CS4450

Computer Networks:
Architecture and Protocols

Lecture 13
Path-Vector Protocol (BGP)

Rachit Agarwal

Goals for Today's Lecture

- **Dive deeper into Inter-domain routing: Border-Gateway Protocol**
    - One of the most non-intuitive protocols
    - Driven by "business goals", rather than "performance goals"
        - I will try to provide as much intuition as possible
        - But, for the above reasons, BGP is one of the harder protocols
- Keep sanity: very different from everything we have seen so far

Recap from last lecture

Recap: Three requirements for addressing

- **Scalable routing**
  - How must state must be stored to forward packets?
    - Desired: Small #routing entries (less than one entry per host per switch)
  - How much state needs to be updated upon host arrival/departure?
    - Desired: Small #updates (less than one update per switch per host change)
- **Efficient forwarding**
  - How quickly can one locate items in routing table?
- **Host must be able to recognize packet is for them**

Recap: Using L2 (MAC) names does not enable scalable routing

- **Scalable routing**
  - How much state to forward packets?
    - One entry per host (at each switch)
  - How much state updated for each arrival/departure?
    - One entry per host (at each switch)
- **Efficient forwarding**
  - Exact match lookup on MAC addresses (exact match is easy!)
- **Host must be able to recognize the packet is for them**
  - MAC address does this perfectly

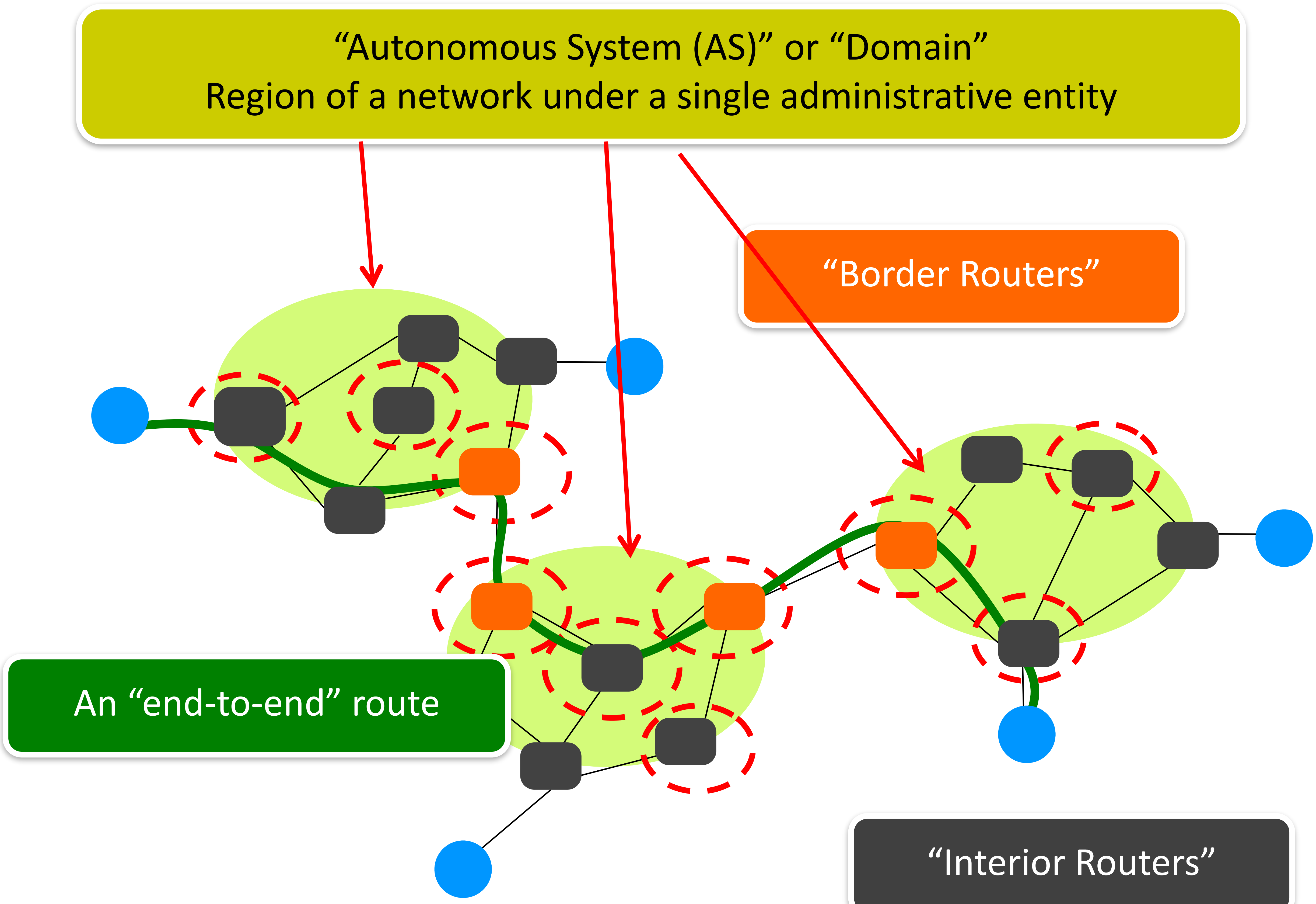
Recap: Today's Addressing (CIDR)

- Classless Inter-domain Routing
- Idea: Flexible division between network and host addresses
- Prefix is **network address**
- Suffix is **host address**
- **Example:**
    - **128.84.139.5/23 is a 23 bit prefix with:**
    - First 23 bits for network address
    - Next 9 bits for host addresses: maximum 2^9 hosts
    - **All hosts within the network have the same first 23 bits (x.y.z.*)**
- **Terminology: "Slash 23"**

Recap: How does CIDR meet our requirements?

- To understand this, we need to understand the routing on the Internet
- And to understand that, we need to understand the Internet

Recap: What does a computer network look like?

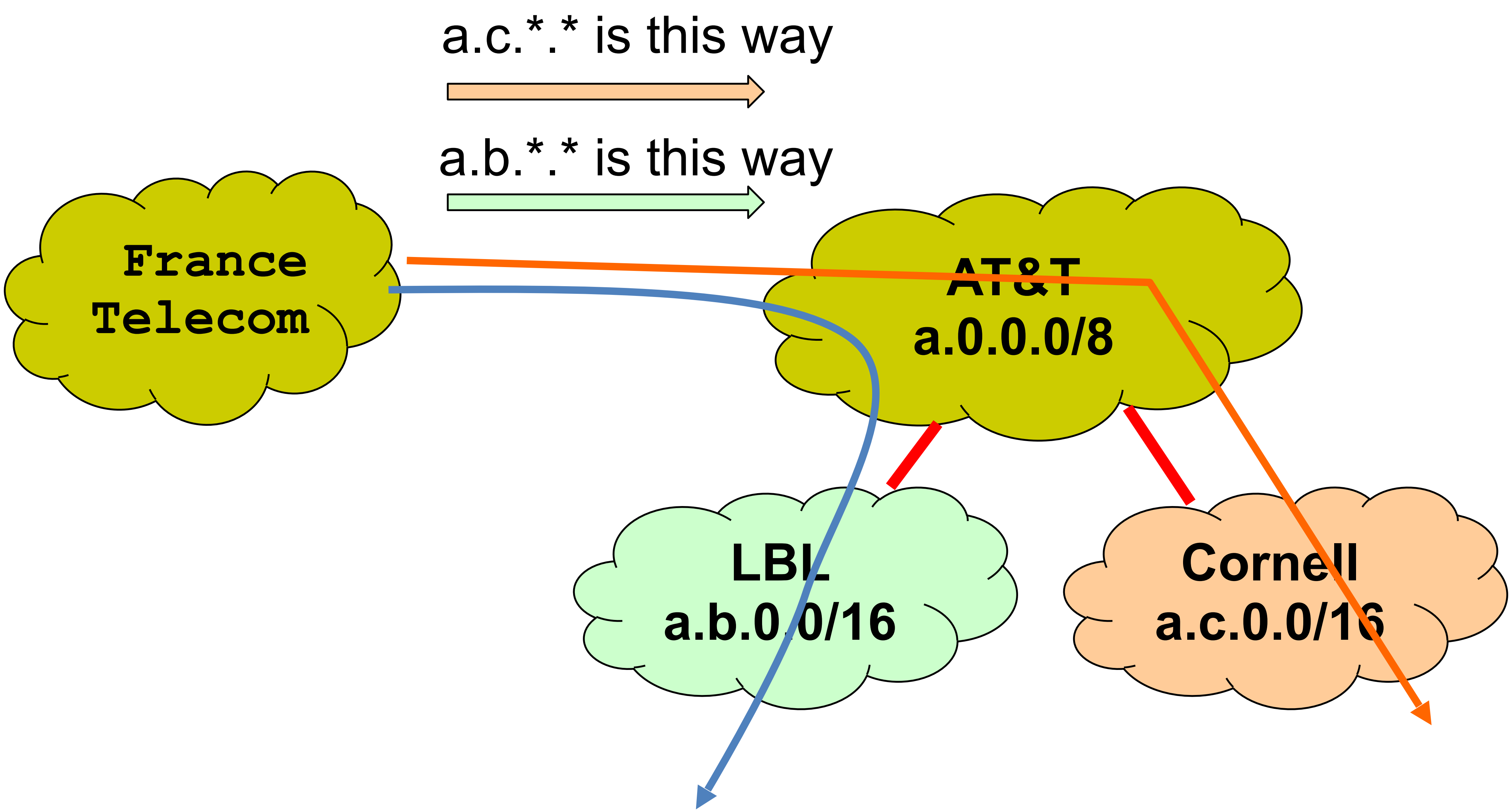


Recap: Autonomous Systems (AS)

- An AS is a network under a single administrative control
  - Currently over 30,000
  - **Example: AT&T, France Telecom, Cornell, IBM, etc.**
  - A collection of routers interconnecting multiple switched Ethernets
  - And interconnections to neighboring ASes
- Sometimes called “Domains”

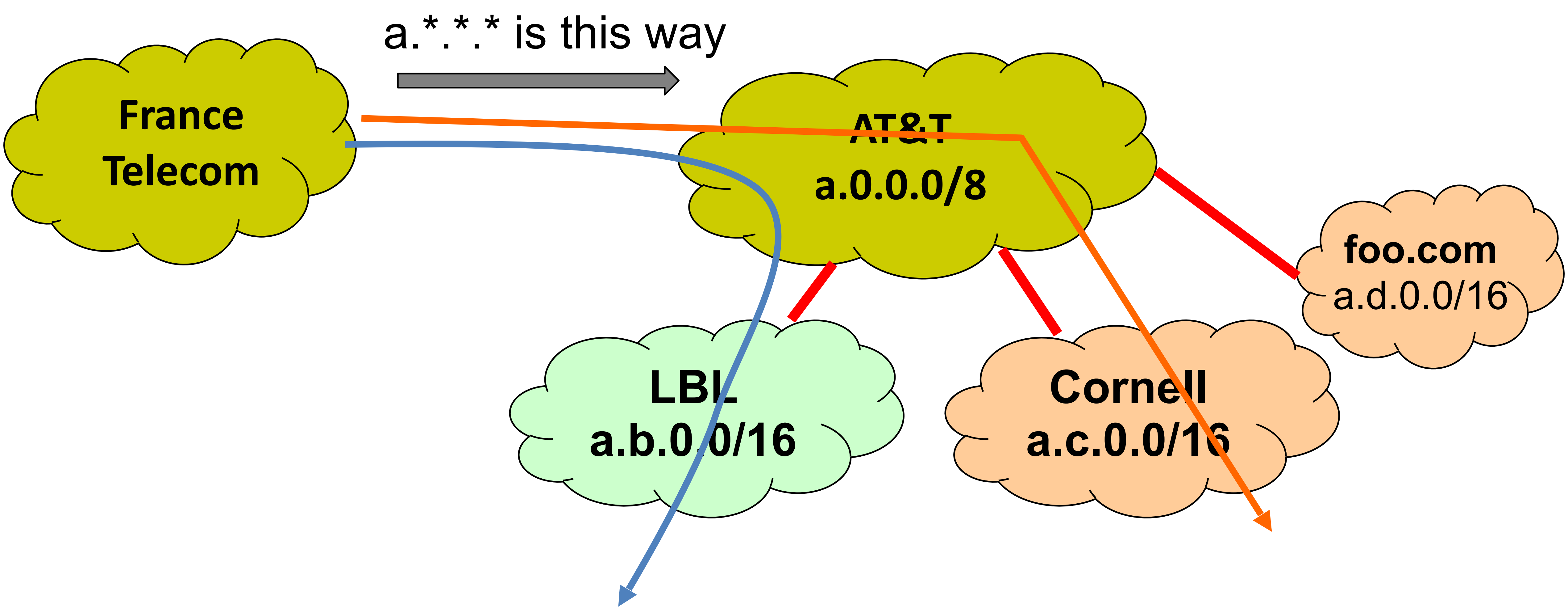
IP addressing

Intuition: IP addressing -> Scalable Routing?



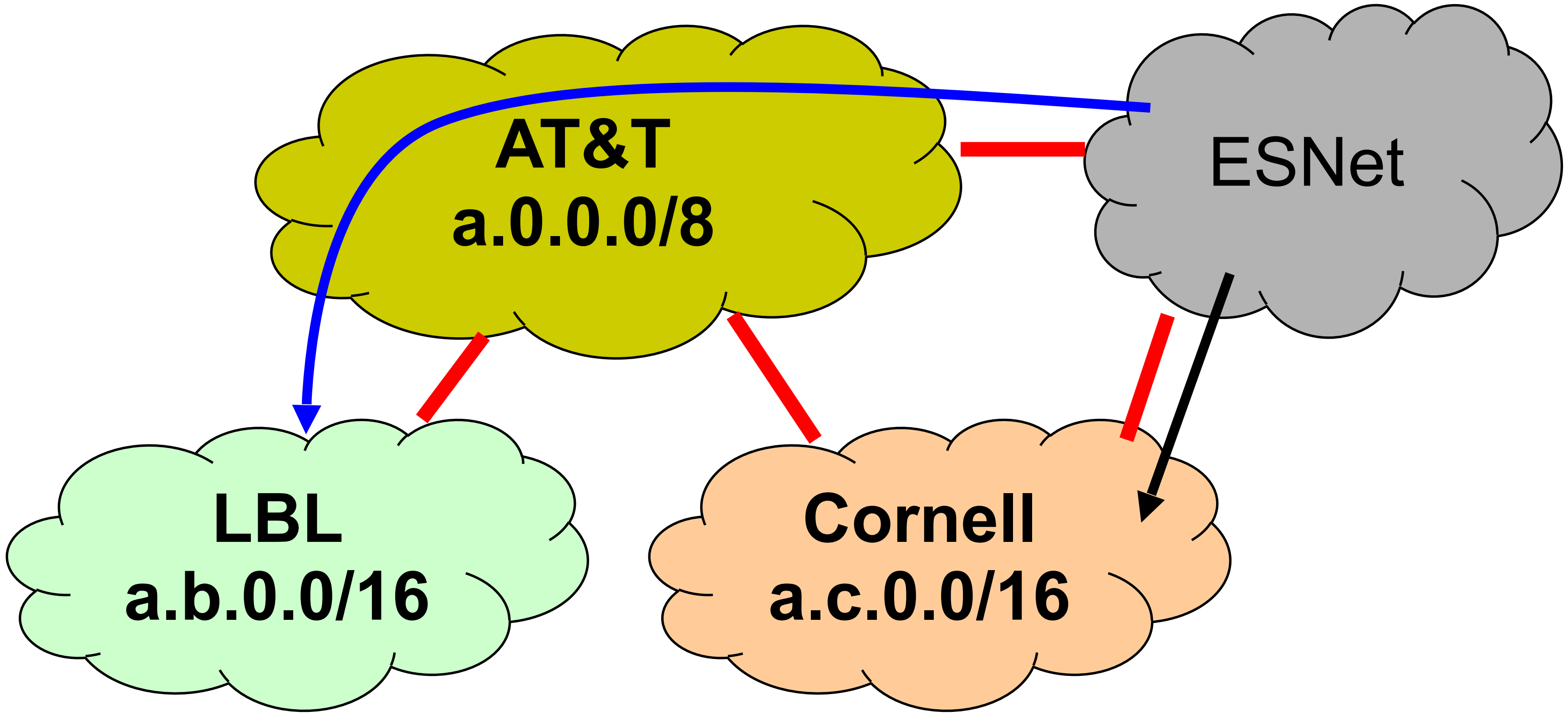
Intuition: IP addressing -> Scalable Routing?

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



Intuition: IP addressing -> Scalable Routing?

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



Given this addressing,

How do we think about Inter-domain routing protocols?

Administrative Structure Shapes Inter-domain Routing

- ASes want freedom to pick routes based on policy
  - *"My traffic can't be carried over my competitor's network!"*
  - *"I don't want to carry A's traffic through my network!"*
  - Cannot be expressed 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 (LS) vs. Distance Vector (DV)

- LS offers no privacy — broadcasts all network information
- LS limits autonomy — need agreement on metric, algorithm

- DV is a decent starting point
  - Per-destination updates by intermediate nodes give us a hook
  - But, wasn't designed to implement policy
  - … and is vulnerable to loops if shortest paths not taken

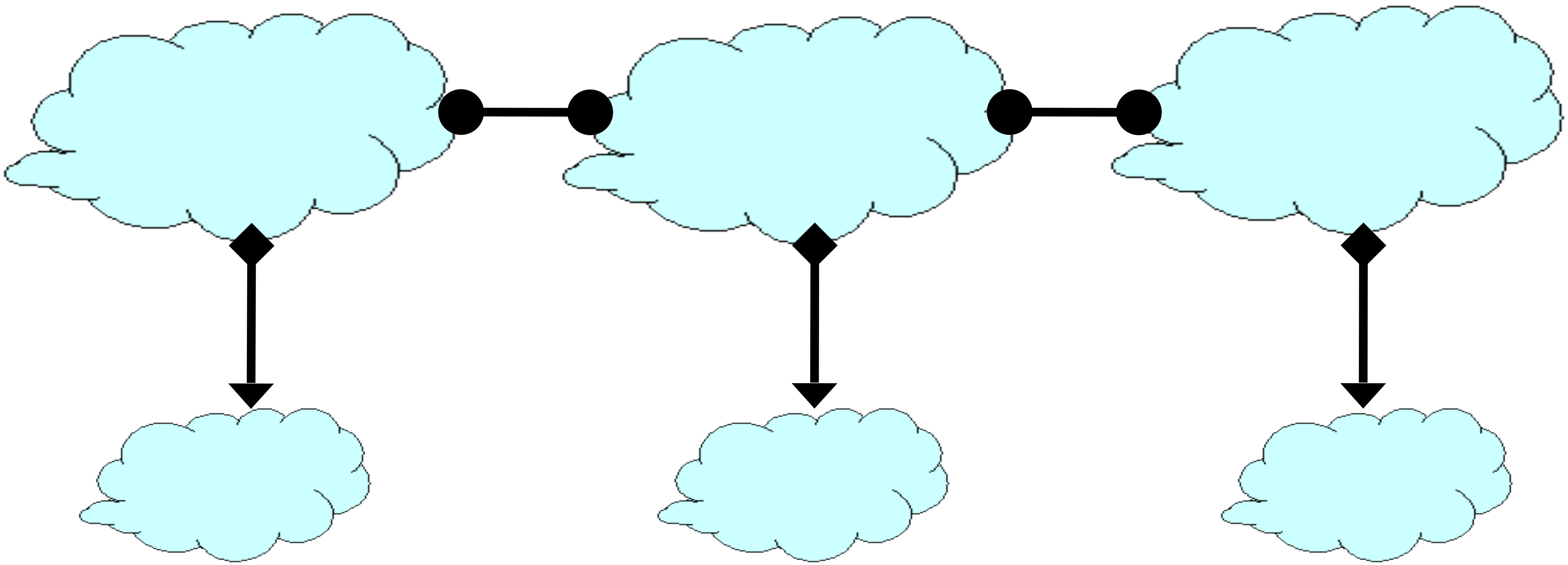
The "Border Gateway Protocol" (BGP) extends Distance-Vector ideas to accomodate policy

Business Relationships Shape Topology and Policy

- 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

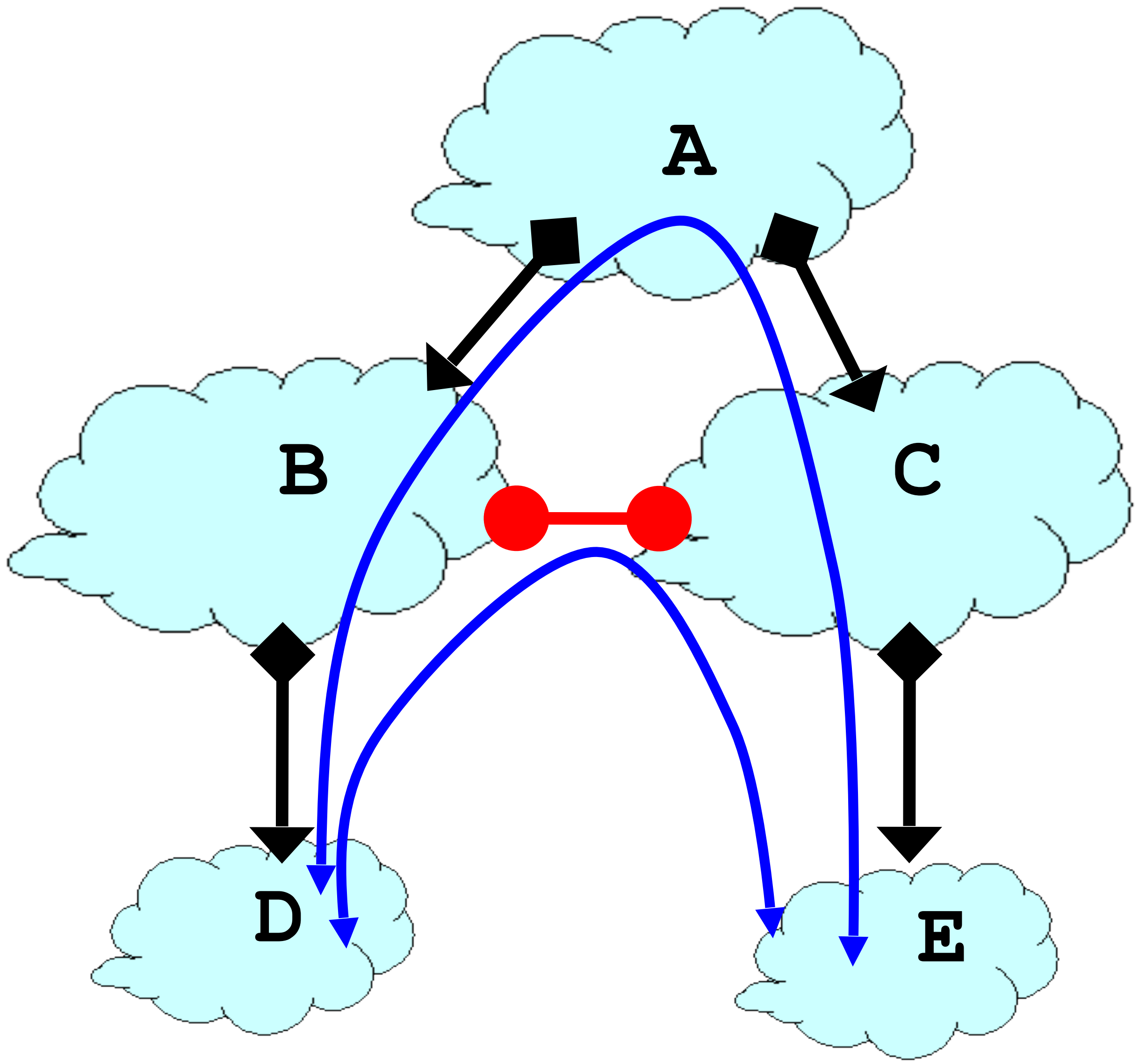
provider ◆———▶ customer

peer ●———● peer

Business Implications

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

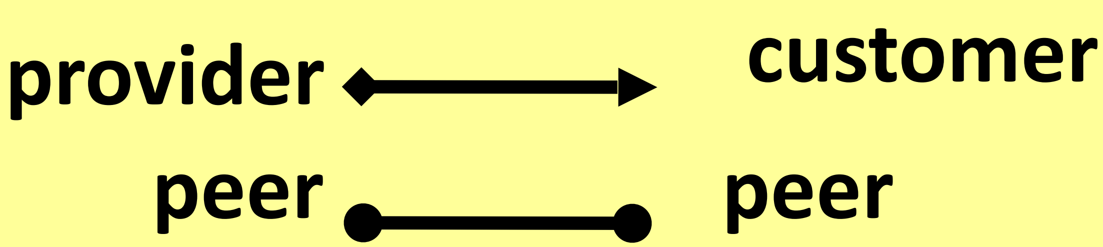
Why Peer?



E.g., D and E talk a lot

Peering saves B *and* C money

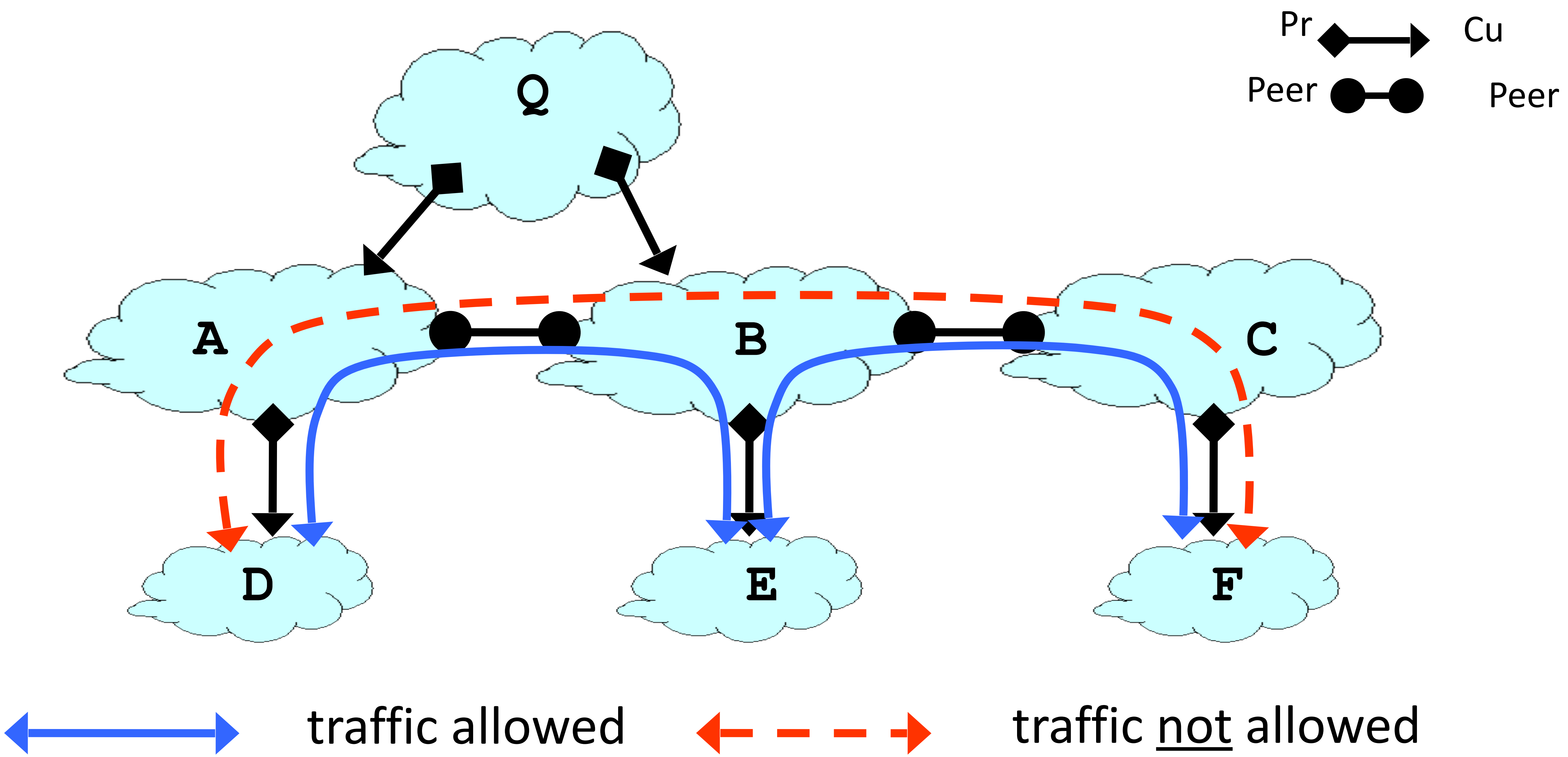
Relations between ASes



Business Implications

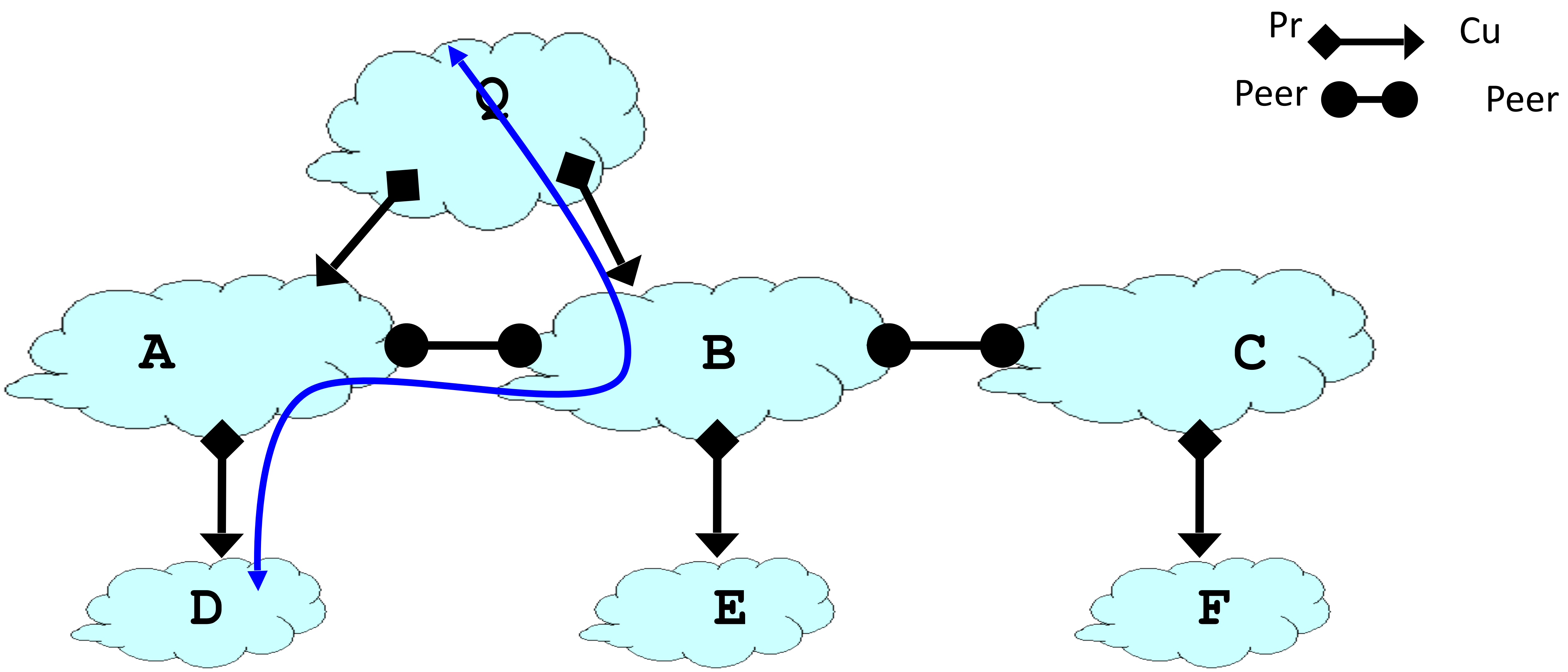
- Customers pay provider
- Peers don't pay each other

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