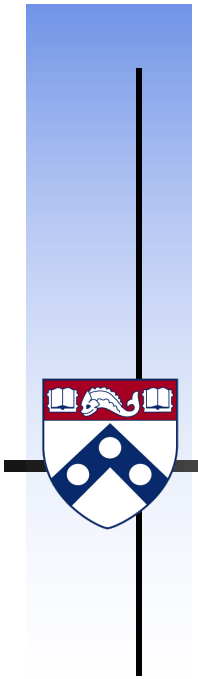


CIS 5530: Networked Systems

Interdomain Routing

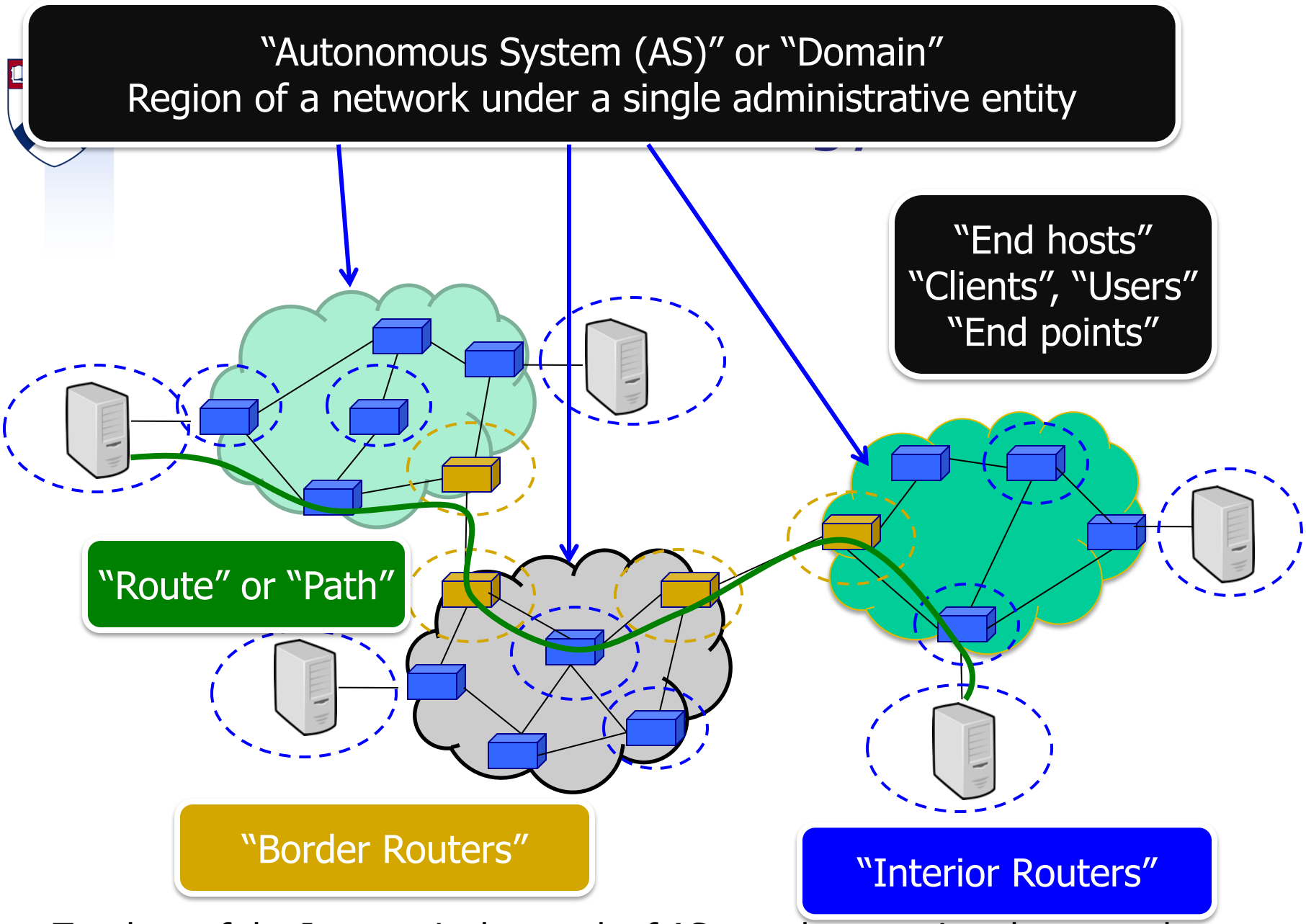
February 8, 2023

Agenda

- Intradomain Routing ✓
  - Distance Vector ✓
  - Link State ✓
- Interdomain Routing NEXT
  - Structure of the Internet
  - Policy
  - BGP
  - Issues with BGP

Link State Versus Distance Vector

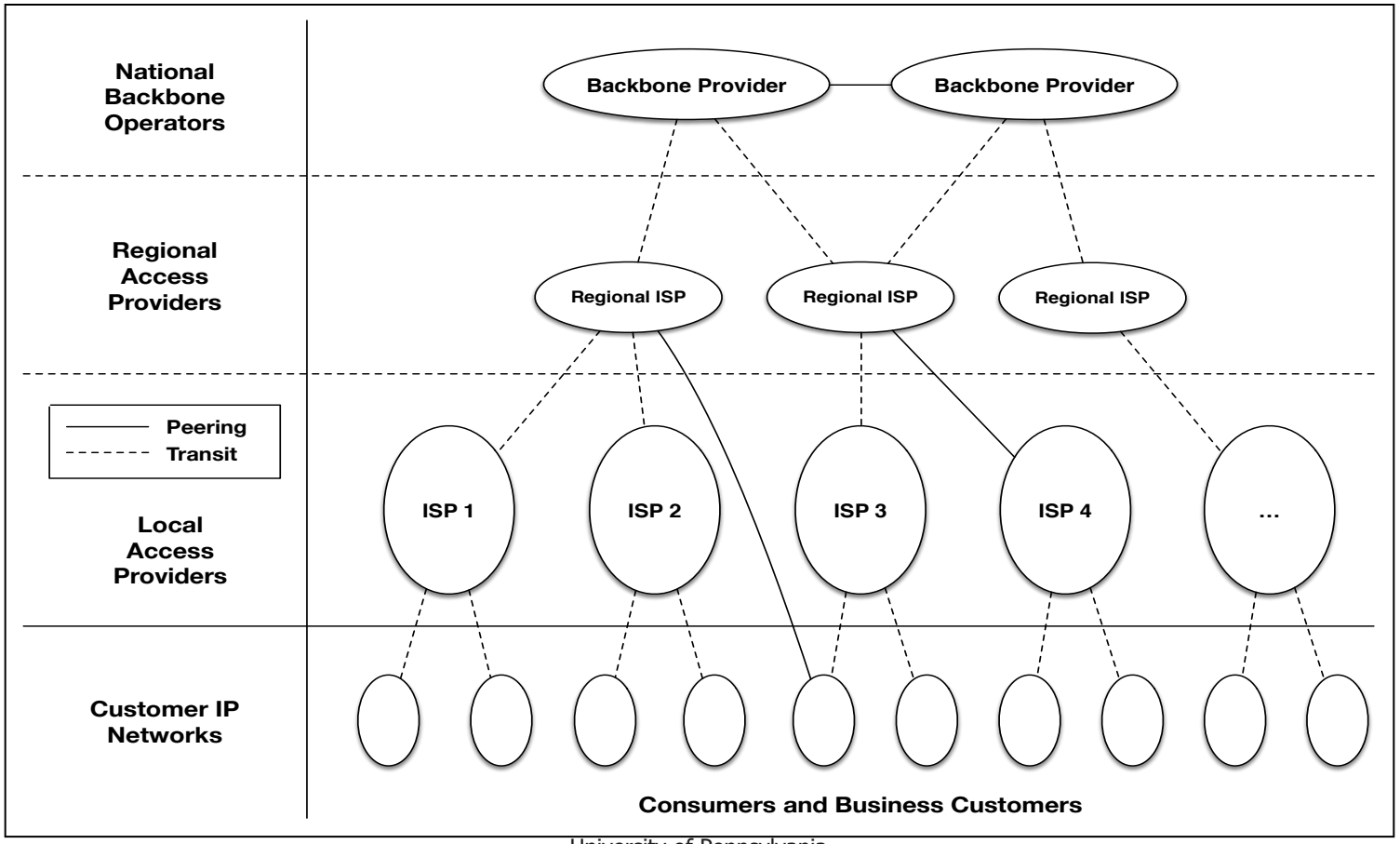
- Storage:
  - LS: Store all links (entire network)
  - DV: Entry for each possible destination/next-hop
- Convergence:
  - LS: Reacts more quickly, in bounded time to connectivity changes
  - DV: Count-to-infinity problem. Slower convergence. Bounded path length (16)
- Global policies:
  - LS: Able to impose global policies in a globally consistent way
  - DV: Harder, since do not have complete network topology. Up next: see a variant called Path Vector used in BGP



Topology of the Internet is the graph of ASes and connections between them

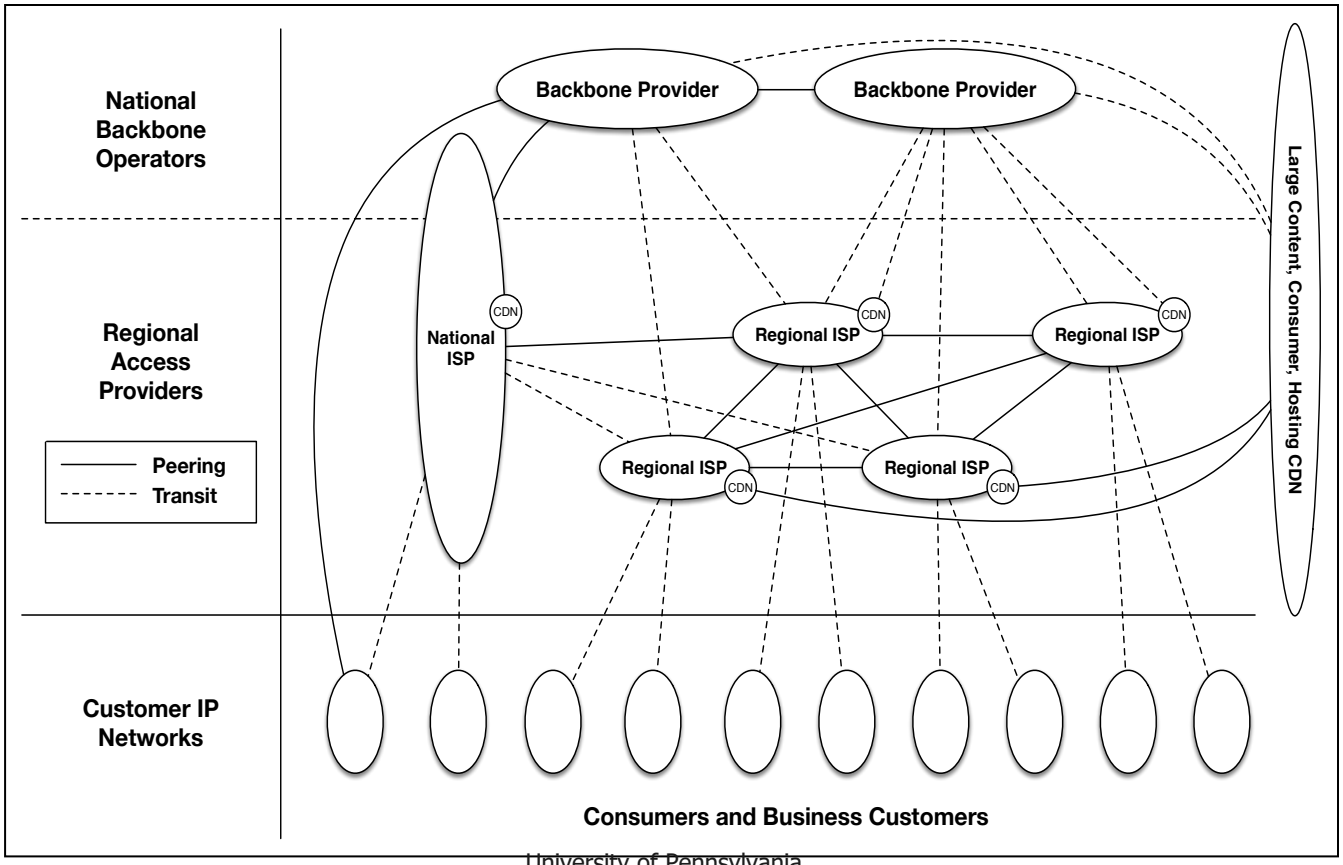
Interconnection Circa 1995

The backbone eventually transitioned from a single government-operated backbone to a federated backbone model comprised of multiple commercial network operators.



Interconnection Today

Interconnection today has evolved into a complex amalgam of models incorporating new connectivity options, delivery options, traffic management requirements and business practices. ("Flattening" of the hierarchy.)



Inter-domain routing: between ASes

- 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
  - Scaling limitations
  - No privacy – broadcasts all network information
  - Limited autonomy – needs agreement on metric, algo
- Distance-vector is what they chose
  - Per-destination updates give some control
  - BUT wasn't designed to implement policy
  - AND is vulnerable to loops

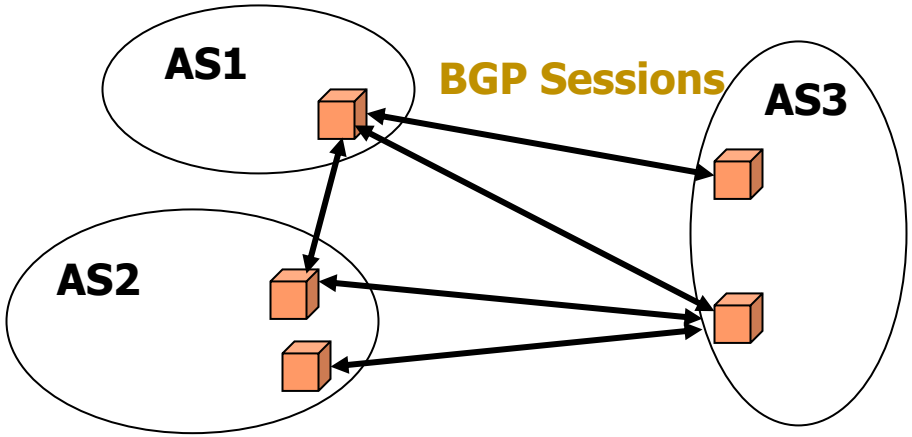
- The "Border Gateway Protocol" (BGP) extends distance-vector to accommodate policy

Path Vector Protocol

- Distance vector algorithm with extra information
  - For each route, store the complete path (ASs)
  - No extra computation, just extra storage
  - Reachability only, no metrics
- Advantages:
  - Can make policy choices based on set of ASs in path
  - Can easily avoid loops

Advertising Routes

- One router can participate in many BGP sessions.
- Initially … node advertises ALL routes it wants neighbor to know (could be > 50K (closer to 200K) routes)
- Ongoing … only inform neighbor of changes



Basic Messages In BGP

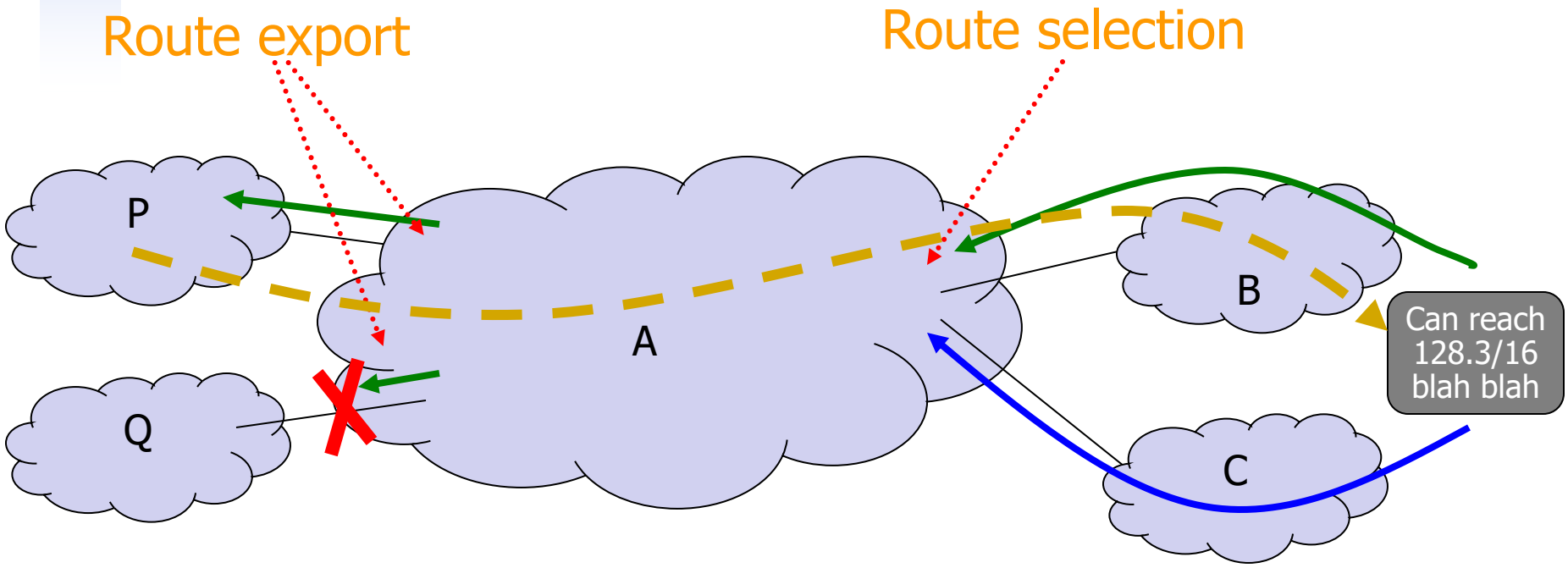
- Open:
  - Establishes BGP session (uses TCP port #179)
  - BGP uses TCP (error control, congestion)
- Update:
  - Inform neighbor of new routes that become active
  - Inform neighbor of old routes that become inactive
- Keepalive:
  - Inform neighbor that connection is still viable

Agenda

- Interdomain Routing 
  - Structure of the Internet 
  - Policy 

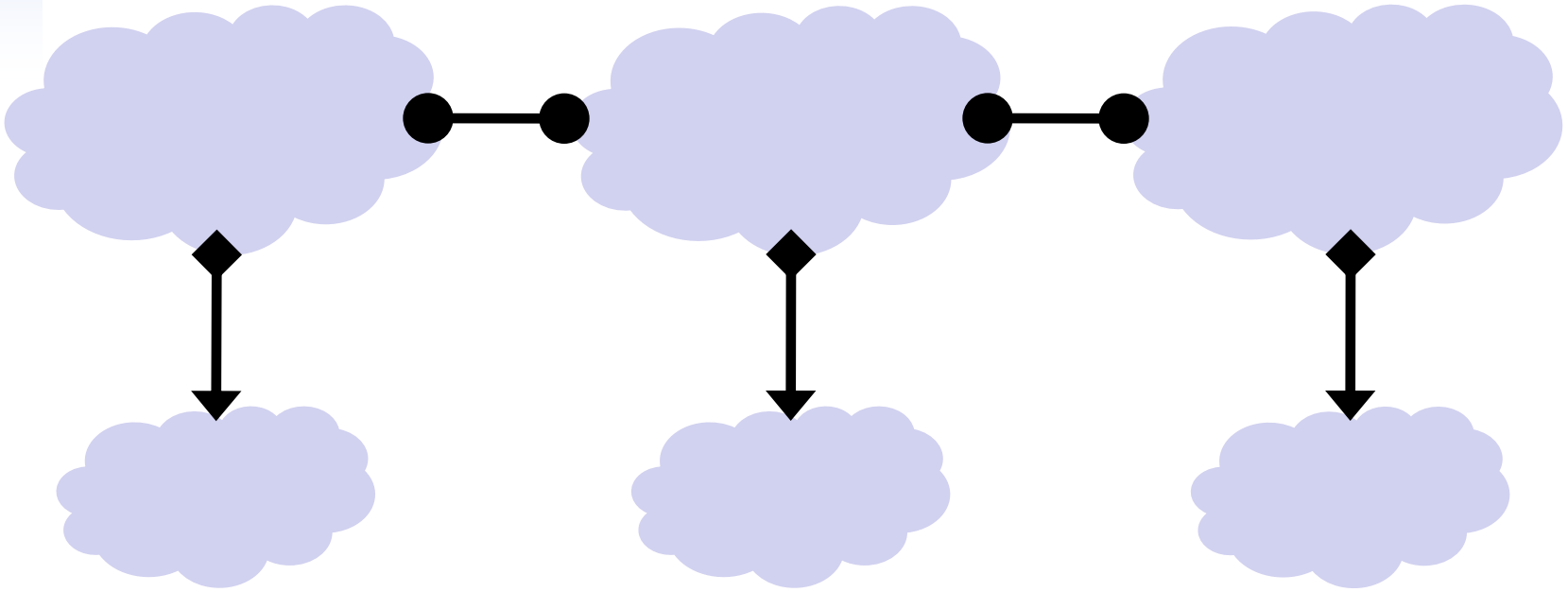
  - BGP
  - Issues with BGP

Policy dictates how routes are "selected" and "exported"



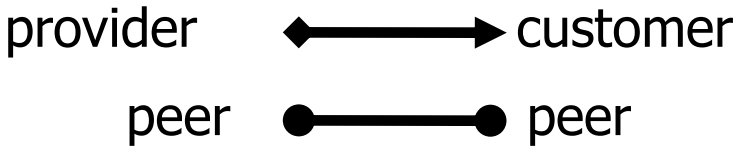
- Selection: Which path to use?
  - Controls whether/how traffic leaves the network
- Export: Which path to advertise?
  - Controls whether/how traffic enters the network

Topology & policy shaped by inter-AS business relationship

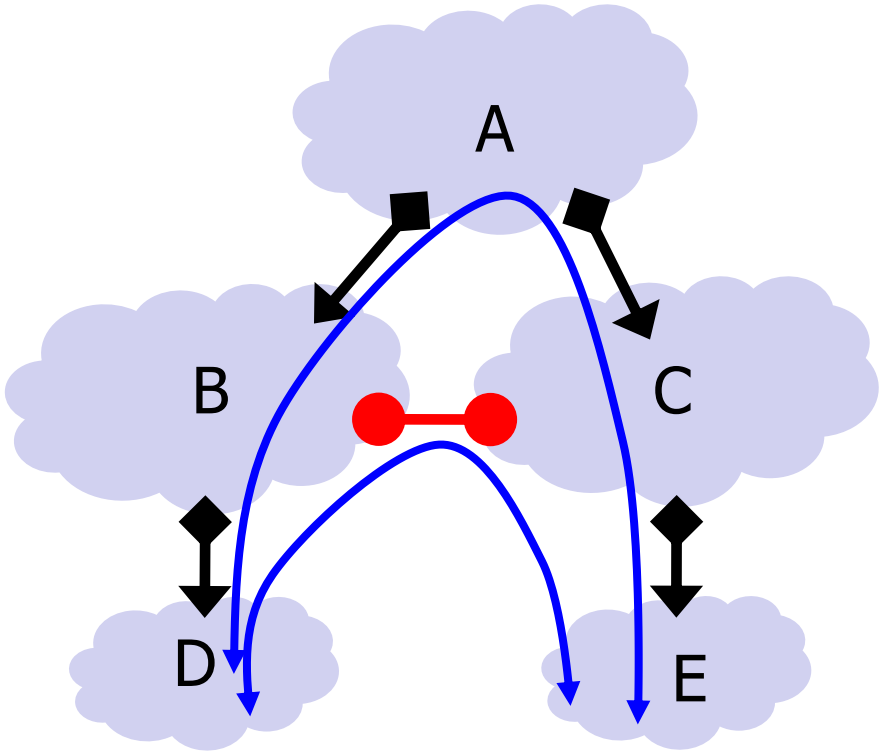


Business implications

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



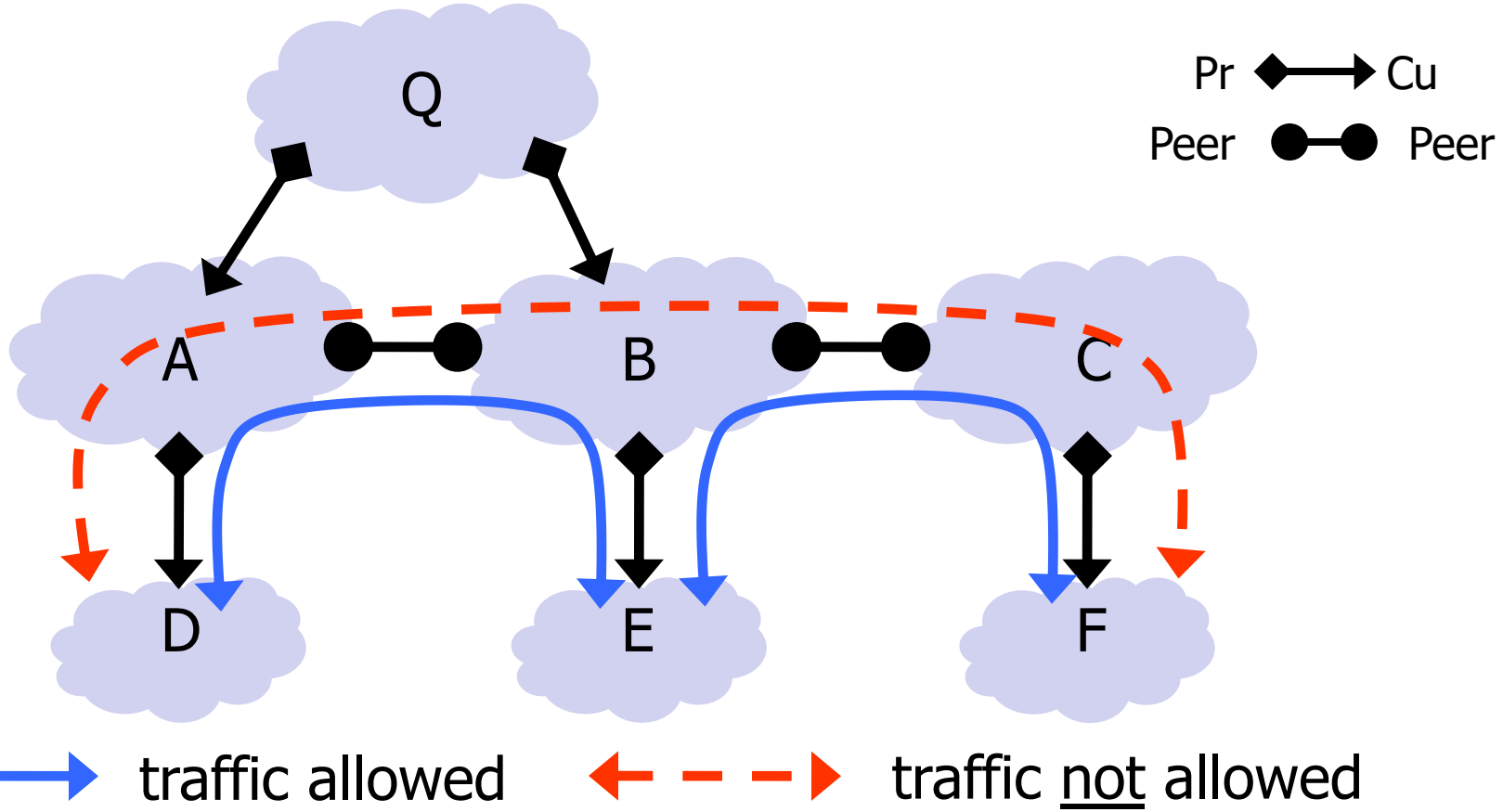
Why peer?



D and E communicate a lot

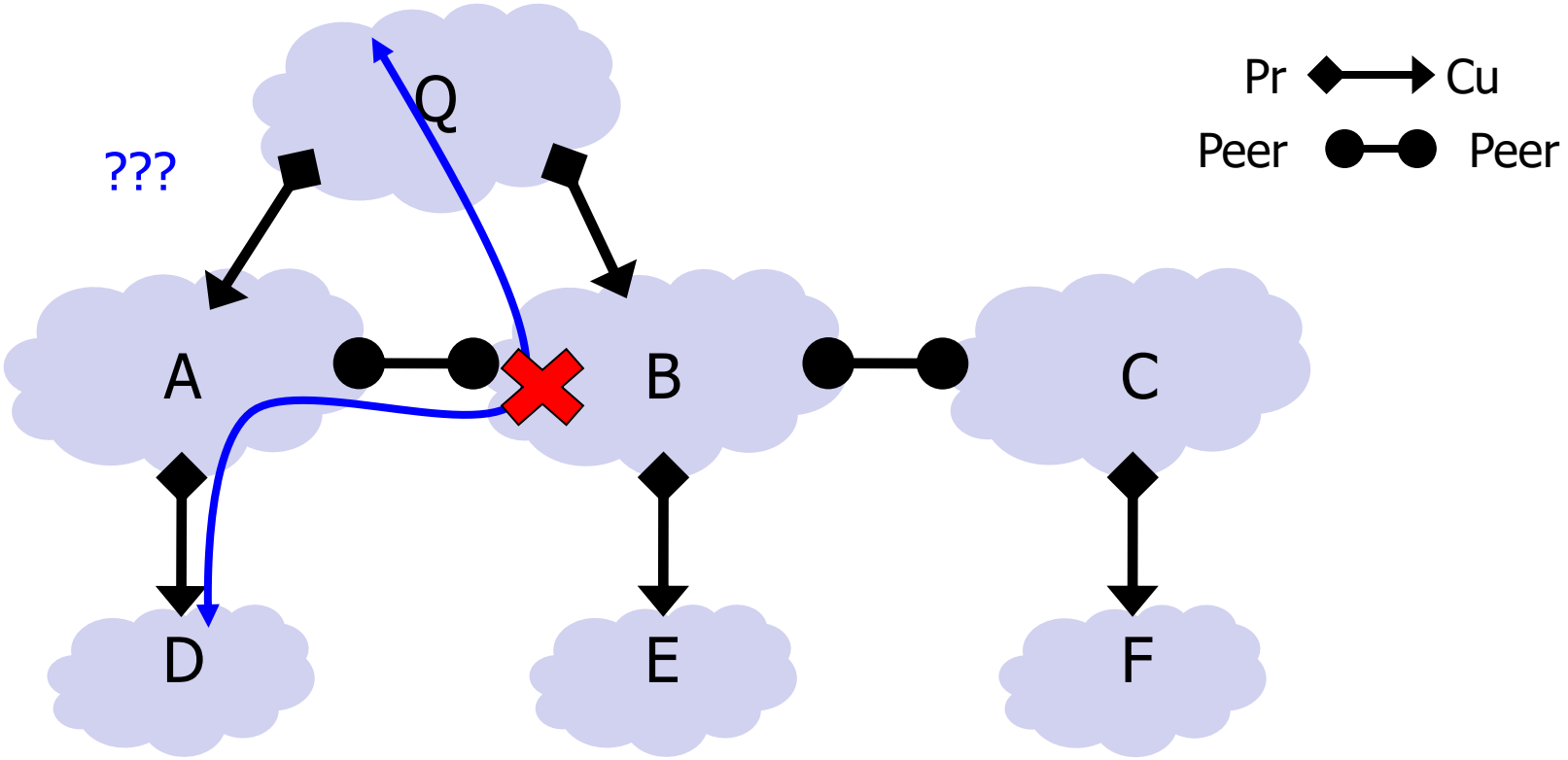
Peering saves B *and* C money

Routing follows the money!



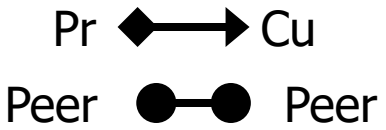
- ASes provide “transit” between their customers
- Peers do not provide transit between other peers

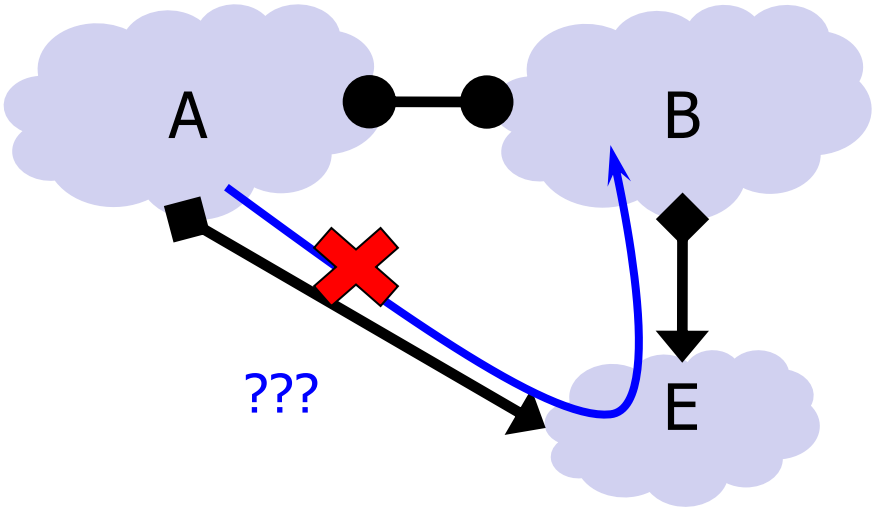
Routing follows the money!



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

Routing follows the money!





2. 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

Typical *selection* policy

- In decreasing order of priority
  - Make/save money (send to customer > peer > provider)
  - Maximize performance (smallest AS path length)
  - Minimize use of my network bandwidth ("hot potato")
  - …

Typical *export* policy

Destination prefix advertised by…	**Export route to…**
Customer	Everyone (providers, peers, other customers)
Peer	Customers
Provider	Customers

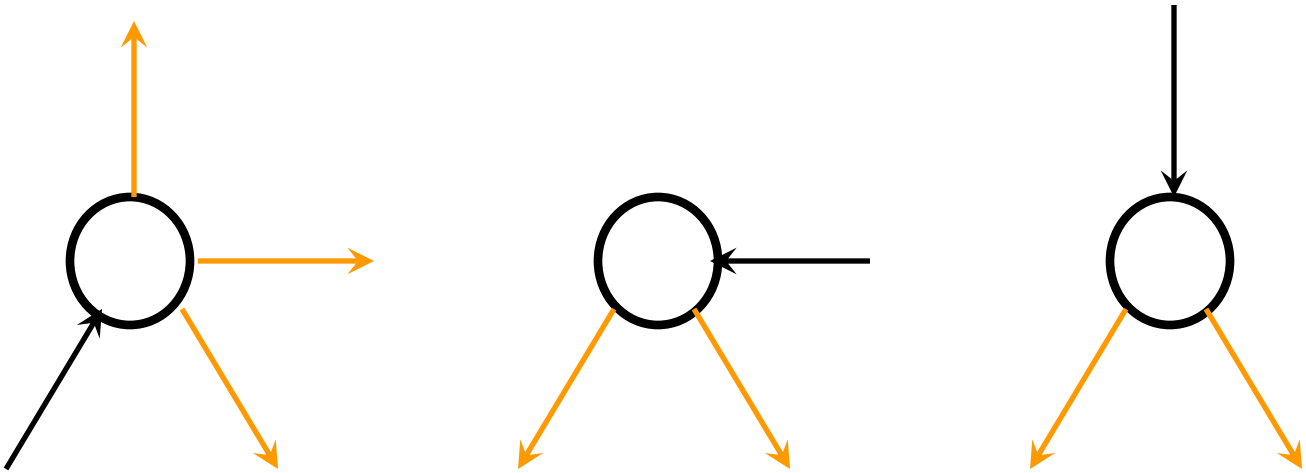
We'll refer to these as the "Gao-Rexford" rules (capture common – but not required! – practice)

Gao-Rexford Rules

Providers

Peers

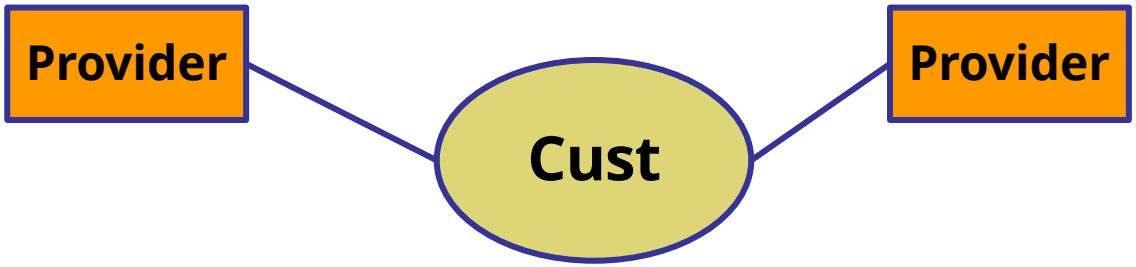
Customers



With Gao-Rexford, the AS policy graph is a DAG (directed acyclic graph) and routes are "valley free"

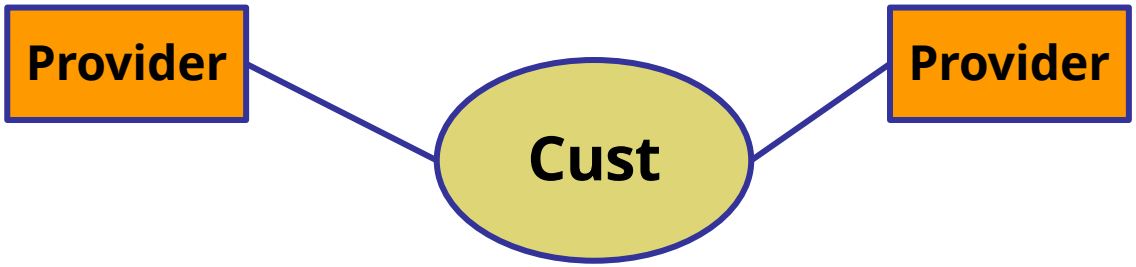
Multi-homing

- Connect to multiple providers for reliability, load sharing
- Customer can choose the best outgoing path from any of the announcements heard from its providers
  - Easy to control outgoing traffic, e.g, for load balancing



Multi-homing

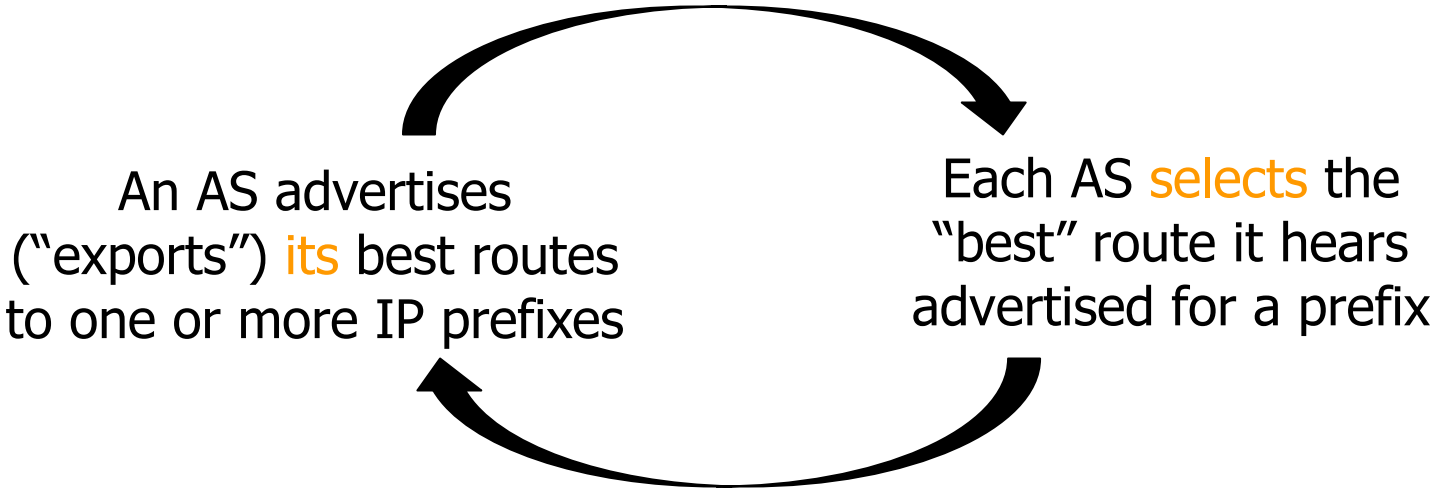
- Less control over what paths other parties will use to reach us
  - Both providers will announce that they can reach to the customer
  - Rest of Internet can choose which path to take to customer
    - Hard for the customer to influence this



Agenda

- Interdomain Routing ✓
  - Structure of the Internet ✓
  - Policy ✓
  - BGP NEXT
  - Issues with BGP

BGP: Basic idea



- Like DV:
  - Per-destination route advertisements
  - No global sharing of network topology information
  - Iterative and distributed convergence on paths
- With four crucial differences!