Routing Protocols: Part 2

Lecture 24
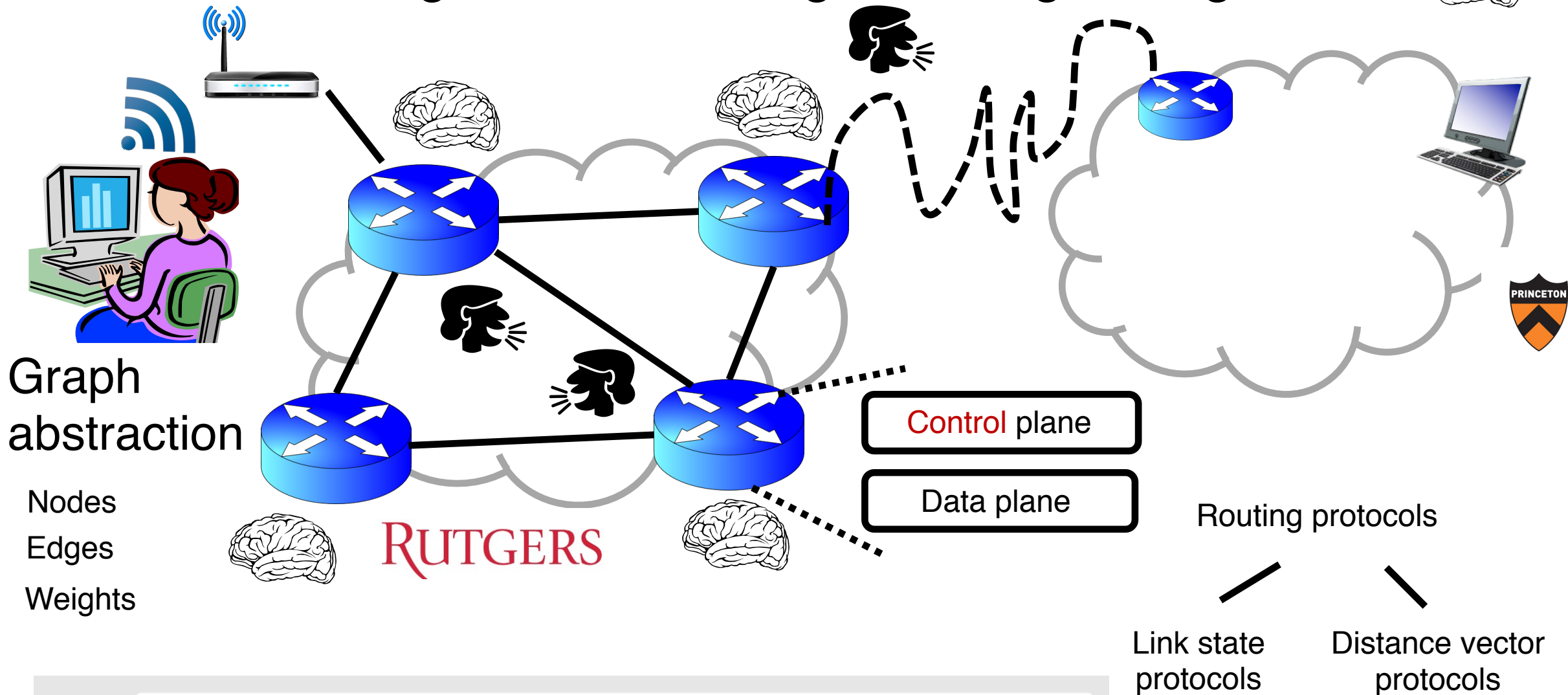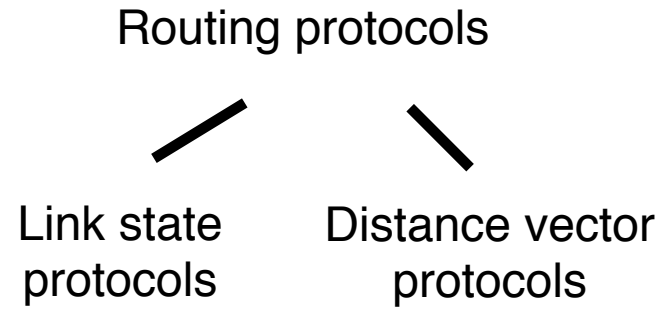### http://www.cs.rutgers.edu/~sn624/352-S22

Srinivas Narayana

RUTGERS
UNIVERSITY | NEW BRUNSWICK

The network layer is all about reachability.

Routing Protocol = Msg Exchange + Algorithm

Net layer

Graph
abstraction

Nodes

Edges

Weights

Control plane

Data plane

Routing protocols

Link state
protocols

Distance vector
protocols

RUTGERS

PRINCETON

www.cs.princeton.edu

Routing protocols

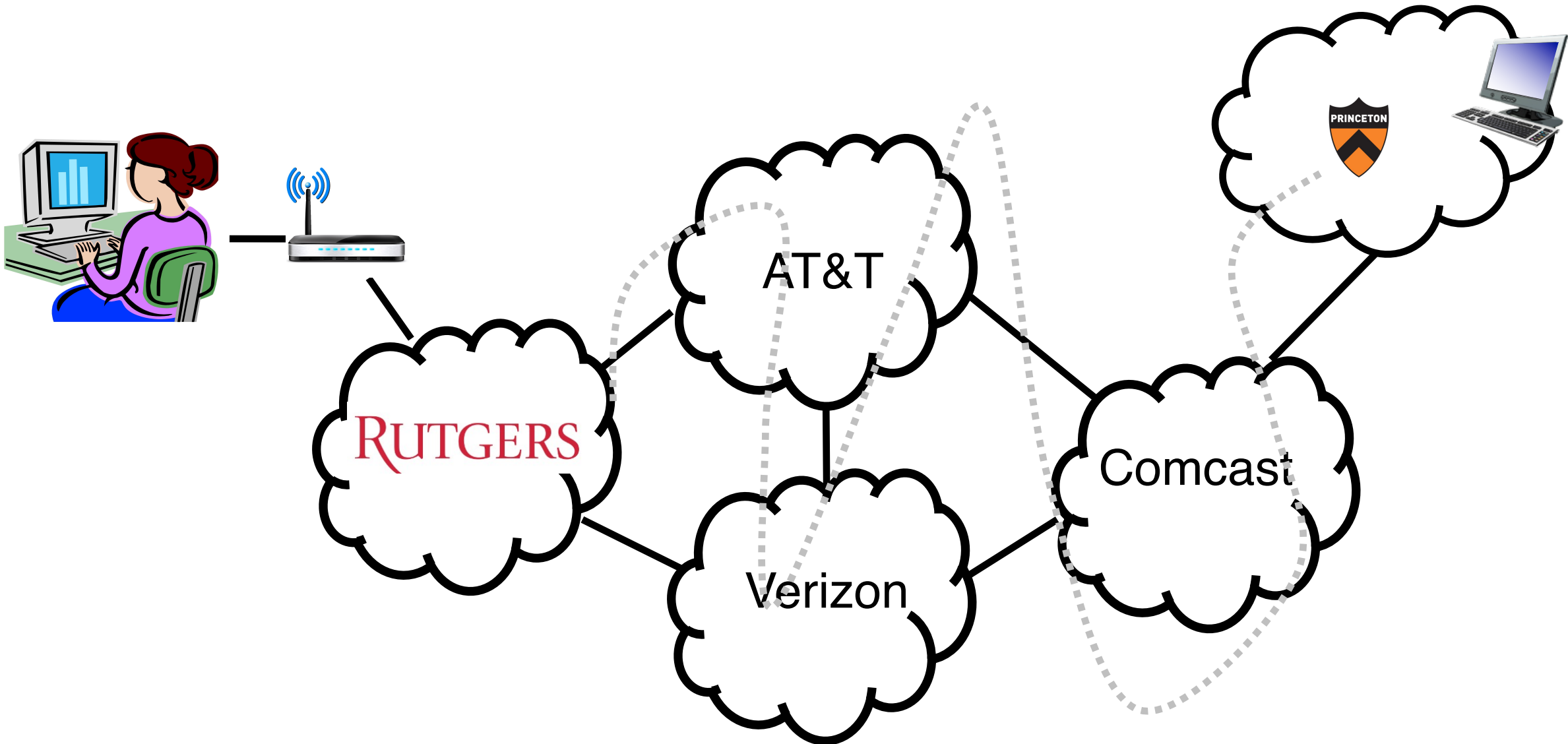Link state protocols

Distance vector protocols

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.
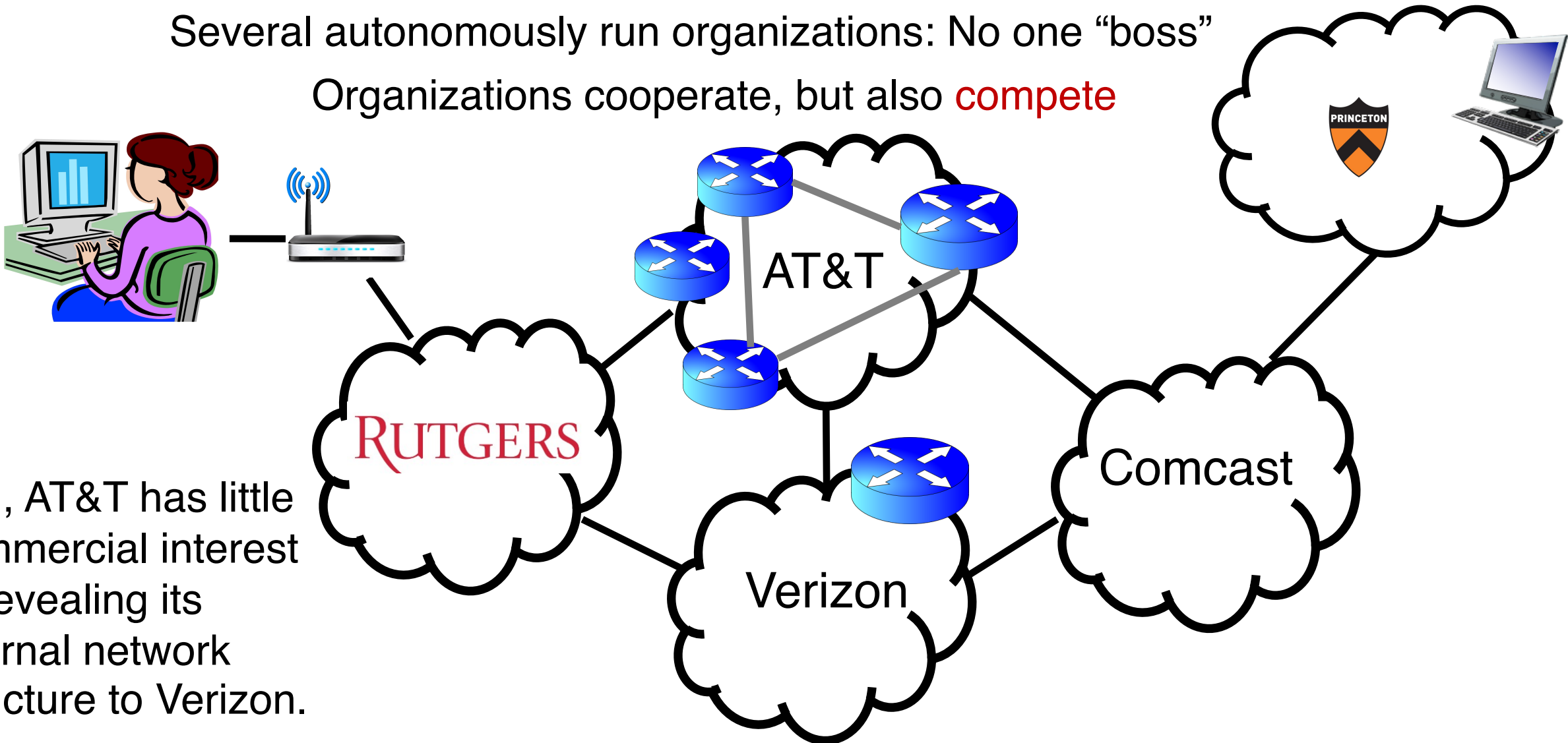
These assumptions do not hold on the Internet.

The Internet is a large federated network

The Internet is a large federated network

Several autonomously run organizations: No one "boss"
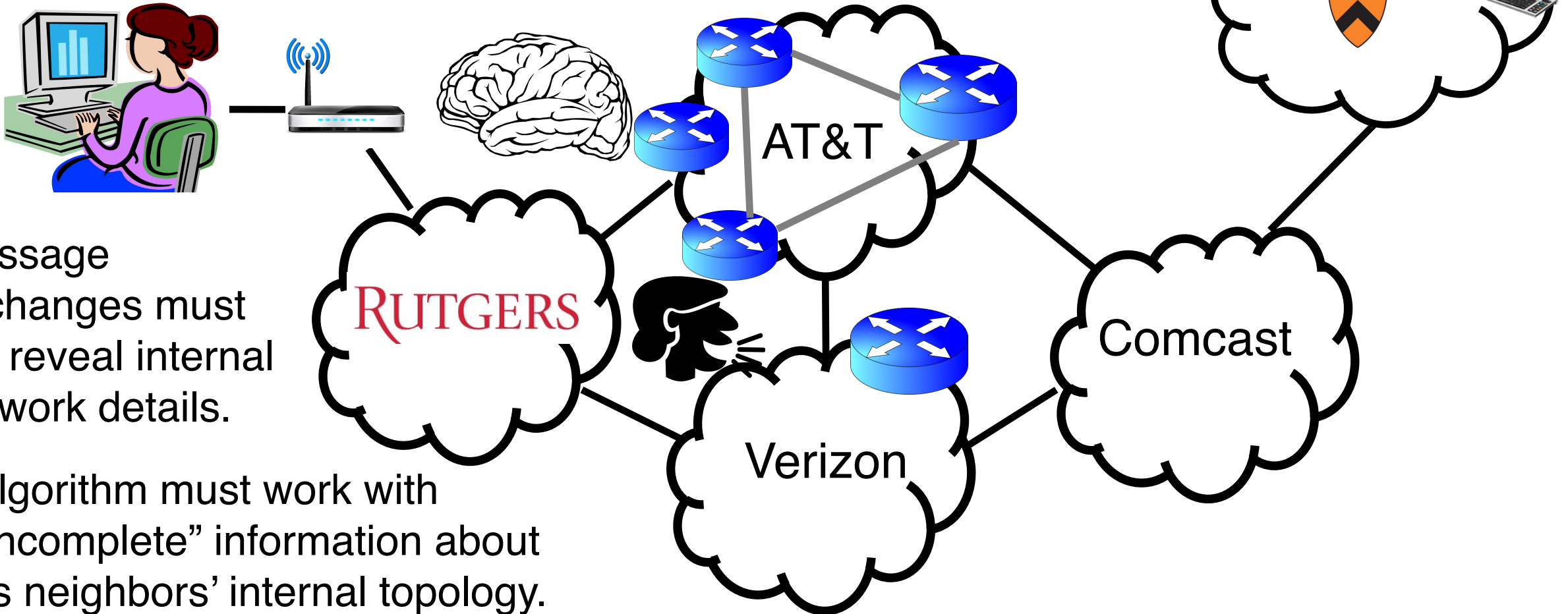
Organizations cooperate, but also compete

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

RUTGERS

AT&T

PRINCETON

Verizon

Comcast

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.

AT&T

RUTGERS

Comcast

Verizon

PRINCETON

The Internet is a large federated network

Internet today: > 70,000 unique autonomous networks

Internet routers: > 800,000 forwarding table entries
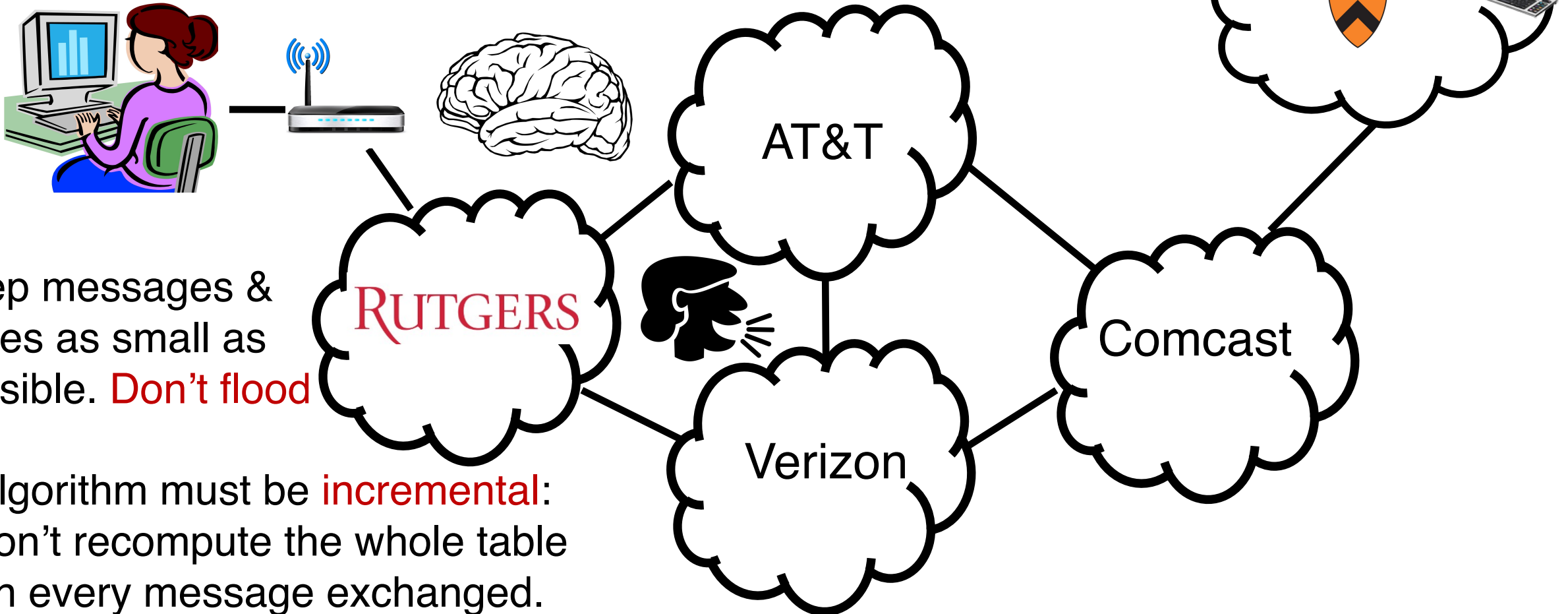
Keep messages & tables as small as possible. Don't flood

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

AT&T

RUTGERS

Comcast

Verizon

PRINCETON

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?

Routing protocols

Link state protocols

Distance vector protocols

Path vector protocols
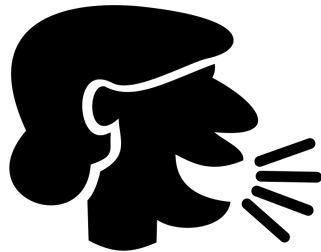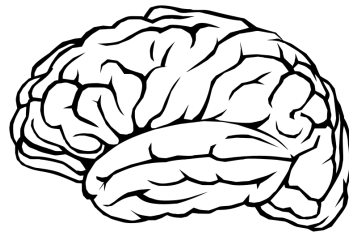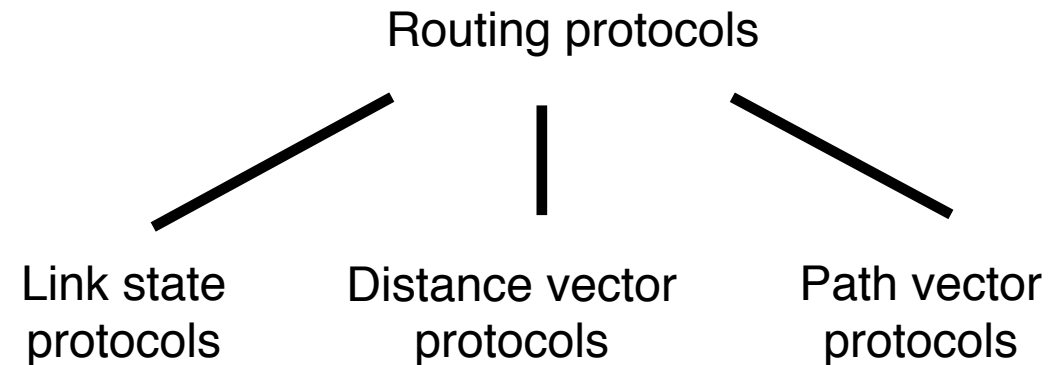
Q1. BGP Messages

Loop detection is easy
(no "count to infinity")

Exchange paths: path vector

- Routing Announcements or Advertisements   No link metrics, distances!
  - "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.

"I can reach X"
Dst: 128.1.2.0/24
AS path: AS2, X

"I am here."
Dst: 128.1.2.0/24
AS path: X

AS 2

1b
1a   1c
1d

2b
2a   2c   X
2d

- An AS promises to use advertised path to reach destination

- Only route changes are advertised after BGP session established
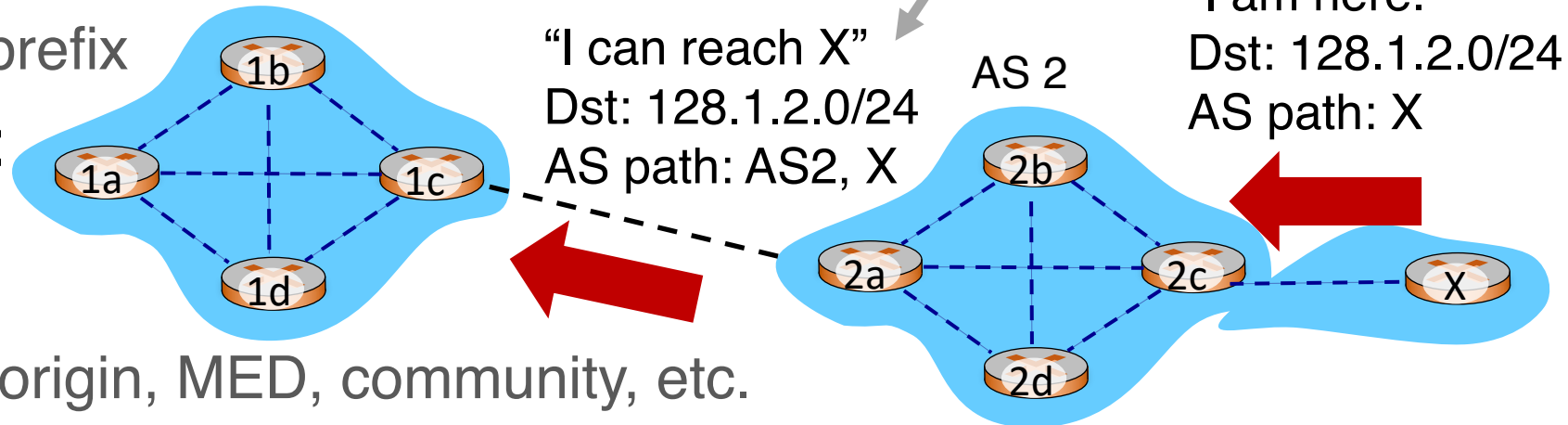
Q1. Next Hop
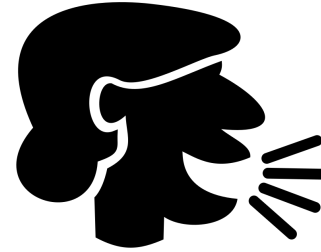
- **Next hop** conceptually denotes the first router interface that begins the AS-level path
  - The meaning of this attribute is context-dependent

- 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**

AS 1

eBGP announcement: AS2, X

AS 2
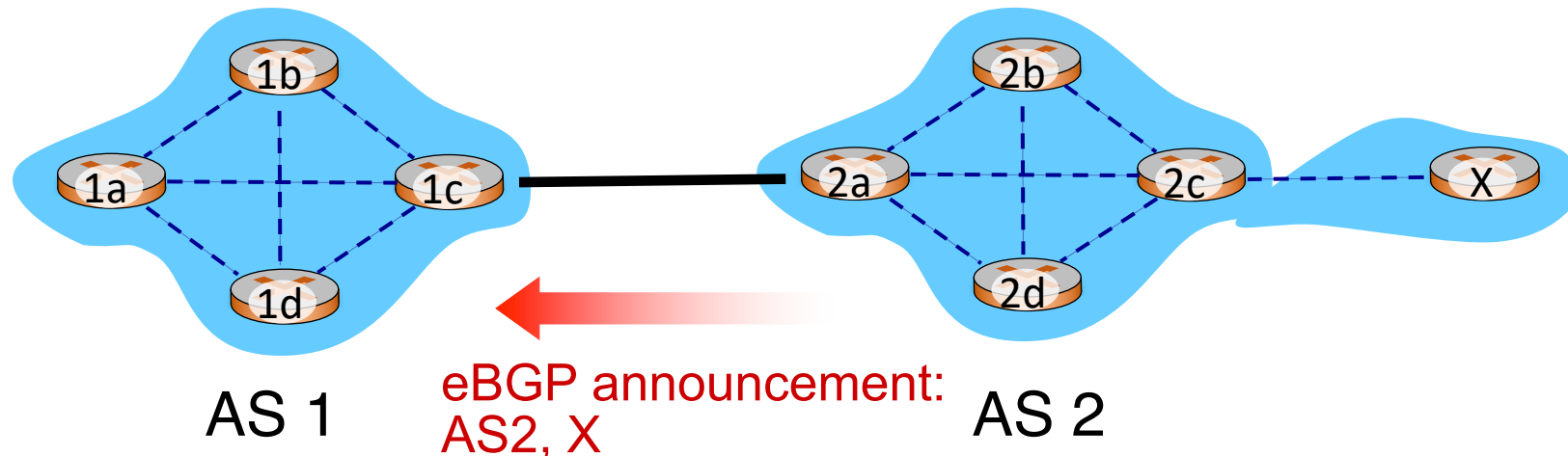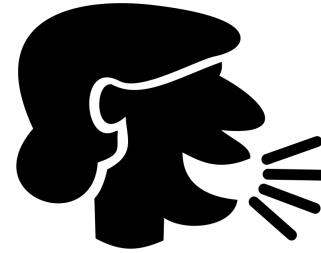
Q1. Next Hop

- 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


AS 1

eBGP announcement:
AS2, X

AS 2
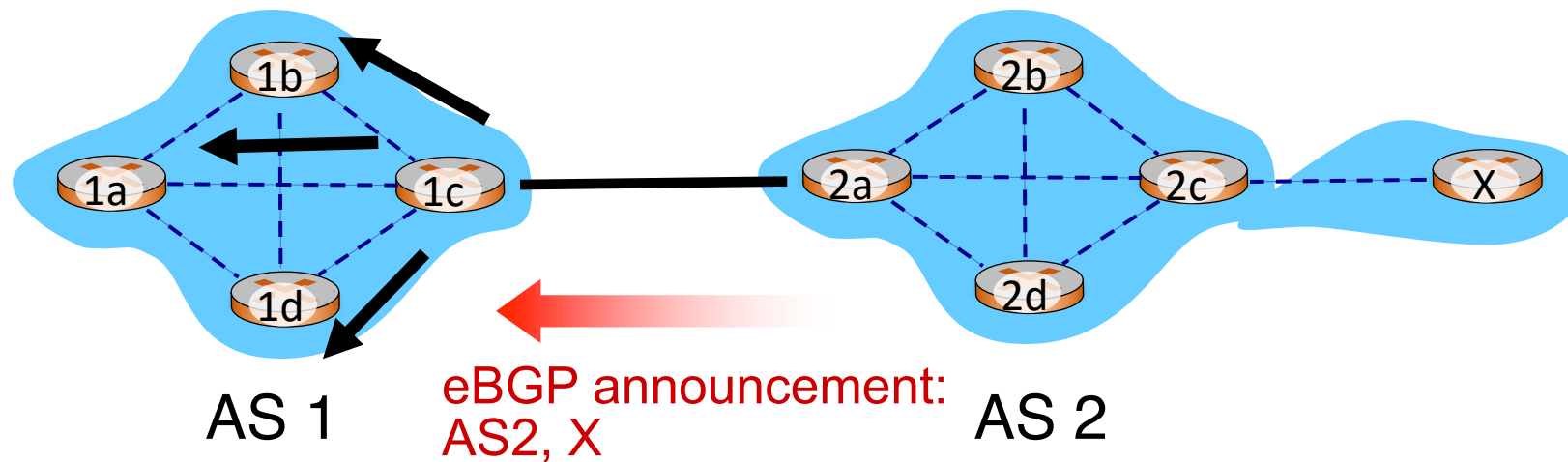
Q2. The algorithm
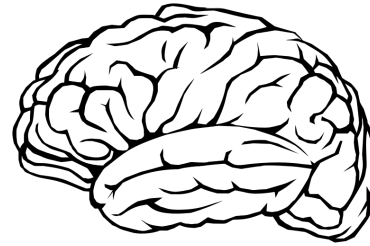
- A BGP router does *not* consider every routing advertisement it receives by default to make routing decisions!
  - An import policy determines whether a route is even considered a candidate

- Once imported, the router performs route selection

Programmed by network operator

- 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

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

Policies in BGP
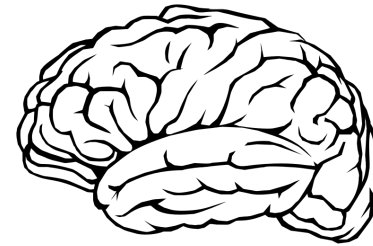
Policy arises from business relationships

- 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)
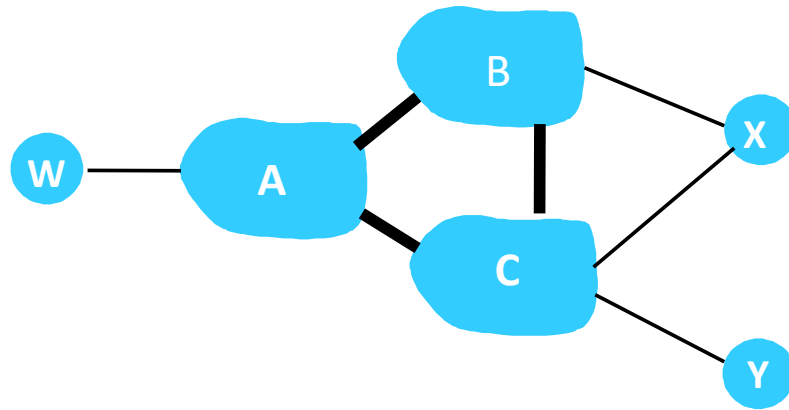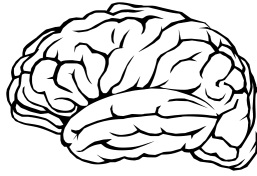
BGP Export Policy
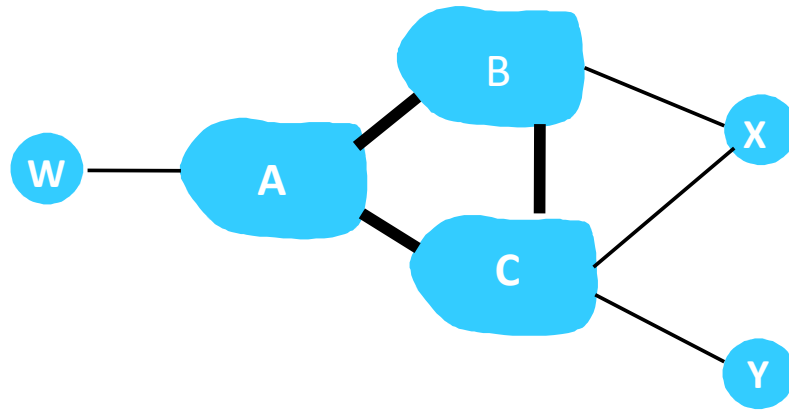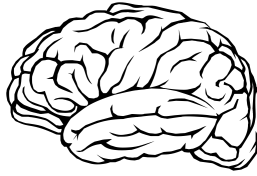
legend:

provider network

customer network:

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

legend:

provider network

customer network:

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

BGP Import Policy
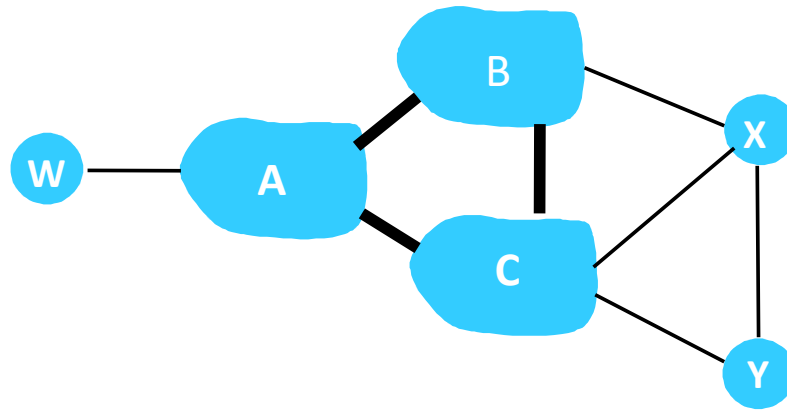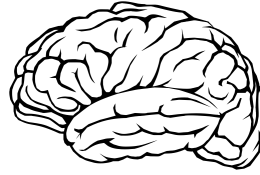
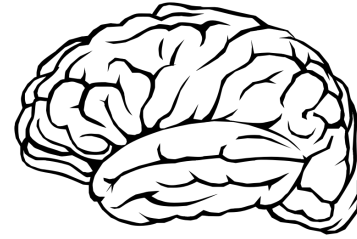

legend:
provider network

customer network:

Suppose an ISP wants to minimize costs by avoiding routing through its providers when possible.

- Suppose C announces path Cy to x
- Further, y announces a direct path ("y") to x
- Then x may choose not to import the path Cy to y since it has a peer path ("y") towards y

Q2. BGP Route Selection

- When a router imports more than one route to a destination IP prefix, it selects route based on:
    1. local preference value attribute (import policy decision -- set by network admin)
    2. shortest AS-PATH
    3. closest NEXT-HOP router
    4. Several additional criteria: You can read up on the full, complex, list of criteria, e.g., at https://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/13753-25.html

Example of route selection

- Suppose AS A and B are connected to each other both in North America (NA) and in Europe (EU)

- A source in NA wants to reach a destination in EU

- There are two paths available
  - *Assume* same local preference
  - Same AS path length

- Closest next hop-router: choose path via B1 rather than B2

NA   EU

Dst-EU

A1

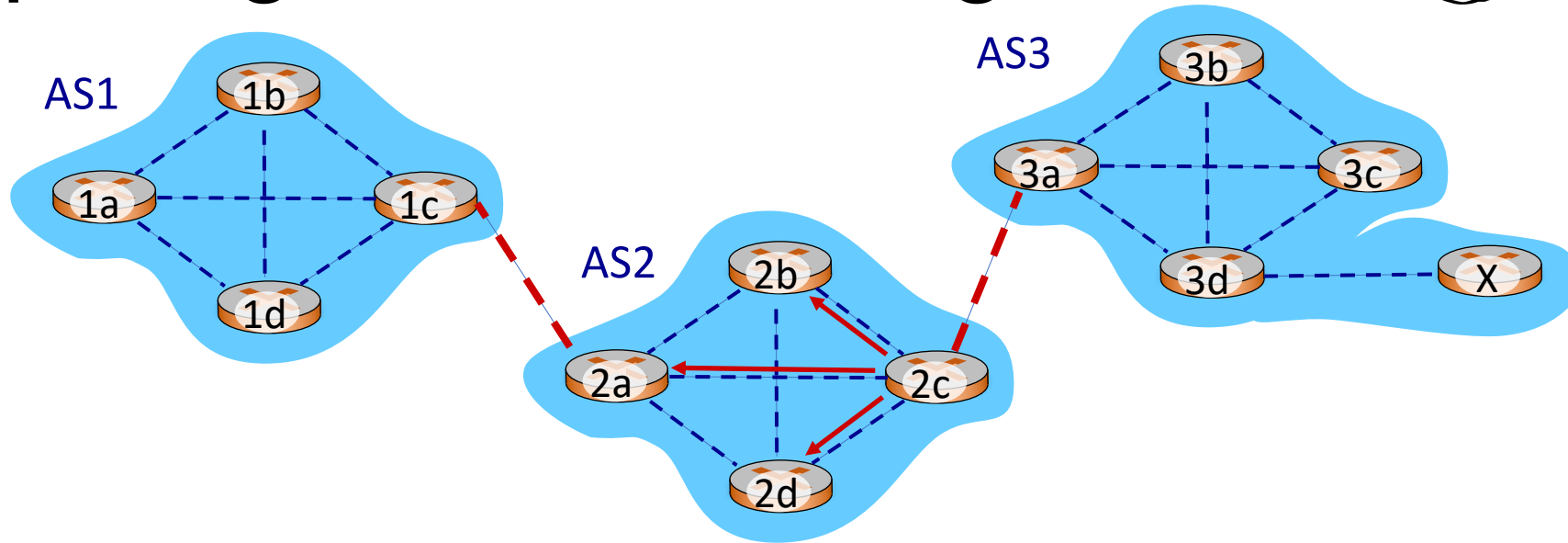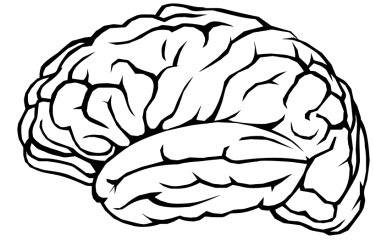A2

AS A

B1

B2

Src-NA

AS B
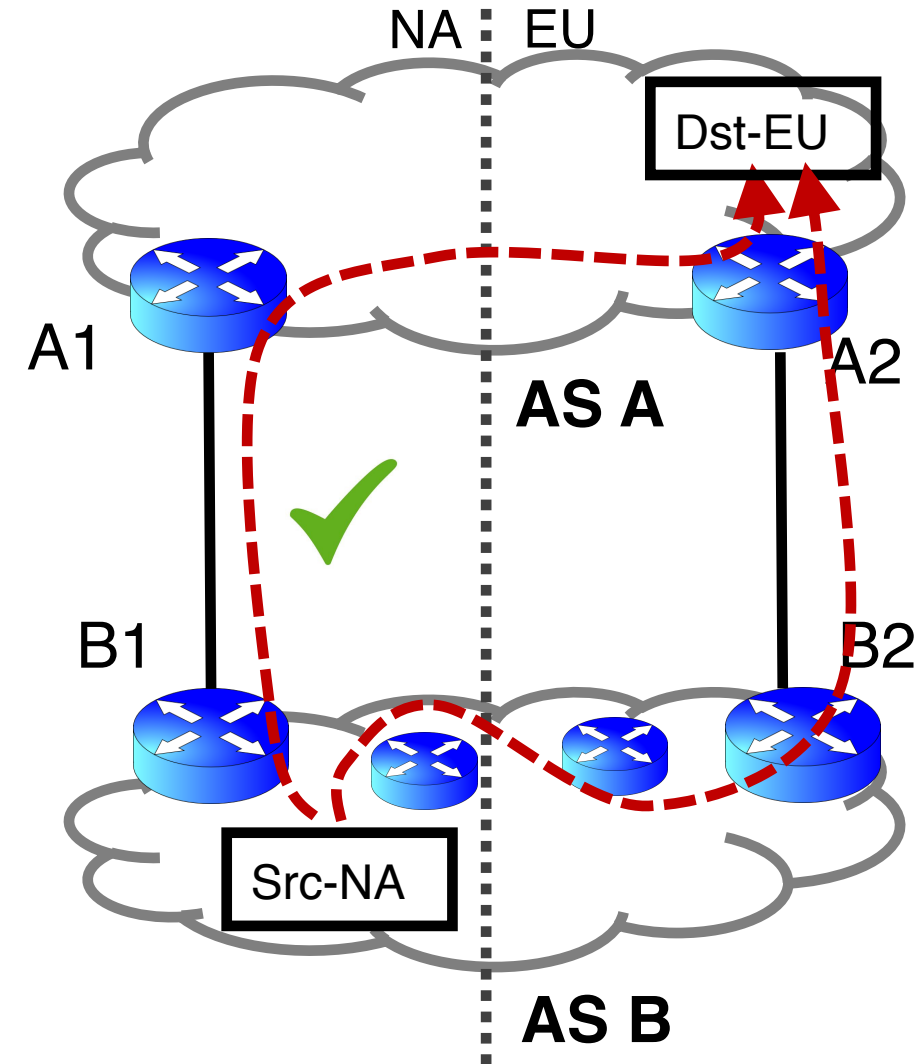
Example of route selection

- Choosing closest next-hop results in **early exit routing**
  - Try to exit the local AS as early as possible
  - Also called **hot potato routing**

- Reduce resource use within local AS
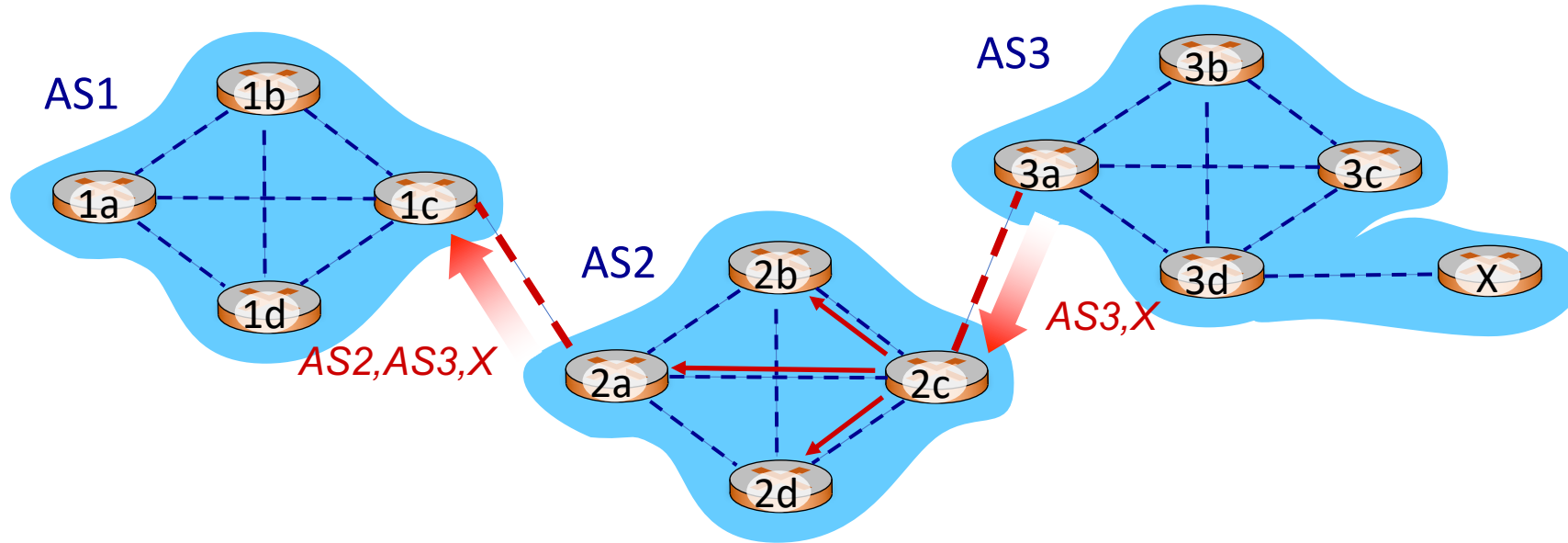 - potentially at the expense of another AS
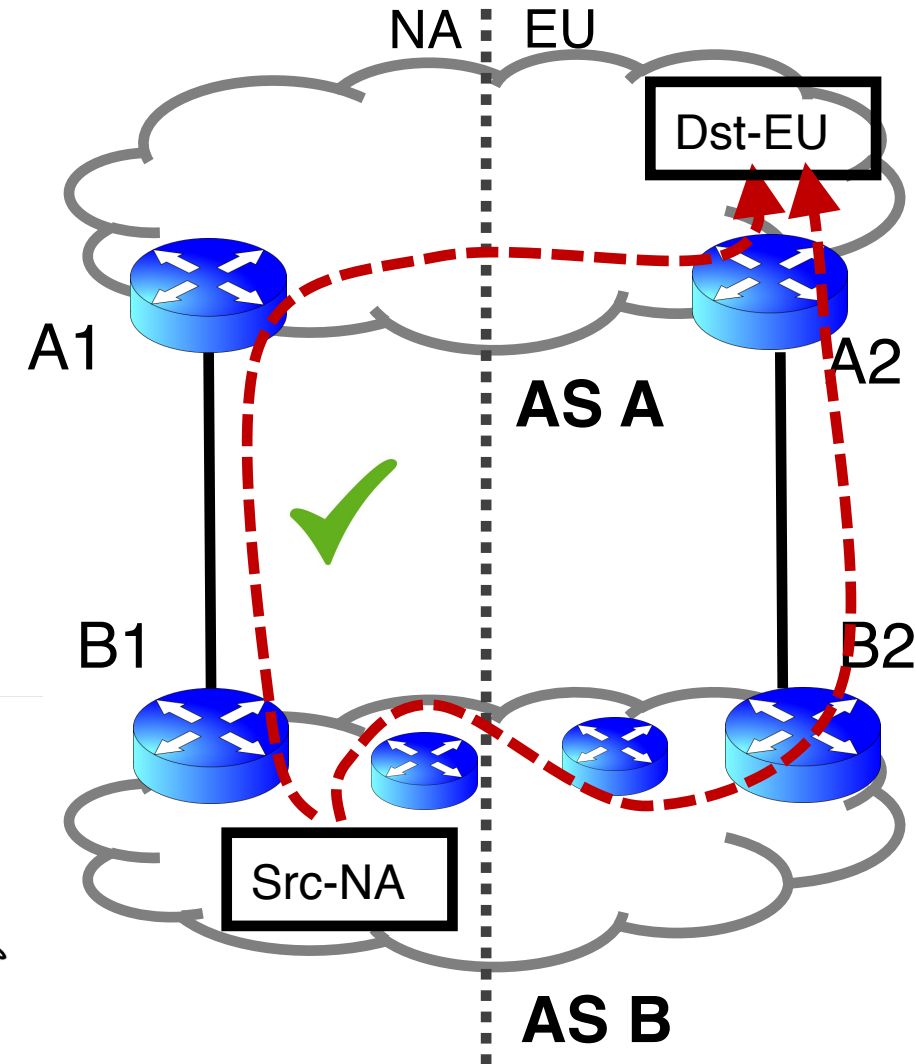
Computing the forwarding table

- Suppose a router in AS1 wants to forward a packet destined to external prefix X.

- How is the forwarding table entry for X at 1d computed?

- How is the forwarding table entry for X at 1c computed?

eBGP and iBGP announcements

- AS2 router 2c receives path announcement AS3,X (via eBGP) from AS3 router 3a

- Based on AS2 import policy, AS2 router 2c imports and selects path AS3,X, propagates (via iBGP) to all AS2 routers

- Based on AS2 export policy, AS2 router 2a announces (via eBGP) path AS2, AS3, X to AS1 router 1c

eBGP and iBGP announcements
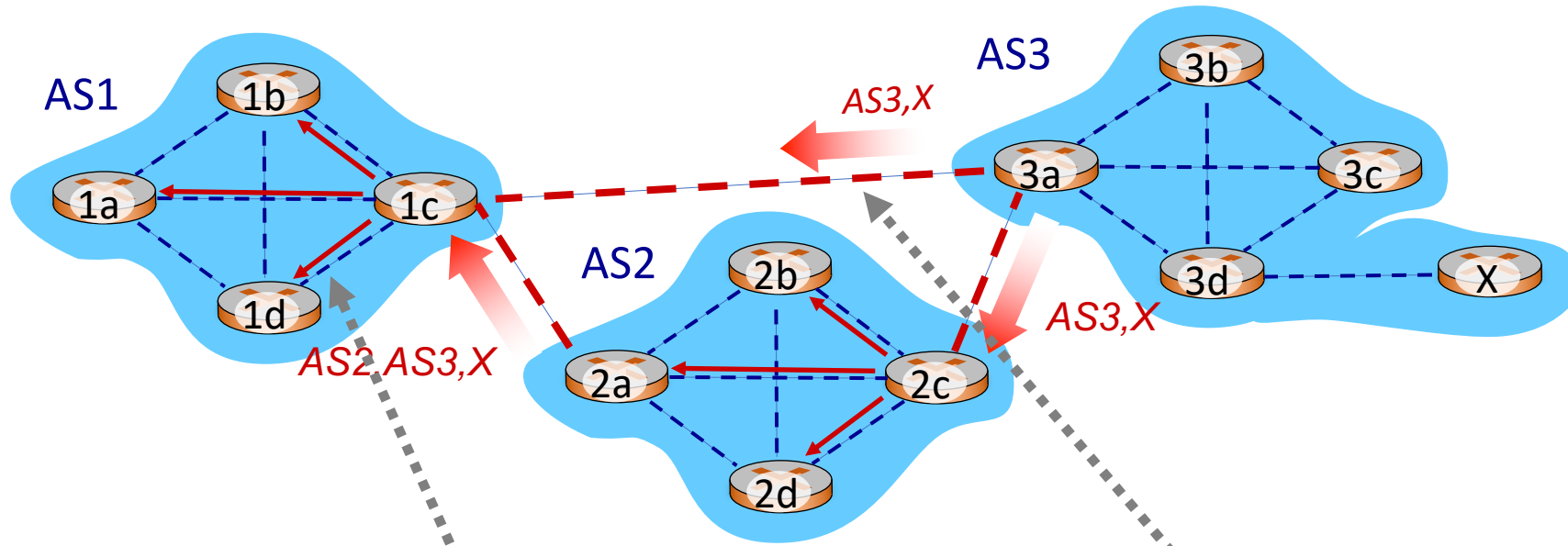
A given router may learn about multiple paths to destination:

- AS1 gateway router 1c learns path AS2,AS3,X from 2a (next hop 2a)
- AS1 gateway router 1c learns path AS3,X from 3a (next hop 3a)
- Through BGP route selection process, AS1 gateway router 1c chooses path AS3,X, and announces path within AS1 via iBGP (next hop 1c)

Setting forwarding table entries
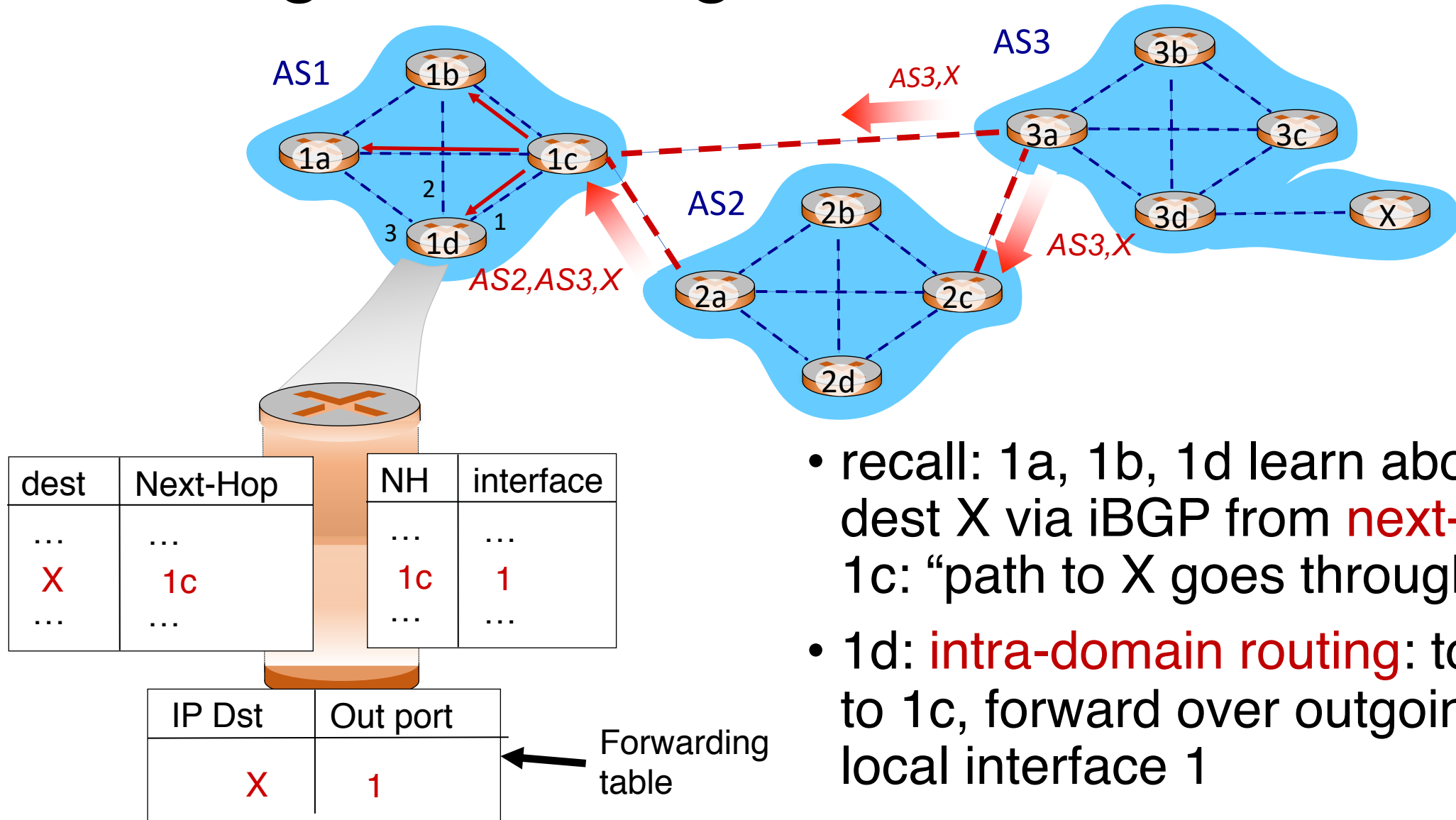
- recall: 1a, 1b, 1d learn about dest X via iBGP from next-hop 1c: "path to X goes through 1c"

- 1d: intra-domain routing: to get to 1c, forward over outgoing local interface 1

Setting forwarding table entries
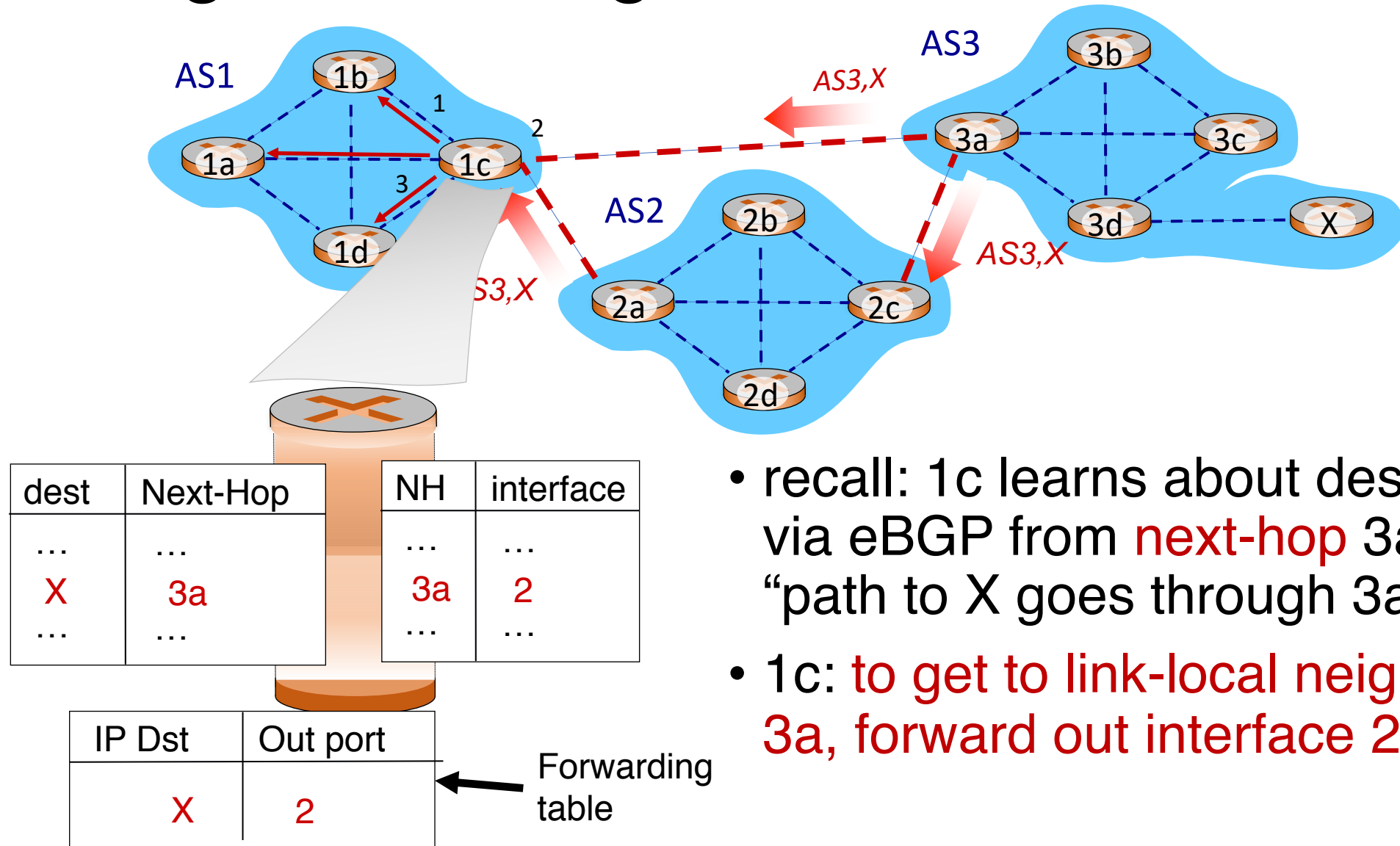
- recall: 1c learns about dest X via eBGP from next-hop 3a: "path to X goes through 3a"

- 1c: to get to link-local neighbor 3a, forward out interface 2

Summary: Inter-domain routing

- **Federation** and **scale** introduce new requirements for routing on the Internet

- **BGP** is *the* protocol that handles Internet routing

- **Path vector**: exchange paths to a destination with attributes

- **Policy-based** import of routes, route selection, and export