Routing (part 3)

Lecture 25

http://www.cs.rutgers.edu/~sn624/352-F24

Srinivas Narayana

RUTGERS
UNIVERSITY | NEW BRUNSWICK

The network layer enables reachability. We'll see protocols that solve subproblems.

How does an endpoint get an address?

DHCP

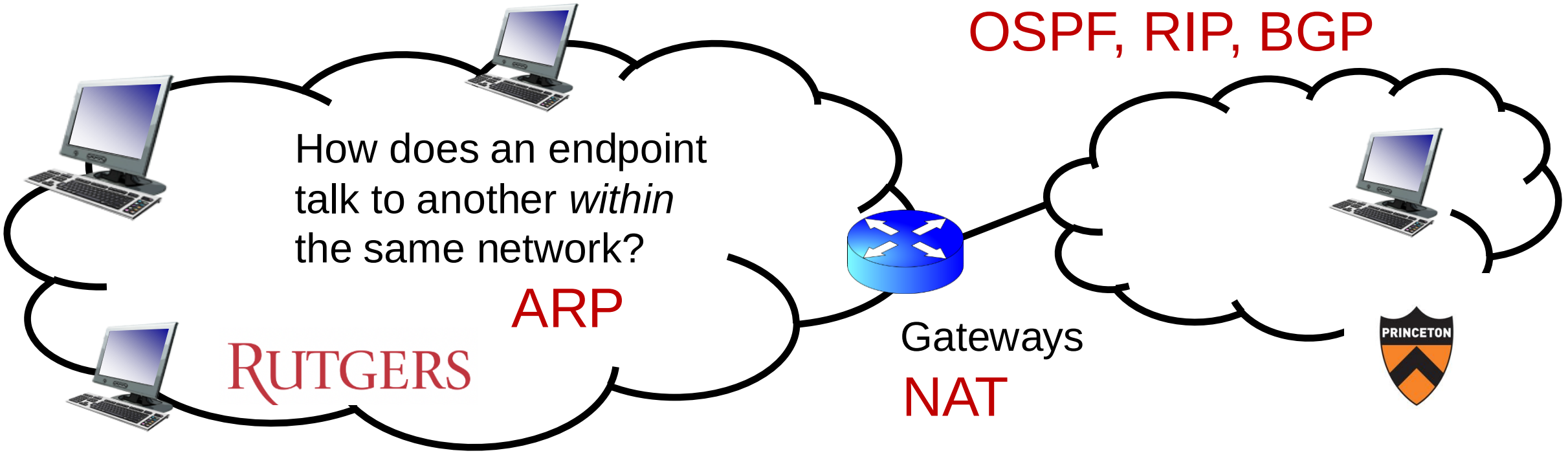
Debugging?

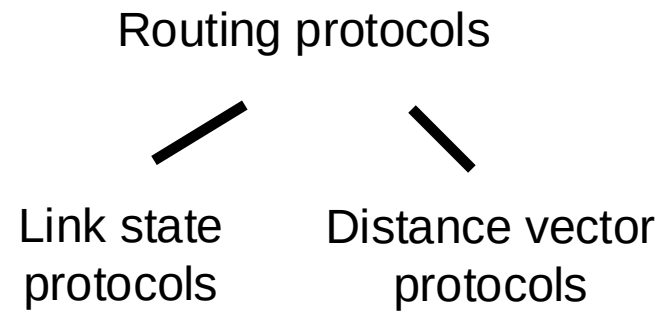
ICMP

How does an endpoint talk to another *outside* its network?

Routing protocols

OSPF, RIP, BGP





Every router is aware of the existence of every other router.

Messages reveal information on the full network (graph) structure.

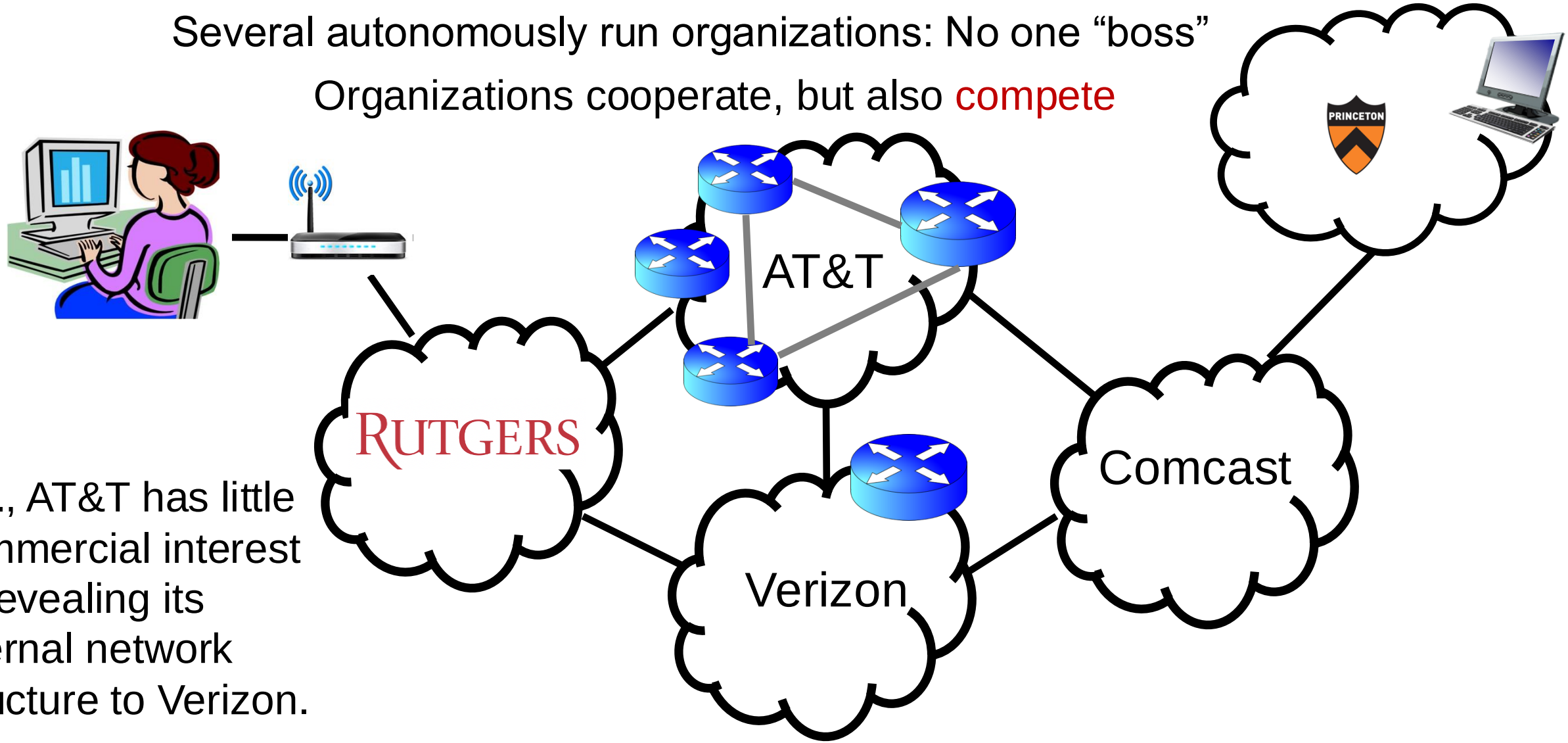
Message exchange and forwarding tables scale with network size.

These assumptions/settings cannot work on the Internet.

The Internet is a large federated network

Several autonomously run organizations: No one “boss”

Organizations cooperate, but also compete

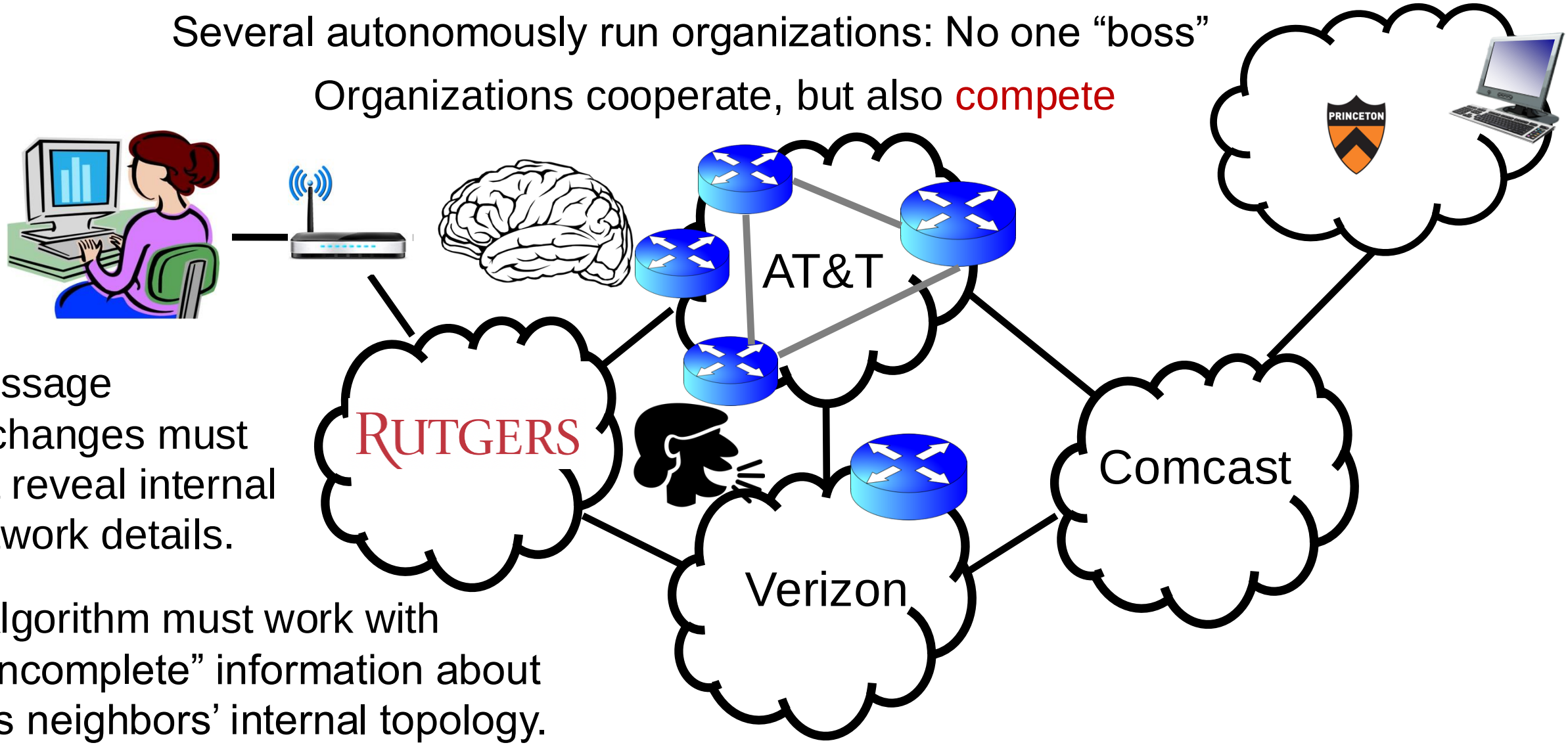


e.g., AT&T has little commercial interest in revealing its internal network structure to Verizon.

The Internet is a large federated network

Several autonomously run organizations: No one “boss”

Organizations cooperate, but also compete



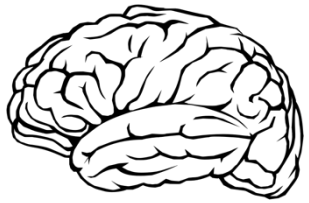
Message exchanges must not reveal internal network details.

Algorithm must work with “incomplete” information about its neighbors’ internal topology.

The Internet is a large federated network

Internet today: > 70,000 unique autonomous networks

Internet routers: > 800,000 forwarding table entries



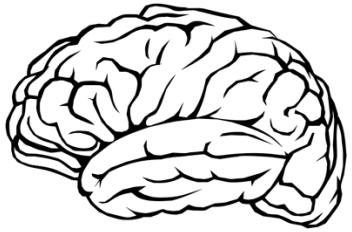
Keep messages & tables as small as possible. Don't flood

Algorithm must be incremental: don't recompute the whole table on every message exchanged.

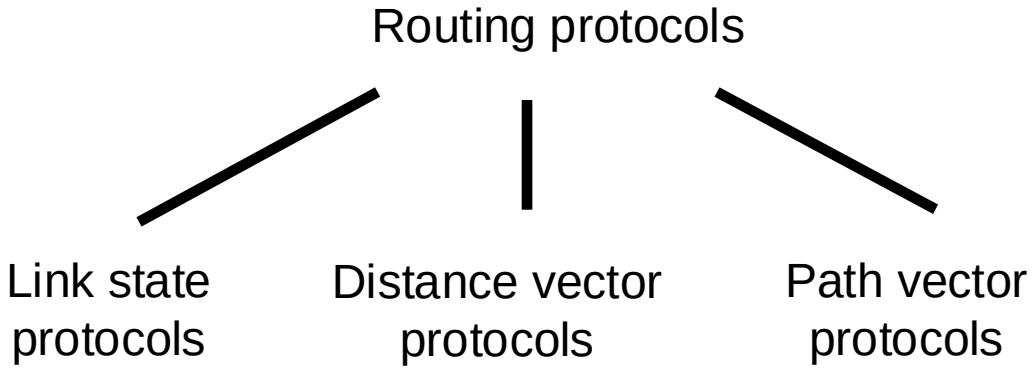
Inter-domain Routing

- Routing approaches so far (LS + DV) are applicable within one autonomous system (AS), e.g., Rutgers
  - Called intra-domain routing protocols
- The Internet uses Border Gateway Protocol (BGP)
- All AS'es speak BGP. It is the glue that holds the Internet together
- BGP is a path vector protocol

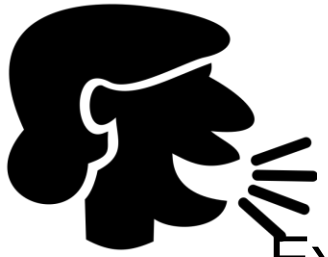
Messages?



Algorithm?



Q1. BGP Messages

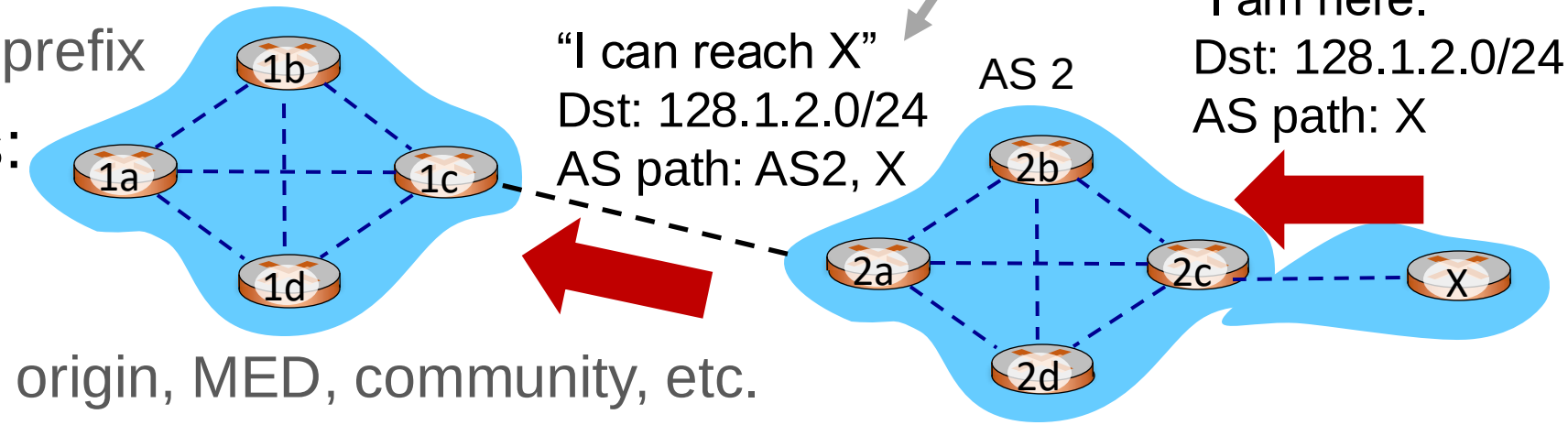


Loop detection is easy
(no “count to infinity”)

Exchange paths: path vector

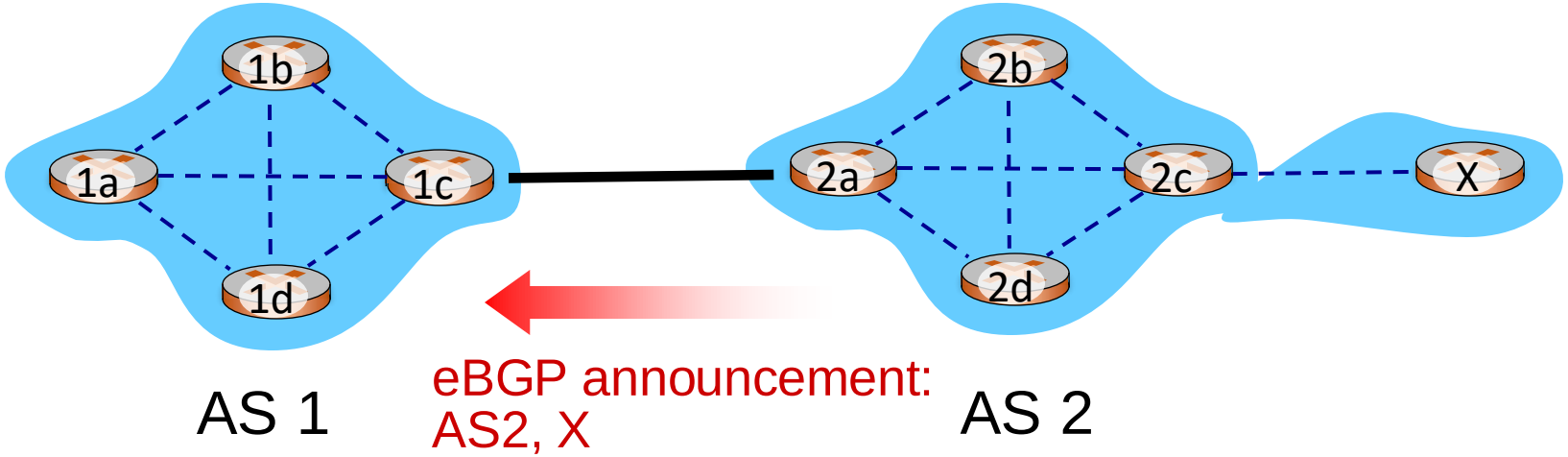
- Routing Announcements or Advertisements
  - “I am here” or “I can reach here”
  - Occur over a TCP connection (BGP session) between routers
- Route announcement = destination + attributes
  - Destination: IP prefix
- Route Attributes:
  - AS-level path
  - Next hop
  - Several others: origin, MED, community, etc.
- An AS promises to use advertised path to reach destination
- Only route changes are advertised after BGP session established

No link metrics, distances!



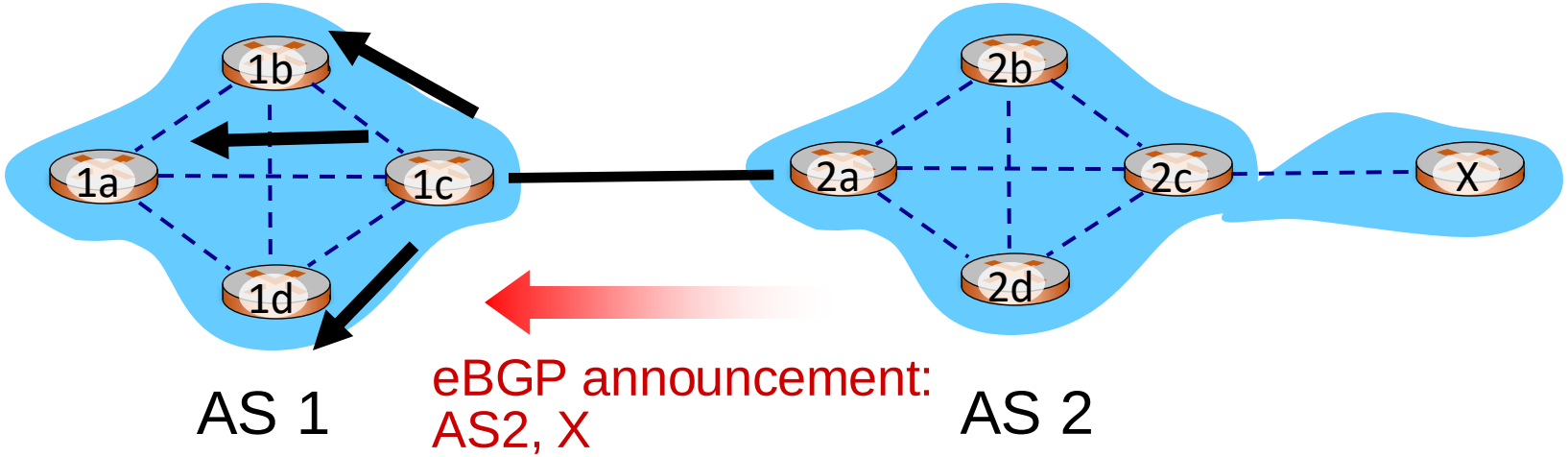
Q1. Next Hop

- Next hop conceptually denotes the first router interface that begins the AS-level path
  - The meaning of this attribute is context-dependent
- In an announcement arriving from a different AS (eBGP), next hop is the router in the next AS which sent the announcement
  - Example: Next Hop of the eBGP announcement reaching 1c is 2a

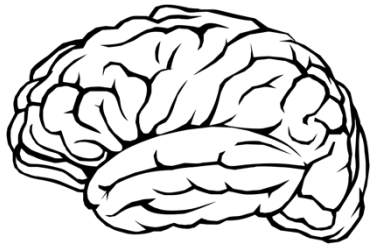


Q1. Next Hop

- Suppose router 1c imports the path (more on this soon)
- Router 1c will propagate the announcement inside the AS using iBGP
- The next hop of this (iBGP) announcement is set to 1c
  - In particular, the next hop is an AS1 internal address



Q2. The algorithm

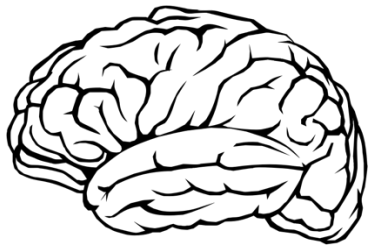


- A BGP router does *not* consider every routing advertisement it receives by default to make routing decisions!
  - An import policy determines whether a route is even considered a candidate
- Once imported, the router performs route selection
- A BGP router does *not* propagate its chosen path to a destination to all other AS'es by default!
  - An export policy determines whether a (chosen) path can be advertised to other AS'es and routers

Programmed by network operator

Policy considerations make BGP very different from intra-domain (LS / DV) protocols

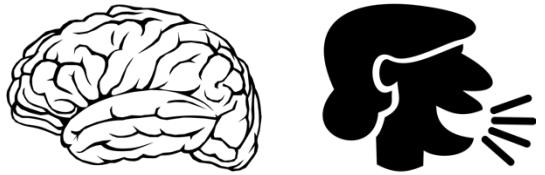
Policies in BGP

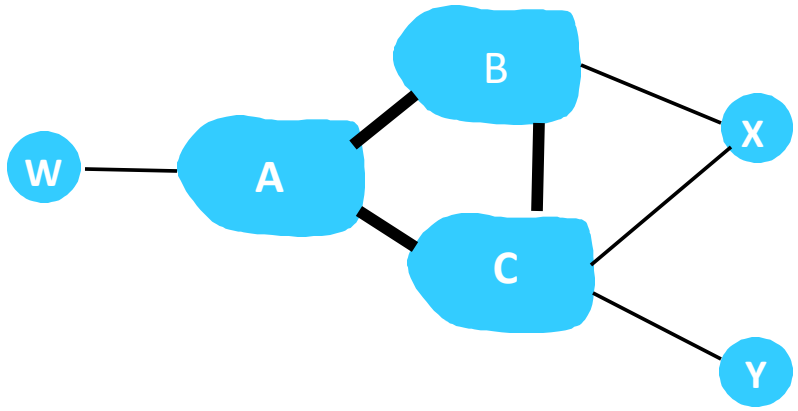


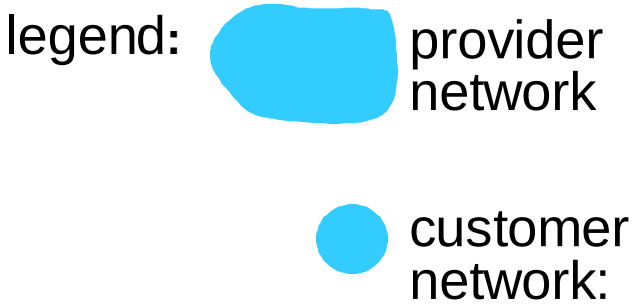
Policy arises from business relationships

- Customer-provider relationships:
  - E.g., Rutgers is a customer of AT&T
- Peer-peer relationships:
  - E.g., Verizon is a peer of AT&T
- Business relationships depend on where connectivity occurs
  - “Where”, also called a “point of presence” (PoP)
  - e.g., customers at one PoP but peers at another
  - Internet-eXchange Points (IXPs) are large PoPs where ISPs come together to connect with each other (often for free)

BGP Export Policy



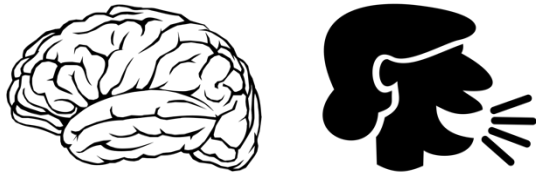


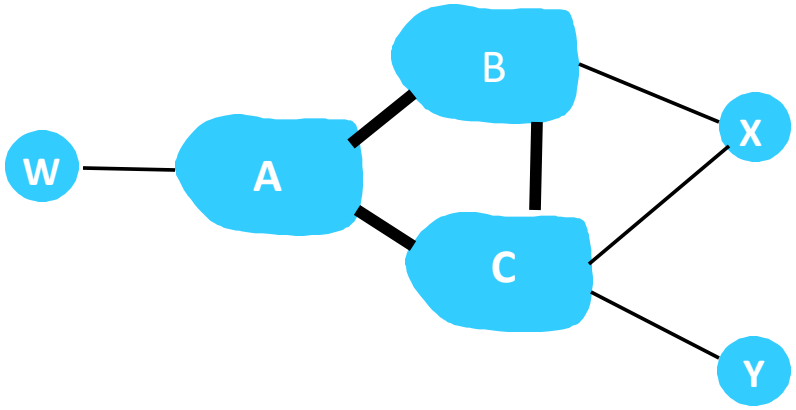


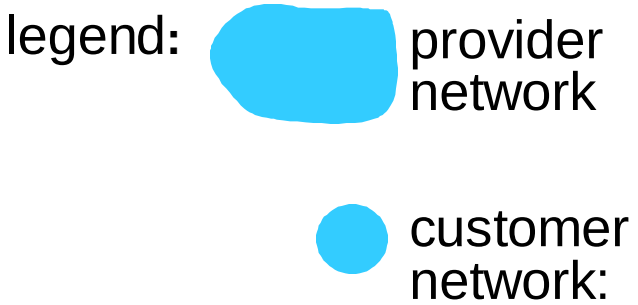
Suppose an ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs)

- A,B,C are provider networks
- X,W,Y are customers (of provider networks)
- X is dual-homed: attached to two networks
- policy to enforce: X does not want to route from B to C via X
  - So, X will not announce to B a route to C

BGP Export Policy







Suppose an ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs)

- A announces path Aw to B and to C
- B will not announce BAw to C:
  - B gets no “revenue” for routing CBAw, since none of C, A, w are B’s customers
- C will route CAw (not using B) to get to w

Thumb rules for export policy

- Export customer-learned routes to the rest of the Internet
  - Financially attractive if the Internet uses your AS to get to a customer
  - OK to export customer-learned route to a provider since the customer must be reachable from the rest of the Internet
- Do not export a provider-learned route to another provider
  - Don’t lose money on both sides
- Do not export peer-learned route to another peer
  - The resources of your AS are used but no financial gains from free “transit” service
- OK to export provider-learned route to a customer
  - Customer must be able to reach the rest of the Internet
  - Also financially attractive