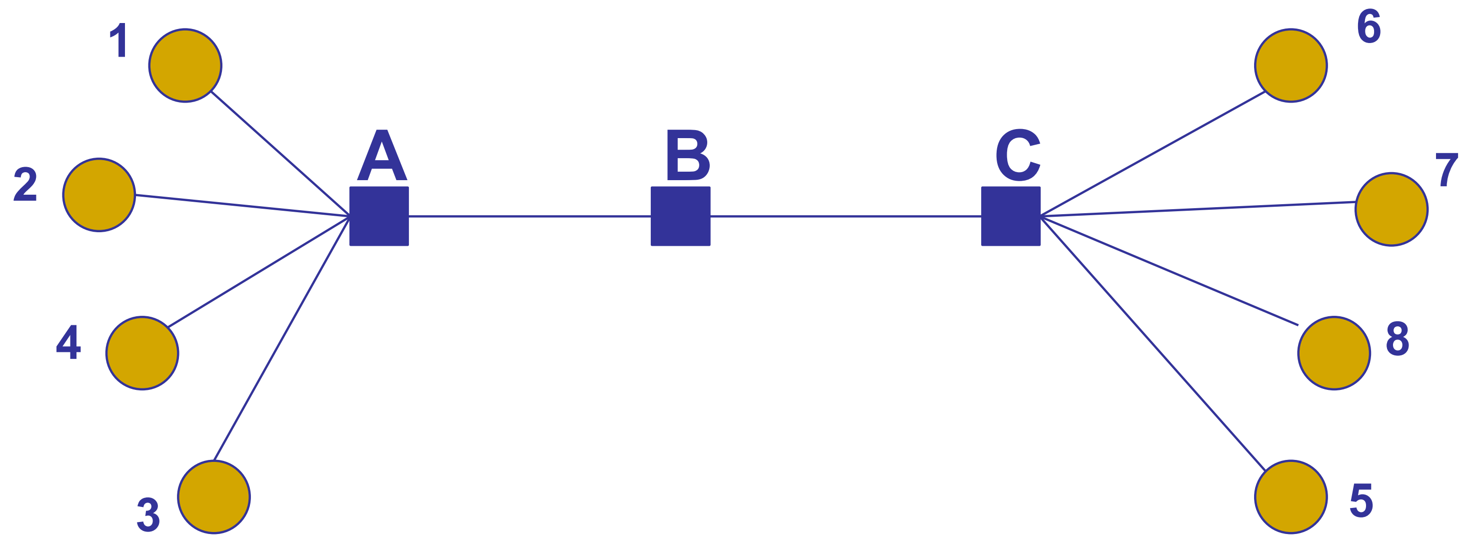
- The "Border Gateway Protocol" (BGP) extends distance-vector ideas to accommodate policy

IP ADDRESSING

Goal of addressing: Scalable routing

- State: Small forwarding tables at routers
  - Much less than the number of hosts
- Churn: Limited rate of change in routing tables
- Ability to aggregate addresses is crucial for both

Aggregation works if...



- Groups of destinations reached via the same path
- These groups are assigned contiguous addresses
- These groups are relatively stable
- Few enough groups to make forwarding easy

IP addressing is hierarchical

- Hierarchical address structure
- Hierarchical address allocation
- Hierarchical addresses and routing scalability

IP addresses (IPv4)

- Unique 32-bit number associated with a host

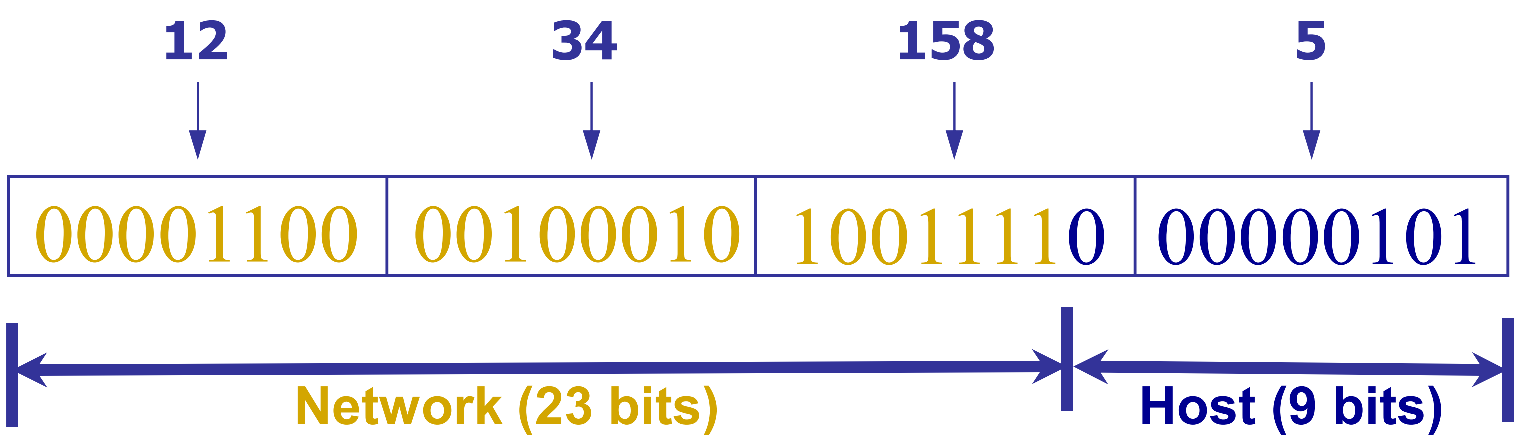
00001100 00100010 10011110 00000101

- Represented with the “dotted-decimal” notation
  - e.g., 12.34.158.5

12	34	158	5
00001100	00100010	10011110	00000101

Hierarchy in IP addressing

- 32 bits are partitioned into a prefix and suffix components
- Prefix is the network component; suffix is the host component



- Inter-domain routing operates on network prefix

CIDR: Classless inter-domain routing

- Flexible division between network and host addresses
- Offers a better tradeoff between size of the routing table and efficient use of the IP address space

CIDR example

- Suppose a network has 50 computers
  - Allocate 6 bits for host addresses ($2^5 < 50 < 2^6$)
  - Remaining 32 - 6 = 26 bits as network prefix
- Flexible boundary means the boundary must be explicitly specified with the network address!
  - Informally, “slash 26” → 128.23.9/26
  - Formally, prefix represented with a 32-bit mask: 255.255.255.192, where all network prefix bits set to “1” and host suffix bits to “0”
  - Also known as subnet mask (a group of machines with the same prefix are in the same subnet)

Before CIDR: Classful addressing

- Three classes
  - 8-bit network prefix (Class A),
  - 16-bit network prefix (Class B), or
  - 24-bit network prefix (Class C)
- Example: an organization needs 500 addresses.
  - A single class C address is not enough (<500 hosts)
  - Instead, a class B address is allocated (~65K hosts)
    - Huge waste!

5-MINUTE BREAK!

IP addressing is hierarchical

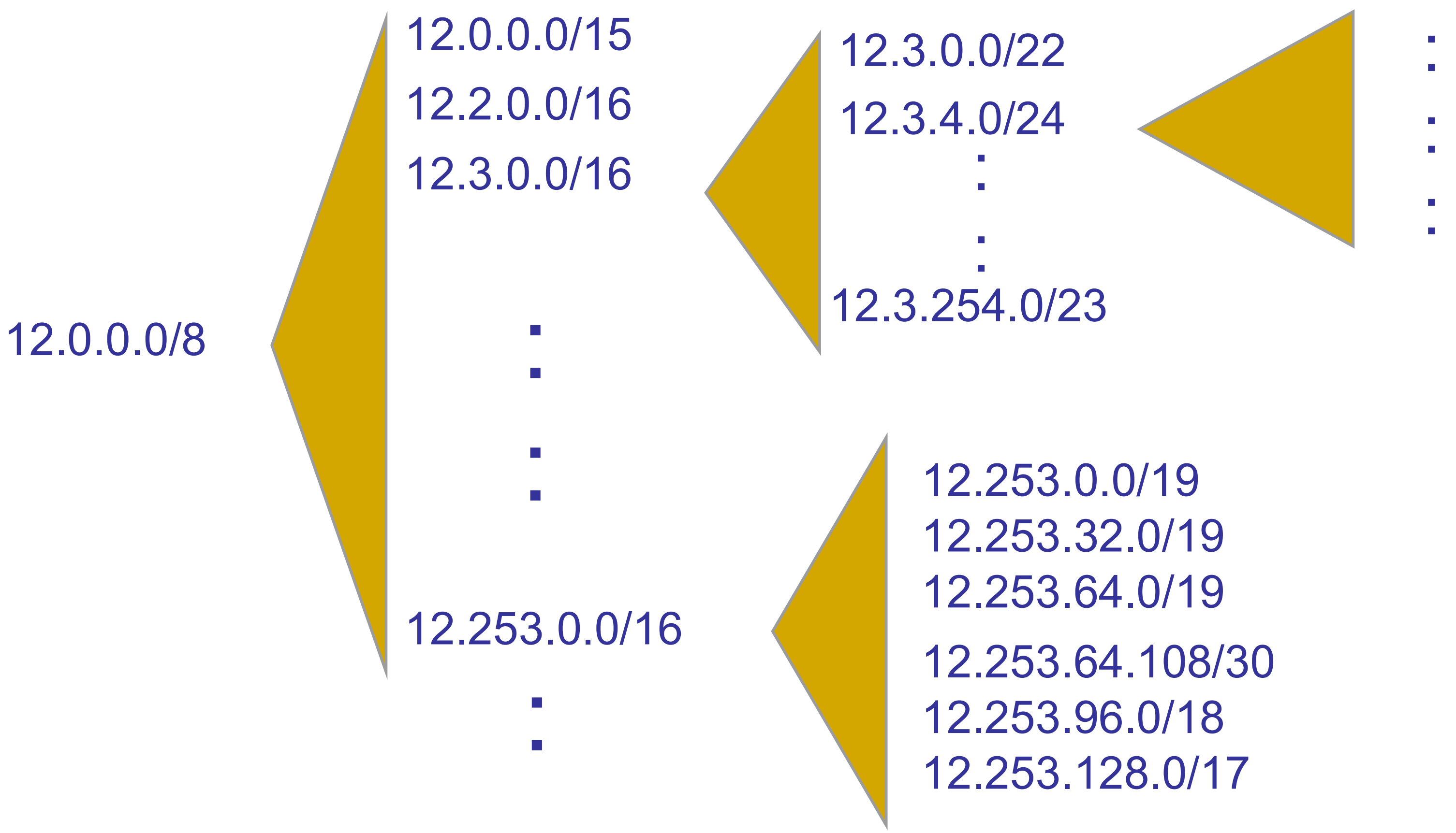
- Hierarchical address structure
- Hierarchical address allocation
- Hierarchical addresses and routing scalability

Allocation done hierarchically

- Internet Corporation for Assigned Names and Numbers (ICANN) gives large blocks to…
- Regional Internet Registries, such as the American Registry for Internet Names (ARIN), which give blocks to…
- Large institutions (ISPs), which give addresses to…
- Individuals and smaller institutions
- FAKE Example:
  - ICANN → ARIN → AT&T → UMICH → EECS

CIDR: Addresses allocated in contiguous prefix chunks

- Recursively break down chunks as get closer to host



FAKE example in more detail

- ICANN gave ARIN several /8s
- ARIN gave AT&T one /8, **12.0/8**
  - Network Prefix: **00001100**
- AT&T gave UMICH a /16, **12.34/16**
  - Network Prefix: **0000110000100010**
- UMICH gave EECS a /24, **12.34.56/24**
  - Network Prefix: **000011000010001000111000**
- EECS gave me specific address **12.34.56.78**
  - Address: **00001100001000100011100001001110**

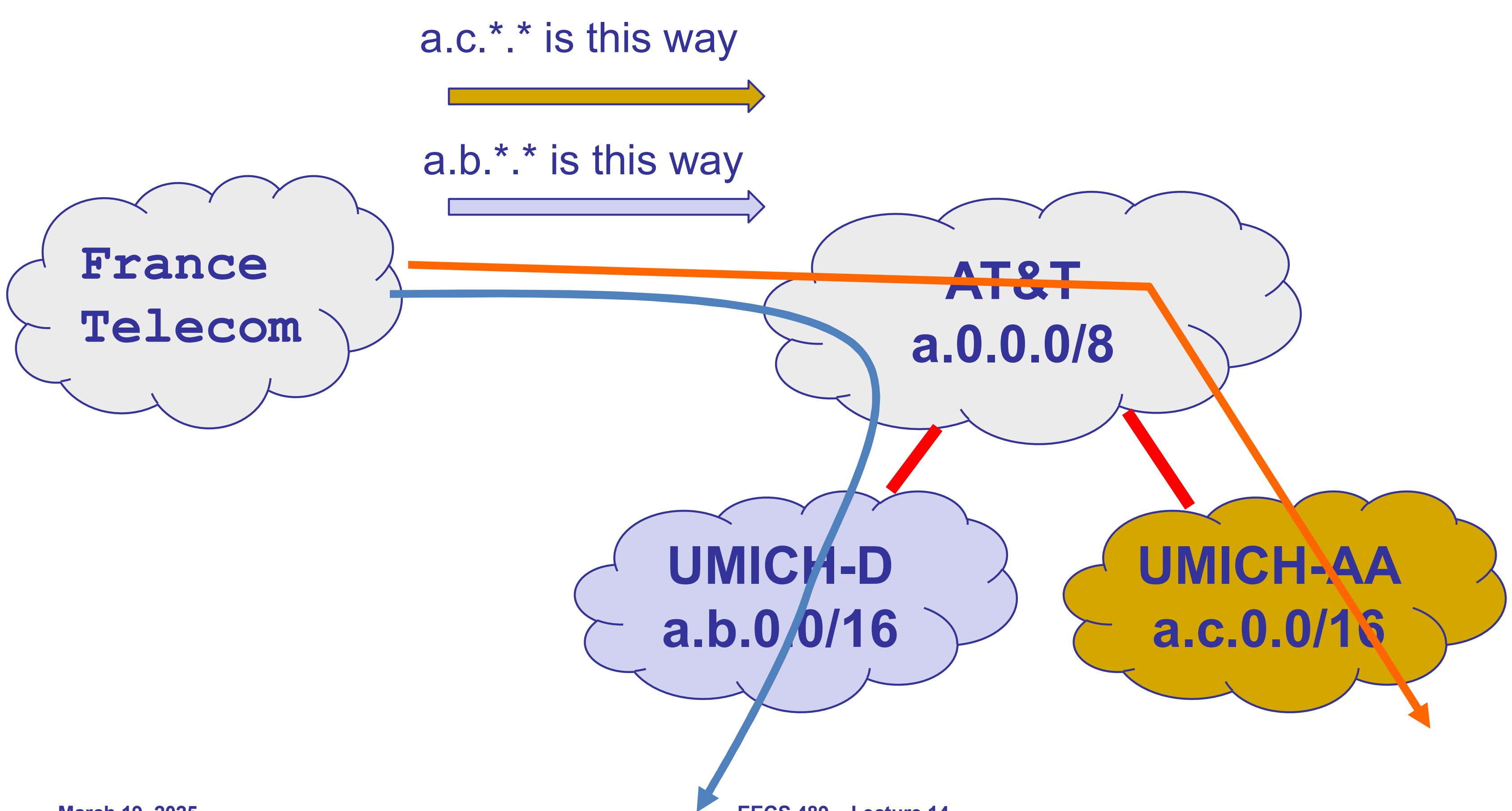
IP addressing is hierarchical

- Hierarchical address structure
- Hierarchical address allocation
- Hierarchical addresses and routing scalability

IP addressing → Scalable routing?

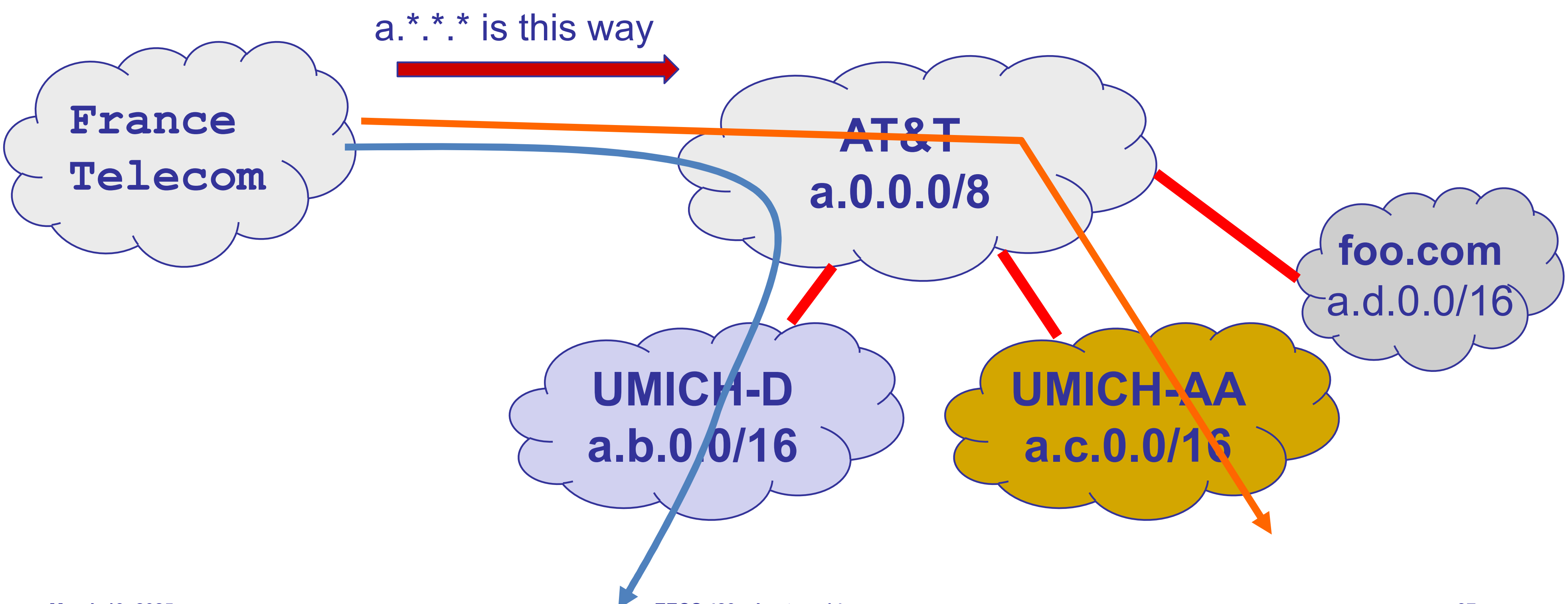
- Hierarchical address allocation only helps routing scalability if allocation matches topological hierarchy

IP addressing → Scalable routing?



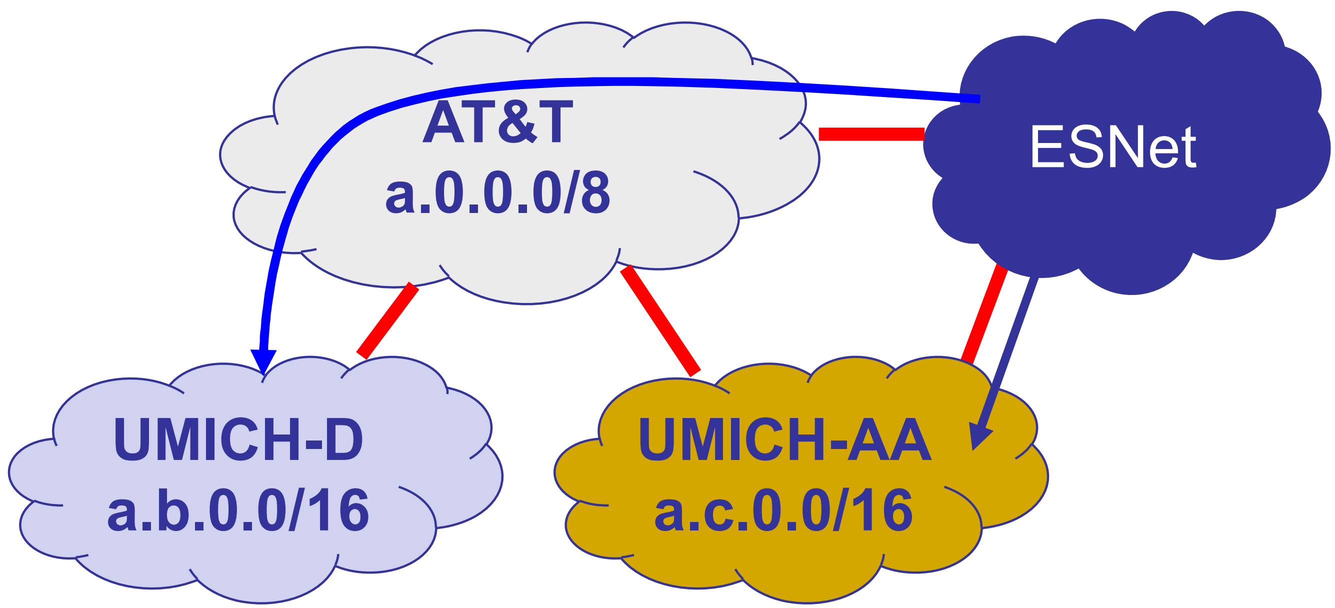
IP addressing → Scalable routing?

Can add new hosts/networks without updating the routing entries at France Telecom



IP addressing → Scalable routing?

ESNet must maintain routing entries for both a.*.*.* and a.c.*.*



IP addressing → Scalable routing?

- Hierarchical address allocation only helps routing scalability if allocation matches topological hierarchy
- May not be able to aggregate addresses for "multi-homed" networks
  - A multi-homed network is connected to more than one ASes for fault-tolerance, load balancing, etc.

BGP: BORDER GATEWAY PROTOCOL

BGP (Today)

- The role of policy
  - What we mean by it
  - Why we need it

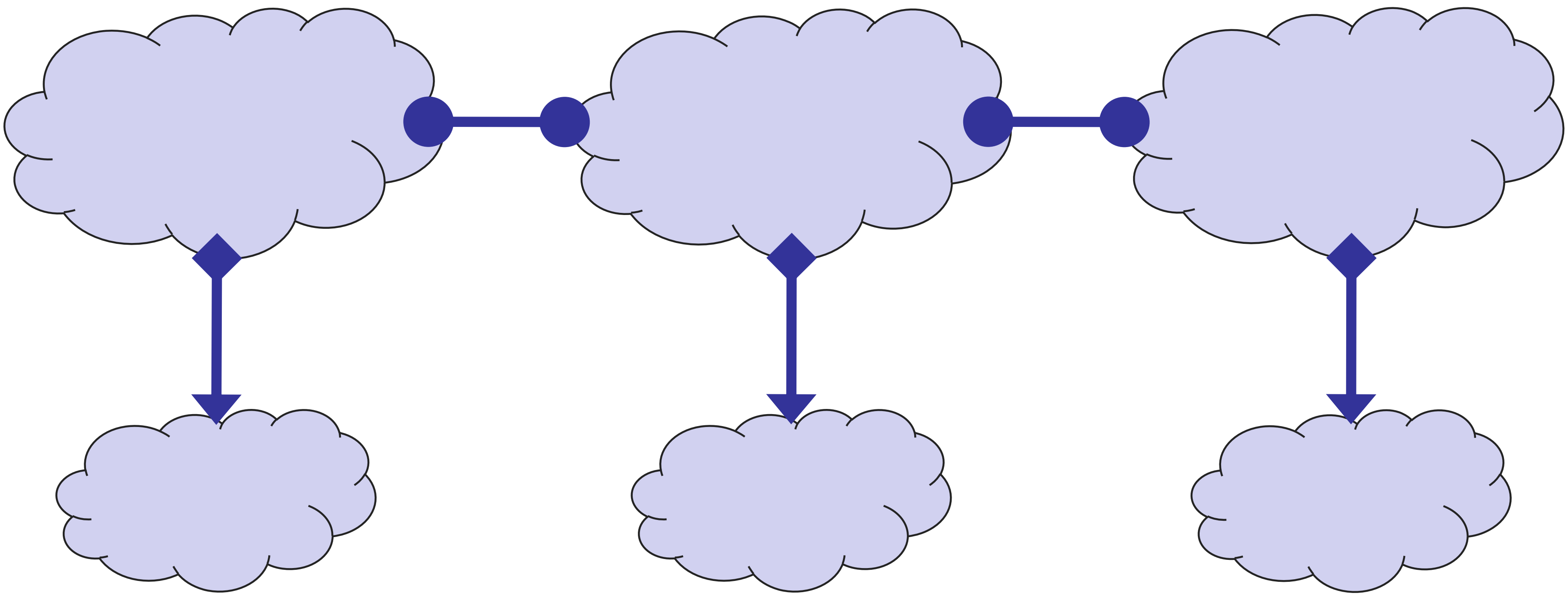
Administrative structure shapes Inter-domain routing

- ASes want freedom to pick routes based on policy
- ASes want autonomy
- ASes want privacy

Topology & policy shaped by inter-AS business relationship

- Three basic kinds of relationships between ASes
  - AS A can be AS B's customer
  - AS A can be AS B's provider
  - AS A can be AS B's peer
- Business implications
  - Customer pays provider
  - Peers don't pay each other
    - Exchange roughly equal traffic

Business relationships



Relations between ASes

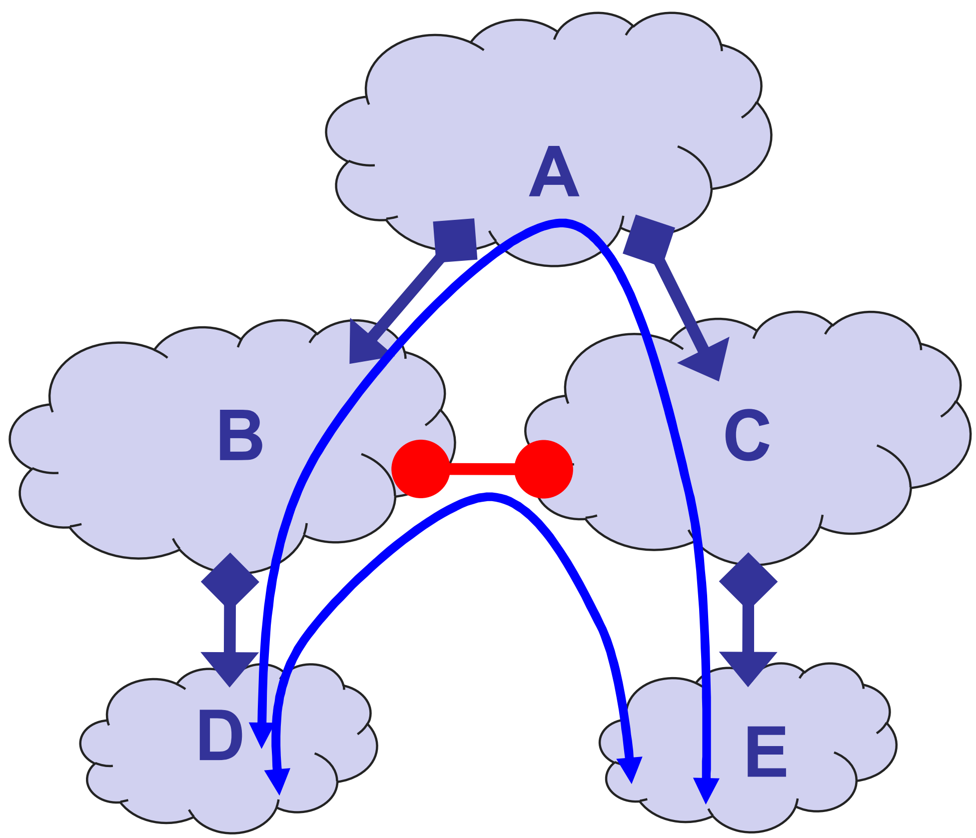
provider ◆——▶ customer

peer ●——● peer

Business implications

- Customers pay provider
- Peers don't pay each other

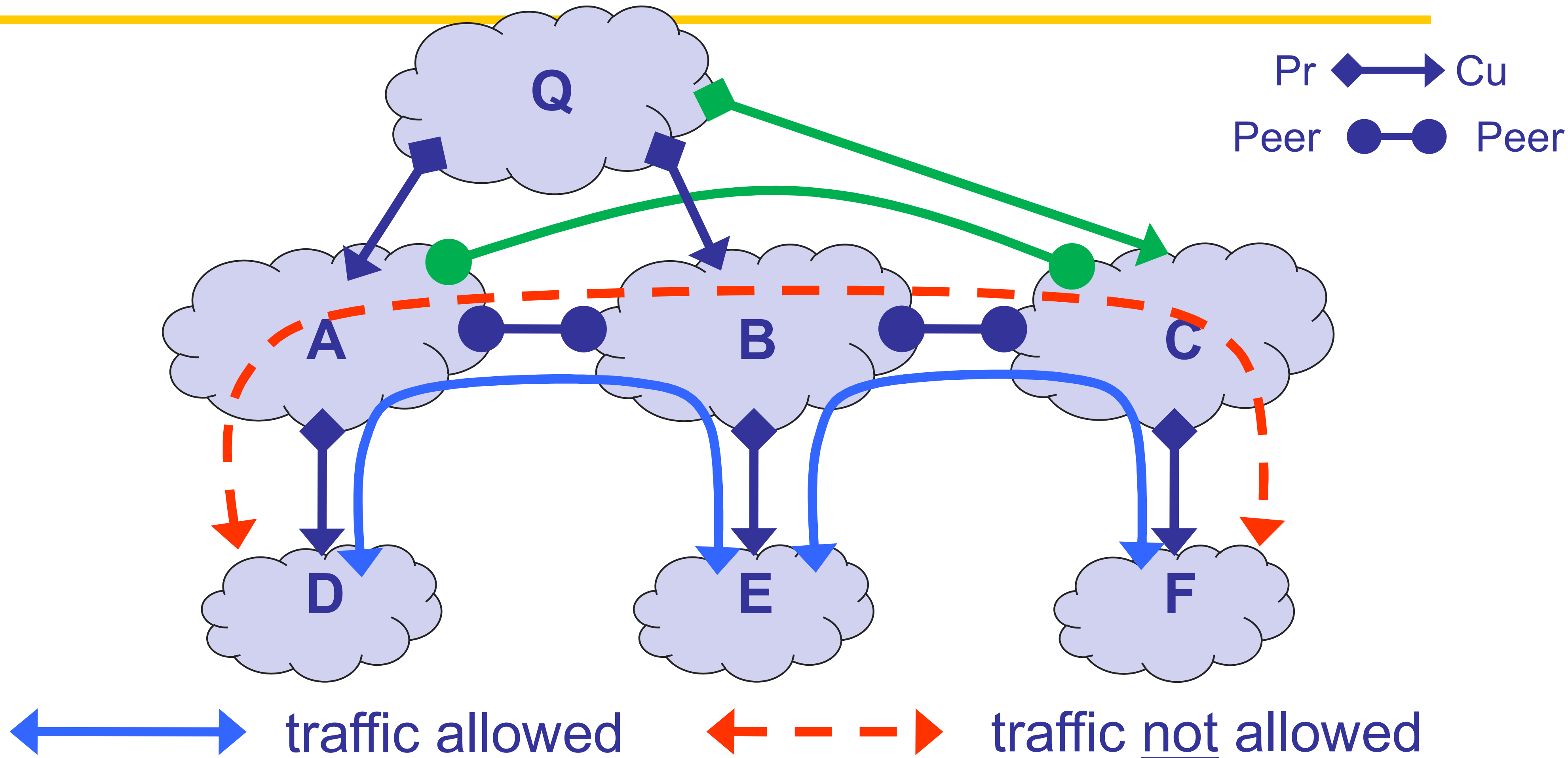
Why peer?



D and E communicate a lot

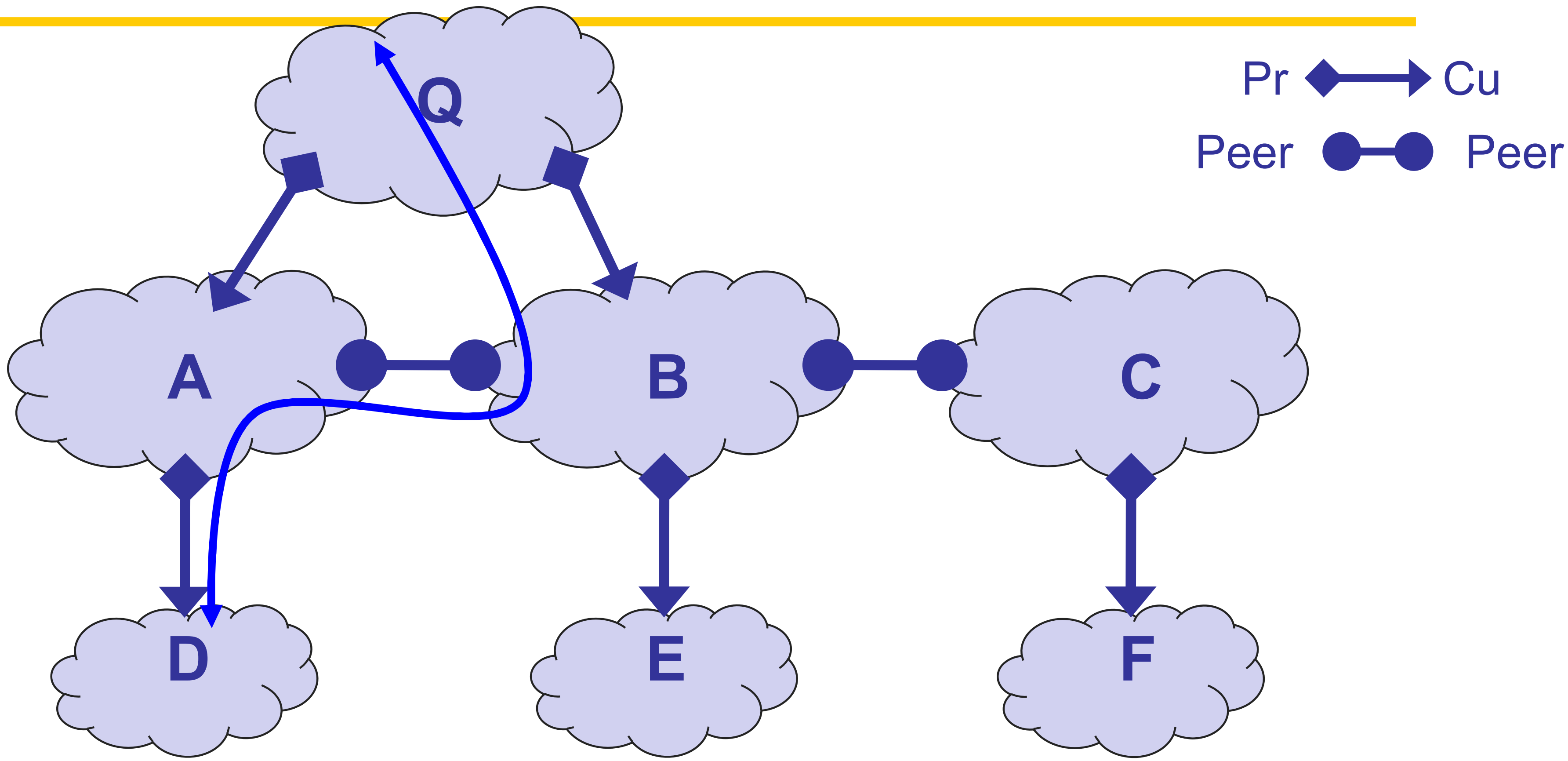
Peering saves B *and* C money

Routing follows the money!



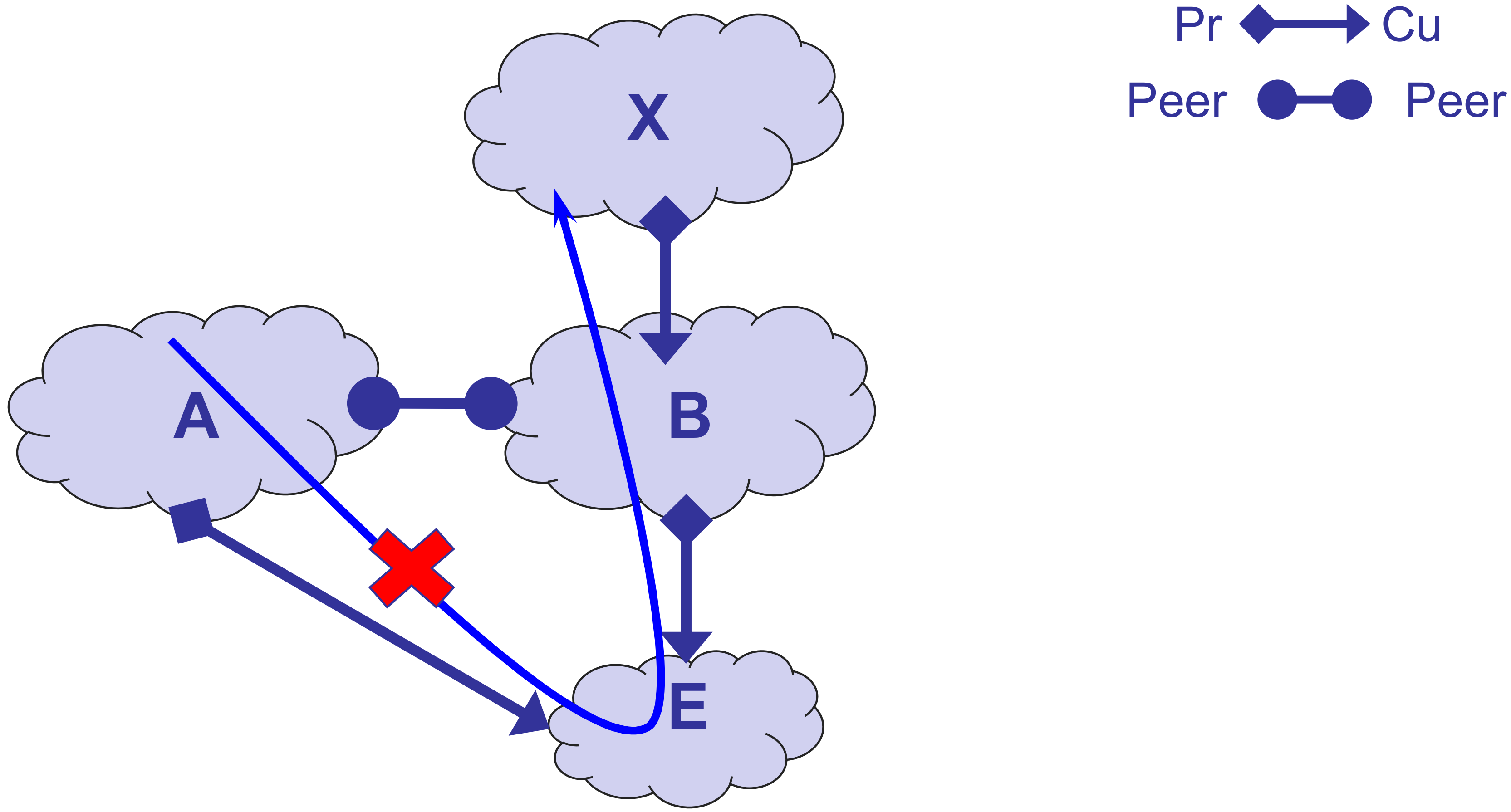
- ASes provide "transit" between their customers
- Peers do not provide transit between other peers

Routing follows the money!



- An AS only carries traffic to/from its own customers over a peering link

Routing follows the money!



- Routes are “valley” free (more details later)

In short

- AS topology reflects business relationships between ASes
- Business relationships between ASes impact which routes are acceptable

Summary

- Two key challenges in inter-domain routing
  - Scaling (Addressing)
  - Administrative structure (BGP)
    - Issues of autonomy, policy, privacy

- Next lecture: BGP policies, protocol, and challenges