UNIVERSITY OF WATERLOO

CS 456/656 Computer Networks

Lecture 2: Introduction – Part 2

Mina Tahmasbi Arashloo and Uzma Maroof

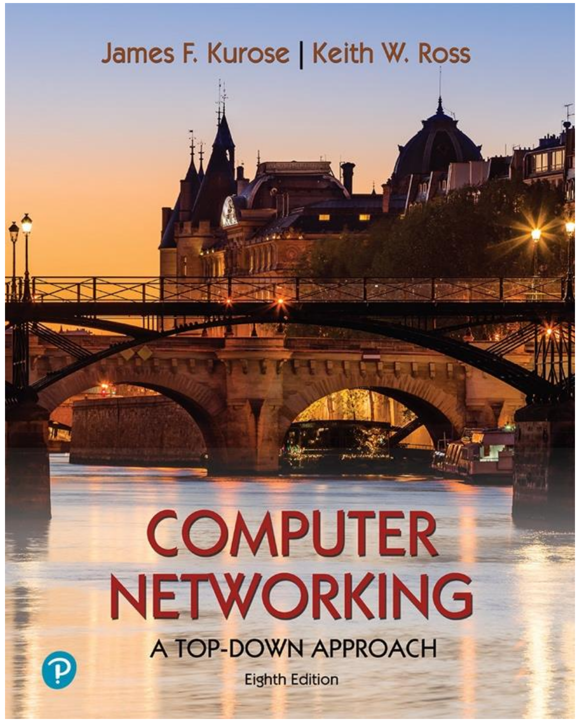
Fall 2025

Today's agenda

- Challenges of communicating over a shared network – continued
- Network performance measures
- A high-level overview of the Internet
  - The "nuts-and-bolts" view of the Internet
    - Internet structure
    - Network edge: access networks
    - Network core
  - The service view of the Internet

A note on the slides

Adapted from the slides that accompany this book.



Computer Networking: A Top-Down Approach
8th edition
Jim Kurose, Keith Ross
Pearson, 2020

Recap: Communicating over a shared network

- To make that happen, networking people have to solve several challenging problems:
  - How to decide when a sender gets to transmit data?
  - How to pick good paths for getting data from its source to its destination?
  - How to adapt when a switch/router or a link fails?
  - …

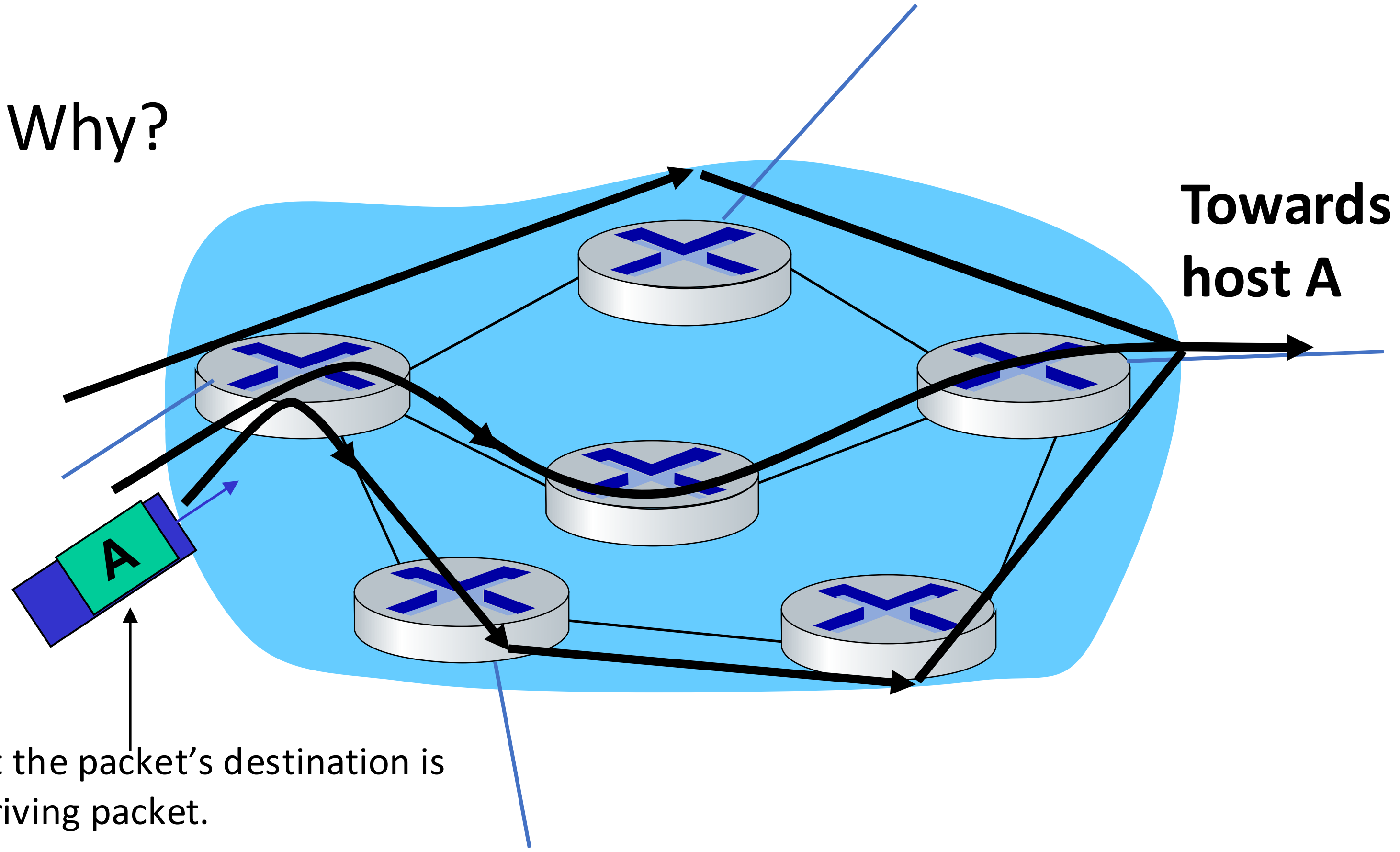
This is where we ended last time

Communicating over a shared network

- To make that happen, networking people have to solve several challenging problems:
  - How to decide when a sender gets to transmit data?
  - How to pick good paths for getting data from its source to its destination?
  - How to adapt when a switch/router or a link fails?
  - …

Routing: Picking the “best route”

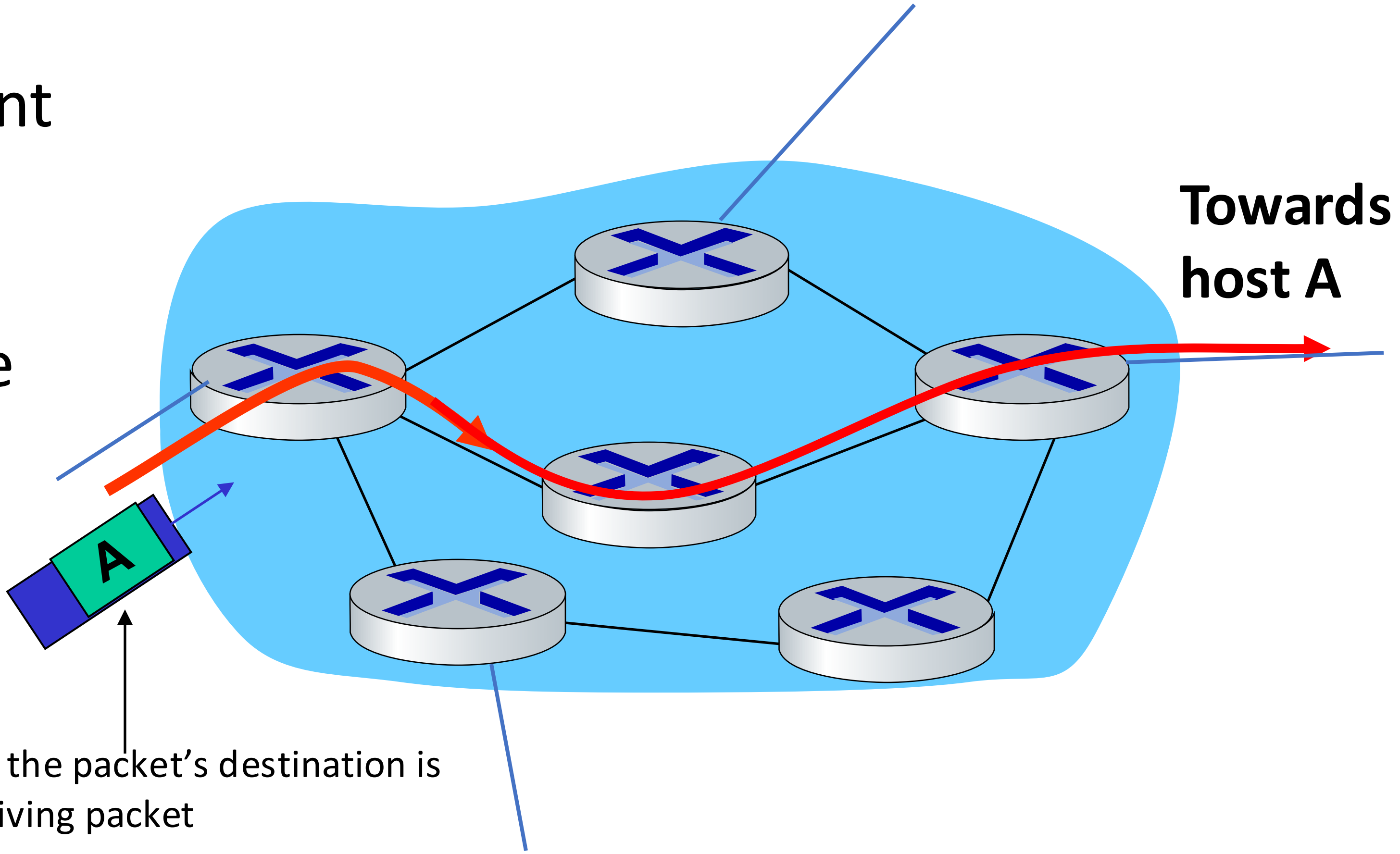
- Many different potential paths to the destination.
- Which one would you pick? Why?



Information about the packet’s destination is included in the arriving packet.

Routing: Picking the "best route"

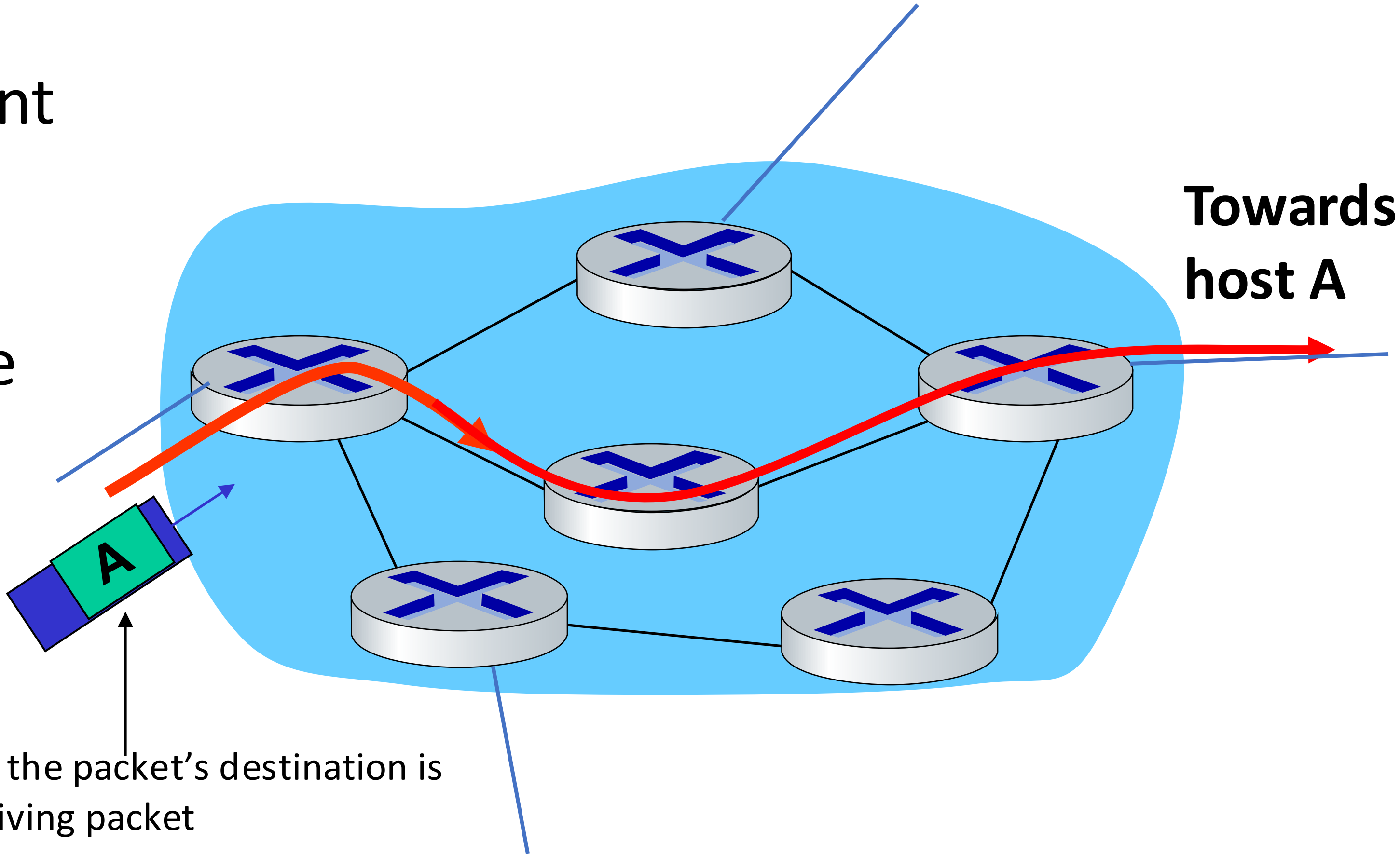
- Suppose we want the network to pick the middle path.
- Each device is an independent entity, though
- How do they coordinate to make sure to pick the middle path for packets going to A?



Information about the packet's destination is included in the arriving packet

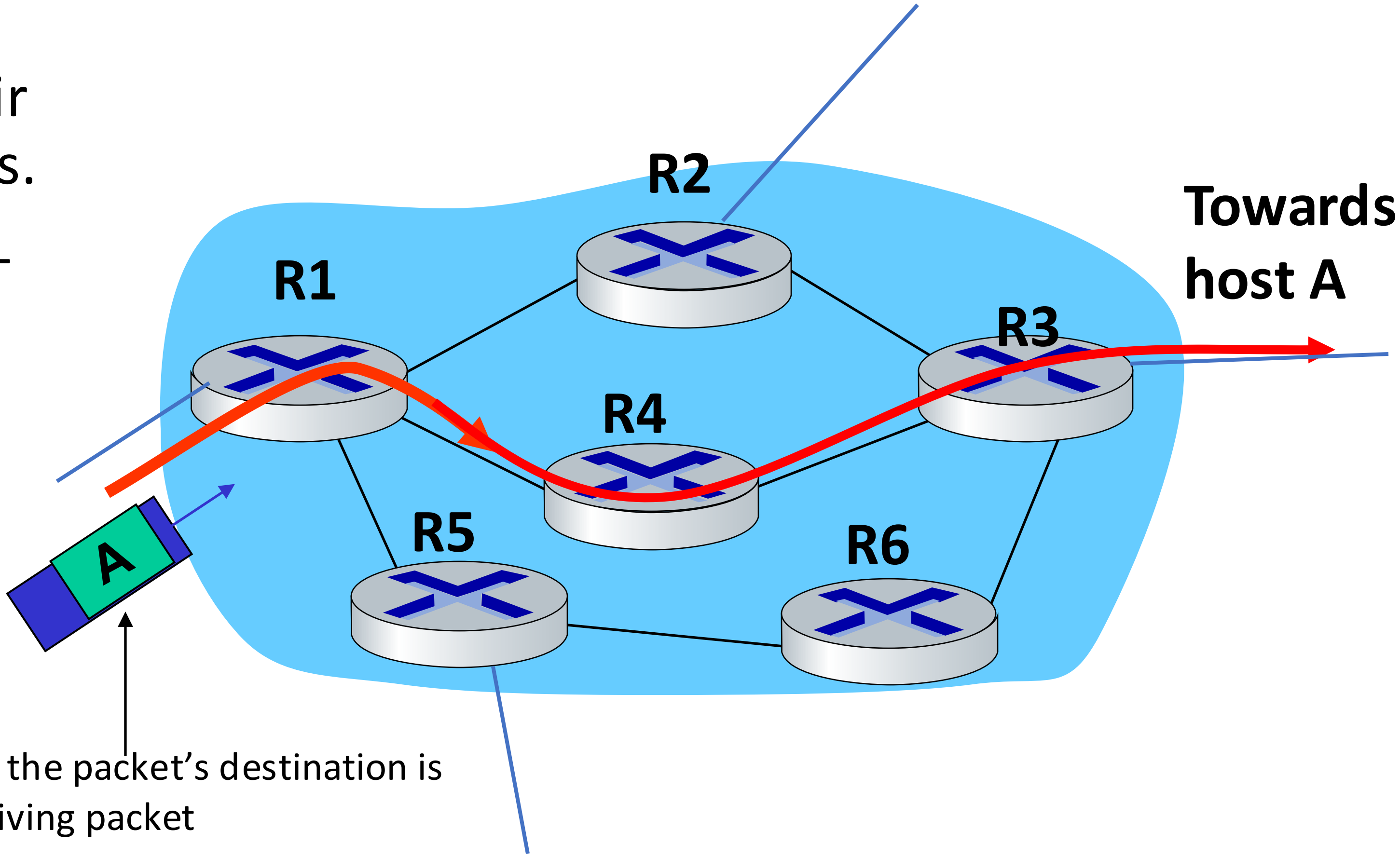
Routing: Picking the "best route"

- Suppose we want the network to pick the middle path.
- Each device is an independent entity, though
- How do they coordinate to make sure to pick the middle path for packets going to A?
  - Routing protocols!



(Distributed) Routing Protocols

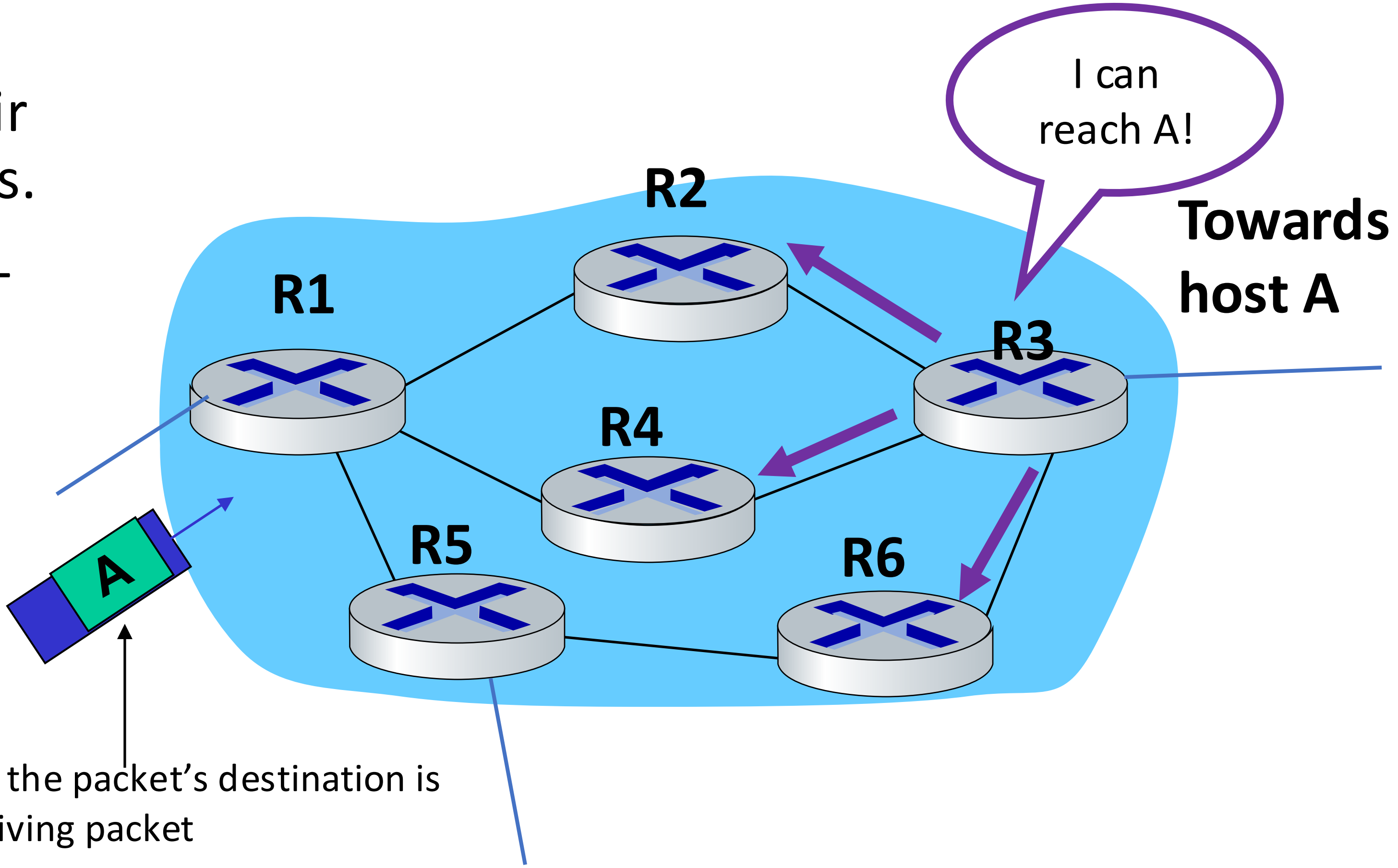
- Through routing protocols, network devices
  - exchange messages about their routes to different destinations.
  - coordinate on a (good) end-to-end path.



Information about the packet's destination is included in the arriving packet

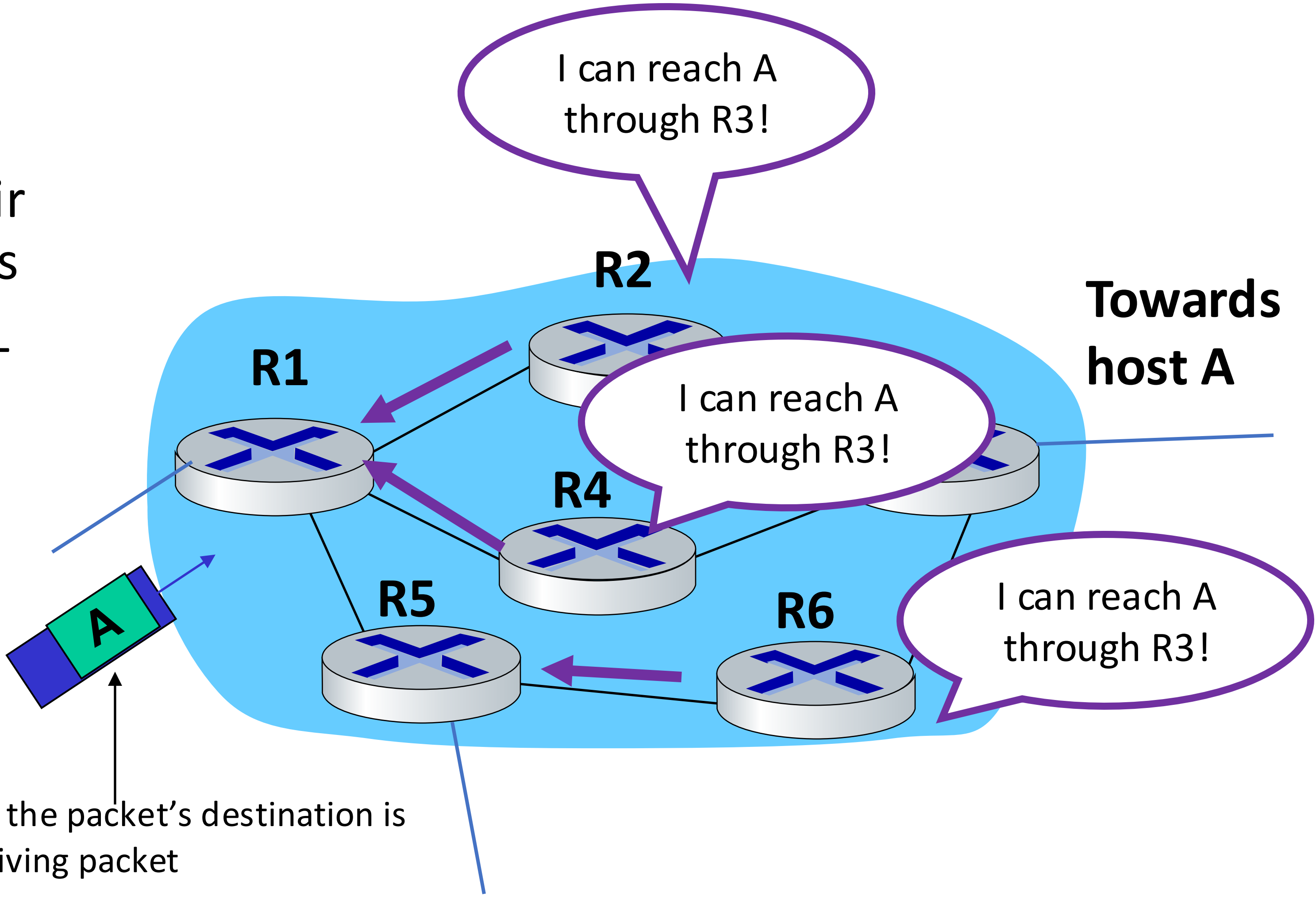
(Distributed) Routing Protocols

- Through routing protocols, network devices
  - exchange messages about their routes to different destinations.
  - coordinate on a (good) end-to-end path.



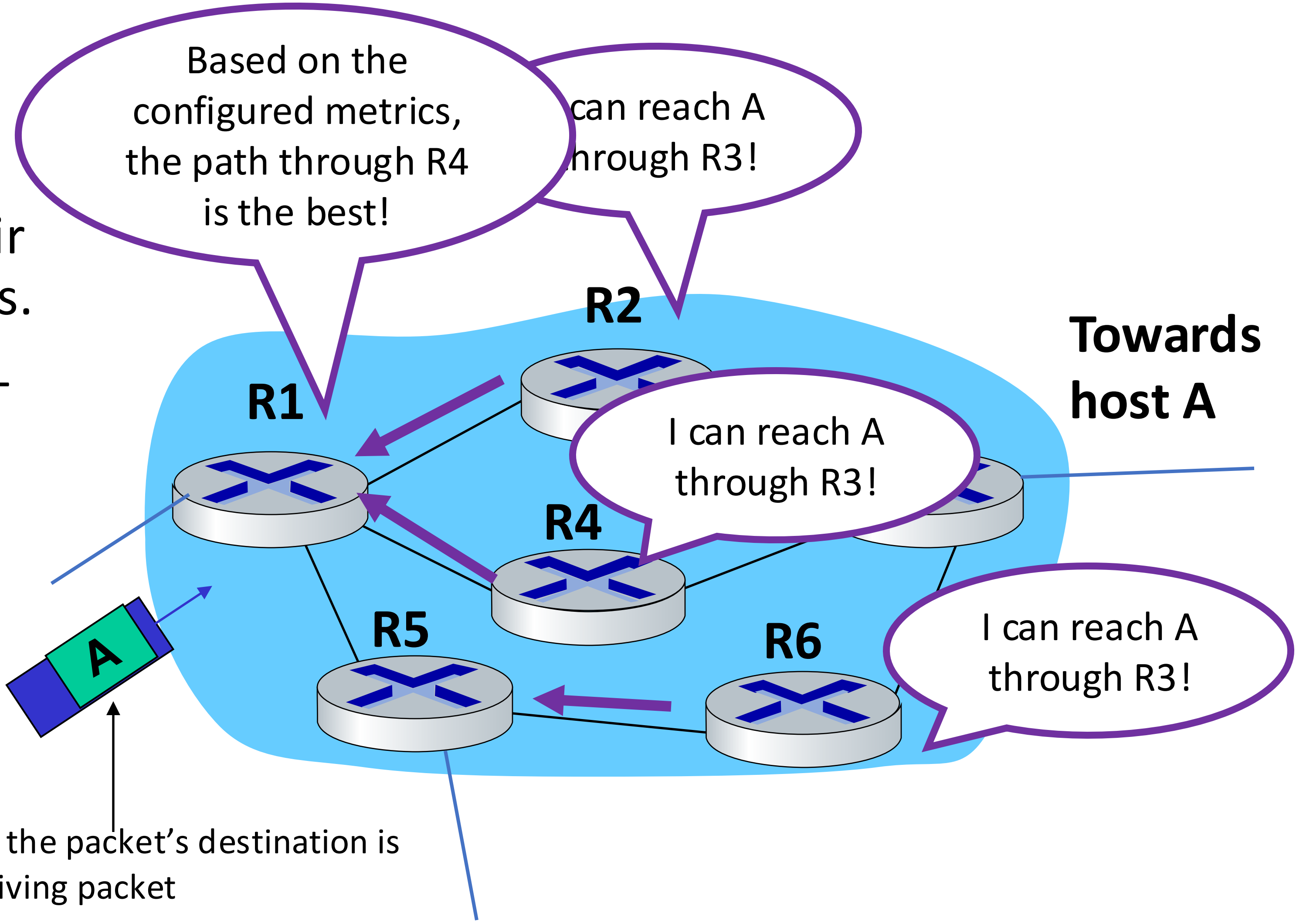
(Distributed) Routing Protocols

- Through routing protocols, network devices
  - exchange messages about their routes to different destinations
  - coordinate on a (good) end-to-end path.



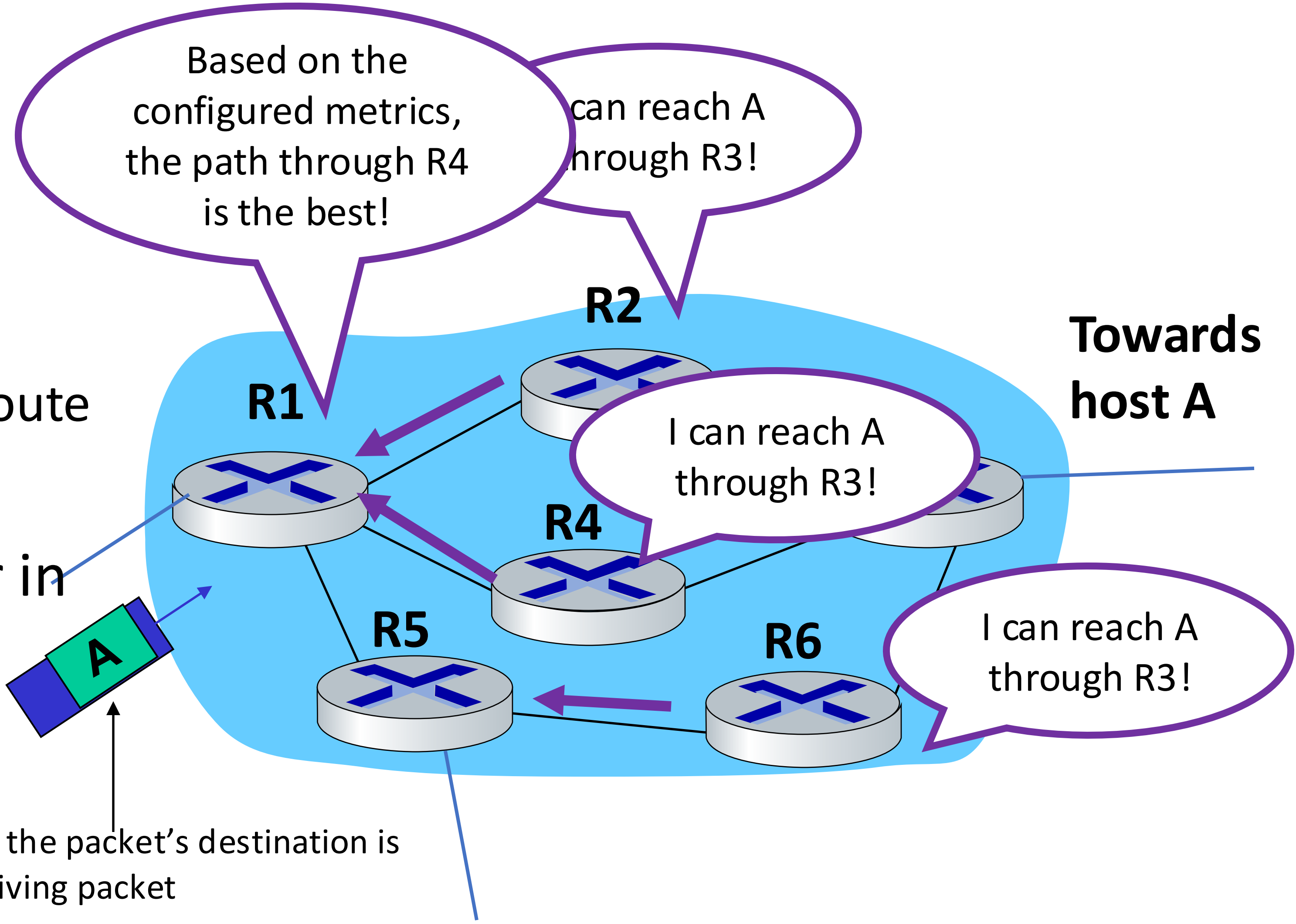
(Distributed) Routing Protocols

- Through routing protocols, network devices
  - exchange messages about their routes to different destinations.
  - coordinate on a (good) end-to-end path.



Routing protocols are difficult to design

- There are many factors to consider
  - Message overhead
  - Convergence time
  - Metrics to decide the “best” route
  - …
- We’ll learn about these later in the course!



Communicating over a shared network

- To make that happen, networking people have to solve several challenging problems:
  - How to decide when a sender gets to transmit data?
  - How to pick good paths for getting data from its source to its destination?
  - How to adapt when a switch/router or a link fails?
  - …

Communicating over a shared network

- To make that happen, networking people have to solve several challenging problems:
  - How to decide when a sender gets to transmit data?
  - How to pick good paths for getting data from its source to its destination?
  - How to adapt when a switch/router or a link fails?
  - …
- We will discuss these challenges, along with other challenging problems and some of the well-known solutions in the course

What is a “good” network?

- So far, we have discussed
  - What a computer network is
  - The fact that a computer network is typically shared among many traffic flows
  - Some of the challenges in designing a shared network infrastructure
- Now, suppose we have come up with some solutions to the above challenges and have designed a computer network.
- How do we know it is a “good” network?
  - High performance? How do we measure performance?
  - Highly secure?
  - …?

What is a “good” network?

- So far, we have discussed
  - What a computer network is
  - The fact that a computer network is typically shared among many traffic flows
  - Some of the challenges in designing a shared network infrastructure
- Now, suppose we have come up with some solutions to the above challenges and have designed a computer network.
- How do we know it is a “good” network?
  - High performance? How do we measure performance?
  - Highly secure?
  - …?

Network performance metrics

End-to-End Delay

how long does it take for a piece of data to go from one end of the network to the other?

Loss

how often does the network lose data during transfer?

Throughput

How much data per second can the network transfer?

Network performance metrics

End-to-End Delay

how long does it take for a piece of data to go from one end of the network to the other?

Loss

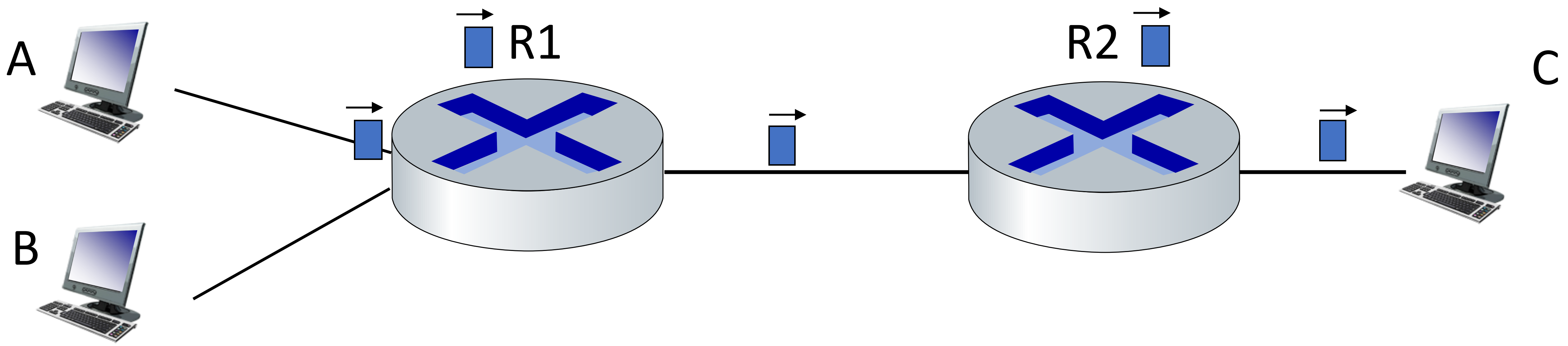
how often does the network lose data during transfer?

Throughput

How much data per second can the network transfer?

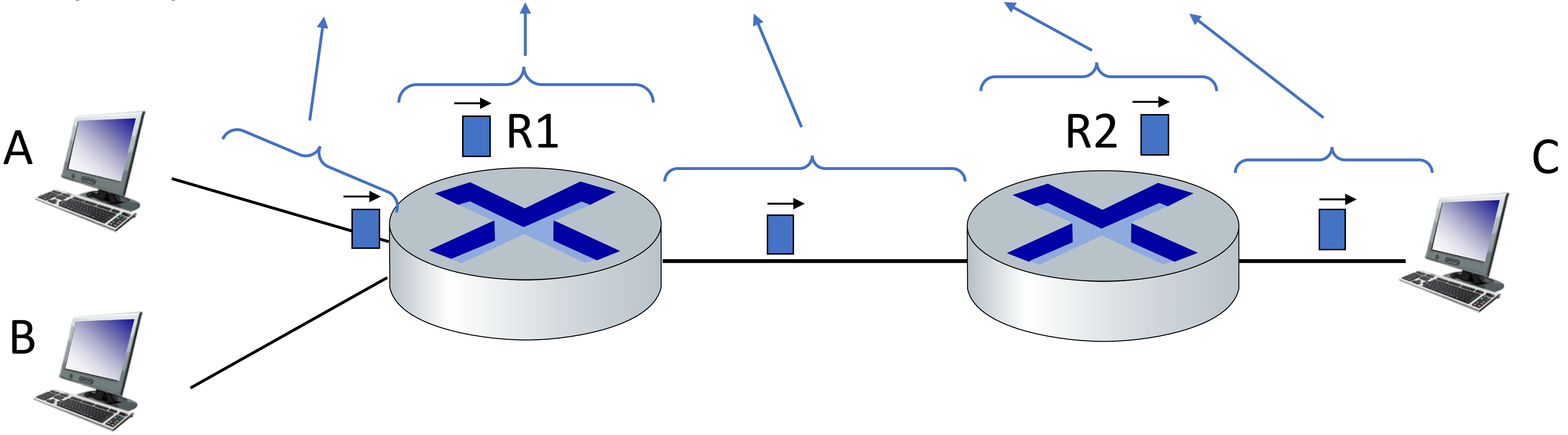
End-to-End delay

- End-to-end delay: how long does it take for a piece of data to go from one end of the network to the other, e.g., from A to C?



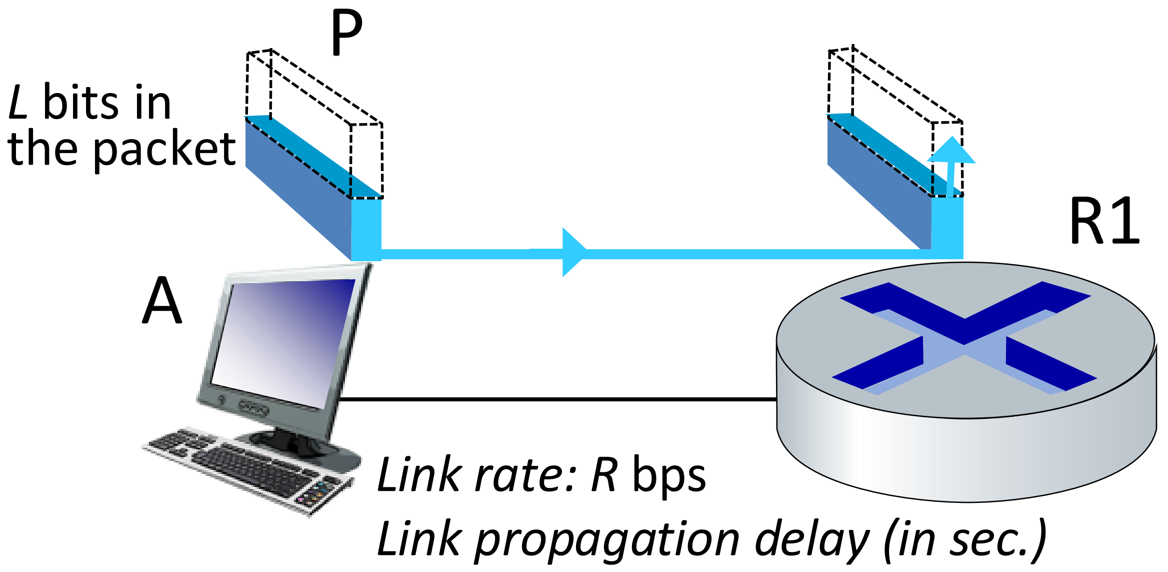
Network Delay

- D(x, y): time it takes for packet to go from x to y.
- D(x): time it takes for the packet to go through x
- D(A, C) = D(A, R1) + D(R1) + D(R1, R2) + D(R2) + D(R2, C)



Network delay over a link

- $d_{\text{trans}}(A, R_1, P)$: The time it takes for all the L bits in the packet to go from A onto the link (it depends on the packet P's length)
- $d_{\text{prop}}(A, R_1)$: The time it takes for one bit to transfer from A to R1
- $D(A, R_1) \;=\; d_{\text{trans}}(A, R_1, P) + \; d_{\text{prop}}(A, R_1)$



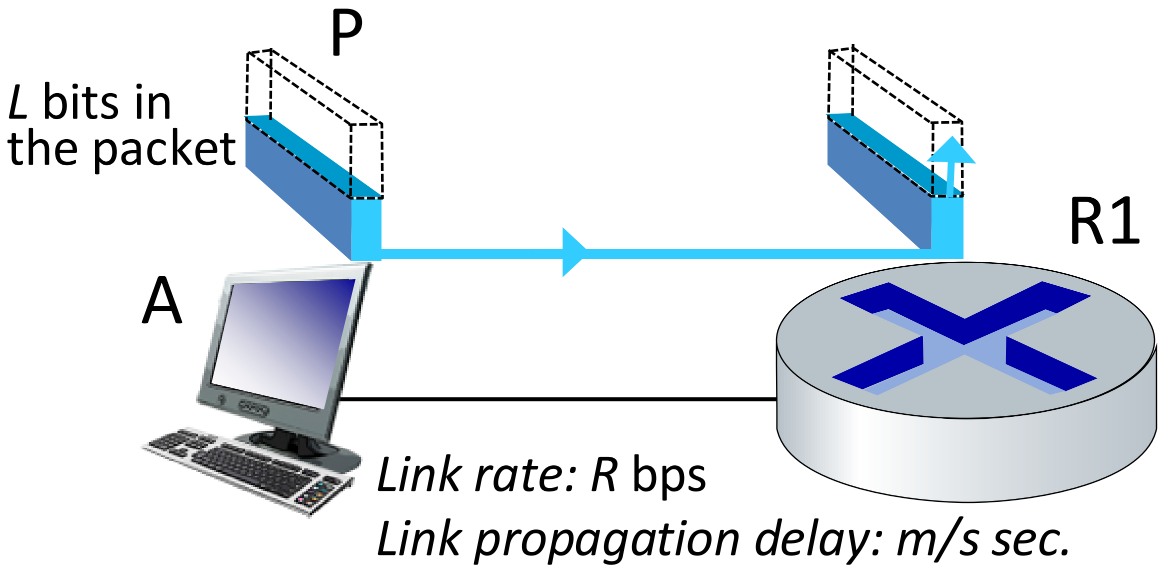
Network delay over a link

d_{trans}: transmission delay:

- L: packet length (bits)
- R: link *transmission rate (bps)*
- $d_{trans}(A, R1, P) = L/R$

d_{prop}: propagation delay:

- m: length of physical link
- s: propagation speed (~2x10^8 m/sec)
- d_{prop} (A, R1) = m/s



Network delay over a link

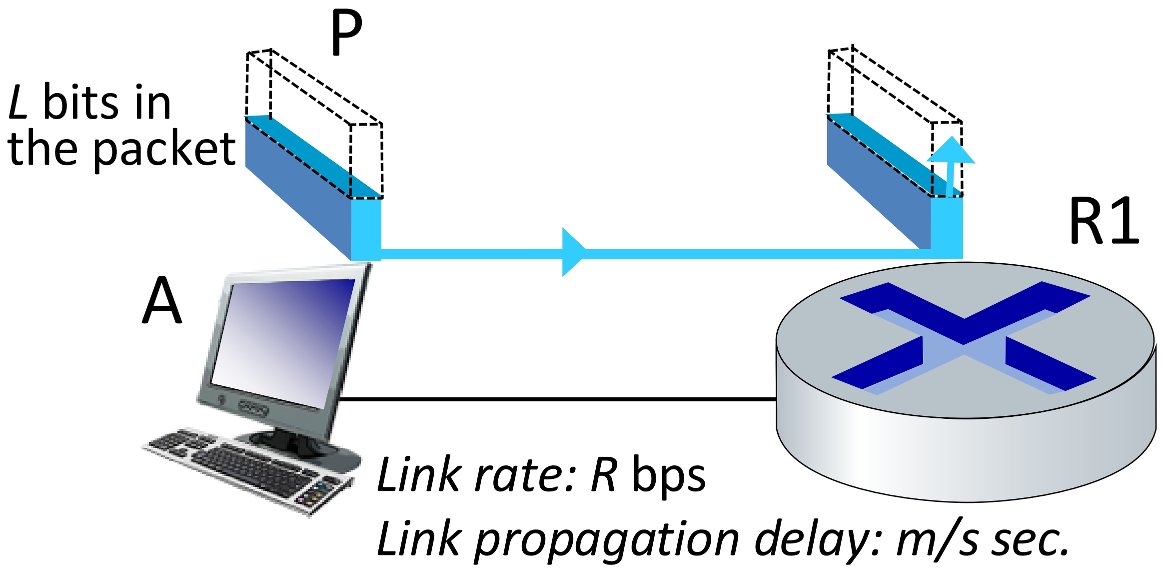
d_{trans}: transmission delay:

- *L*: packet length (bits)
- *R*: link *transmission rate (bps)*
- d_{trans} *(A, R1, P) = L/R*

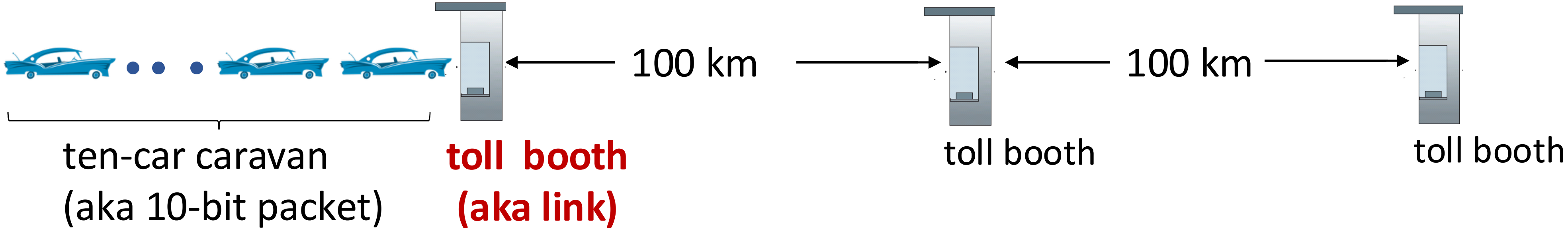
d_{prop}: propagation delay:

- *m*: length of physical link
- *s*: propagation speed (~2x10^8 m/sec)
- d_{prop} (A, R1) = *m/s*

d_{trans} and d_{prop} *very* different



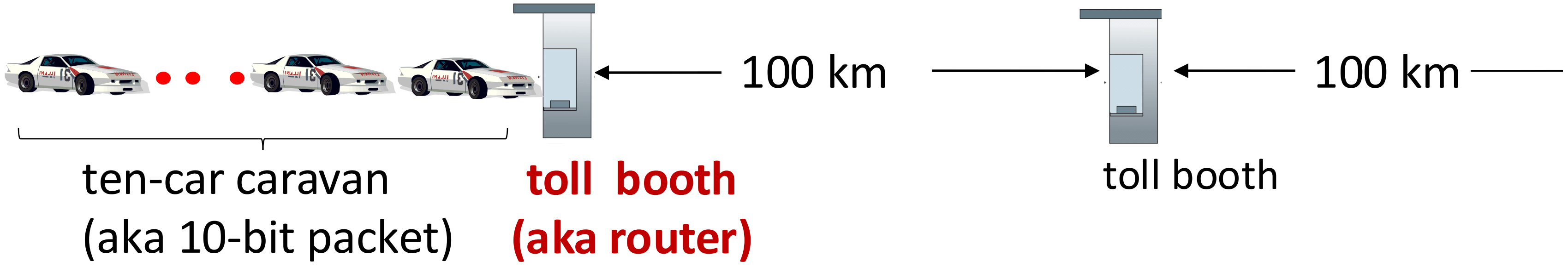
Caravan analogy #1



- car ~ bit; caravan ~ packet; toll service ~ link transmission
- toll booth takes 12 sec to service a car (bit transmission time)
- "propagate" at 100 km/hr
- *Q:* How long until caravan is lined up before 2nd toll booth?

- time to "push" entire caravan through toll booth onto highway = 12*10 = 120 sec
- time for last car to propagate from 1st to 2nd toll both: 100km/(100km/hr) = 1 hr
- *A:* 62 minutes

Caravan analogy #2



- suppose cars now "propagate" at 1000 km/hr
- and suppose toll booth now takes one min to service a car
- *Q: Will cars arrive to 2nd booth before all cars serviced at first booth?*

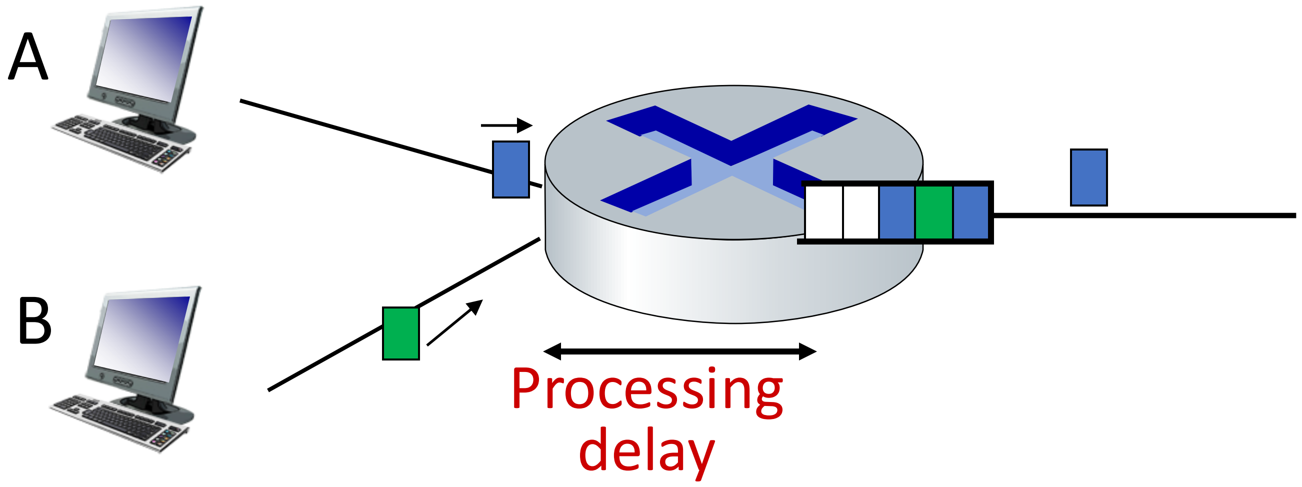
A: Yes! after 7 min, first car arrives at second booth; three cars still at first booth

1st bit of packet can arrive at 2nd router before packet is fully transmitted at 1st router!

Network delay within a node

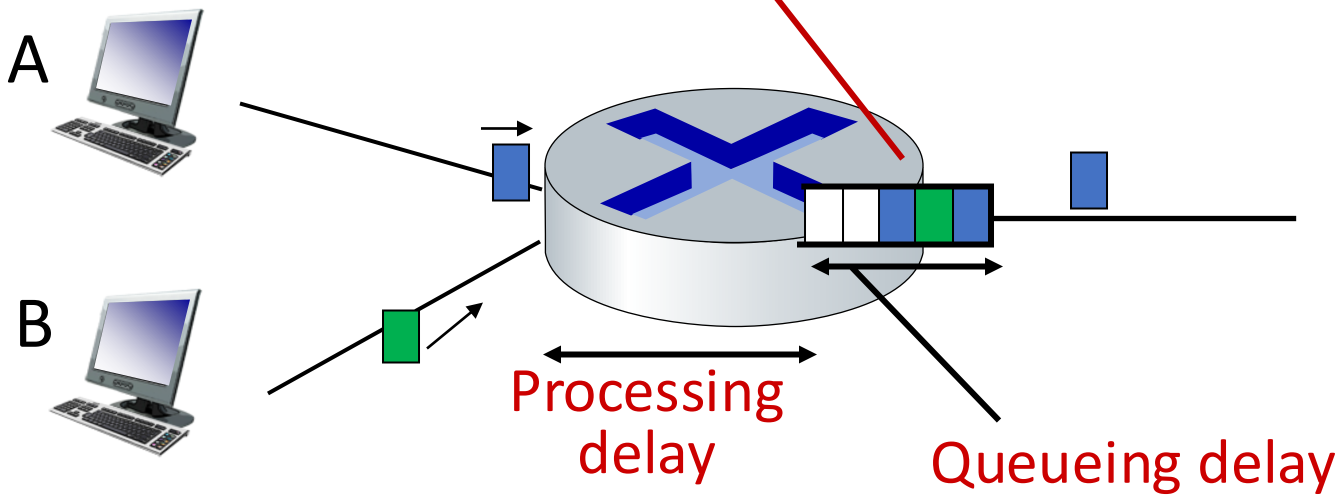
d_{proc}: processing delay

- Time it takes to process the packet ☺
  - check bit errors
    - What are some possible causes of bit errors?
  - determine output link
  - …
- typically < microsecs



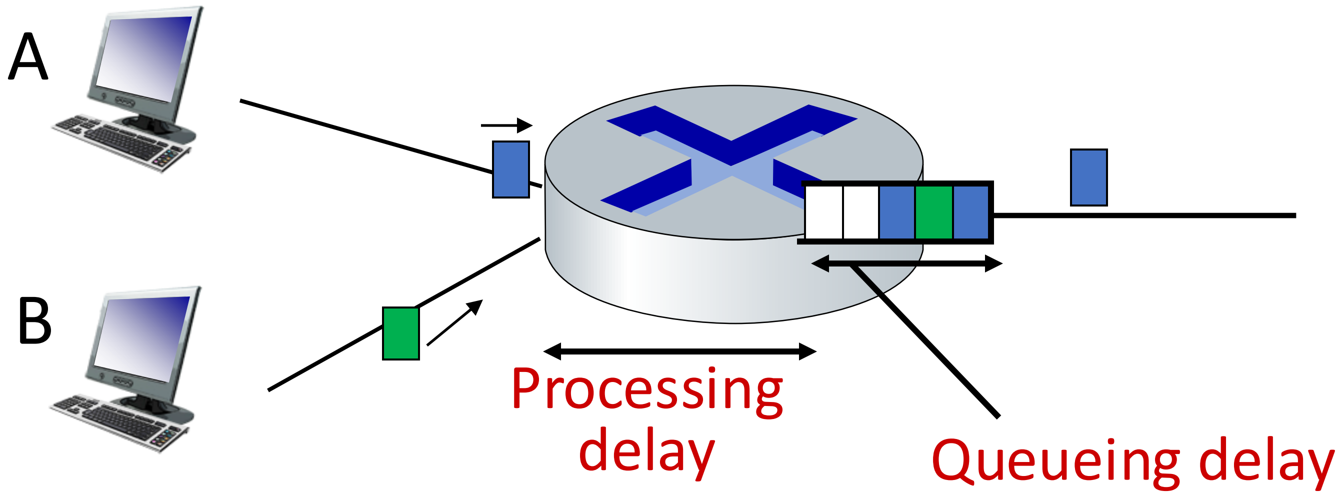
Network delay within a node

- packets *queue up* in network device buffers, waiting for turn for transmission



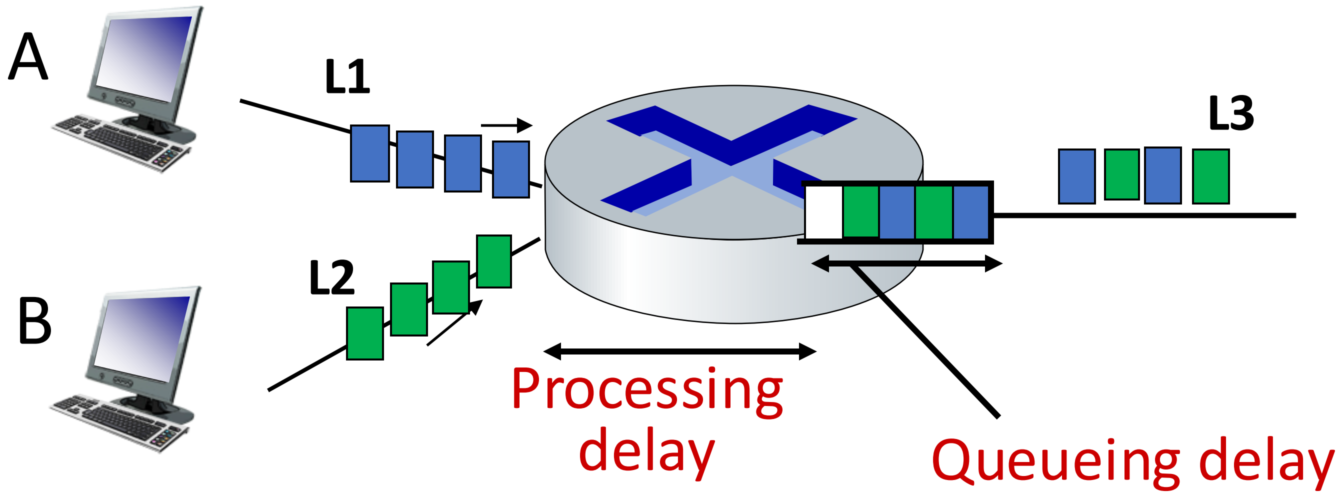
Network delay within a node

- packets *queue up* in network device buffers, waiting for turn for transmission
  - Queue length grows when arrival rate to link (temporarily) exceeds output link capacity
  - When could this happen?



Example of queue build-up in a network device

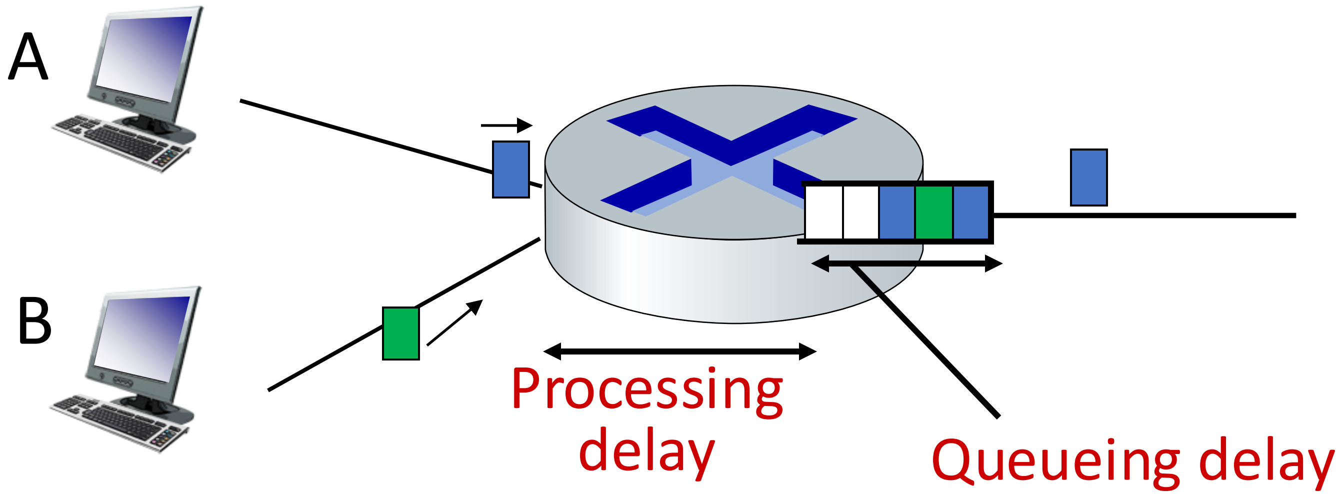
- Suppose each link can transfer 4 packets per second in each direction
- Suppose A and B each send 4 packets on L1 and L2, respectively, in one second, and all 8 packets are supposed to go out L3.
- L3 can only send 4 packets out in one second
- The next 4 should wait in the queue another second before they are transmitted.



Network delay within a node

d_{queue}: queueing delay

- time waiting at output link for transmission
- depends on how ”congested” that link is.

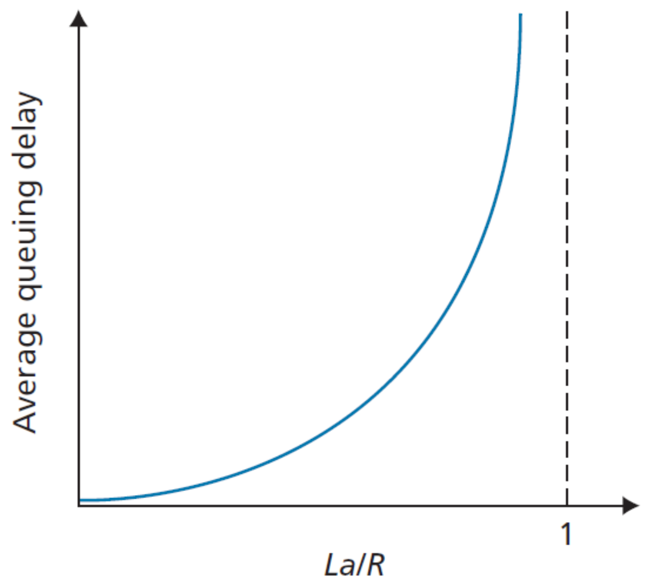


Network delay within a node

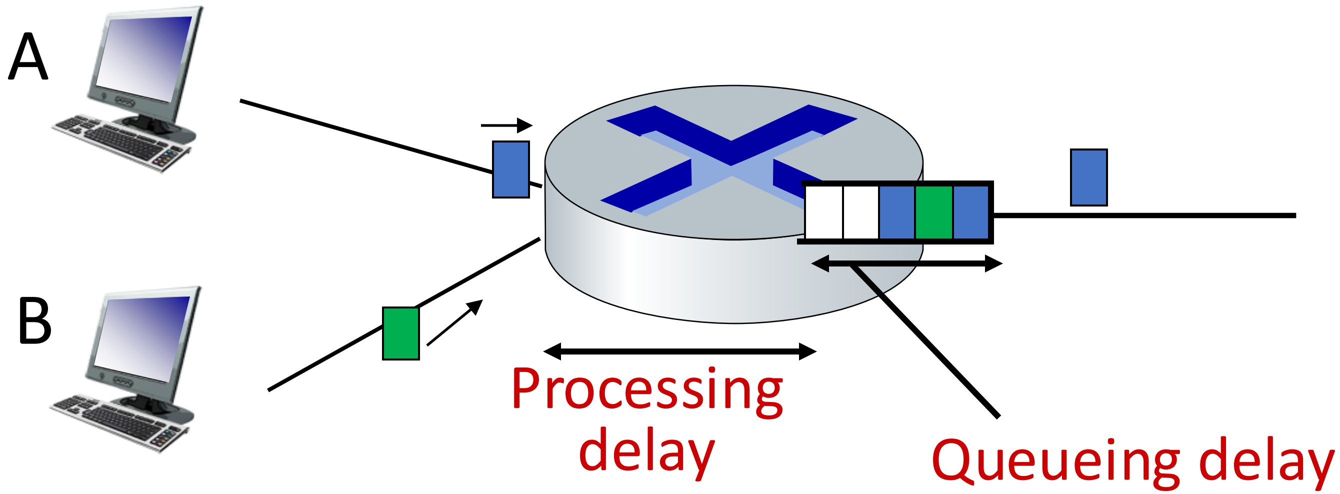
d_{queue}: queueing delay

- time waiting at output link for transmission
- depends on how ”congested” that link is.

See page 40 in the book for a more ”formal” intuition.



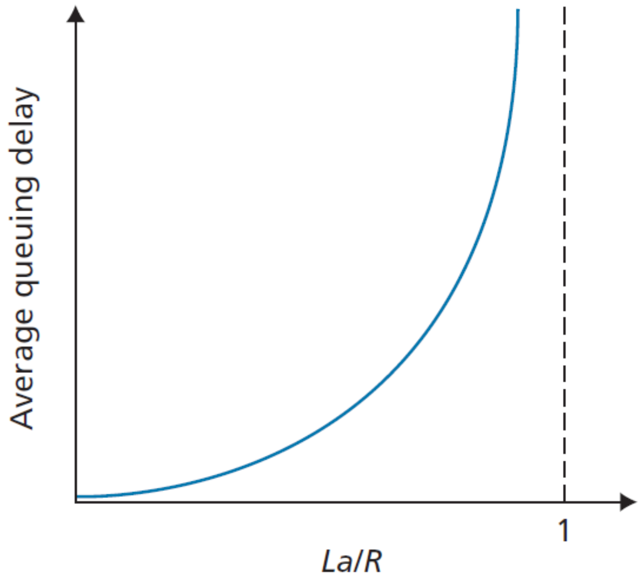
Traffic intensify La/R. a is the average packet arrival rate.



Queueing delay (revisited)

- R=link bandwidth (bps)
- a=average packet arrival rate
- L=packet length (bits)
- La is average bits arrival rate

traffic intensity = La/R



- La/R ~ 0: ?
  - average queueing delay small
- La/R -> 1:
  - delays become large
- La/R > 1:
  - more "work" arriving than can be serviced, average delay infinite!

Putting all the delay components together

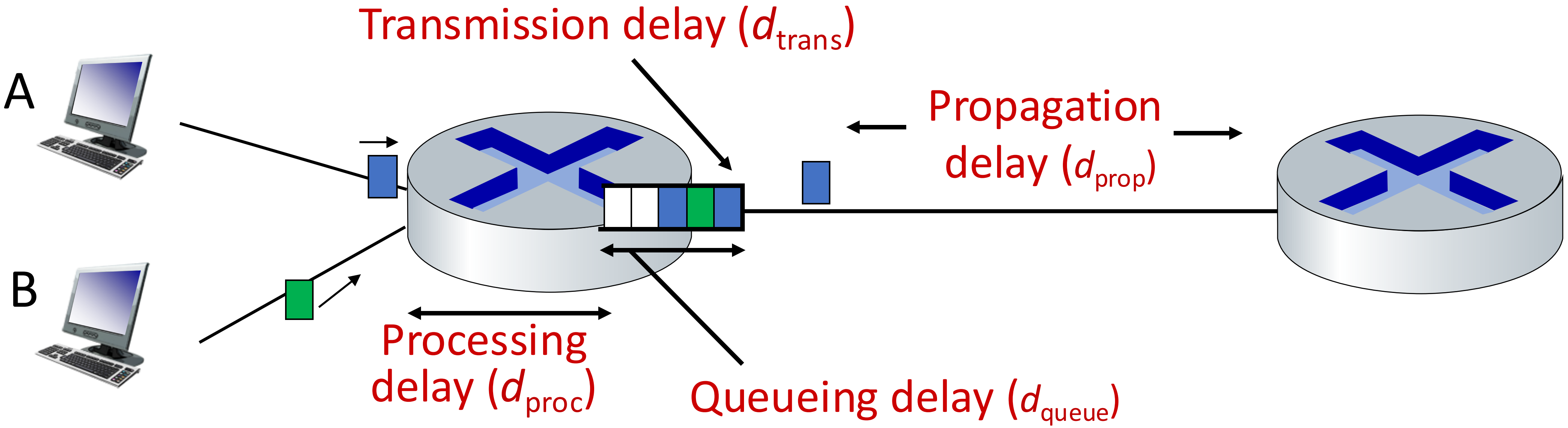
$$\boxed{d_{\text{nodal}} = d_{\text{proc}} + d_{\text{queue}} + d_{\text{trans}} + d_{\text{prop}}}$$

- d_{proc} = processing delay
  - typically a few microsecs or less
- d_{queue} = queuing delay
  - depends on congestion
- d_{trans} = transmission delay
  - = L/R, significant for low-speed links
- d_{prop} = propagation delay
  - a few microsecs to hundreds of msecs

Putting all the delay components together

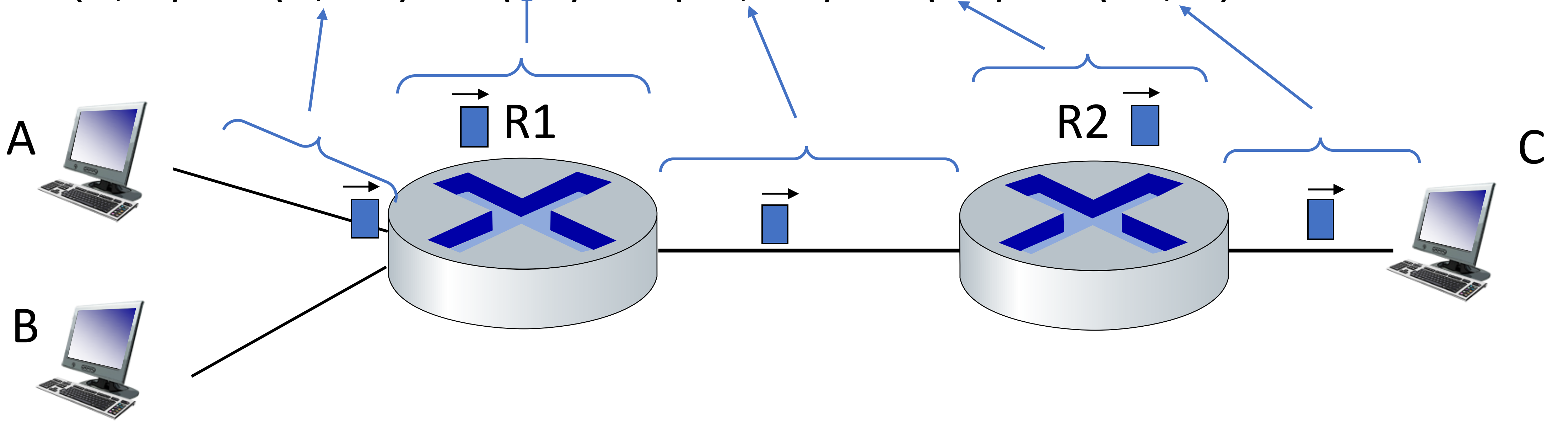
The end-to-end delay for a packet:

the sum of d_{proc} , d_{queue}, d_{trans} , and d_{prop} for all the components (devices and links) along the path.



Putting all the delay components together

- D(x, y): transmission delay (d_{trans}) + propagation delay (d_{prop})
- D(x): processing delay (d_{proc}) + queueing delay (d_{queue})
- D(A, C) = D(A, R1) + D(R1) + D(R1, R2) + D(R2) + D(R2, C)



Network performance metrics

End-to-End Delay

how long does it take for a piece of data to go from one end of the network to the other?

Loss

how often does the network lose data during transfer?

Throughput

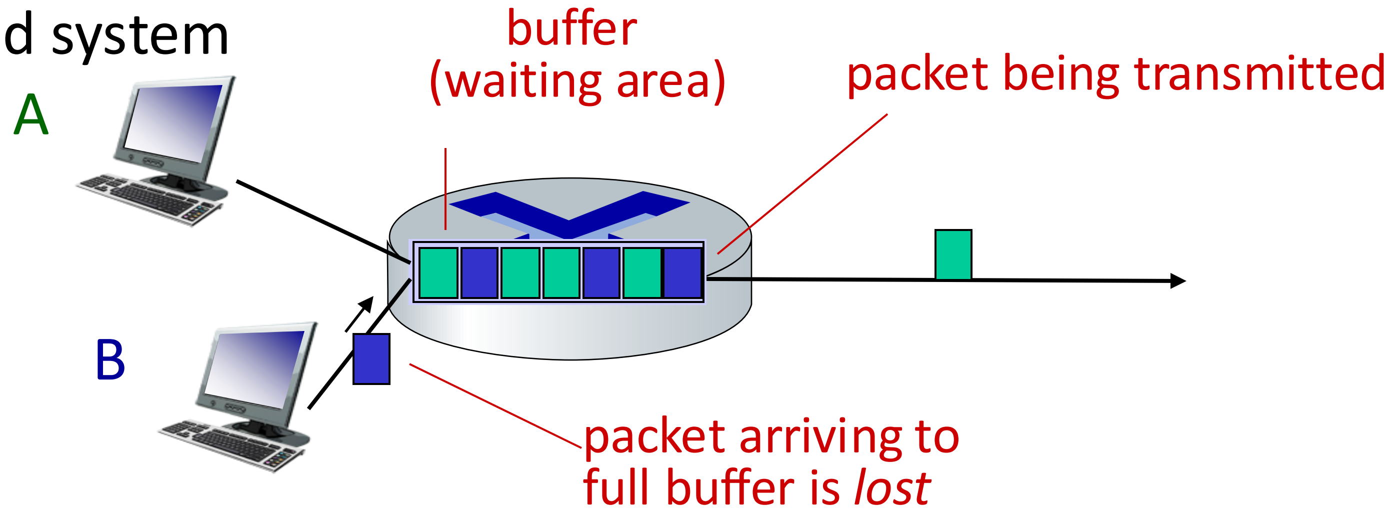
How much data per second can the network transfer?

Packet loss

- Packet loss/drop happens when a packet's transmission over the network is aborted before it reaches the destination.
- Two common causes for packet loss are
- corruption and network congestion
- Loss due to corruption: if, for any reason, packet bits are corrupted during transfer, network devices may be able to detect that and drop that packet.

Packet loss due to congestion

- queues (aka buffers) in network devices have finite capacity
- packet *loss* occurs when memory to hold queued packets fills up
- That is, packets arriving to full queue are dropped (aka lost)
- lost packet may be retransmitted by?
  - previous node
  - source end system
  - not at all



* Check out the Java applet for an interactive animation (on publisher's website) of queuing and loss

Network performance metrics

End-to-End Delay

how long does it take for a piece of data to go from one end of the network to the other?

Loss

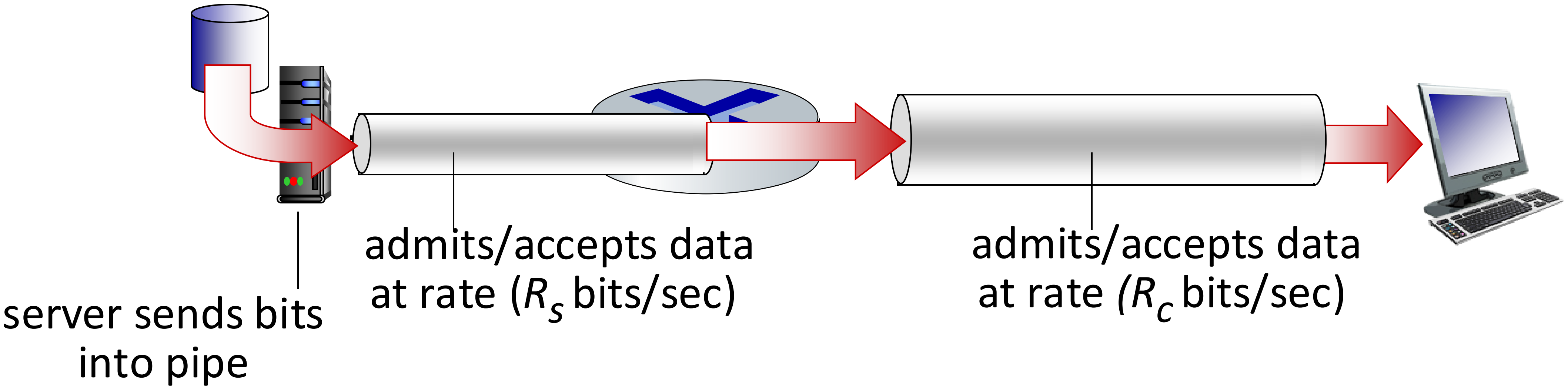
how often does the network lose data during transfer?

Throughput

How much data per second can the network transfer?

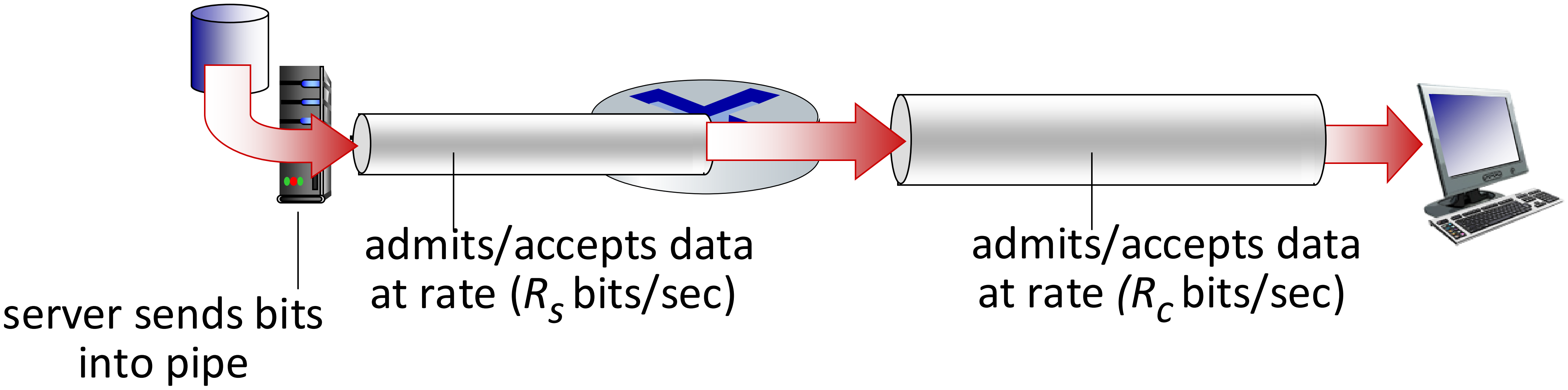
Throughput

- Throughput →rate (bits/time unit)
- Suppose we abstract a segment of a network as a data pipe.
- To find the throughput of that segment, we measure how many bits exit the end of the pipe within one time unit.



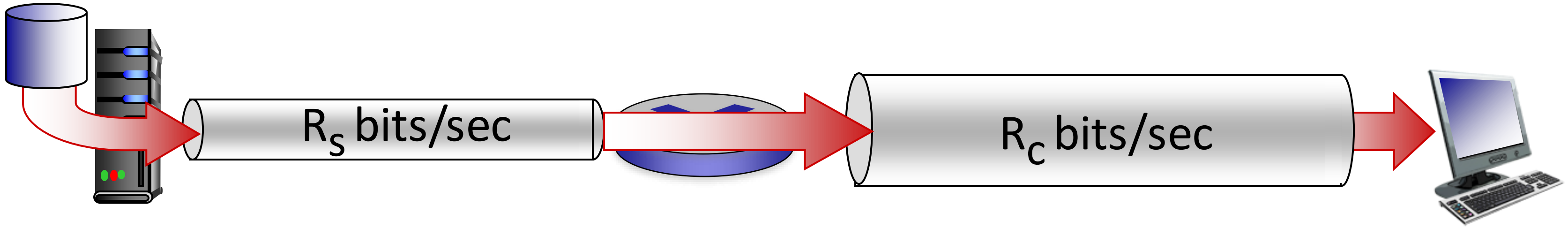
Throughput vs Delay

- Suppose the "pipe"'s throughput is R bits/sec and its propagation delay is d sec.
- If we look at the end of the pipe, R bits of data exit (or enter) the pipe in one second.
- It takes each bit d sec. to go from one end of the pipe to the other.

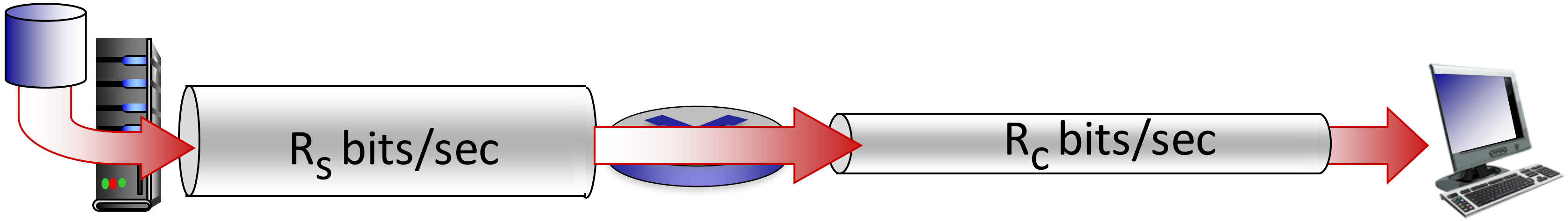


Throughput

$R_s < R_c$ What is end-end throughput?



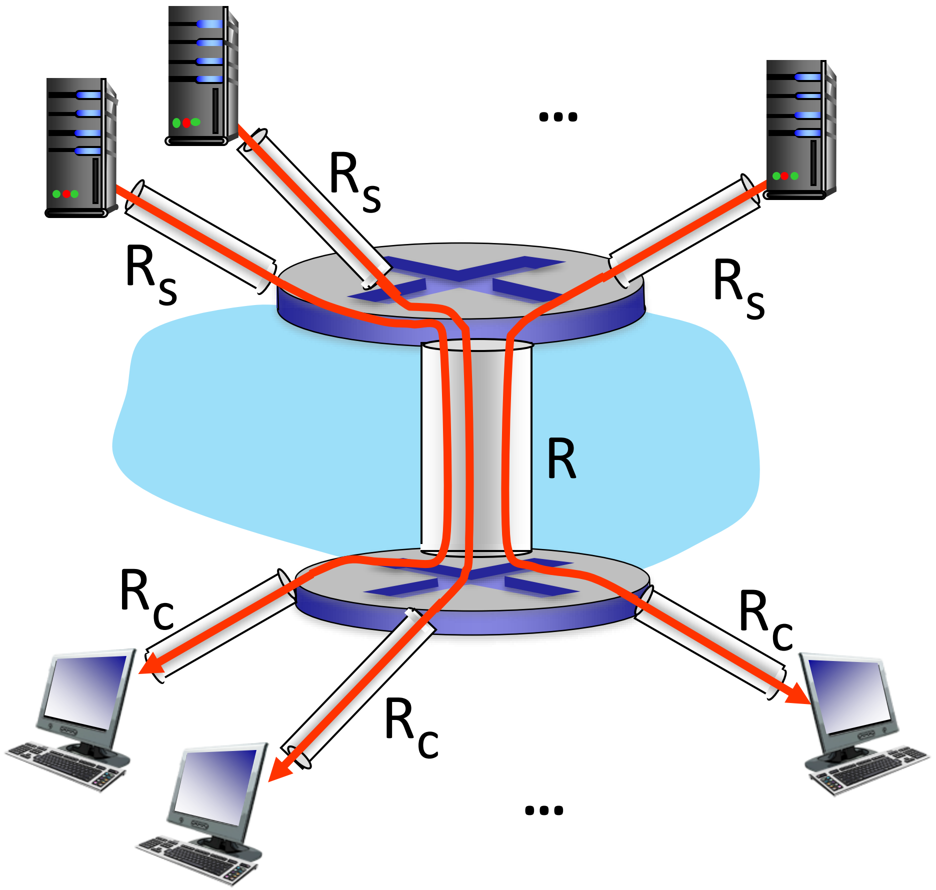
$R_s > R_c$ What is end-end throughput?



bottleneck link

link on end-end path that constrains end-end throughput

Throughput: network scenario



Suppose 10 connections (fairly) share backbone link *R* bits/sec

- per-connection end-end throughput:?
- *min(R_c,R_s,R/10)*
- in practice: R_c or R_s is often smaller than R/N
  - Bottleneck link is often closer to the edge of the network.

* Check out the online interactive exercises for more examples: http://gaia.cs.umass.edu/kurose_ross/

Throughput units: bps

- bps = bits per second
  - 1 byte = 8 bits or 1 B = 8 b
- In networking, K/M/G are powers of 10
  - K or Kilo: 10^3
  - M or Mega: 10^6
  - G or Giga: 10^9
  - …
- Q: 1 MBps = ? Bps = ? bps

Network performance metrics

End-to-End Delay

how long does it take for a piece of data to go from one end of the network to the other?

Loss

how often does the network lose data during transfer?

Throughput

How much data per second can the network transfer?

What you need to know about network performance metrics

- Given a scenario, you should be able to compute different kinds of delay, loss, and throughput.
- You should know the difference between throughput and delay
- There will be some questions in this week’s quiz for practice, and we will solve some of them in class next week.

Recap: networks are complex!

- They have many pieces
  - Hosts, routers/switches (network devices), links, protocols, …
- They can get quite large
  - Thousands if not millions of hosts and devices
- They are often shared among many traffic flows
- Making it challenging to design efficient and effective networks
- There are various metrics to measure their performance
- …

Introduction

What is a computer network? ✔

What is the Internet?

Introduction: 1-49

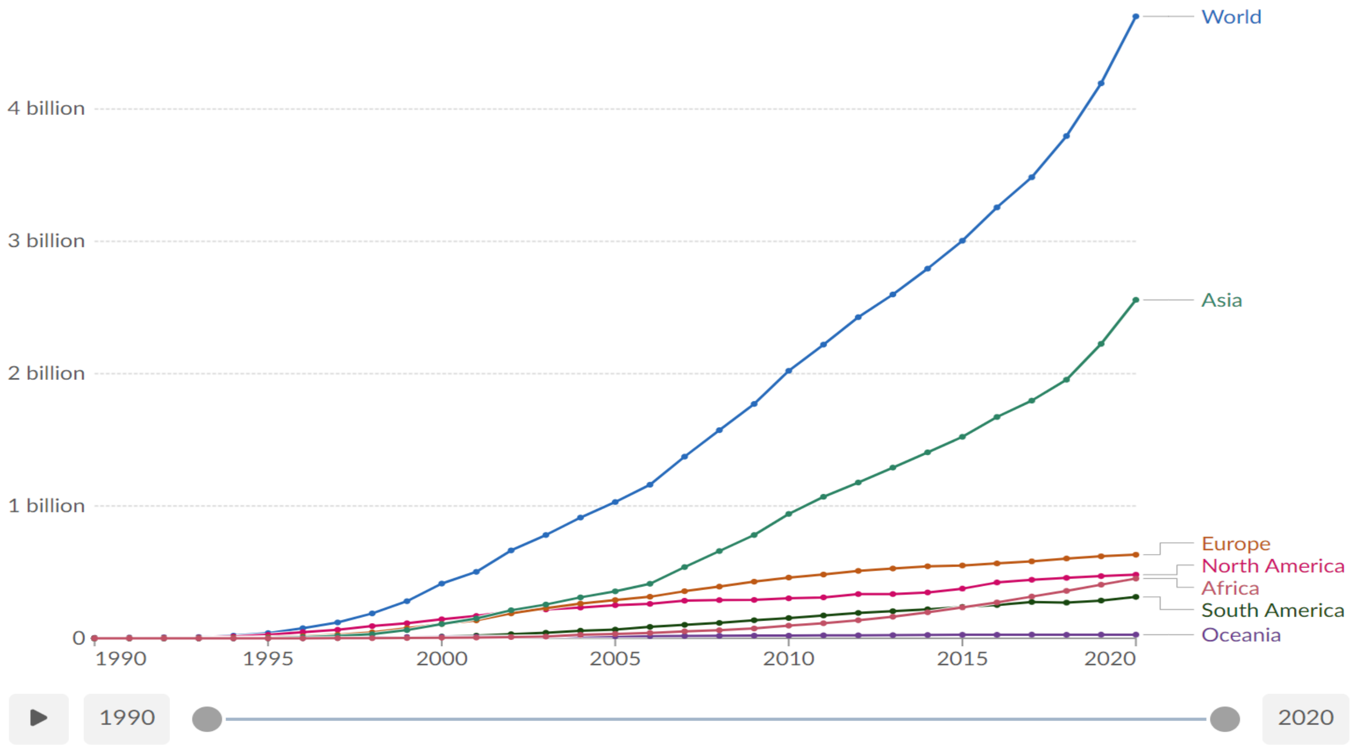
Introduction

What is a computer network? ✔

What is the Internet?

Introduction: 1-50

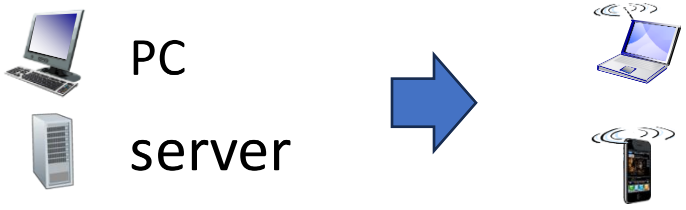
Internet – connecting billions of users



Tweet-a-watt: monitor energy use

Autonomous vehicles

Internet refrigerator



PC

server

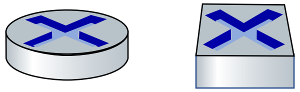
laptop

cellphone

Internet of Things

[Data source: OWID based on International Telecommunication Union (via World Bank) and UN (2022)]

The Internet: a “nuts-and-bolts” view



routers/switches

wired links

access points

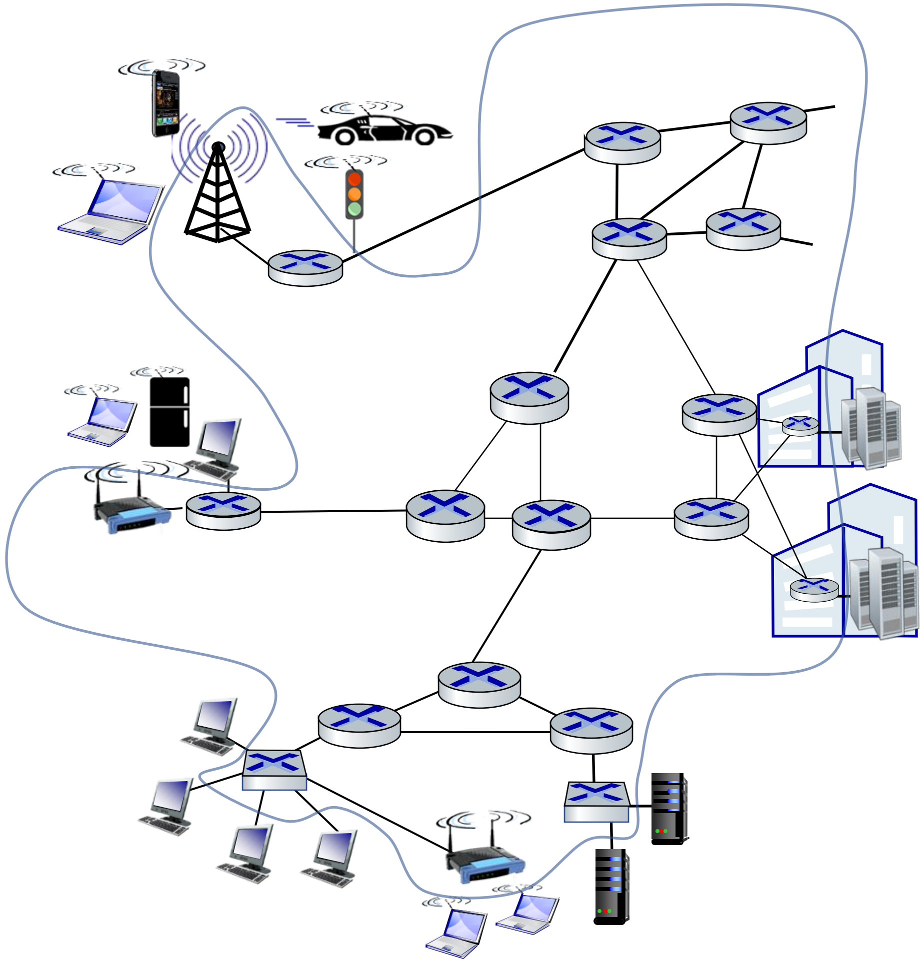
Billions of connected computing *devices*

- running applications

Switches/routers that forward packets (chunks of data)

Communication links:

- fiber, copper, radio, satellite



ISP: internet service provider

The Internet: A network of networks

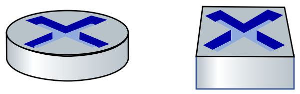

PC
server
laptop
cellphone

Billions of connected computing *devices*
- running applications

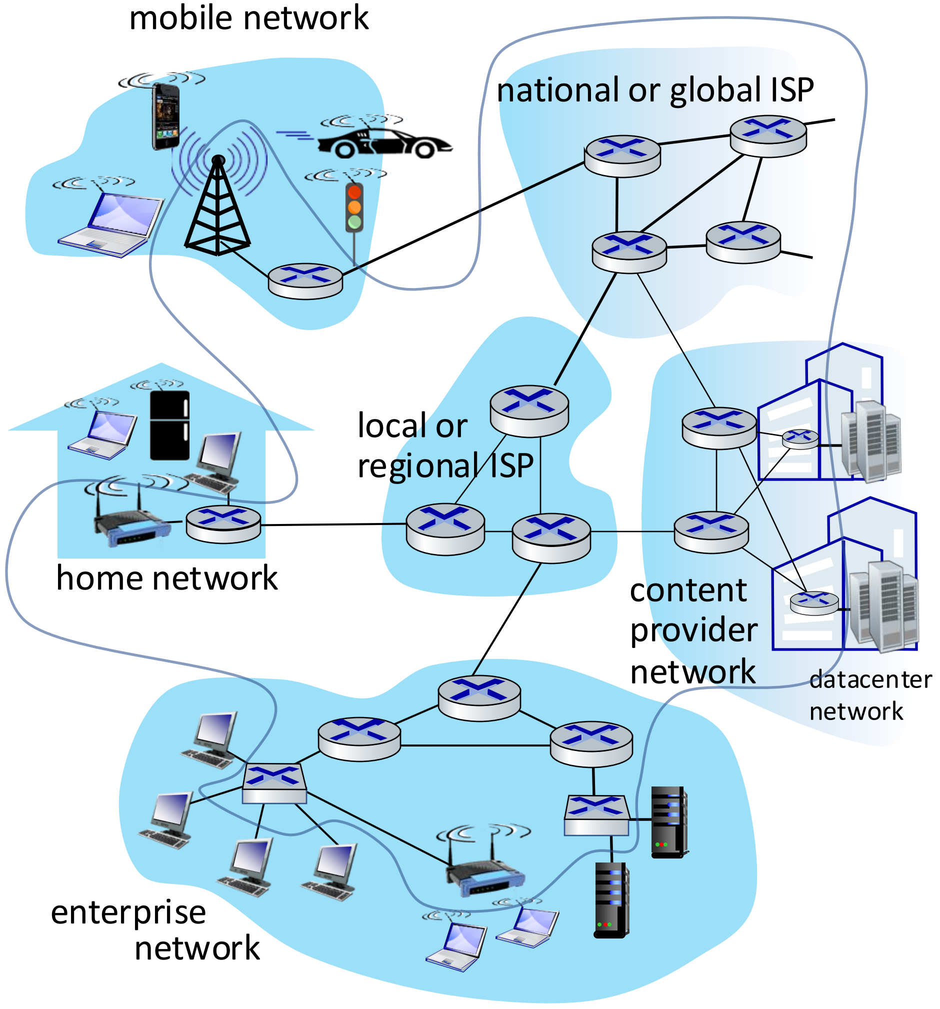

wired links


access points

Communication links:
- fiber, copper, radio, satellite

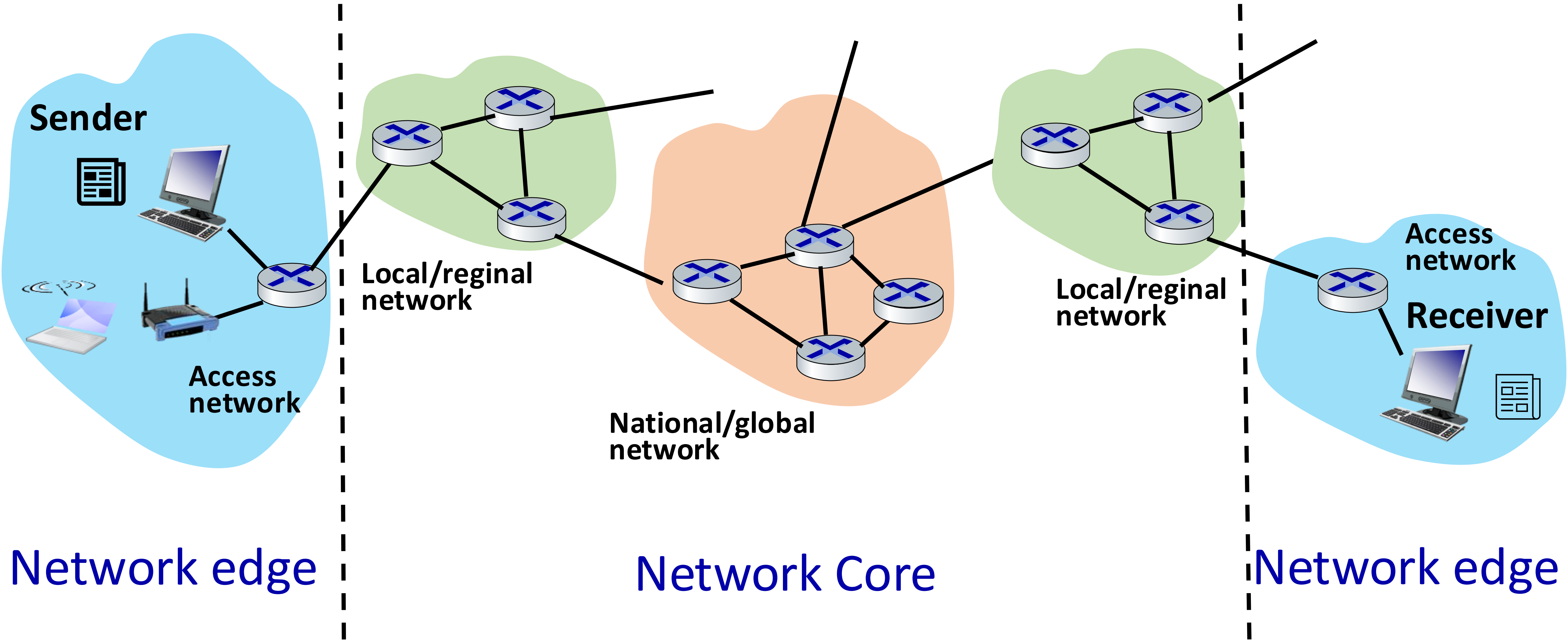

routers/switches

Switches/routers that forward packets (chunks of data)

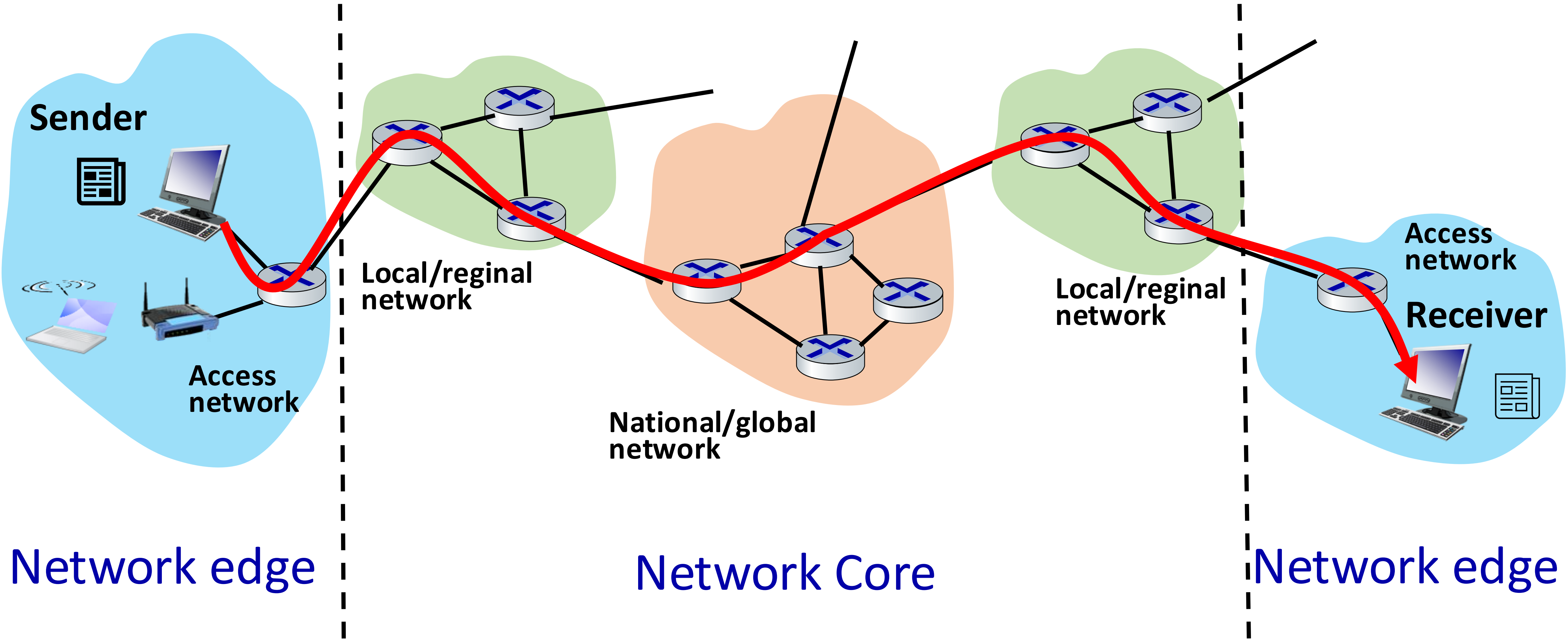


ISP: internet service provider

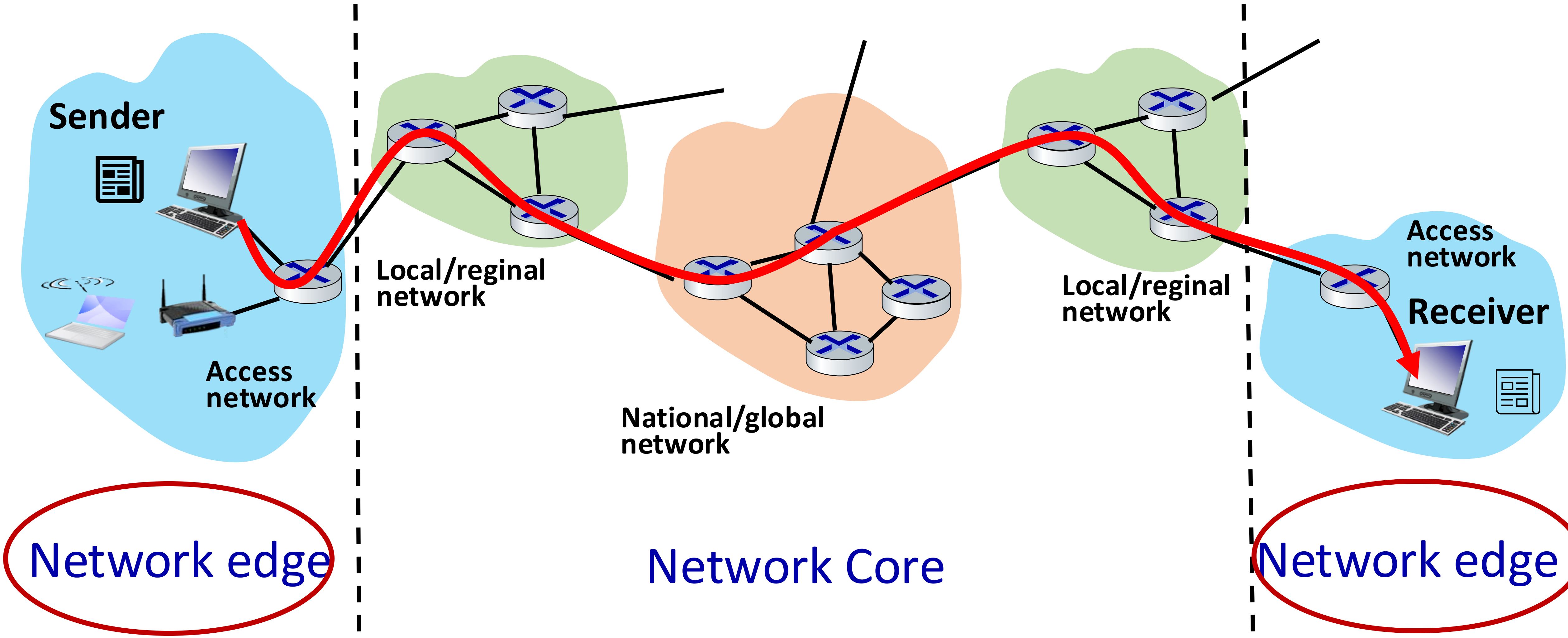
A packet passes through many networks



A packet passes through many networks



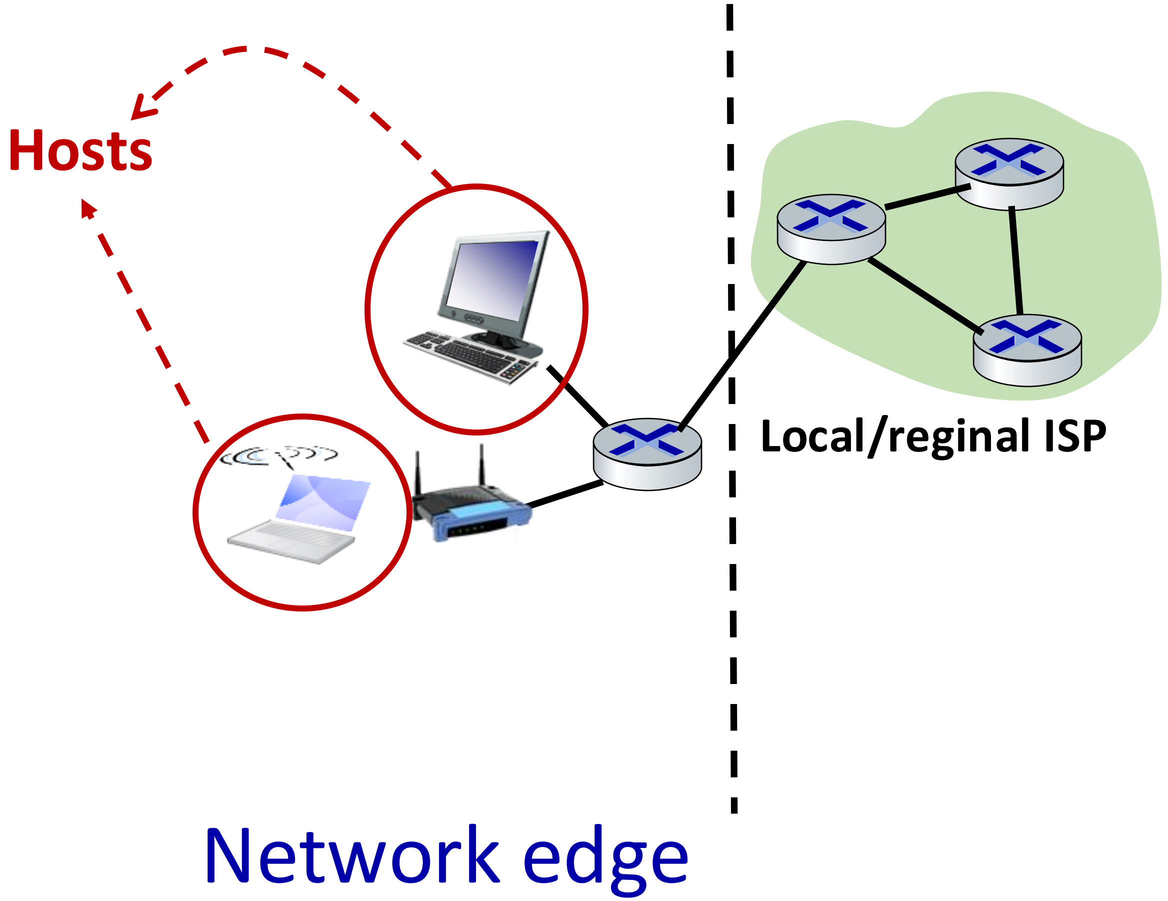
A packet passes through many networks



The network edge

End systems = hosts

- connected computing devices
- running applications that communicate with applications on other end systems



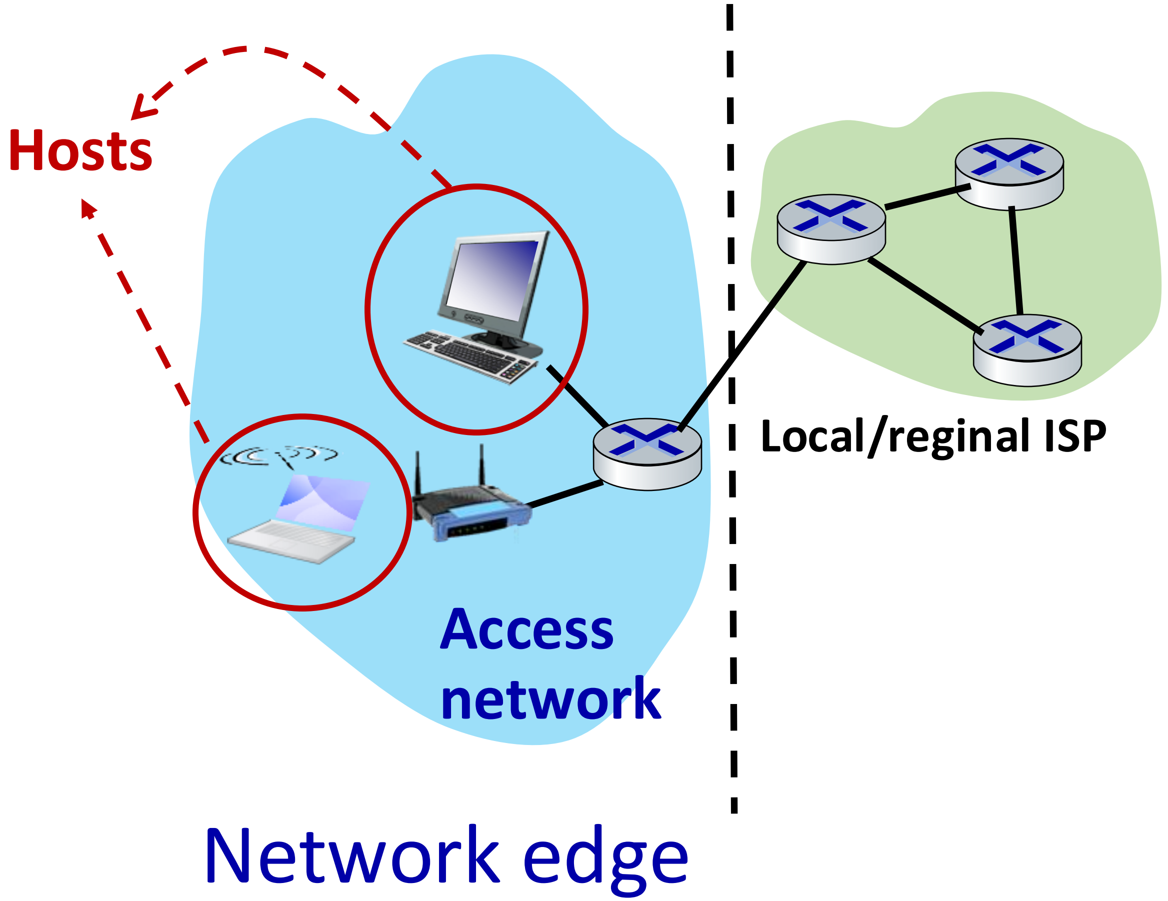
The network edge

End systems = hosts

- connected computing devices
- running applications that communicate with applications on other end systems

Access networks

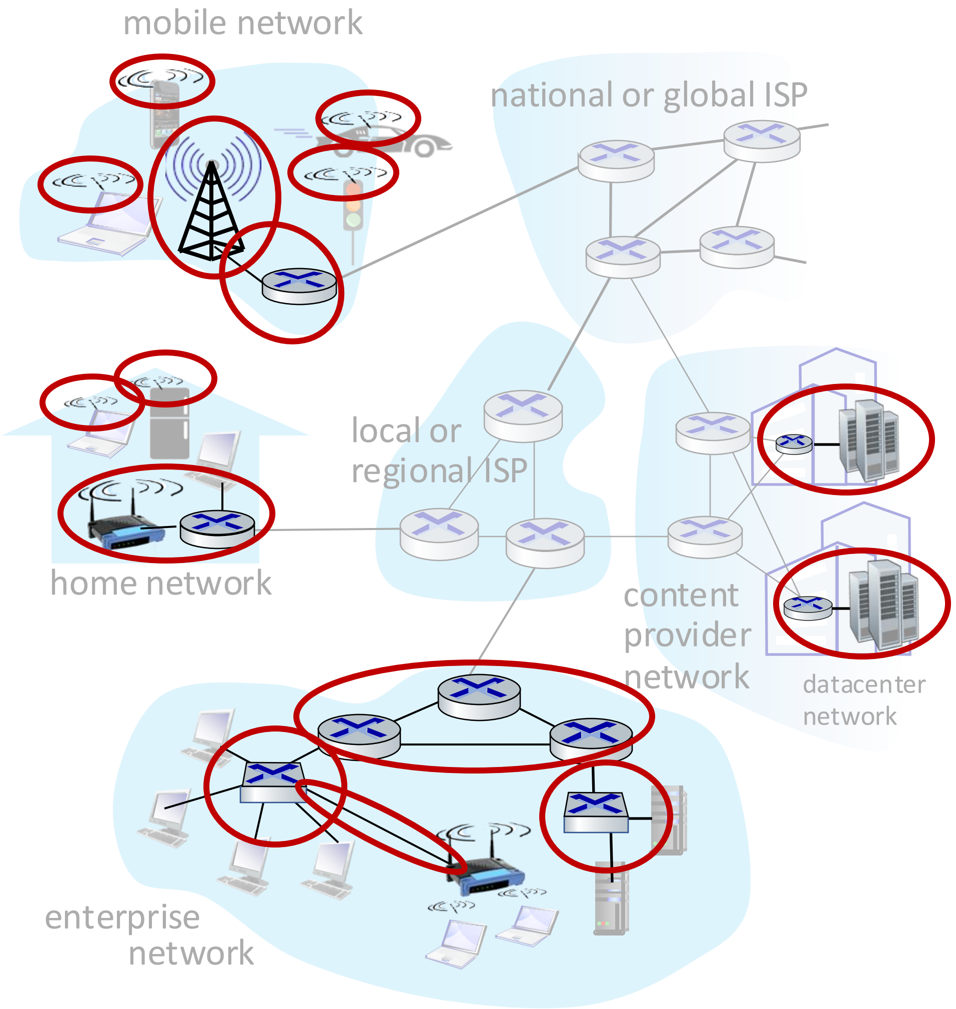
- access networks connect end systems to the edge routers
- edge router = entry point to the Internet
- wired or wireless



Network edge - access networks

There are different kinds of access networks:

- residential access networks
- wireless access networks (WiFi, 4G/5G)
- enterprise (Institutional) access networks
  - school, company, etc.
- data center networks
- …



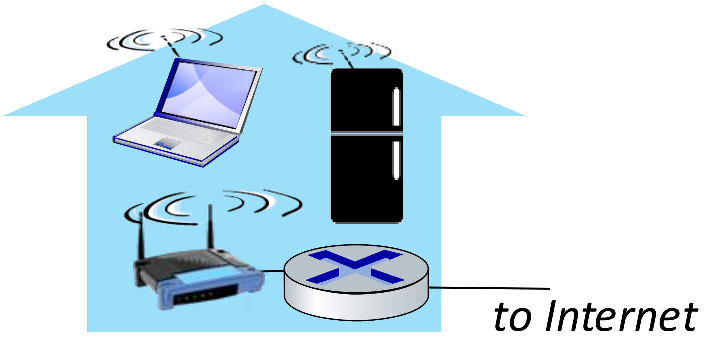
Wireless access networks

Shared *wireless* access network connects end system to router

- via base station aka “access point”

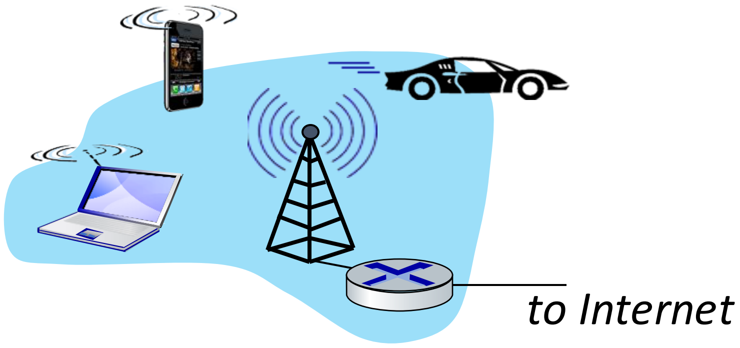
Wireless local area networks (WLANs)

- typically within or around building (~30 m)



Wide-area cellular access networks

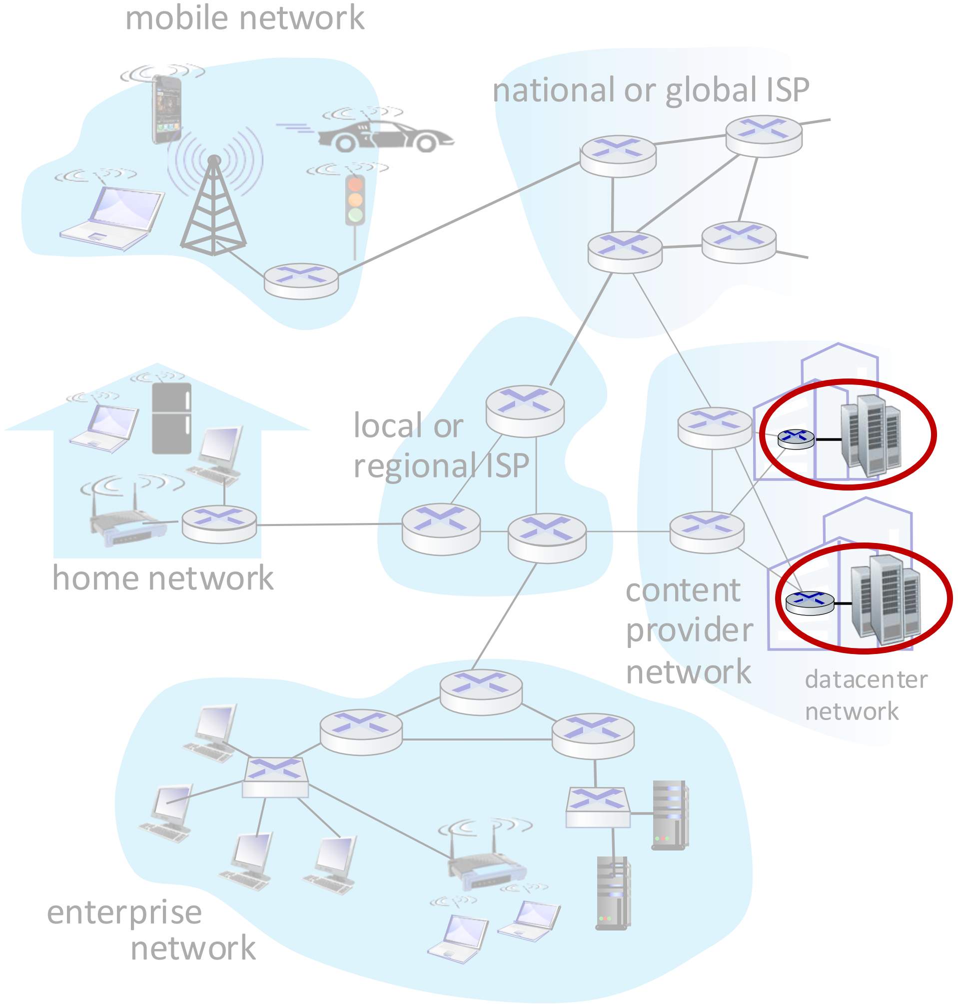
- provided by mobile, cellular network operator (10’s km)
- 4G/5G cellular networks



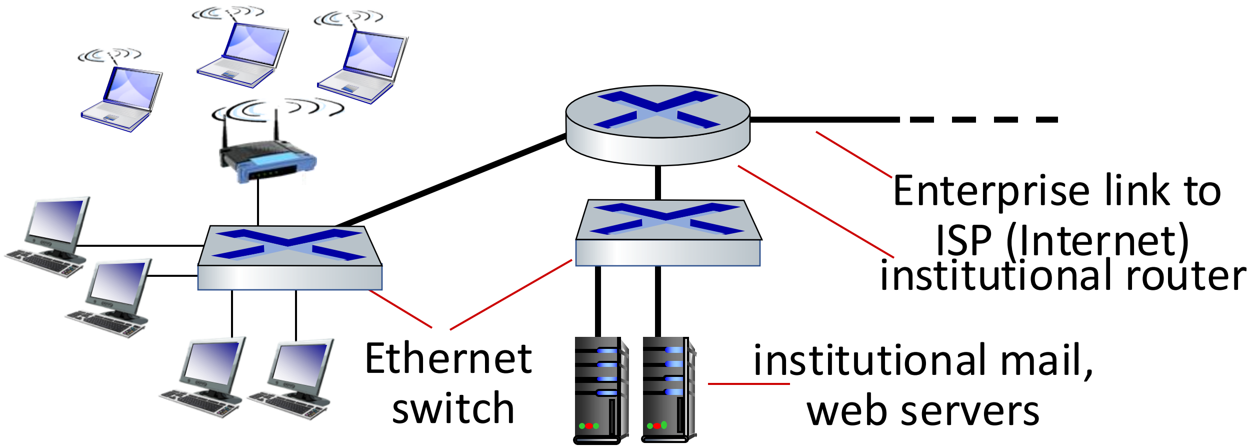
Data center networks

- high-bandwidth links (10s to 100s Gbps) connect hundreds to thousands of servers together, and to Internet

Courtesy: Massachusetts Green High Performance Computing Center (mghpcc.org)

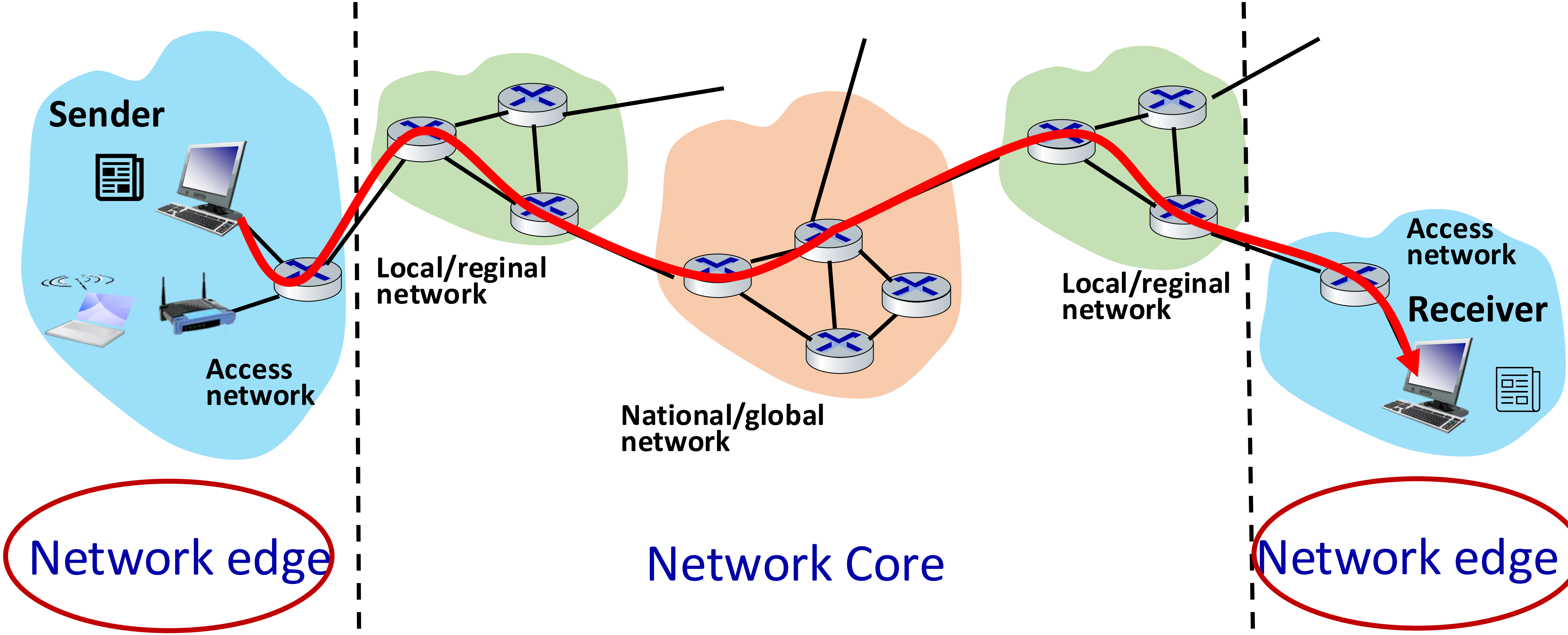


Enterprise access networks

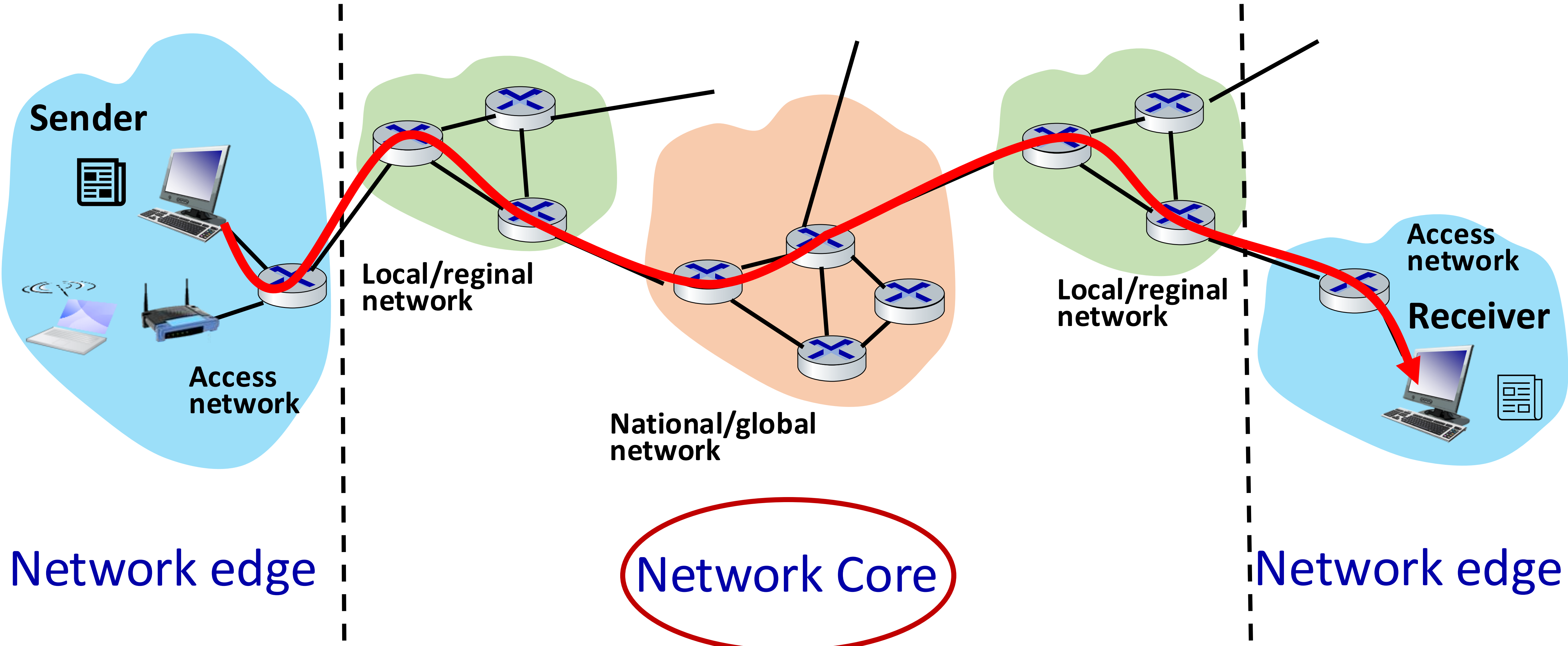


- companies, universities, etc.
- mix of wired, wireless link technologies, connecting a mix of switches and routers.

A packet passes through many networks

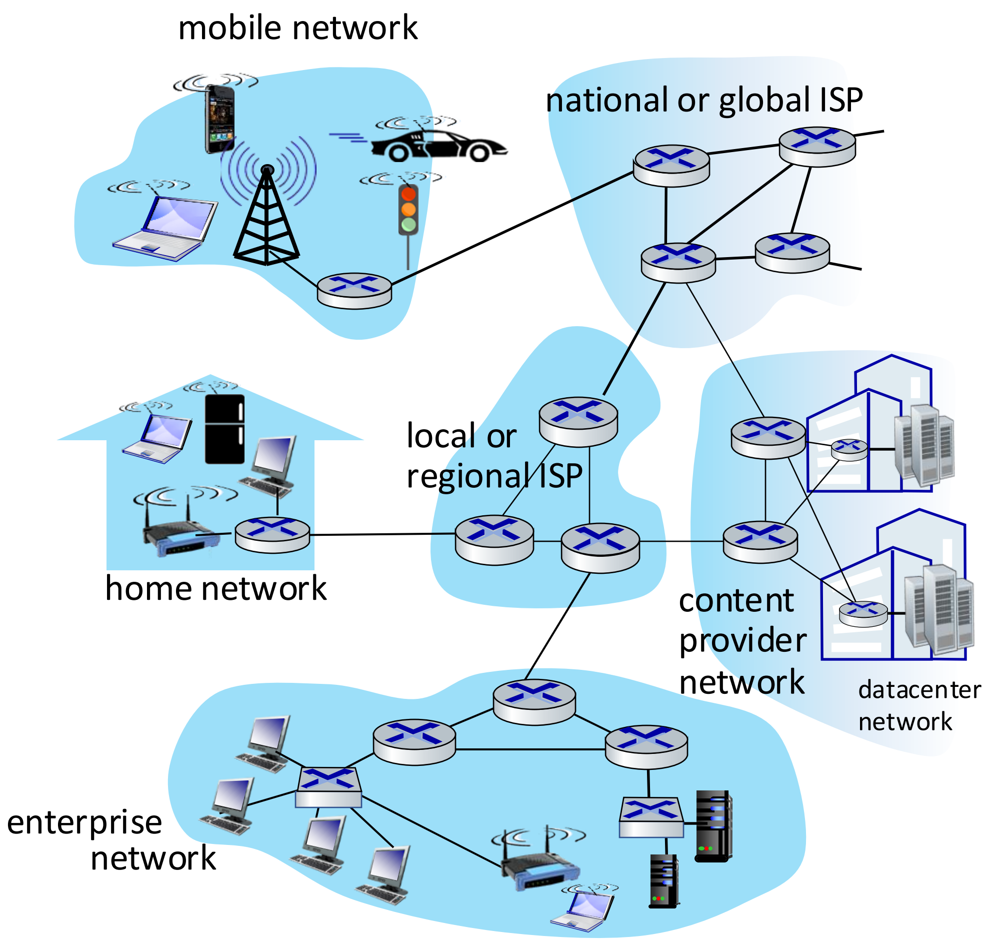


A packet passes through many networks



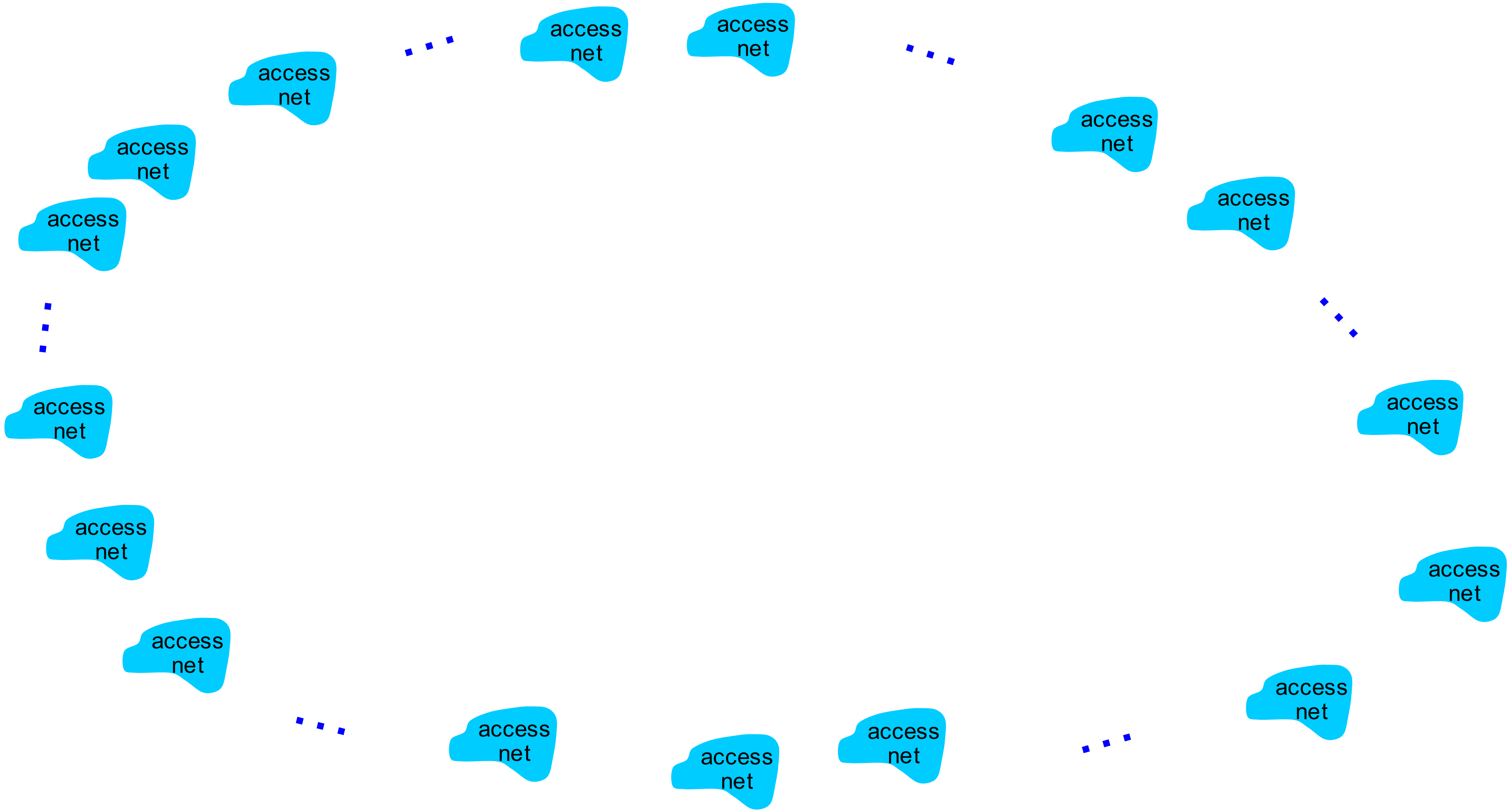
Network core → Connecting access networks

- hosts connect to Internet via access networks
- access networks in turn must be interconnected
  - so that *any* two hosts *(anywhere!)* can send packets to each other
- resulting network of networks (i.e., network core) is very complex
  - evolution driven by economics, national policies



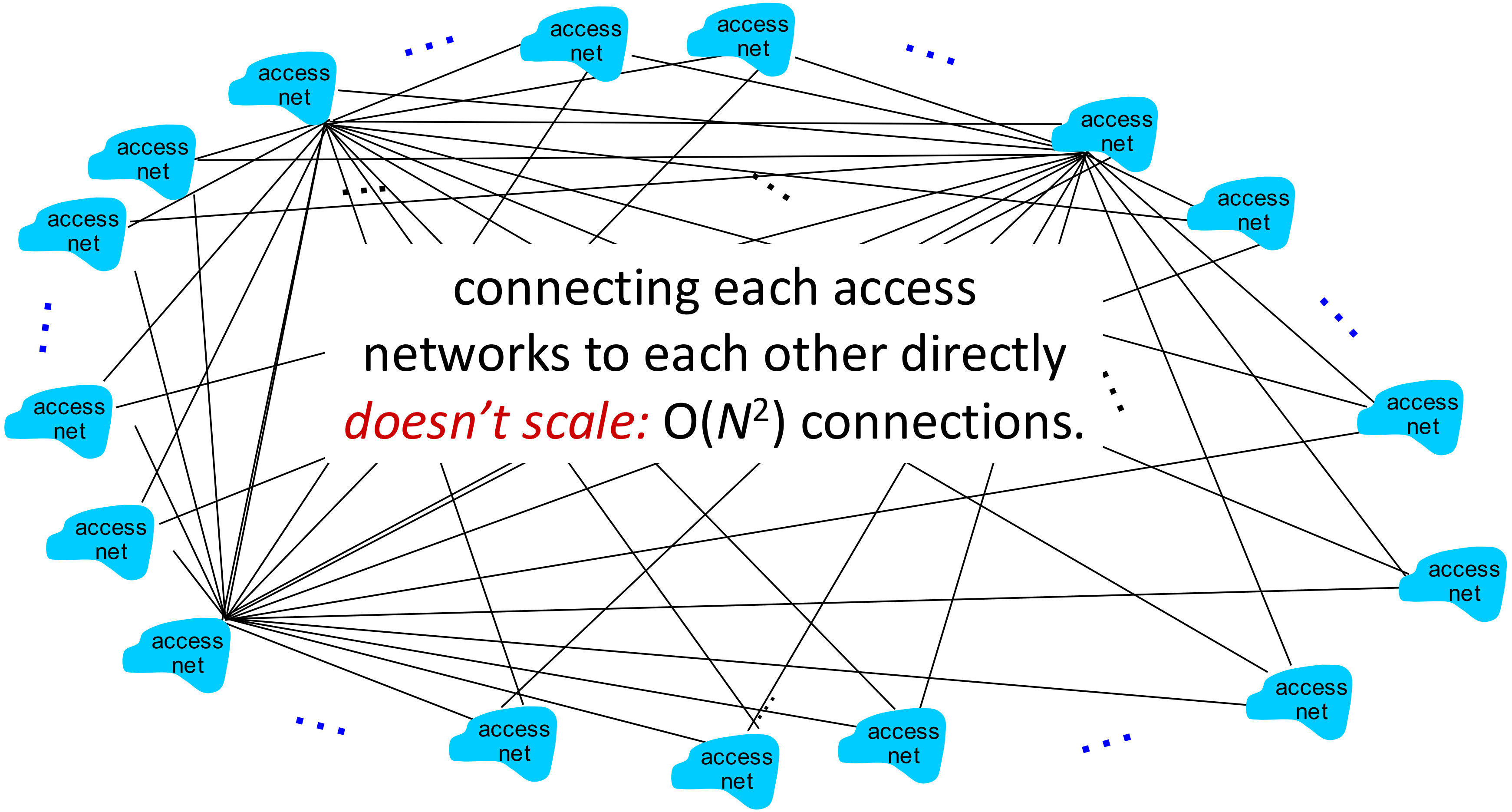
Network core → Connecting access networks

Question: given *millions* of access networks, how to connect them together?



Network core → Connecting access networks

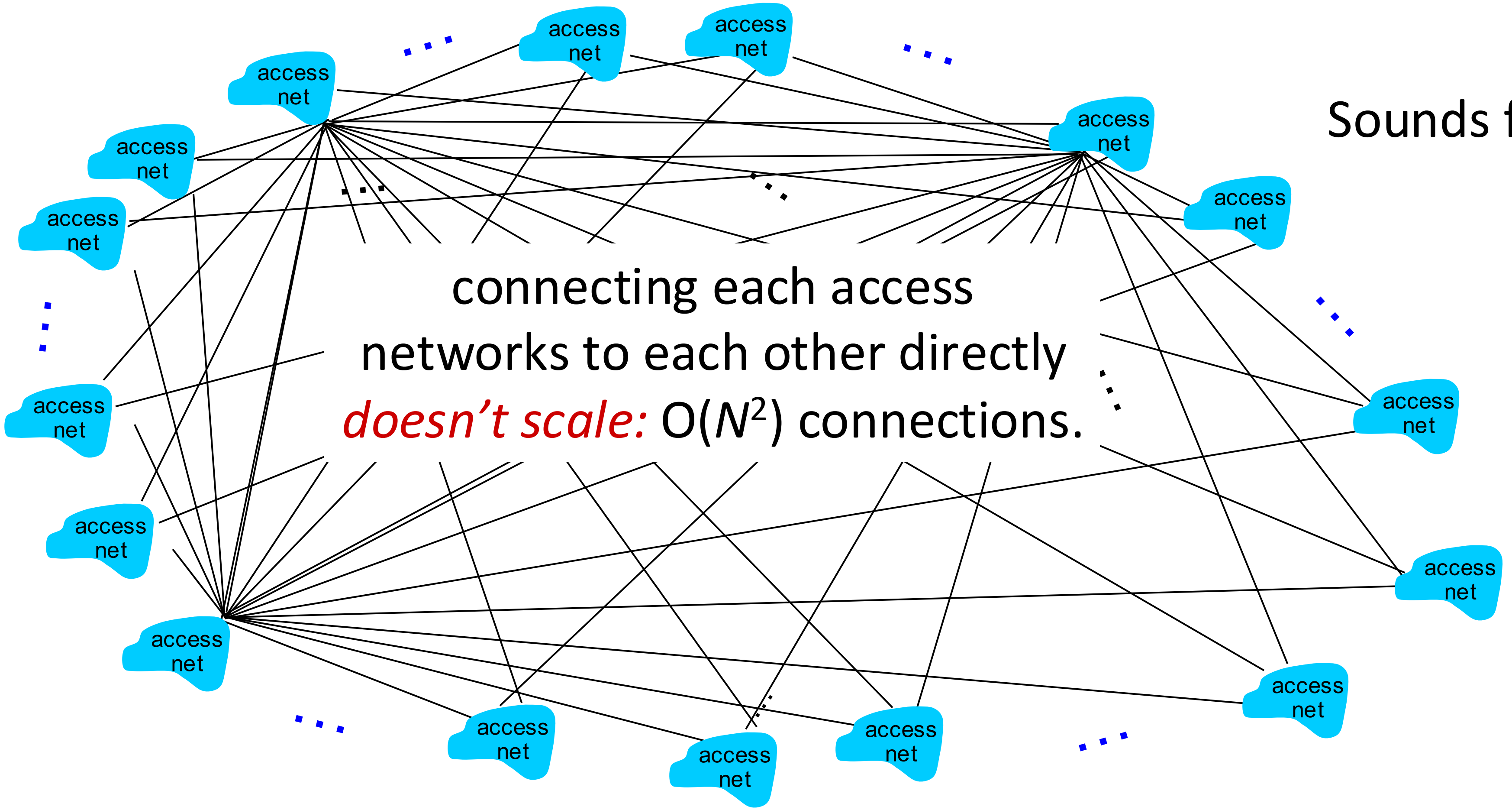
Question: given *millions* of access networks, how to connect them together?



connecting each access networks to each other directly *doesn't scale:* O(N^2) connections.

Network core → Connecting access networks

Question: given *millions* of access networks, how to connect them together?

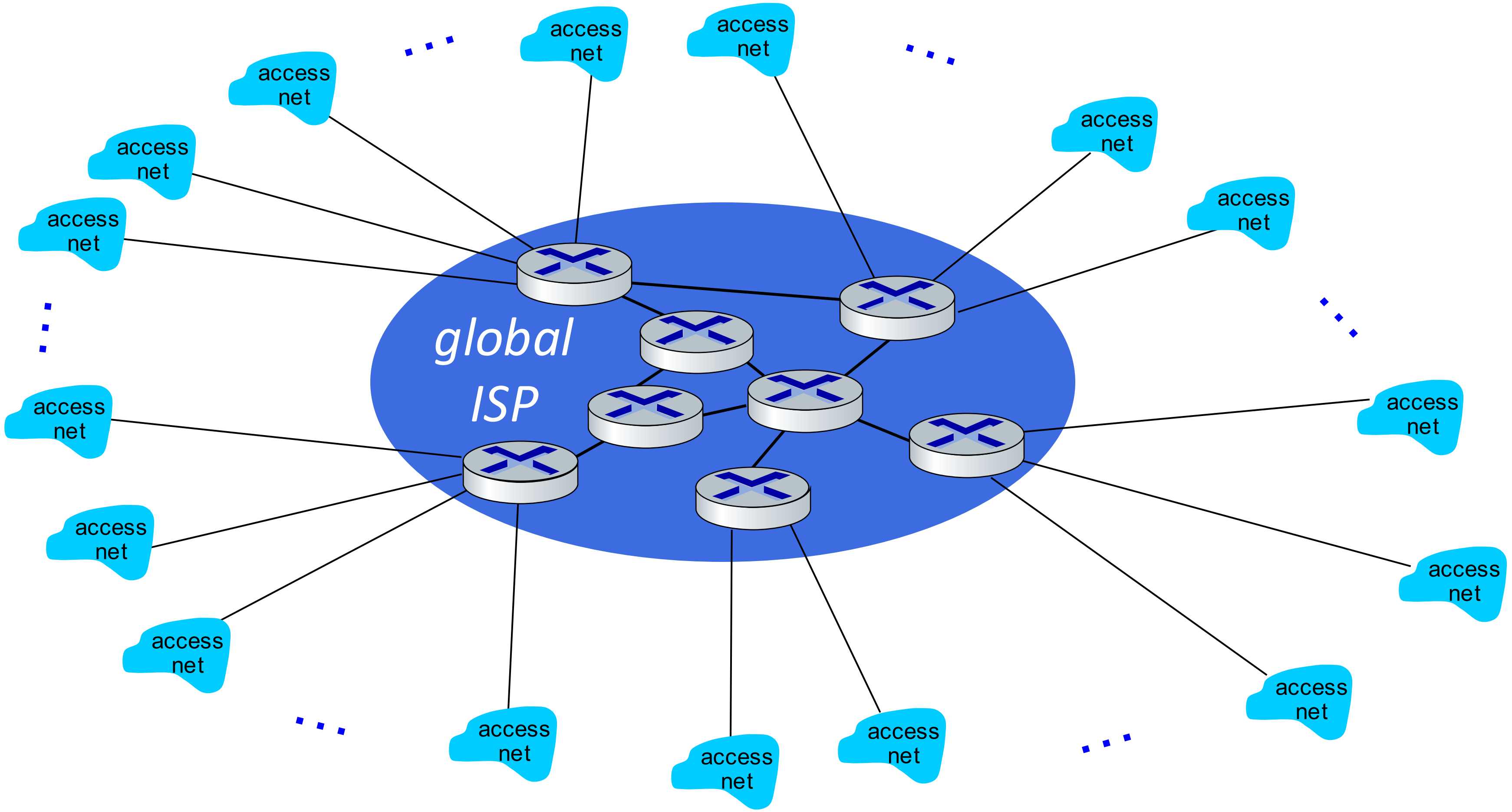


connecting each access networks to each other directly *doesn't scale:* O(N^2) connections.

Sounds familiar?

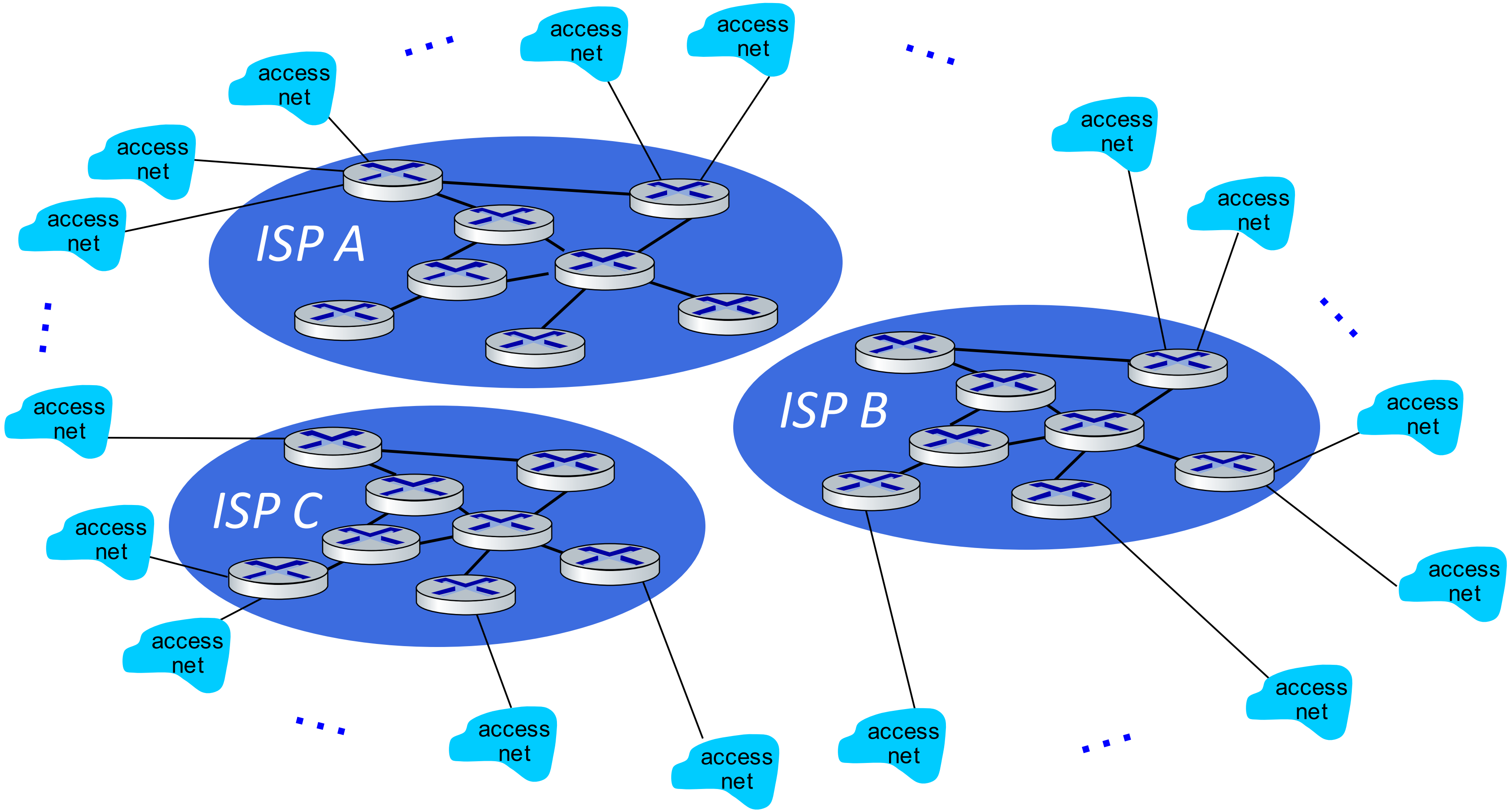
Network core → Connecting access networks

Option: connect each access network to one global transit ISP?
Customer and provider have economic agreement.



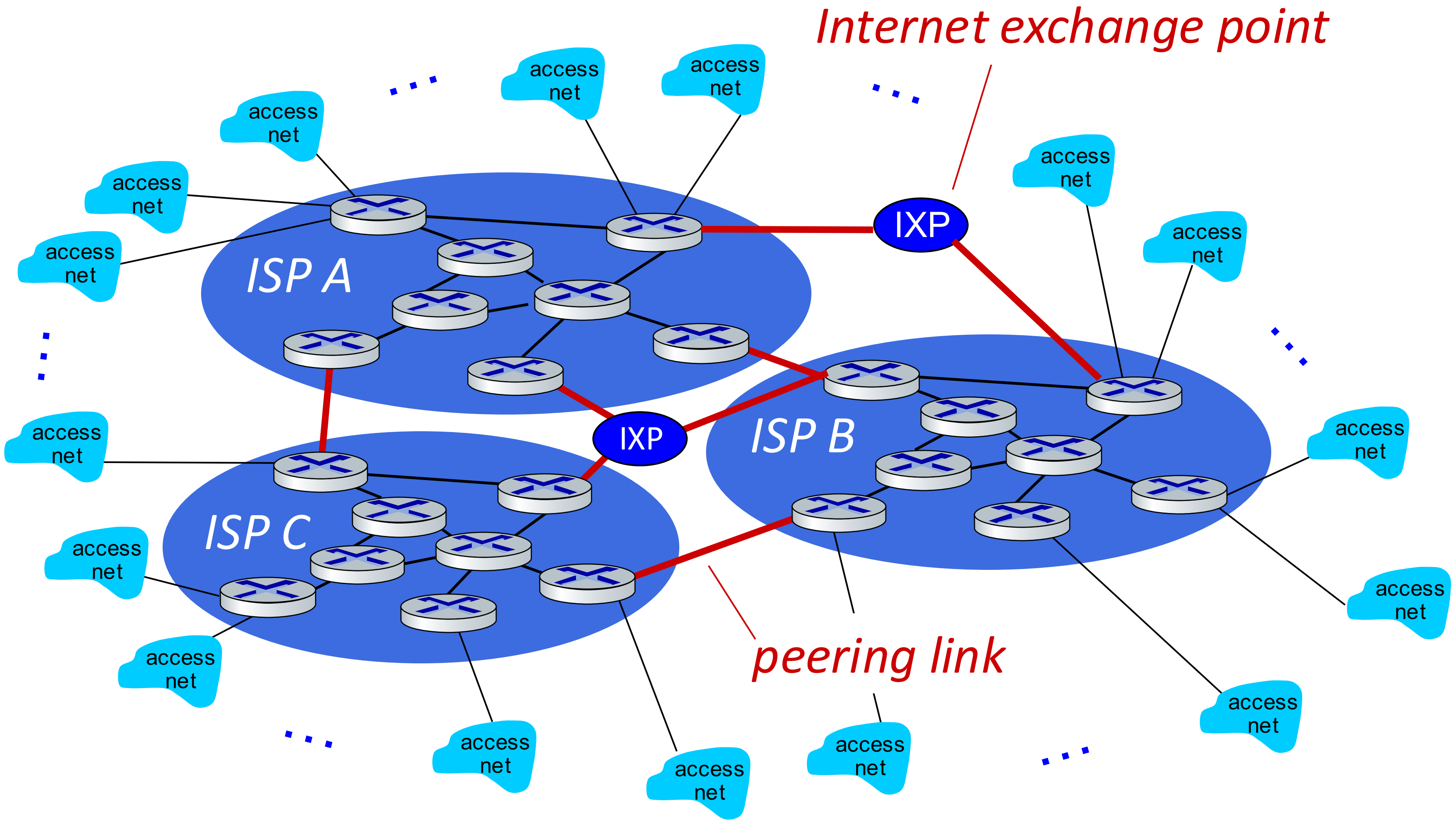
Network core → Connecting access networks

But if one global ISP is viable business, there will be competitors ….



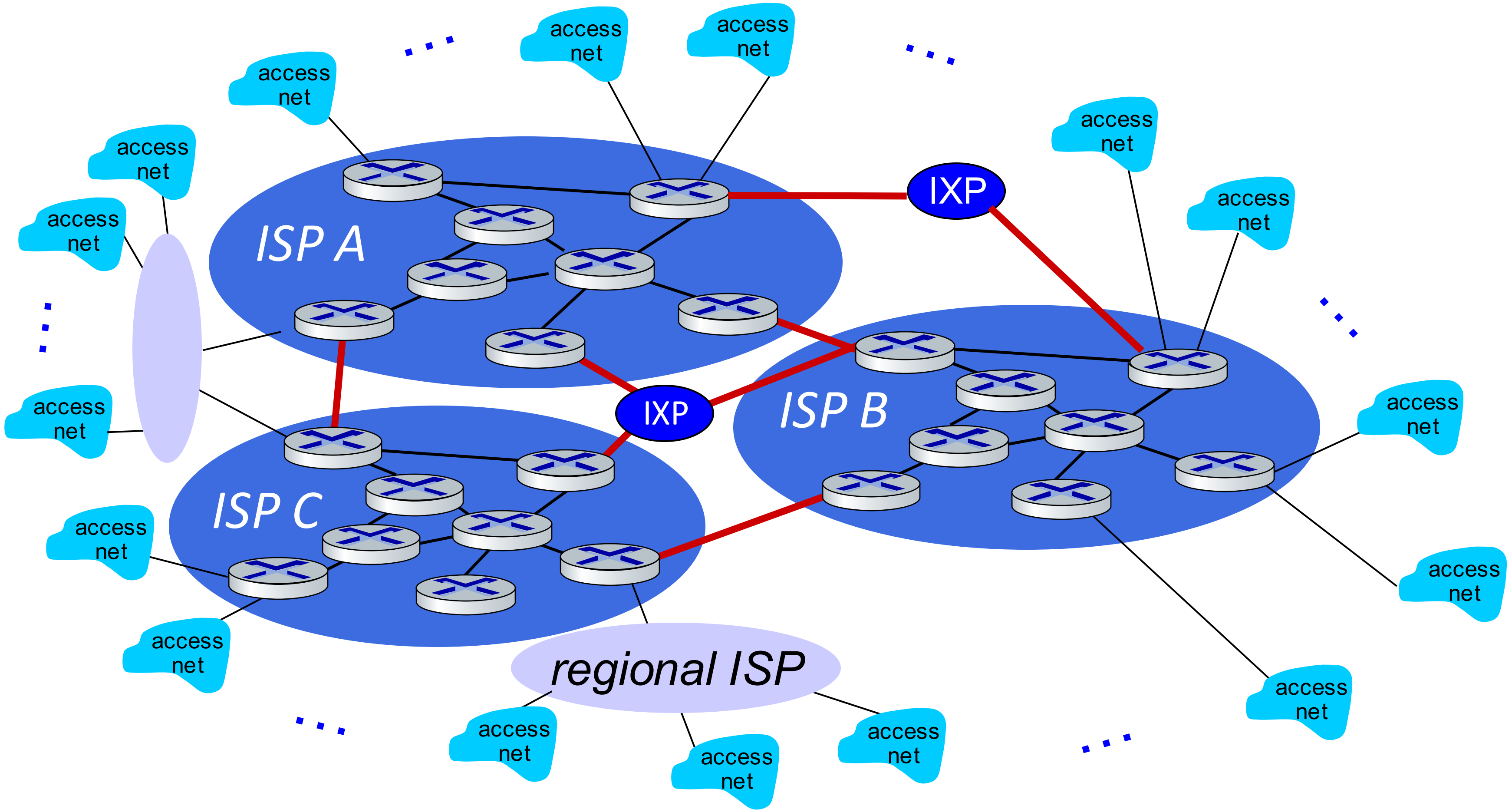
Network core → Connecting access networks

But if one global ISP is viable business, there will be competitors …. who will want to be connected



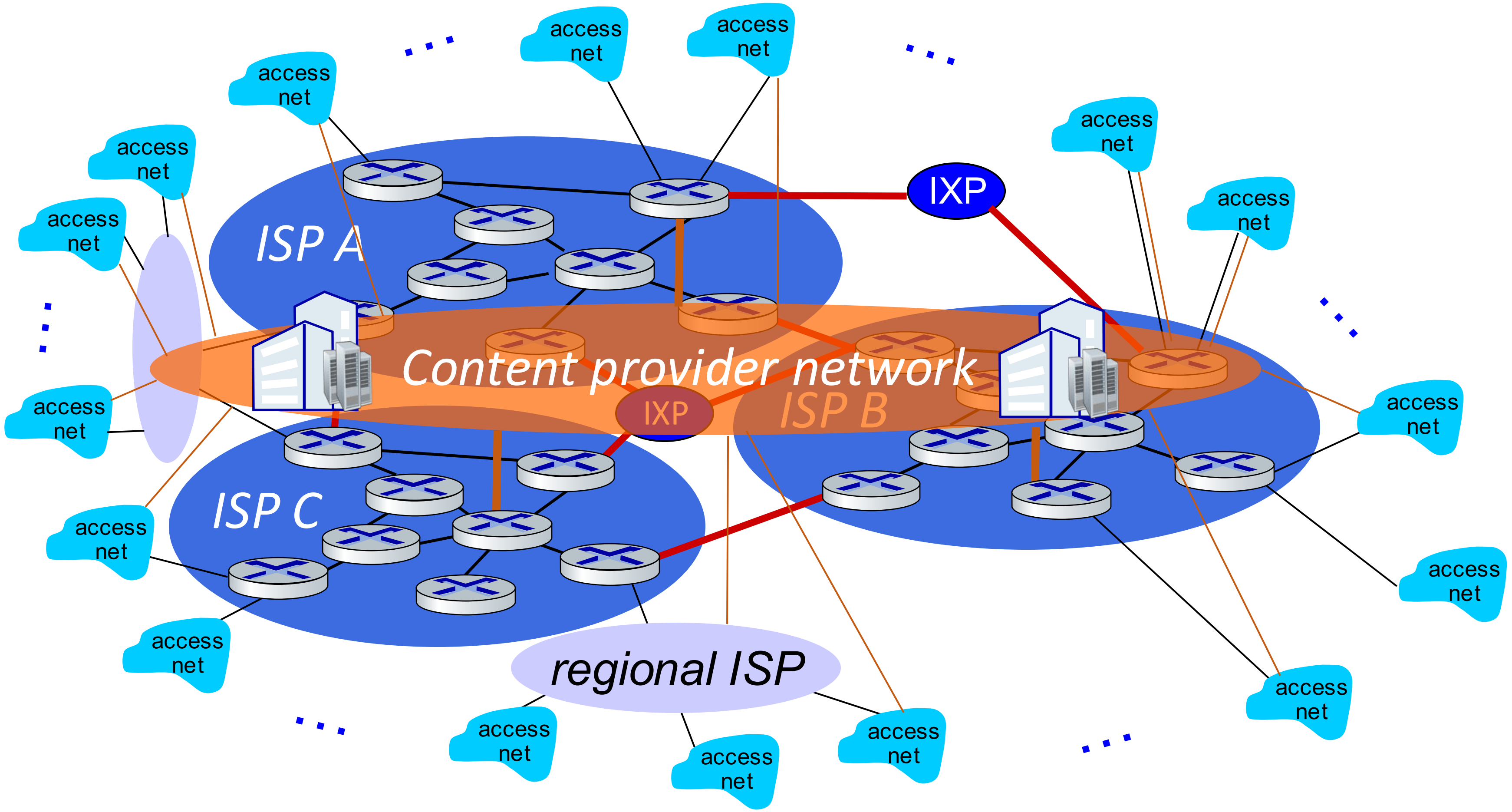
Network core → Connecting access networks

… and regional networks may arise to connect access nets to ISPs



Network core → Connecting access networks

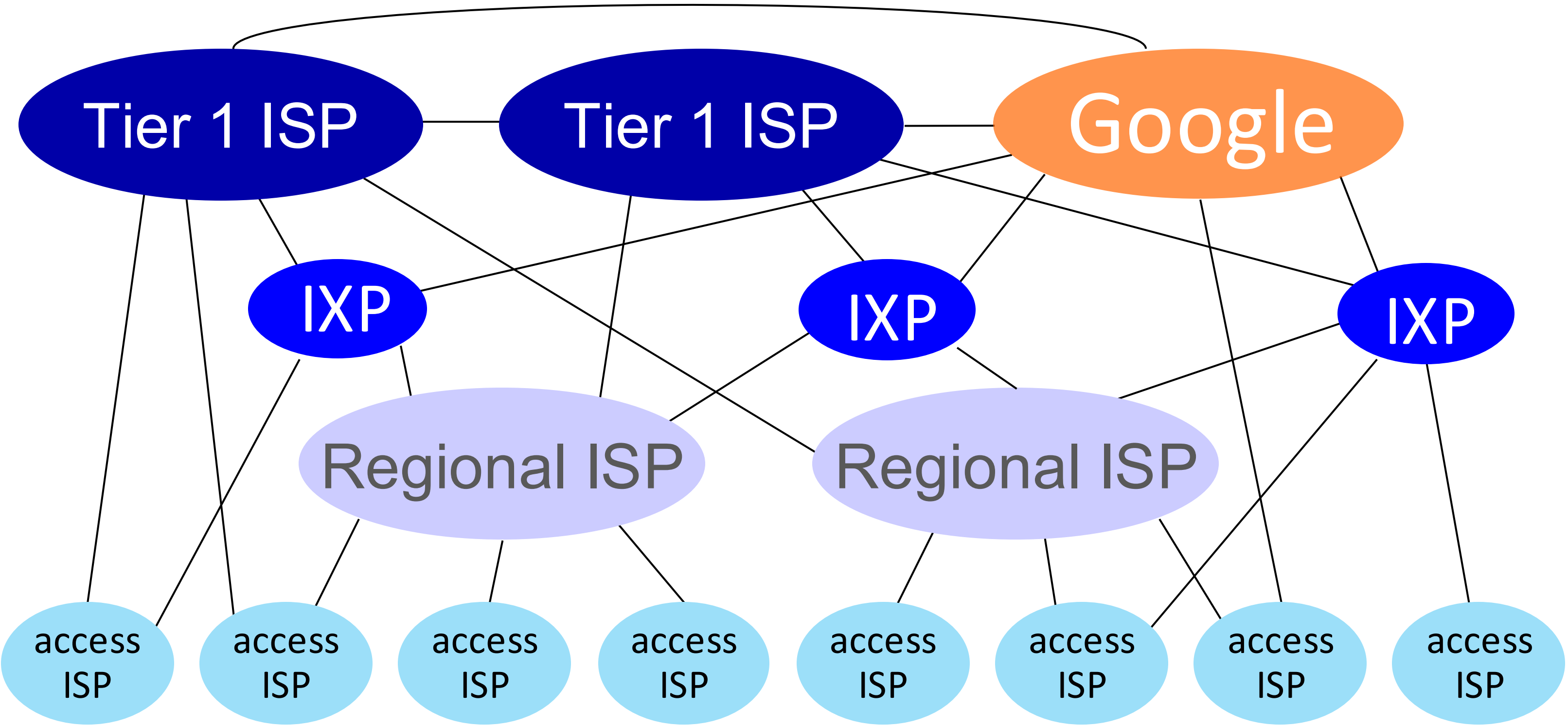
… and content provider networks (e.g., Google, Microsoft, Akamai) may run their own network, to bring services and content close to end users



Content provider network

- Google : one of the leading examples of content-provider network.
- 19 major data centers distributed across North America, Europe, Asia, South America, and Australia
- Each data center having tens or hundreds of thousands of servers
- All are interconnected via Google's private TCP/IP network, which spans the entire globe, separate from the public Internet
- Google private network only carries traffic to/from Google servers

Network core → Connecting access networks

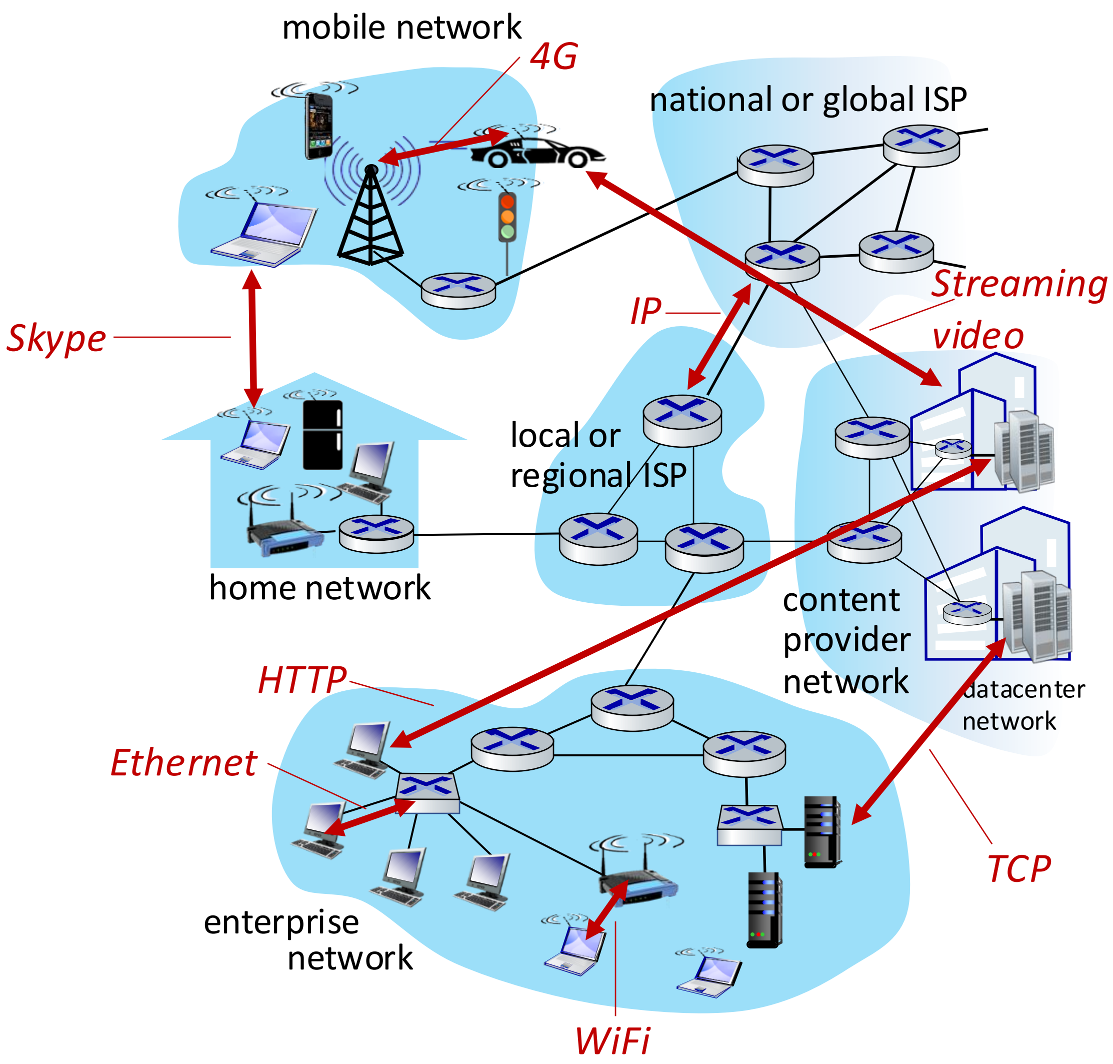


At "center": small # of well-connected large networks

- "tier-1" commercial ISPs (e.g., Level 3, Sprint, AT&T, NTT), national & international coverage
- content provider networks (e.g., Google, Facebook): private network that connects its data centers to Internet, often bypassing tier-1, regional ISPs

The Internet: a “nuts-and-bolts” view

- *Internet: “network of networks”*
  - Interconnected ISPs
- *protocols* are *everywhere*
  - control sending, receiving of messages
  - e.g., HTTP (Web), streaming video, Skype, TCP, IP, WiFi, 4/5G, Ethernet
- *Internet standards*
  - RFC: Request for Comments
  - IETF: Internet Engineering Task Force
  - IEEE 802.11 for WiFi



The Internet: a "nuts-and-bolts" view

PC

server

laptop

cellphone

Billions of connected computing *devices*

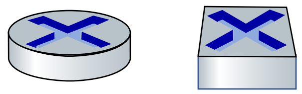
- running applications

wired links

access points

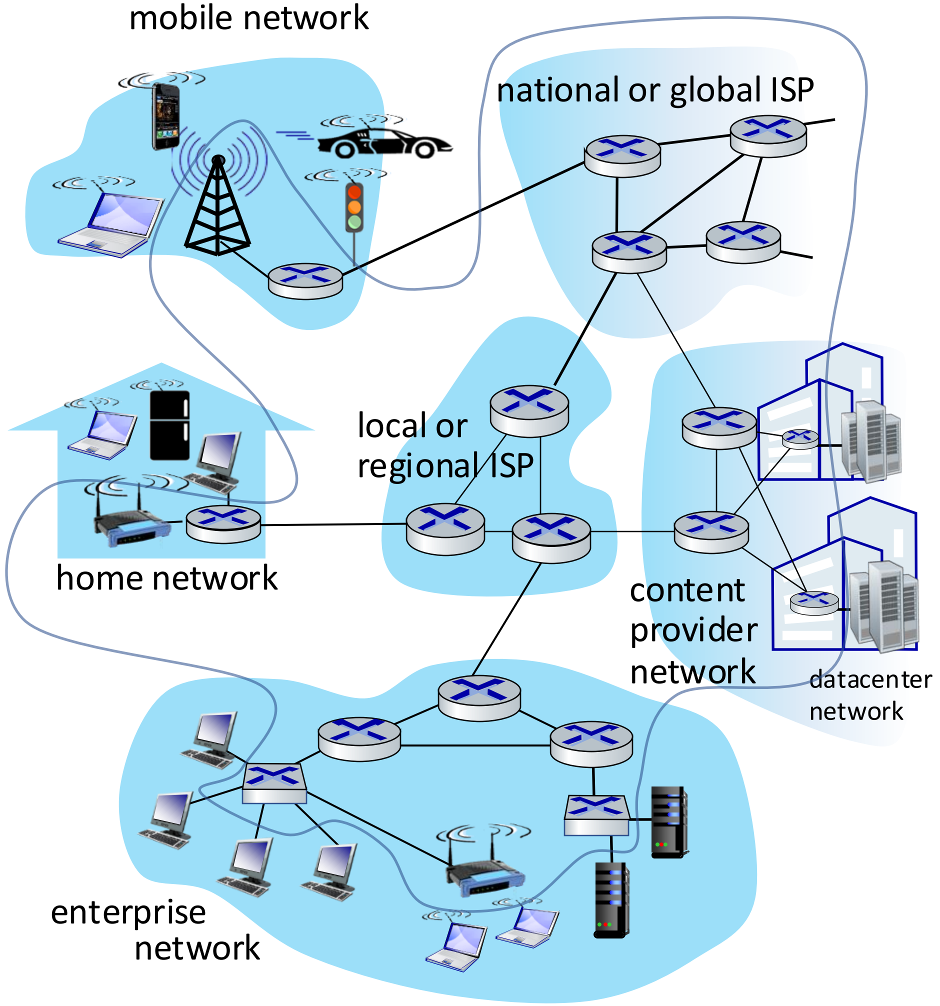
Communication links:

- fiber, copper, radio, satellite



routers/switches

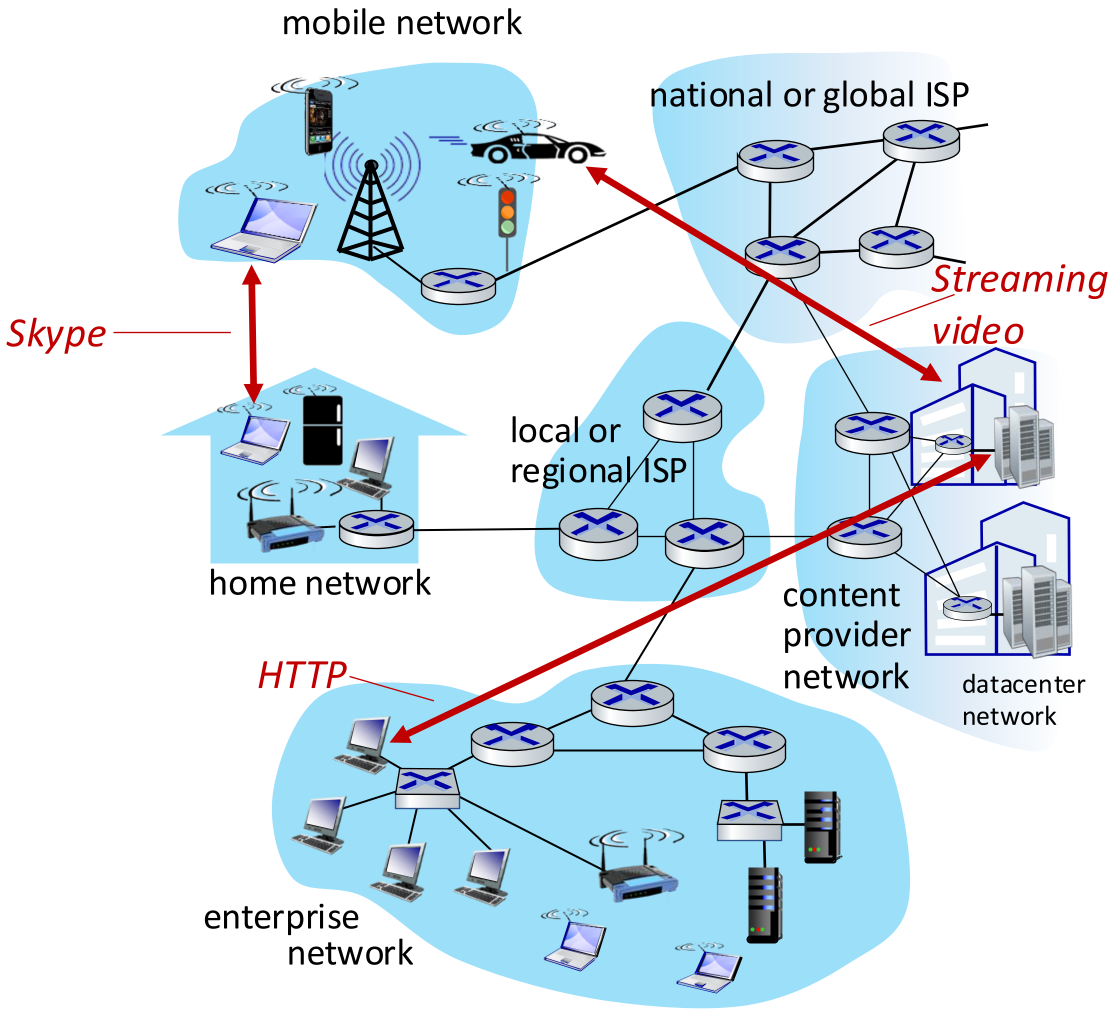
Switches/routers that forward packets (chunks of data)



ISP: internet service provider

The Internet: a "services" view

- *Infrastructure* that provides services to applications:
  - Web, streaming video, multimedia teleconferencing, email, games, e-commerce, social media, inter-connected appliances, …
- provides *programming interface* to distributed applications:
  - "hooks" allowing sending/receiving apps to "connect" to and use Internet transport service
  - provides service options, analogous to postal service



Introduction

What is a computer network? ✔

What is the Internet? ✔

Questions?