

# Computer Networking

Michaelmas/Lent Term

M/W/F 11:00-12:00

LT1 in Gates Building

Slide Set 1

Andrew W. Moore

[andrew.moore@cl.cam.ac.uk](mailto:andrew.moore@cl.cam.ac.uk)

2016-2017

# Computer Networking UROP

- Assessed Practicals for Computer Networking.
  - so supervisors can set/use work
  - so we can have a Computer Networking *tick running over summer 2017*

*Talk to me.*

## Part 2 projects for 16-17

- Fancy doing something at scale or speed?

*Talk to me.*

# Topic 1 Foundation

- Administrivia
- Networks
- Channels
- Multiplexing
- Performance: loss, delay, throughput

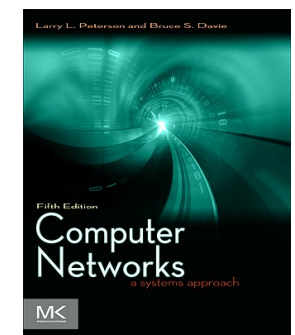
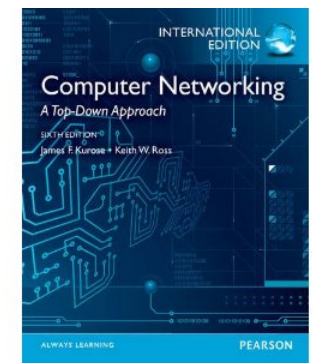
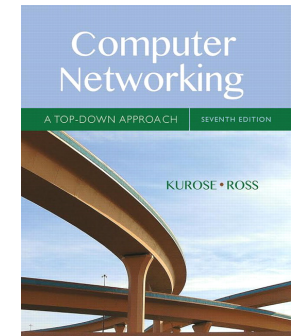
# Course Administration

## Commonly Available Texts

- ❑ Computer Networking: A Top-Down Approach  
Kurose and Ross, 7<sup>th</sup> edition 2016, Addison-Wesley  
(6<sup>th</sup> and 5<sup>th</sup> edition is also commonly available)
- ❑ Computer Networks: A Systems Approach  
Peterson and Davie, 5<sup>th</sup> edition 2011, Morgan-Kaufman

## Other Selected Texts (non-representative)

- ❑ Internetworking with TCP/IP, vol. I + II  
Comer & Stevens, Prentice Hall
- ❑ UNIX Network Programming, Vol. I  
Stevens, Fenner & Rudoff, Prentice Hall

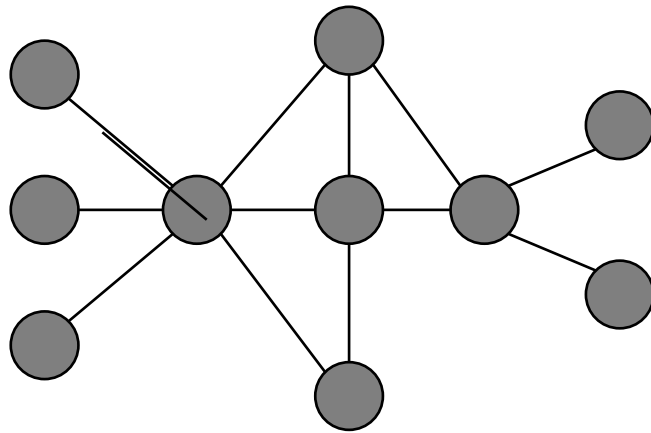


# Thanks

- Slides are a fusion of material from Brad Smith, Ian Leslie, Richard Black, Jim Kurose, Keith Ross, Larry Peterson, Bruce Davie, Jen Rexford, Ion Stoica, Vern Paxson, Scott Shenker, Frank Kelly, Stefan Savage, Jon Crowcroft, Mark Handley, Sylvia Ratnasamy, and Adam Greenhalgh (and to those others I've forgotten, sorry.)
- Supervision material is drawn from Stephen Kell, Andy Rice, and the fantastic TA teams of 144 and 168
- I want practicals too.... ☹️
- Finally thanks to the Part 1b students past and Andrew Rice for all the tremendous feedback.

# What is a network?

- A system of “links” that interconnect “nodes” in order to move “information” between nodes



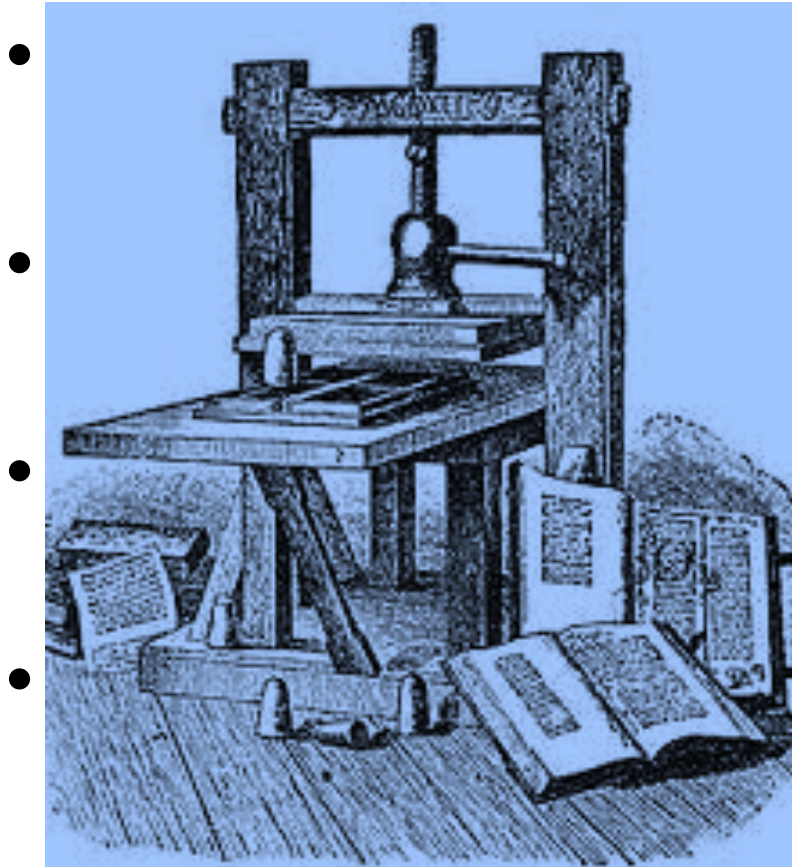
- Yes, this is very vague

# There are *many* different types of networks

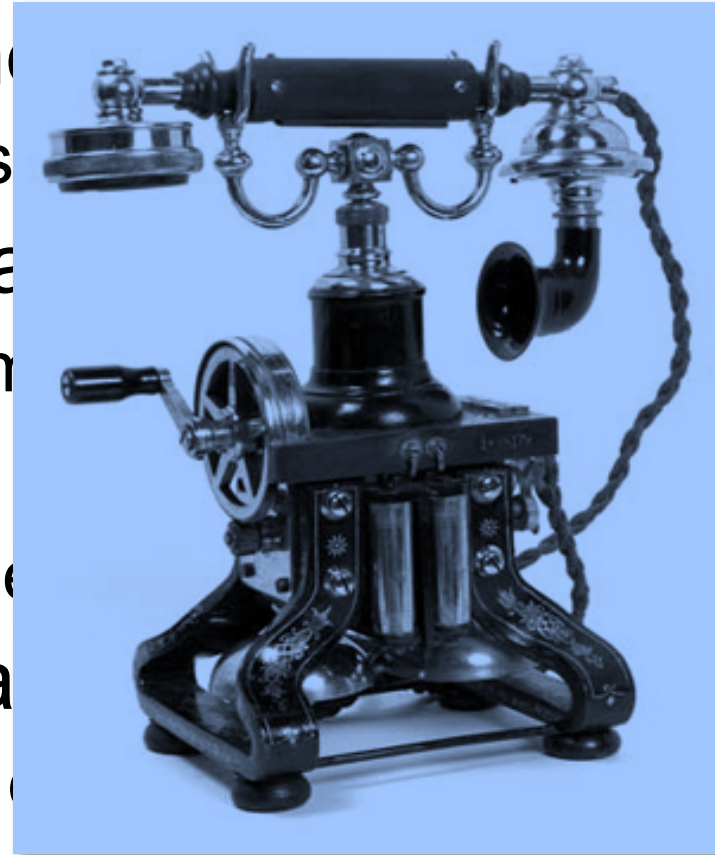
- Internet
- Telephone network
- Transportation networks
- Cellular networks
- Supervisory control and data acquisition networks
- Optical networks
- Sensor networks

**We will focus almost exclusively on the Internet**

# The Internet is transforming everything



since  
artis  
relat  
E-n  
se  
n a  
p, c



KS

Took the dissemination of information to the next level



# The Internet is big business

- Many large and influential networking companies
  - Cisco, Broadcom, AT&T, Verizon, Akamai, Huawei, ...
  - \$132B+ industry (carrier and enterprise alone)
- Networking central to most technology companies
  - Google, Facebook, Intel, HP, Dell, VMware, ...

# Internet research has impact

- **The Internet started as a research experiment!**
- 5 of 10 most cited authors work in networking
- *Many* successful companies have emerged from networking research(ers)

# But why is the Internet *interesting*?

“What’s your formal model for the Internet?” -- *theorists*

“Aren’t you just writing software for networks” – *hackers*

“You don’t have performance benchmarks???” – *hardware folks*

“Isn’t it just another network?” – *old timers at AT&T*

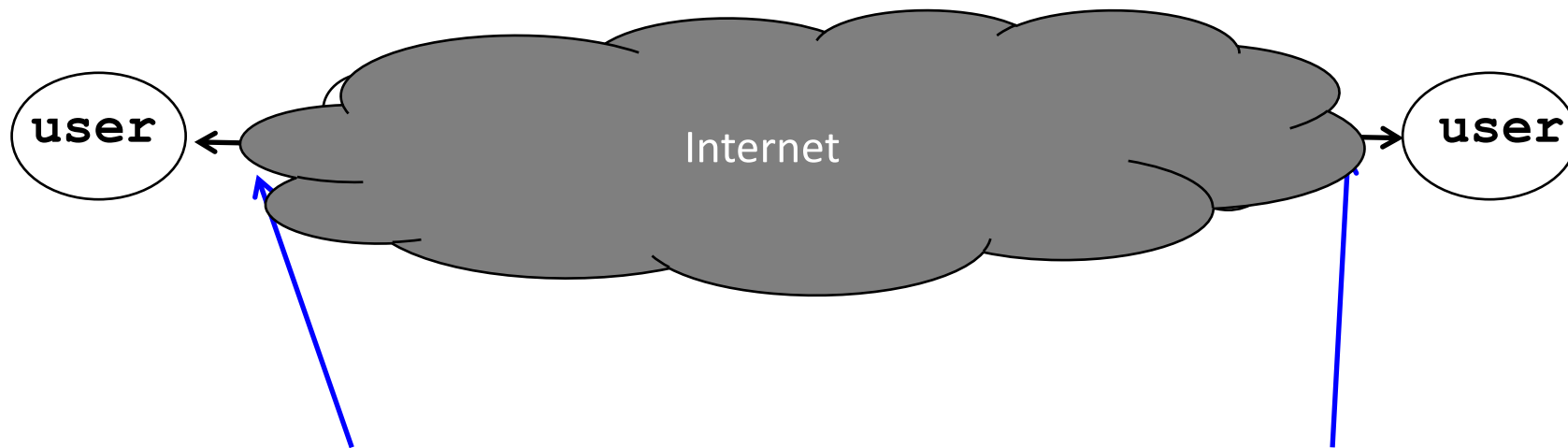
“What’s with all these TLA protocols?” – *all*

“But the Internet seems to be working...” – *my mother*

# A few defining characteristics of the Internet

# A federated system

- The Internet ties together different networks
  - >18,000 ISP networks



Tied together by IP -- the “Internet Protocol” : a single common interface between users and the network and between networks

# A federated system

- The Internet ties together different networks
  - >18,000 ISP networks
- A single, common interface is great for interoperability...
- ...but tricky for business
- Why does this matter?
  - ease of interoperability is the Internet's most important goal
  - practical realities of incentives, economics and real-world trust drive topology, route selection and service evolution

# Tremendous scale

- 3.42 Billion users (46% of world population)
- 1.3+ Trillion unique URLs from 1.1 Trillion servers
- 219 Billion emails sent
- 2+ Billion smartphones
- 1.79 Billion Facebook users
- "Internet Scale" refers to such systems
- 1.6 Billion WhatsApp messages per day
- 1.6 Billion YouTube videos watched per day
- 800 hours of Youtube video added per minute
- Switches that move 300+ Terabits/second
- Network links that carry 1.5 Terabits/second

# Enormous diversity and dynamic range

- Communication latency: microseconds to seconds ( $10^6$ )
- Bandwidth: 1Kbits/second to 100 Gigabits/second ( $10^7$ )
- Packet loss: 0 – 90%
- Technology: optical, wireless, satellite, copper
- **Endpoint devices**: from sensors and cell phones to datacenters and supercomputers
- **Applications**: social networking, file transfer, skype, live TV, gaming, remote medicine, backup, IM
- **Users**: the governing, governed, operators, **malicious**, naïve, savvy, embarrassed, paranoid, addicted, cheap ...



# Constant Evolution

1970s:

- 56kilobits/second “backbone” links
- <100 computers, a handful of sites in the US (and one UK)
- Telnet and file transfer are the “killer” applications

Today

- 100+Gigabits/second backbone links
- 10B+ devices, all over the globe
- *20M* Facebook apps installed per day

# Asynchronous Operation

- Fundamental constraint: **speed of light**
- Consider:
  - How many cycles does your 3GHz CPU in Cambridge execute before it can possibly get a response from a message it sends to a server in Palo Alto?
    - Cambridge to Palo Alto: 8,609 km
    - Traveling at 300,000 km/s: 28.70 milliseconds
    - Then back to Cambridge:  $2 \times 28.70 = 57.39$  milliseconds
    - $3,000,000,000 \text{ cycles/sec} \times 0.05739 = 172,179,999$  cycles!
- Thus, communication feedback is always *dated*

# Prone to Failure

- To send a message, **all** components along a path must function correctly
  - software, modem, wireless access point, firewall, links, network interface cards, switches,...
  - Including **human operators**
- Consider: 50 components, that work correctly 99% of time → 39.5% chance communication will fail
- Plus, recall
  - scale → lots of components
  - asynchrony → takes a long time to hear (bad) news
  - federation (**internet**) → hard to identify fault or assign blame

# An Engineered System

- Constrained by what technology is practical
  - Link bandwidths
  - Switch port counts
  - Bit error rates
  - Cost
  - ...

# Recap: The Internet is...


- A complex federation
- Of enormous scale
- Dynamic range
- Diversity
- Constantly evolving
- Asynchronous in operation
- Failure prone
- Constrained by what's practical to engineer
- Too complex for theoretical models
- “Working code” doesn't mean much
- Performance benchmarks are too narrow

# Performance – not just bits per second

## Second order effects

- Image/Audio quality

## Other metrics...

- Network efficiency (good-put *versus* throughput)
- User Experience? (World Wide Wait) 
- Network connectivity expectation
- Others?

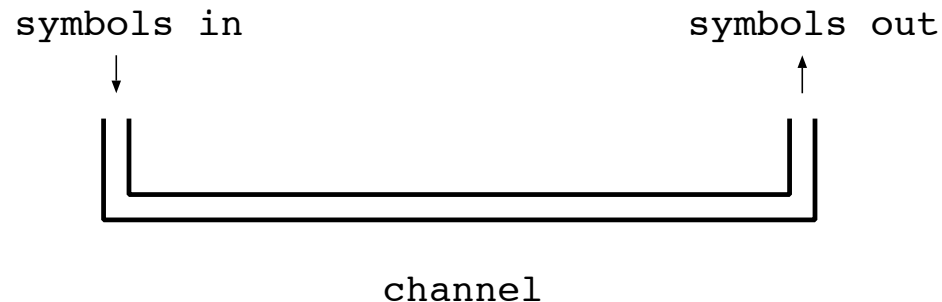


# Channels Concept

(This channel definition is very abstract)

- Peer entities communicate over channels
- Peer entities provide higher-layer peers with higher-layer channels

*A channel is that into which an entity puts symbols and which causes those symbols (or a reasonable approximation) to appear somewhere else at a later point in time.*



# Channel Characteristics

**Symbol type:** bits, packets, waveform

**Capacity:** bandwidth, data-rate, packet-rate

**Delay:** fixed or variable

**Fidelity:** signal-to-noise, bit error rate, packet error rate

**Cost:** per attachment, for use

**Reliability**

**Security:** privacy, unforgability

**Order preserving:** always, almost, usually

**Connectivity:** point-to-point, to-many, many-to-many

Examples:

- Fibre Cable
- 1 Gb/s channel in a network
- Sequence of packets transmitted between hosts
- A telephone call (handset to handset)
- The audio channel in a room
- Conversation between two people



# Example Physical Channels

these example physical channels are also known as *Physical Media*

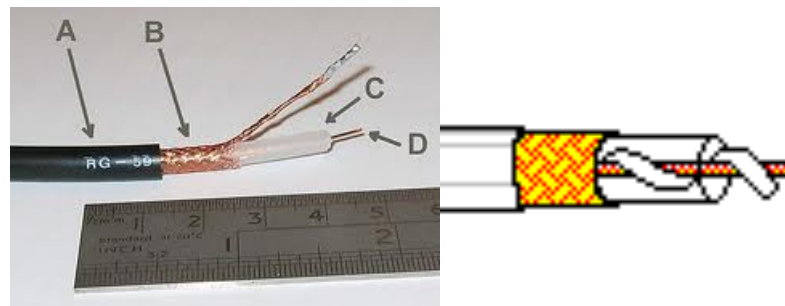
## Twisted Pair (TP)

- two insulated copper wires
  - Category 3: traditional phone wires, 10 Mbps Ethernet
  - Category 6: 1Gbps Ethernet
- Shielded (STP)
- Unshielded (UTP)



## Coaxial cable:

- two concentric copper conductors
- bidirectional
- baseband:
  - single channel on cable
  - legacy Ethernet
- broadband:
  - multiple channels on cable
  - HFC (Hybrid Fiber Coax)



## Fiber optic cable:

- high-speed operation
- point-to-point transmission
- (10' s-100' s Gps)
- low error rate
- immune to electromagnetic noise



# More Physical media: **Radio**

- Bidirectional and multiple access
- propagation environment effects:
  - reflection
  - obstruction by objects
  - interference

## Radio link types:

- ❑ **terrestrial microwave**
  - ❖ e.g. 45 Mbps channels
- ❑ **LAN** (e.g., Wifi)
  - ❖ 11Mbps, 54 Mbps, 200 Mbps
- ❑ **wide-area** (e.g., cellular)
  - ❖ 4G cellular: ~ 4 Mbps
- ❑ **satellite**
  - ❖ Kbps to 45Mbps channel (or multiple smaller channels)
  - ❖ 270 msec end-end delay
  - ❖ geosynchronous versus low altitude



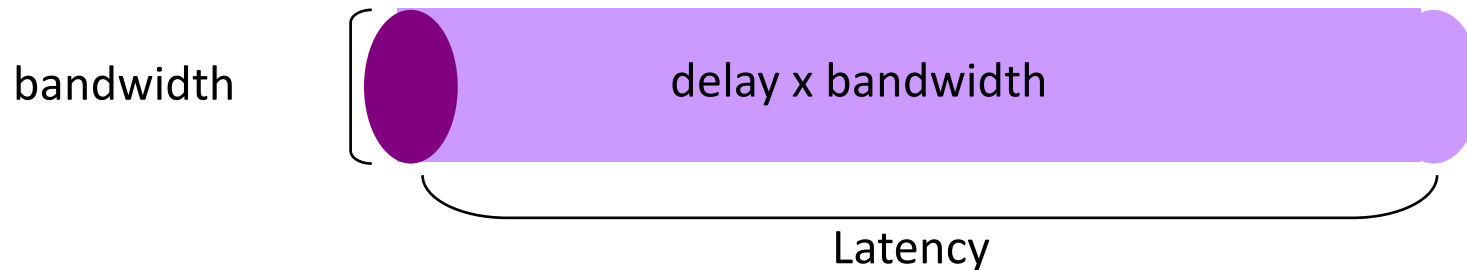
# Nodes and Links



Channels = Links

Peer entities = Nodes

# Properties of Links (Channels)



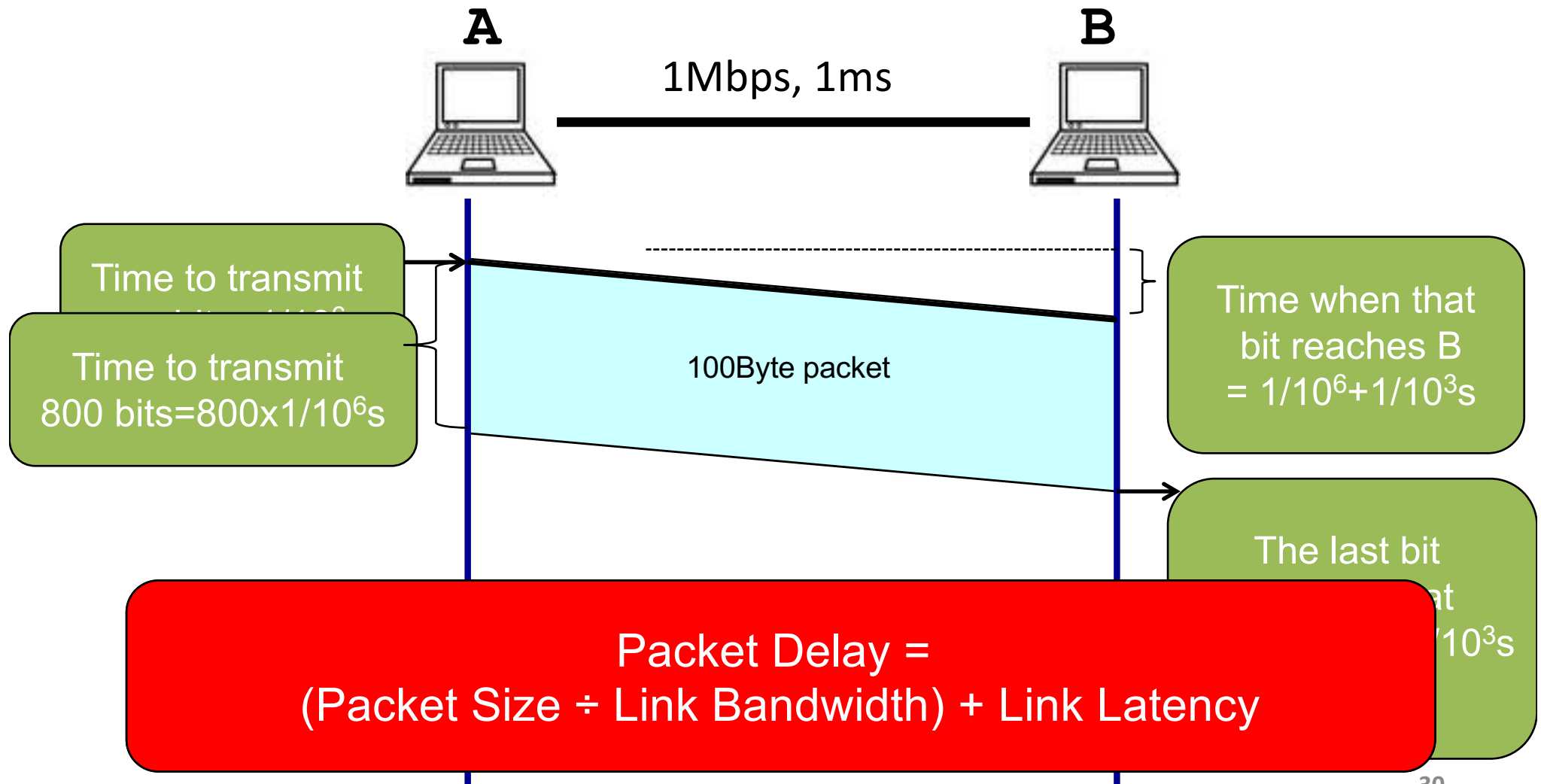
- Bandwidth (capacity): “width” of the links
  - number of bits sent (or received) per unit time (bits/sec or bps)
- Latency (delay): “length” of the link
  - propagation time for data to travel along the link(seconds)
- Bandwidth-Delay Product (BDP): “volume” of the link
  - amount of data that can be “in flight” at any time
  - propagation delay  $\times$  bits/time = total bits in link

# Examples of Bandwidth-Delay

- Same city over a slow link:
  - BW~10Mbps
  - Latency~0.1msec
  - BDP ~  $10^6$ bits ~ 125KBytes
  
- Cross-country over fast link:
  - BW~10Gbps
  - Latency~10msec
  - BDP ~  $10^8$ bits ~ 12.5MBytes

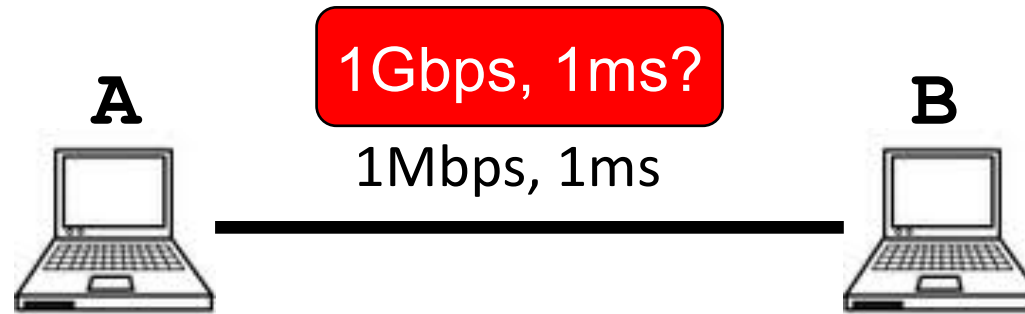
# Packet Delay

*Sending a 100B packet from A to B?*



1GB file in 100B packets **ay**

*Sending a 100B packet from A to B?*



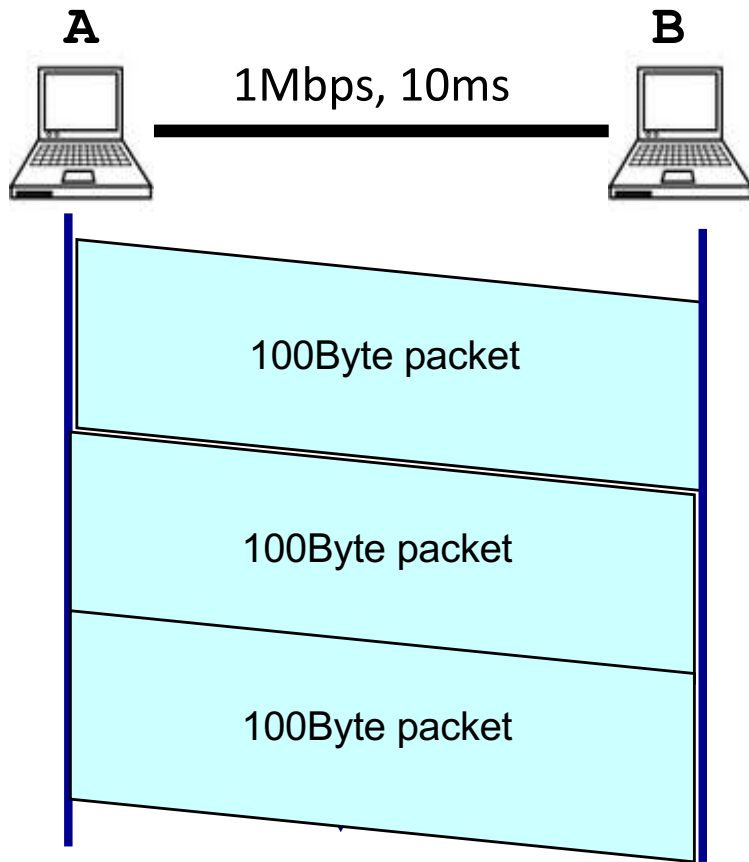
The last bit in the file  
reaches B at  
 $(10^7 \times 800 \times 1/10^9) + 1/10^3\text{s}$   
= 8001ms

The last bit  
reaches B at  
 $(800 \times 1/10^9) + 1/10^3\text{s}$   
= 1.0008ms

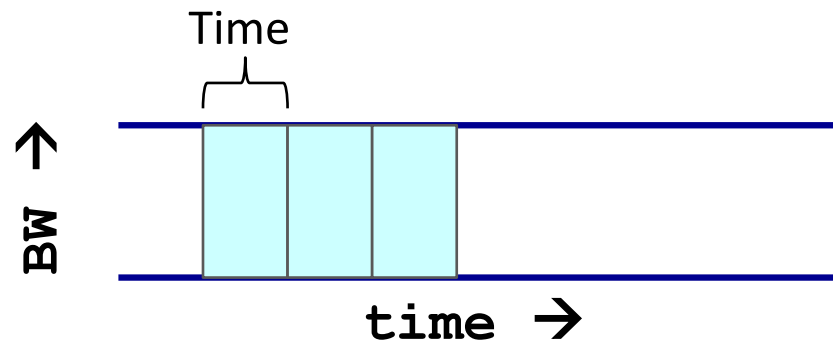
The last bit  
reaches B at  
 $(800 \times 1/10^6) + 1/10^3\text{s}$   
= 1.8ms

# Packet Delay: The “pipe” view

*Sending 100B packets from A to B?*



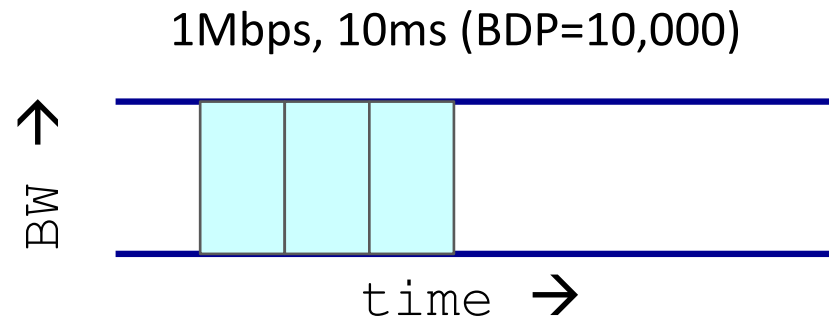
Packet Transmission



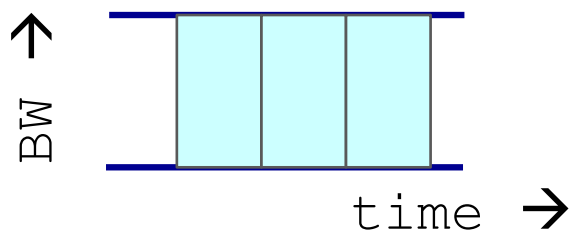


# Packet Delay: The “pipe” view

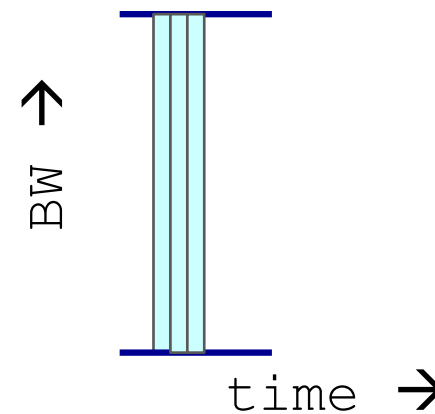
*Sending 100B packets from A to B?*



1Mbps, 5ms (BDP=5,000)

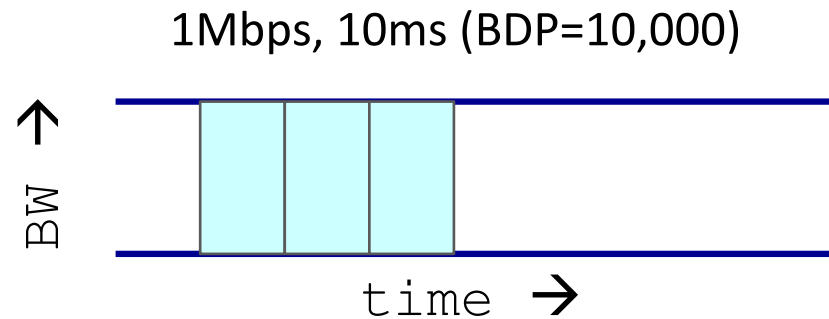


10Mbps, 1ms (BDP=10,000)

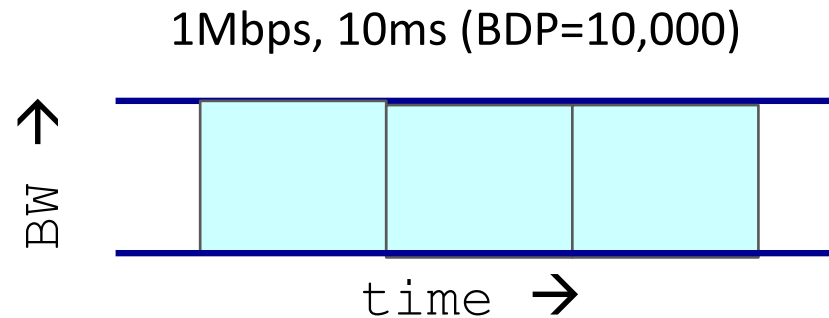


# Packet Delay: The “pipe” view

*Sending 100B packets from A to B?*



What if we used *200Byte packets??*

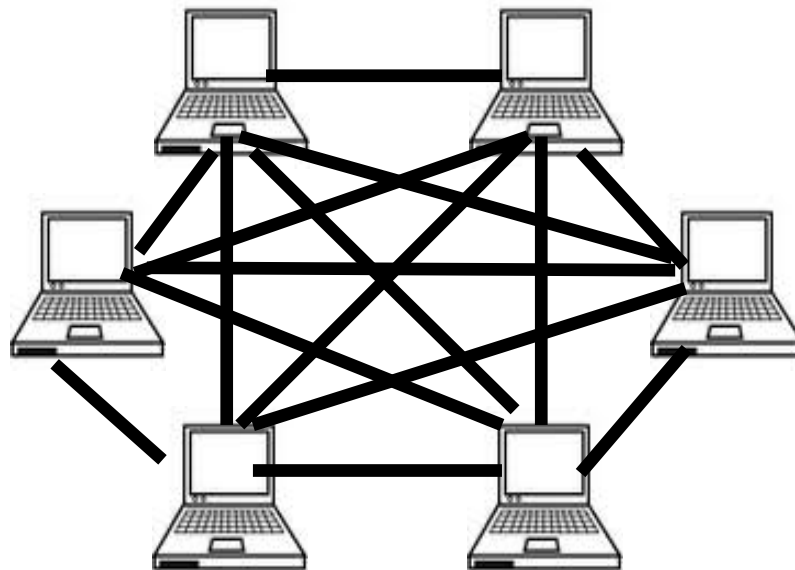


# Recall Nodes and Links



# What if we have more nodes?

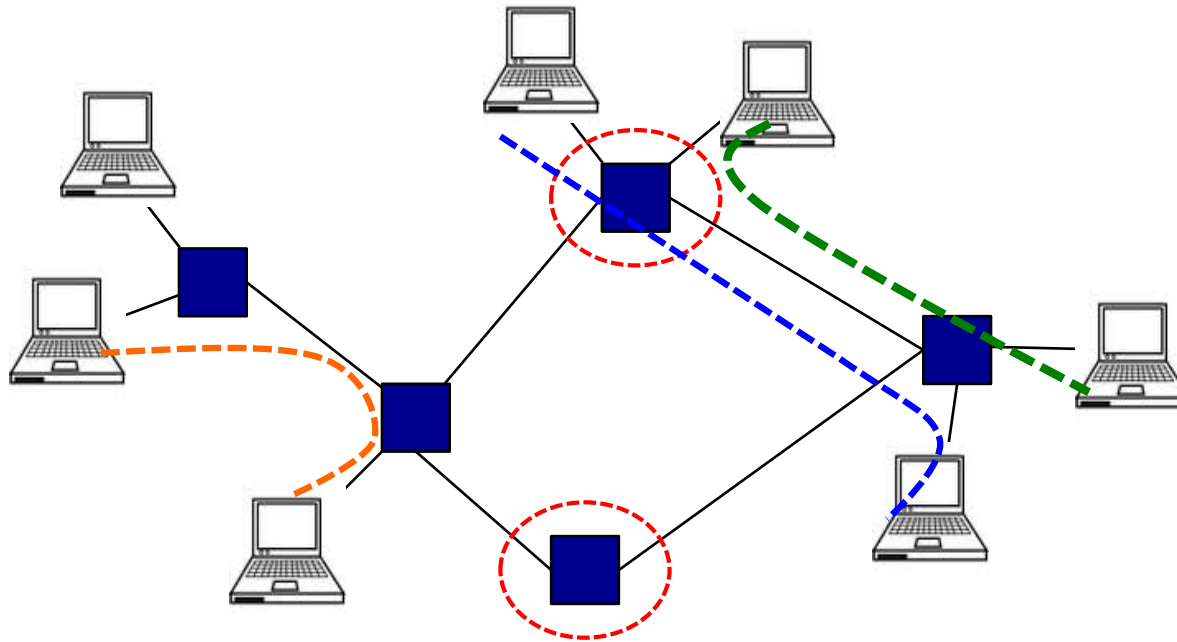
One link for every node?



**Need a scalable way to interconnect nodes**

# Solution: A switched network

Nodes share network link resources



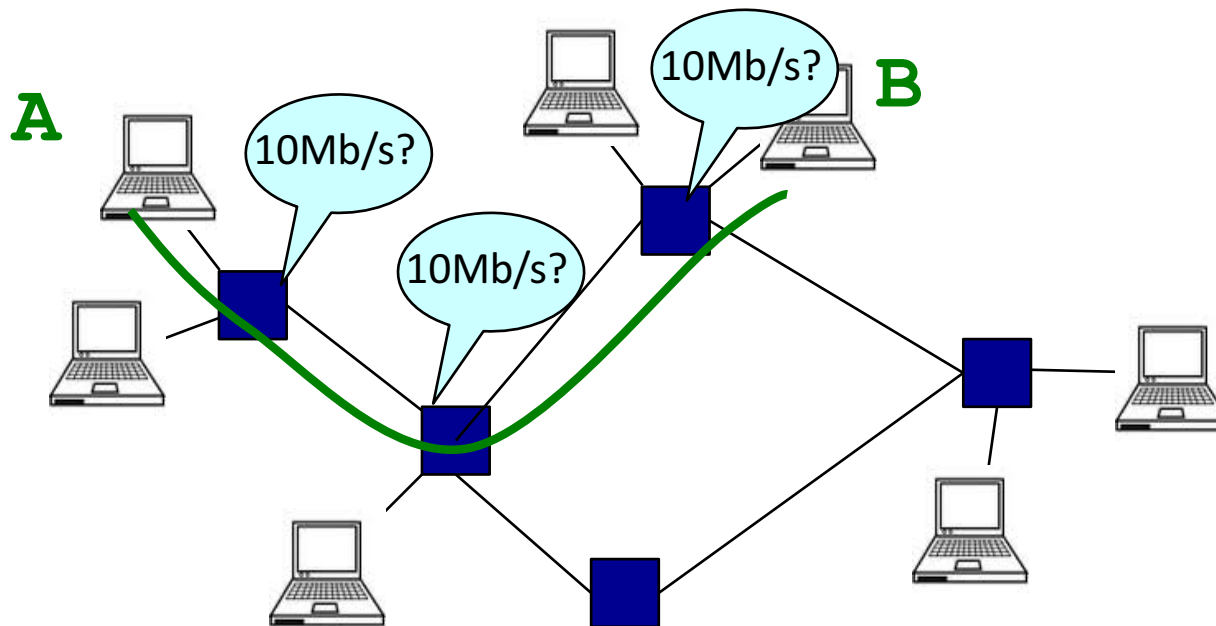
How is this sharing implemented?

# Two forms of switched networks

- Circuit switching (used in the *POTS*: Plain Old Telephone system)
- Packet switching (used in the Internet)

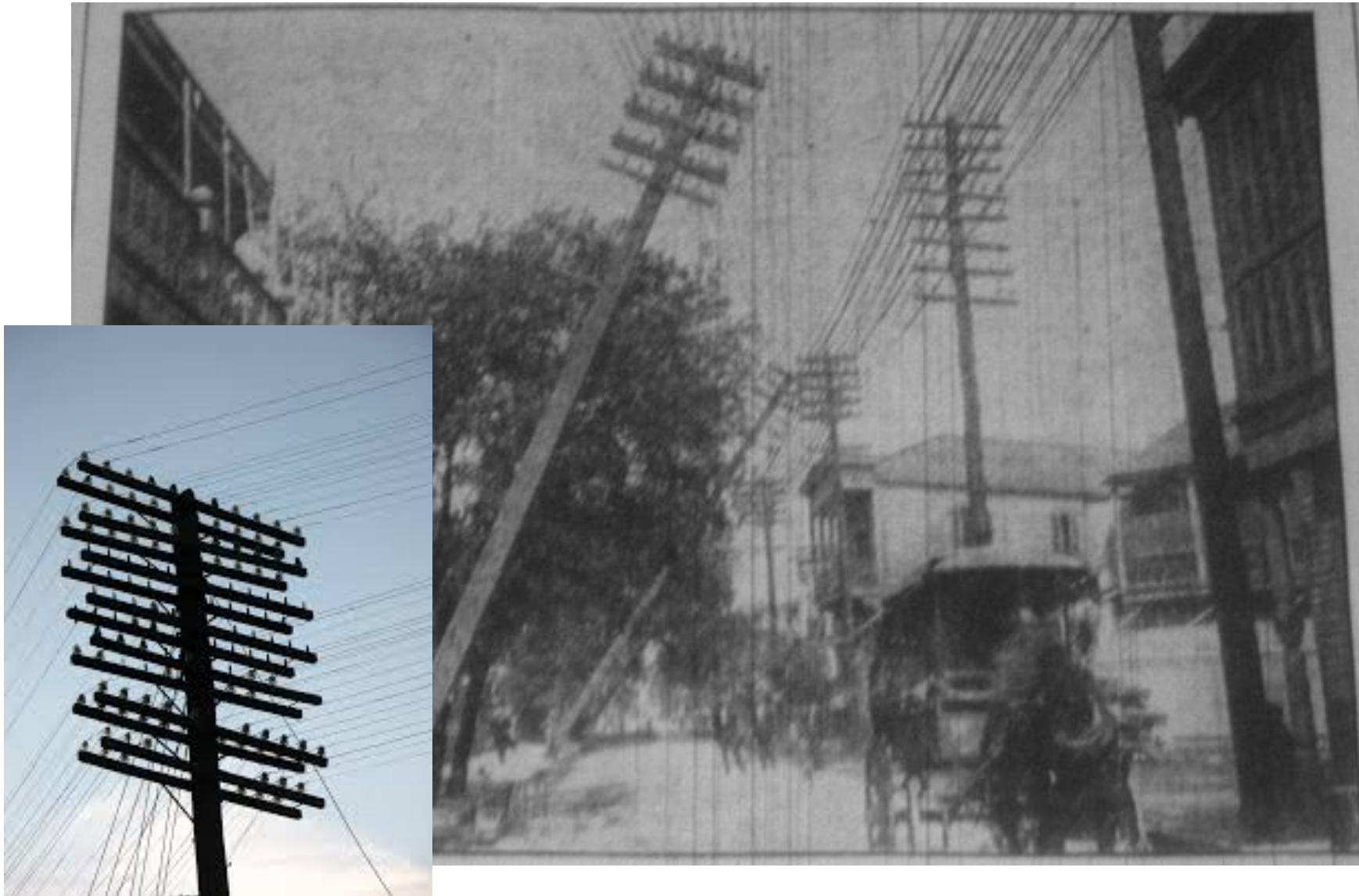
# Circuit switching

Idea: source **reserves** network capacity along a path



- (1) Node A sends a reservation request
- (2) Interior switches establish a connection -- i.e., "circuit"
- (3) A starts sending data
- (4) A sends a "teardown circuit" message

# Old Time Multiplexing





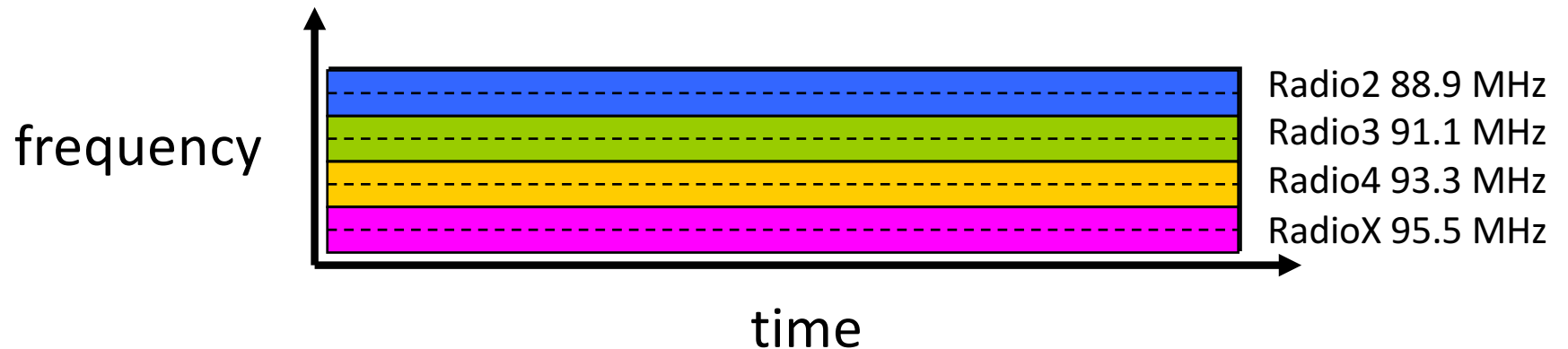
# Circuit Switching: FDM and TDM

Example:

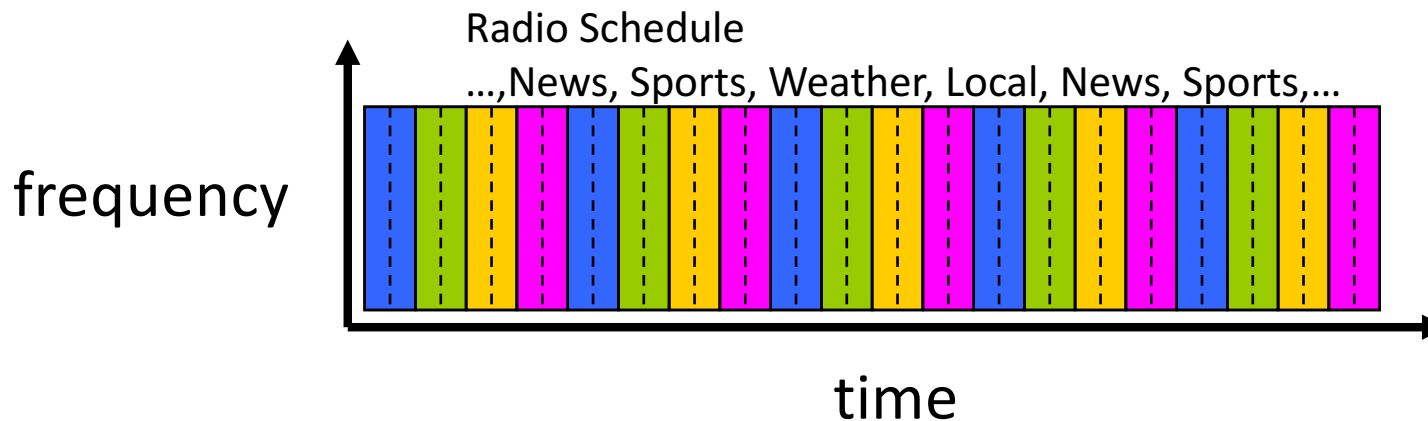
4 users



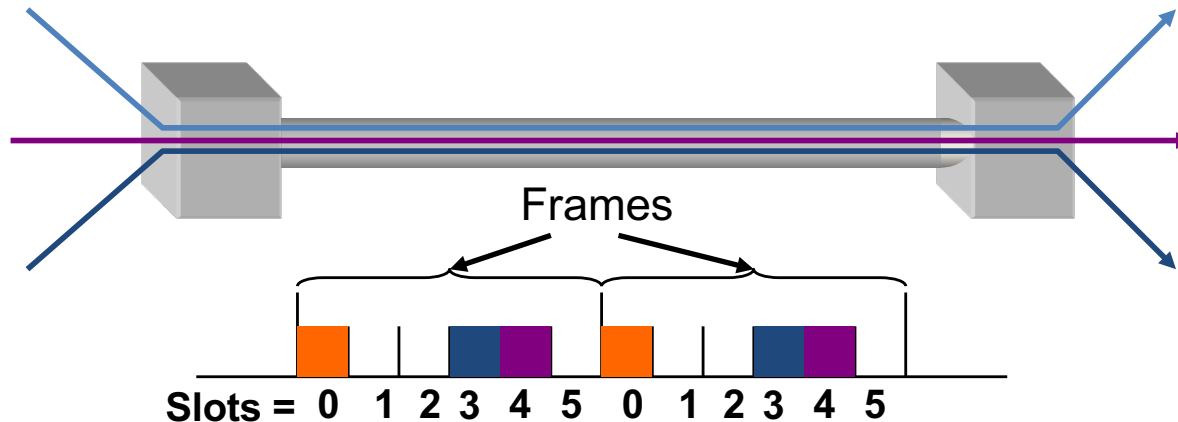
Frequency Division Multiplexing



Time Division Multiplexing

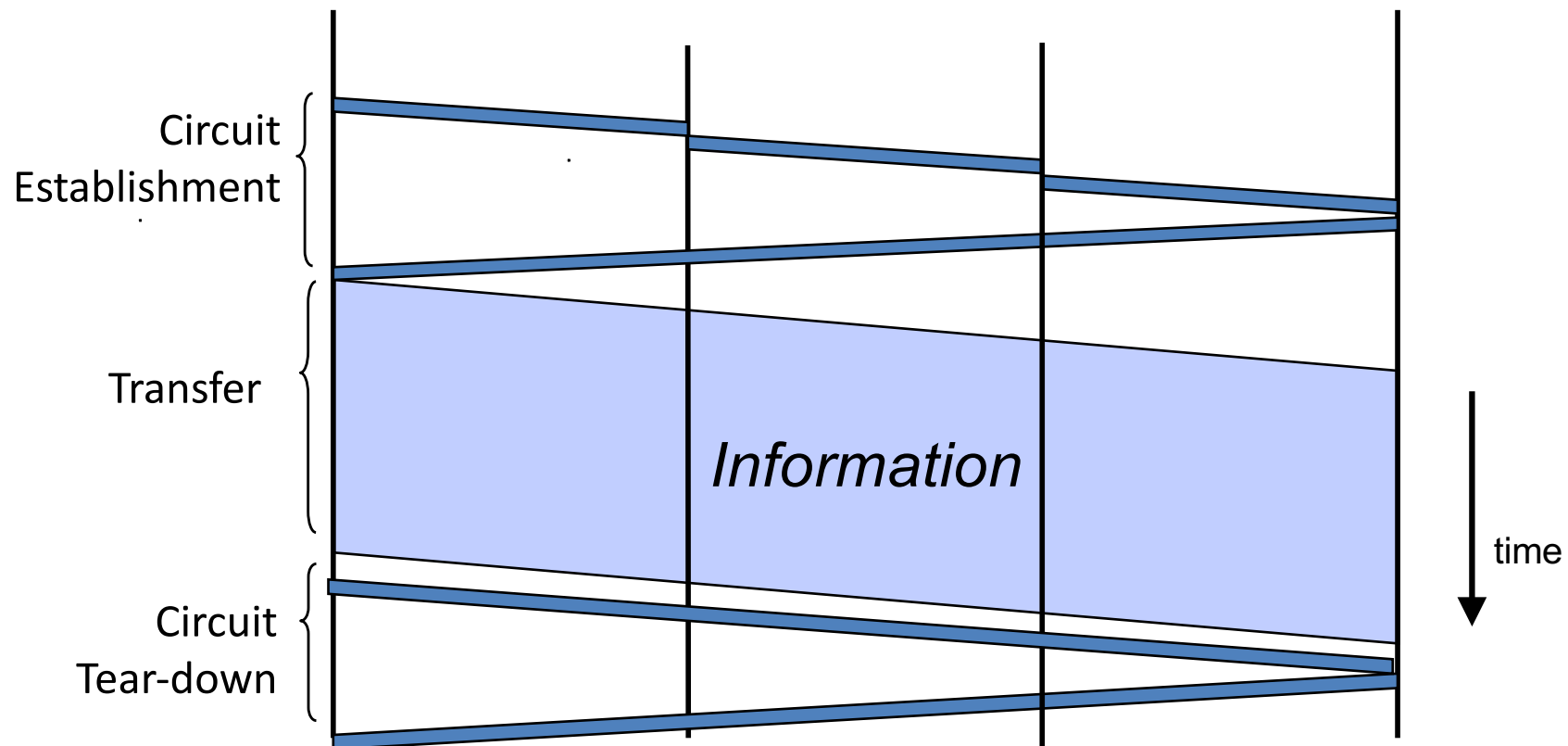
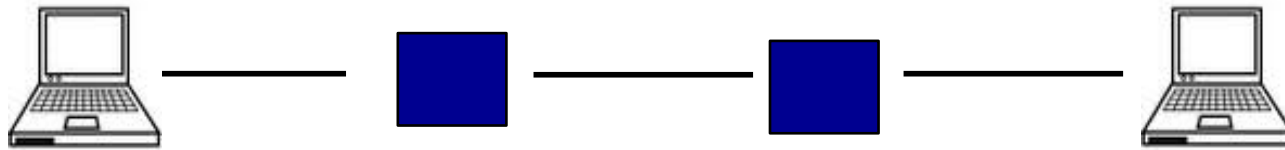


# Time-Division Multiplexing/Demultiplexing



- Time divided into frames; frames into slots
- Relative slot position inside a frame determines to which conversation data belongs
  - e.g., slot 0 belongs to **orange** conversation
- Slots are reserved (released) during circuit setup (teardown)
- If a conversation does not use its circuit **capacity is lost!**

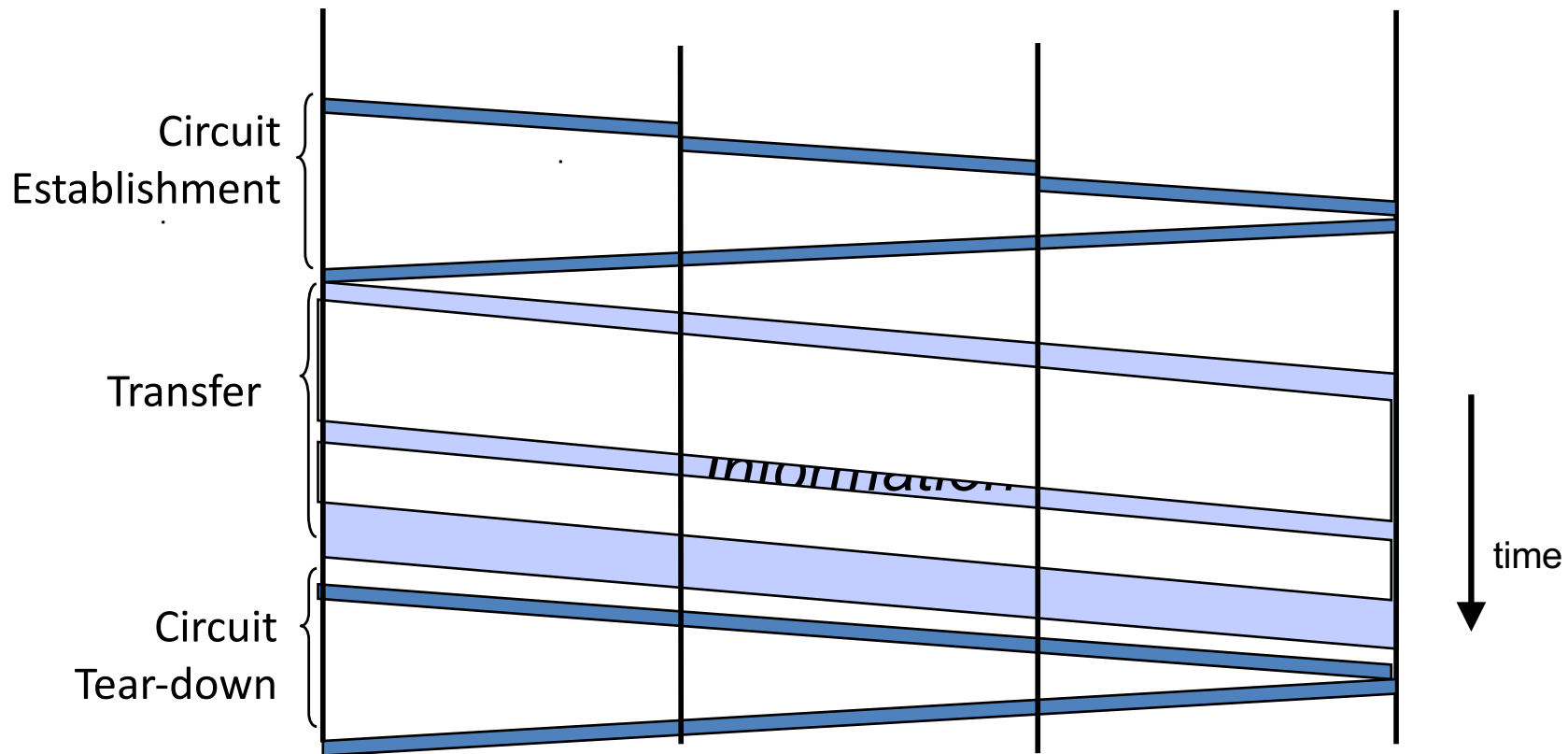
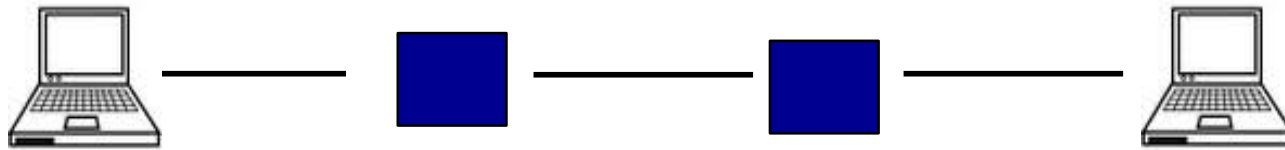
# Timing in Circuit Switching



# Circuit switching: pros and cons

- Pros
  - guaranteed performance
  - fast transfer (once circuit is established)
- Cons

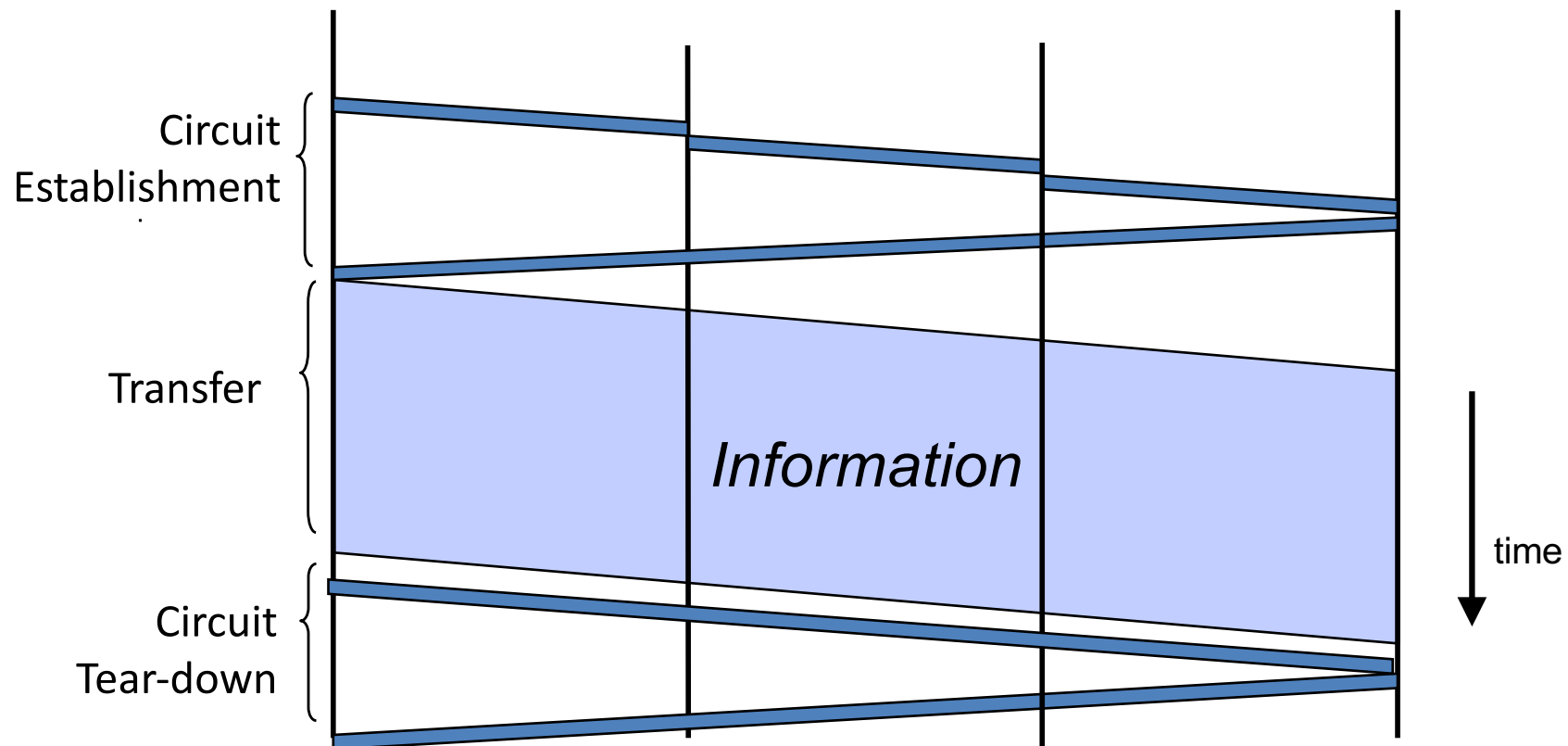
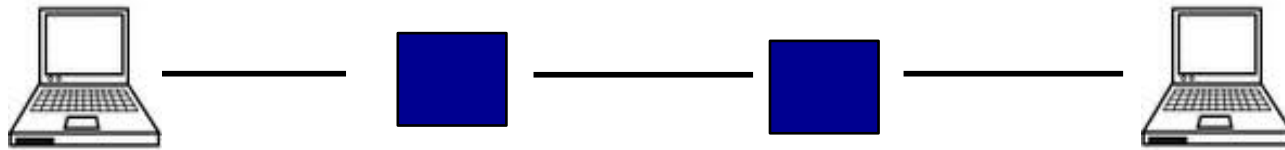
# Timing in Circuit Switching



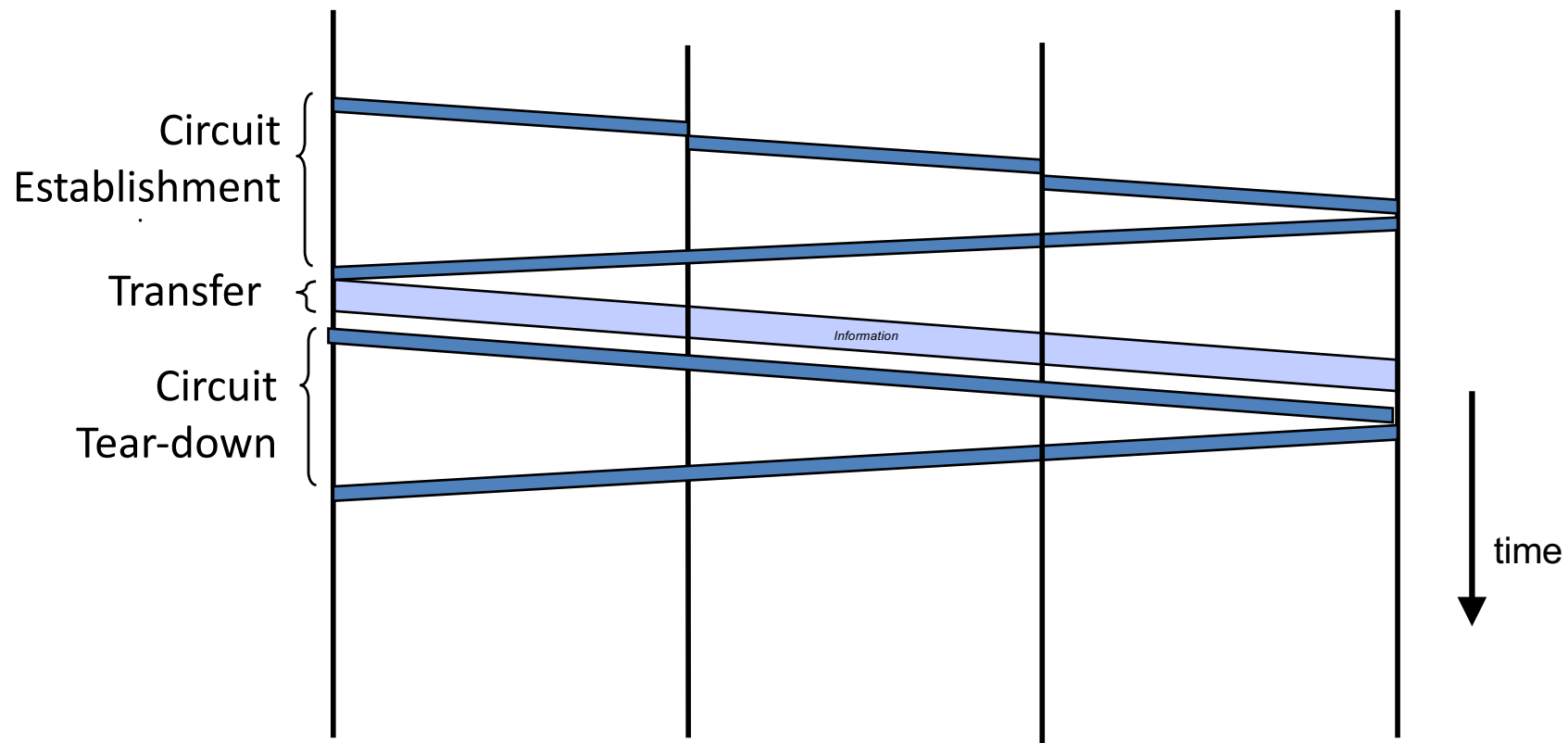
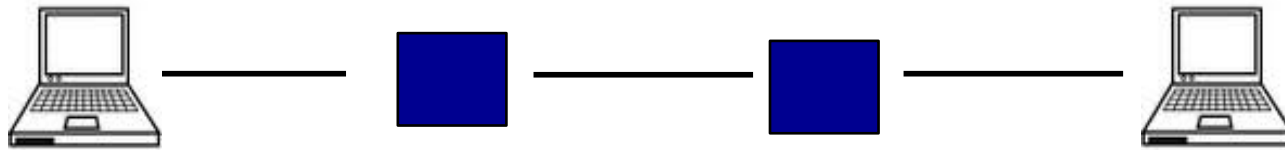
# Circuit switching: pros and cons

- Pros
  - guaranteed performance
  - fast transfer (once circuit is established)
- Cons
  - **wastes bandwidth if traffic is “bursty”**

# Timing in Circuit Switching



# Timing in Circuit Switching

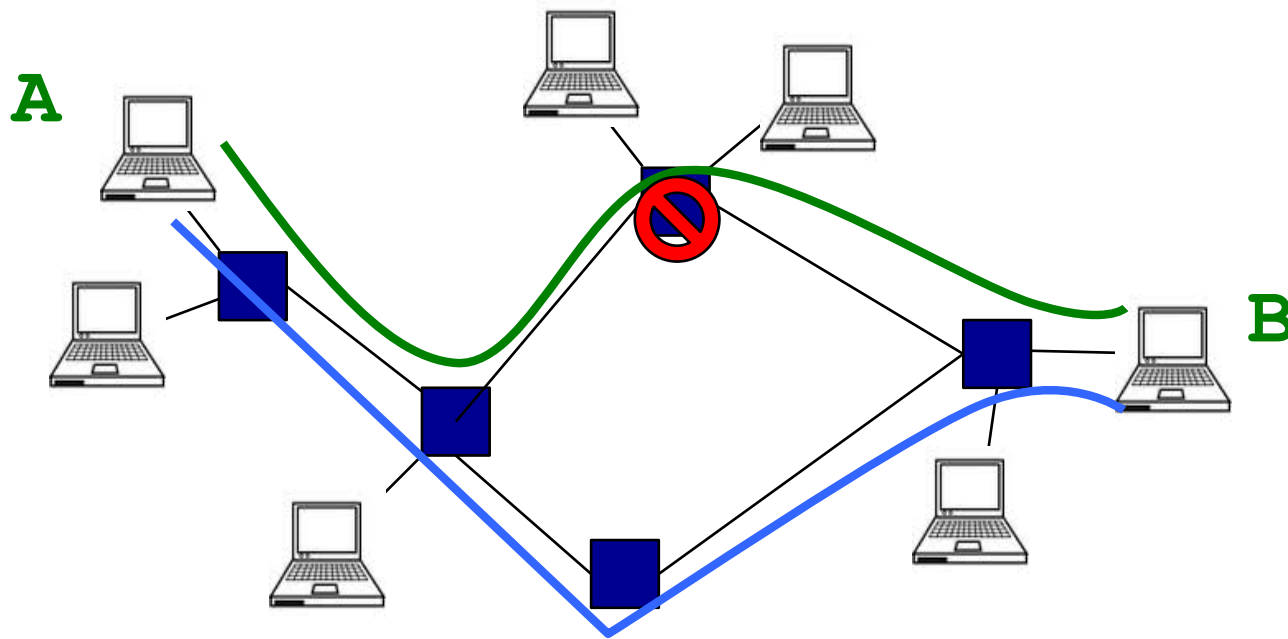




# Circuit switching: pros and cons

- Pros
  - guaranteed performance
  - fast transfers (once circuit is established)
- Cons
  - wastes bandwidth if traffic is “bursty”
  - **connection setup time is overhead**

# Circuit switching



Circuit switching doesn't "route around failure"

# Circuit switching: pros and cons

- Pros
  - guaranteed performance
  - fast transfers (once circuit is established)
- Cons
  - wastes bandwidth if traffic is “bursty”
  - connection setup time is overhead
  - **recovery from failure is slow**

# Numerical example

- How long does it take to send a file of 640,000 bits from host A to host B over a circuit-switched network?
  - All links are 1.536 Mbps
  - Each link uses TDM with 24 slots/sec
  - 500 msec to establish end-to-end circuit

Let's work it out!

# Two forms of switched networks

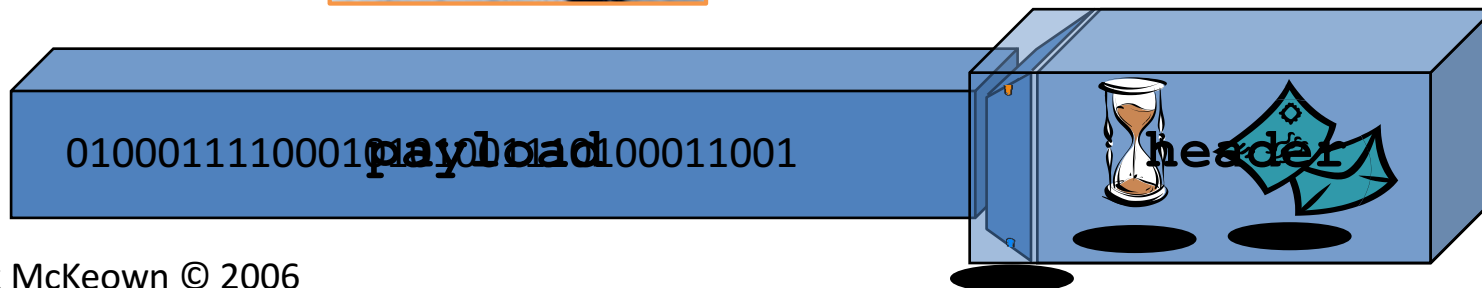
- Circuit switching (e.g., telephone network)
- Packet switching (e.g., Internet)

# Packet Switching

- Data is sent as chunks of formatted bits (**Packets**)
- Packets consist of a “**header**” and “**payload**”\*



1. Internet Address
2. Age (TTL)
3. Checksum to protect header



# Packet Switching

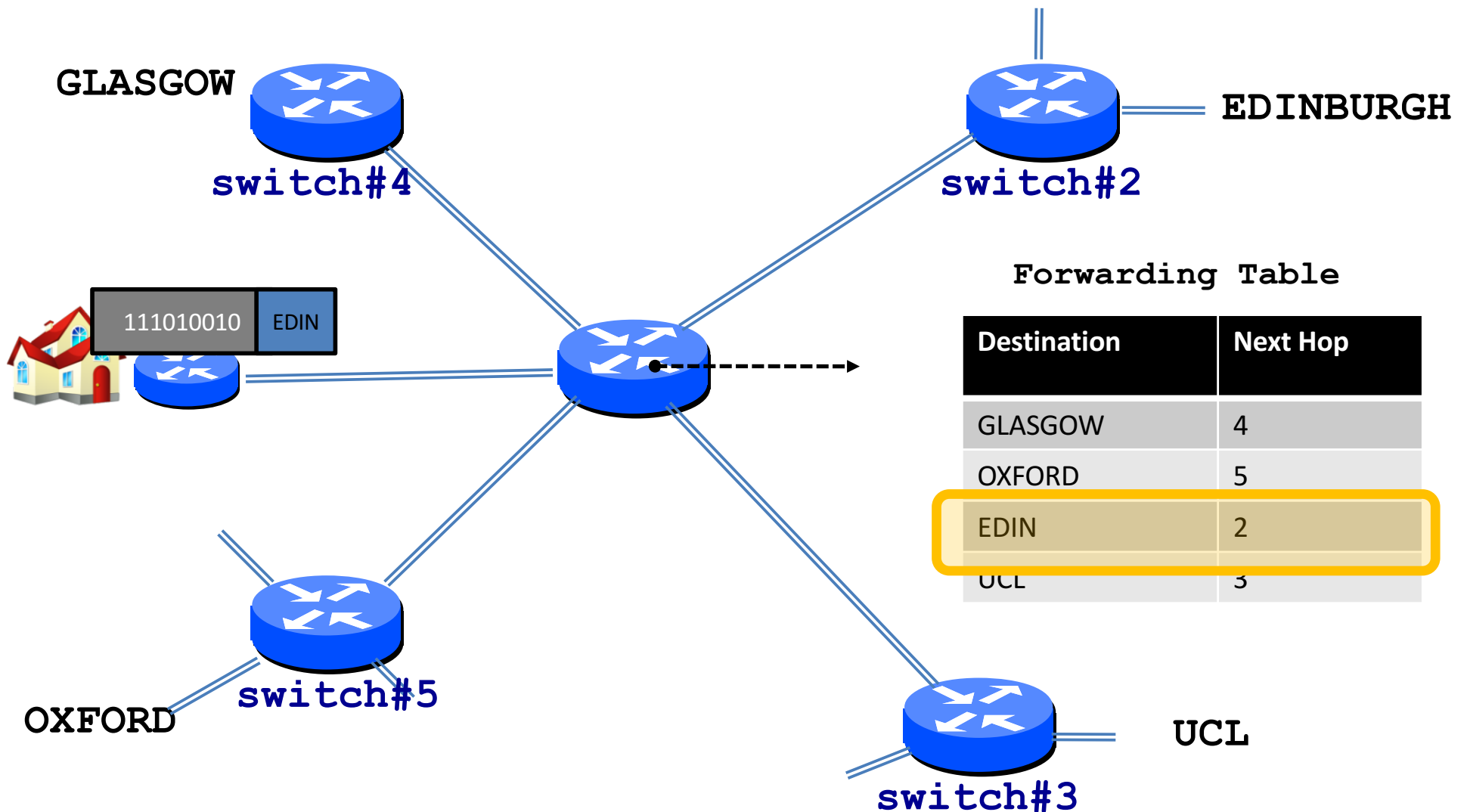
- Data is sent as chunks of formatted bits (**Packets**)
- Packets consist of a “**header**” and “**payload**”<sup>\*</sup>
  - payload is the data being carried
  - header holds instructions to the network for how to handle packet (think of the header as an API)

# Packet Switching

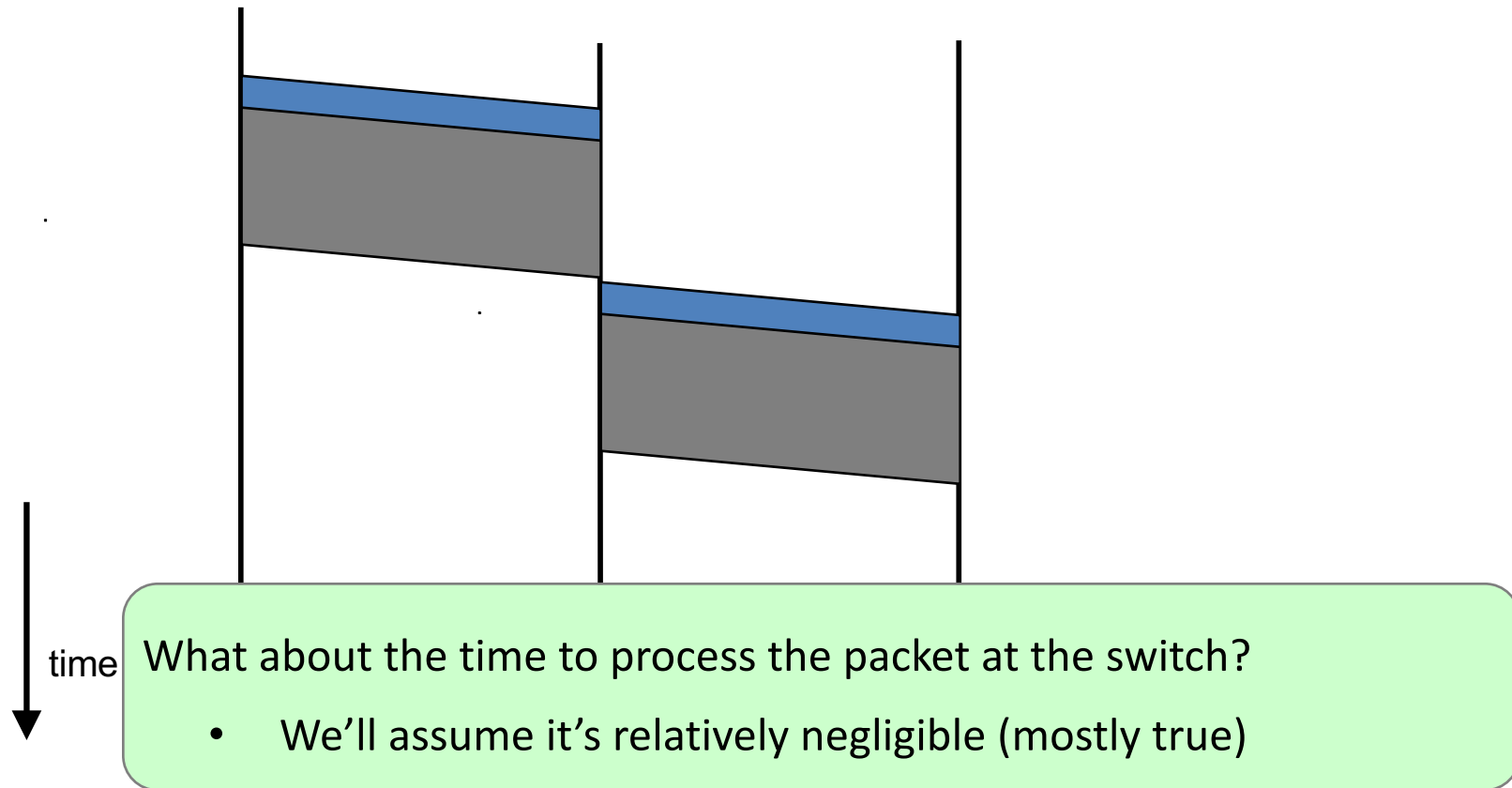
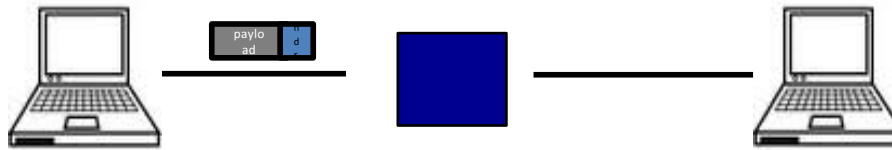
- Data is sent as chunks of formatted bits (Packets)
- Packets consist of a “header” and “payload”
- Switches “forward” packets based on their headers



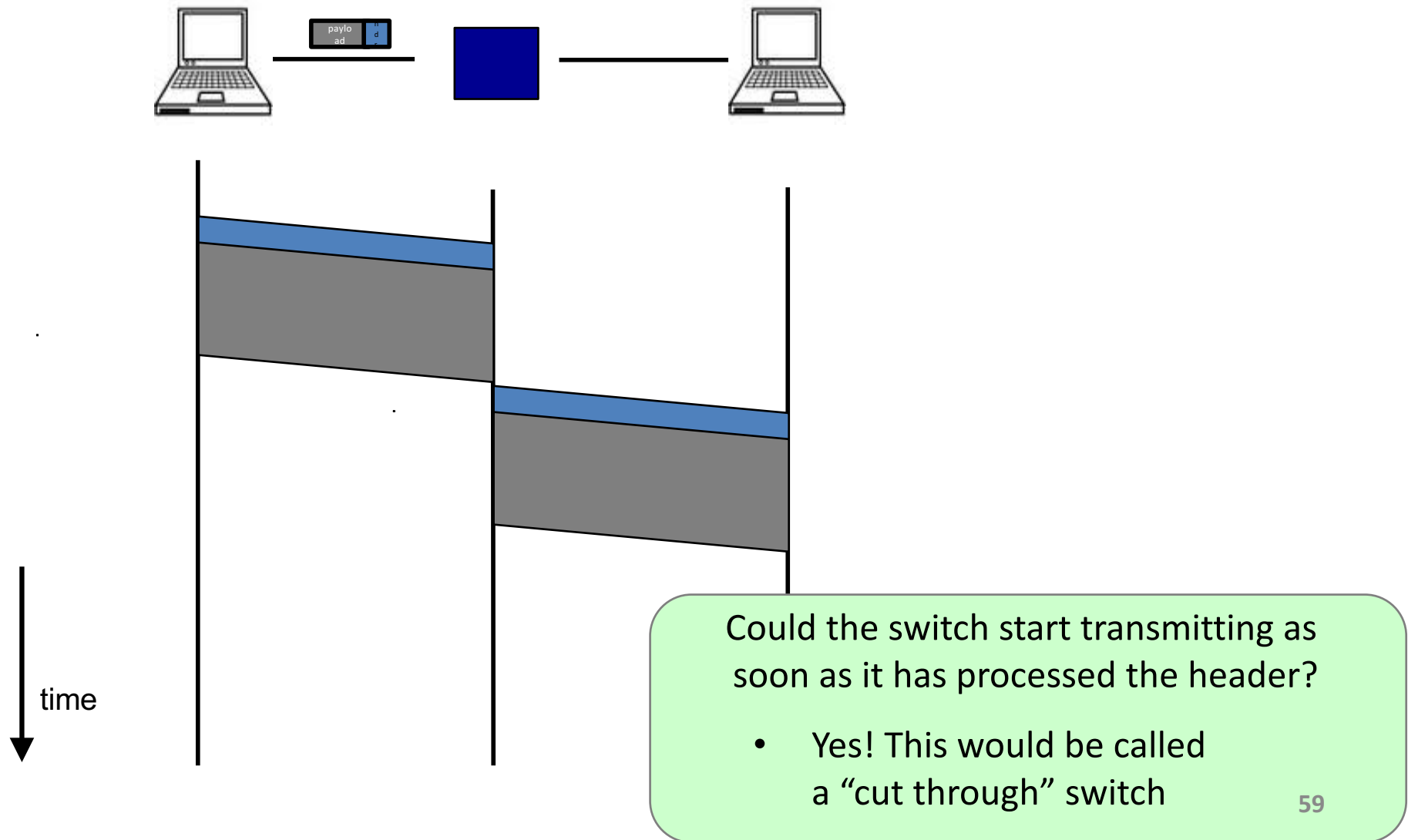
# Switches forward packets



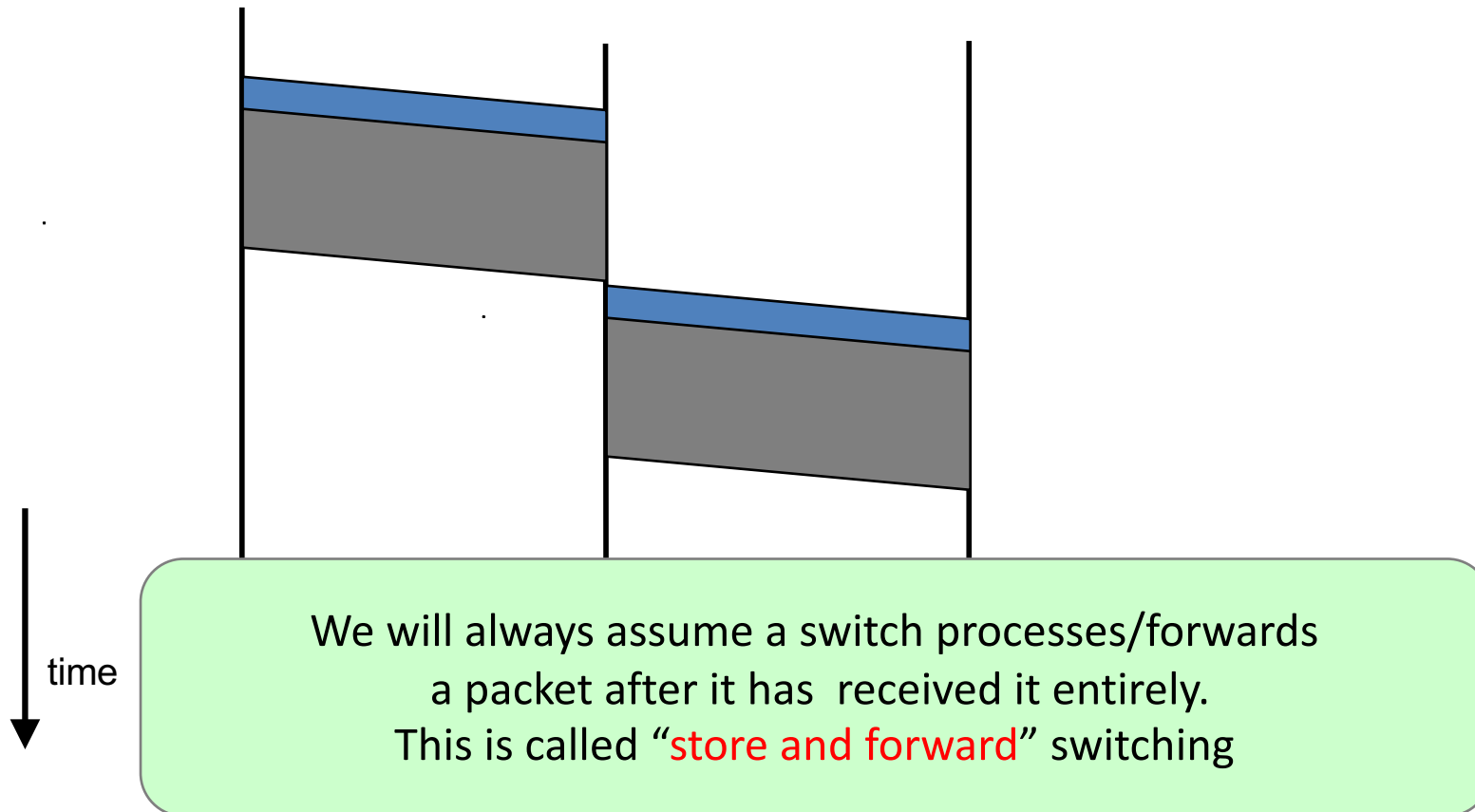
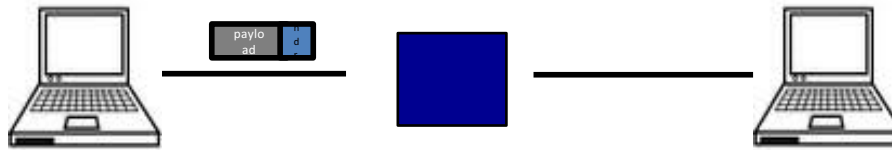
# Timing in Packet Switching



# Timing in Packet Switching



# Timing in Packet Switching



# Packet Switching

- Data is sent as chunks of formatted bits (Packets)
- Packets consist of a “header” and “payload”
- Switches “forward” packets based on their headers

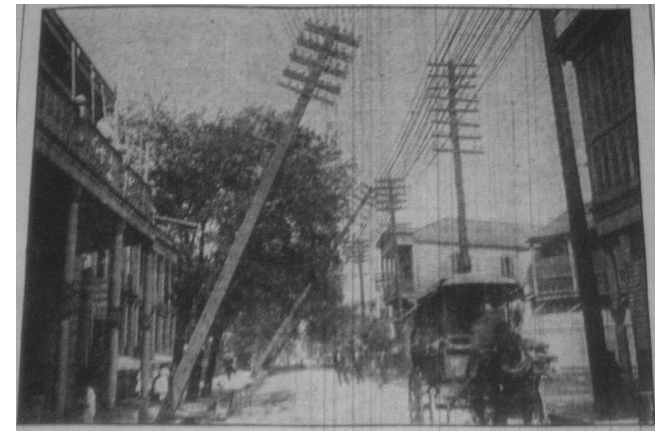
# Packet Switching

- Data is sent as chunks of formatted bits (Packets)
- Packets consist of a “header” and “payload”
- Switches “forward” packets based on their headers
- Each packet travels independently
  - no notion of packets belonging to a “circuit”

# Packet Switching

- Data is sent as chunks of formatted bits (Packets)
- Packets consist of a “header” and “payload”
- Switches “forward” packets based on their headers
- Each packet travels independently
- No link resources are reserved in advance. Instead packet switching leverages **statistical multiplexing** (stat muxing)

# Multiplexing

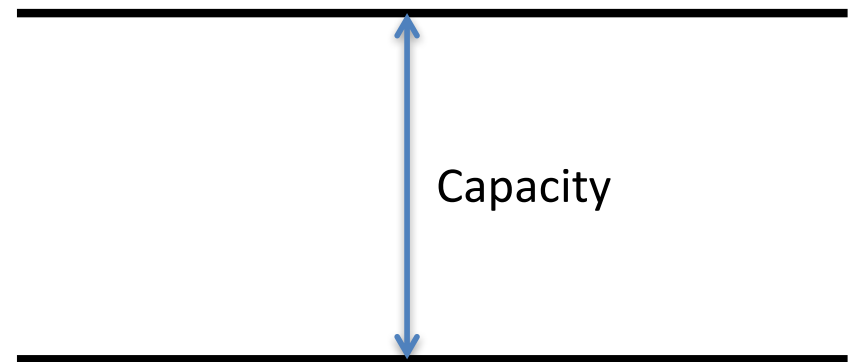
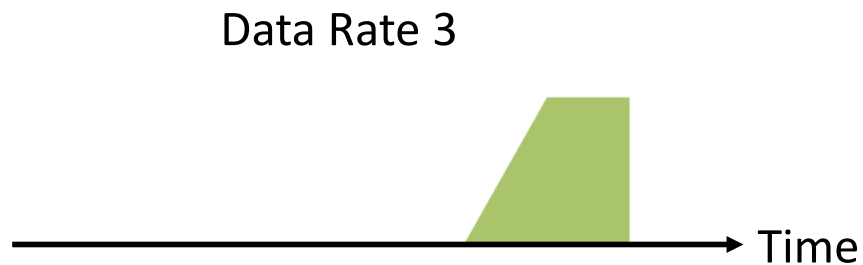
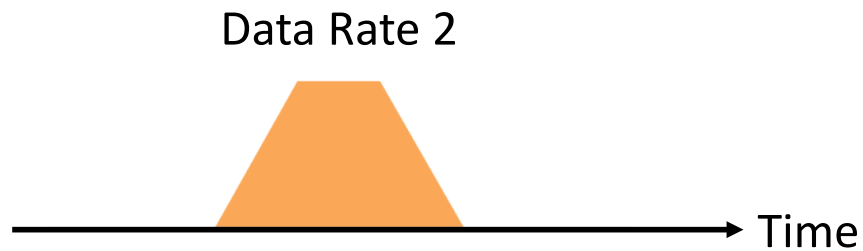
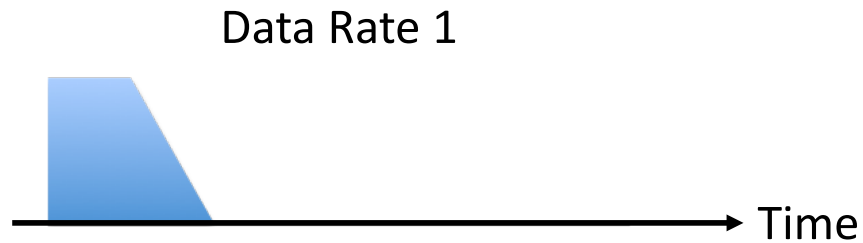


Sharing makes things efficient (cost less)

- One airplane/train for 100's of people
- One telephone for many calls
- One lecture theatre for many classes
- One computer for many tasks
- One network for many computers
- One datacenter many applications

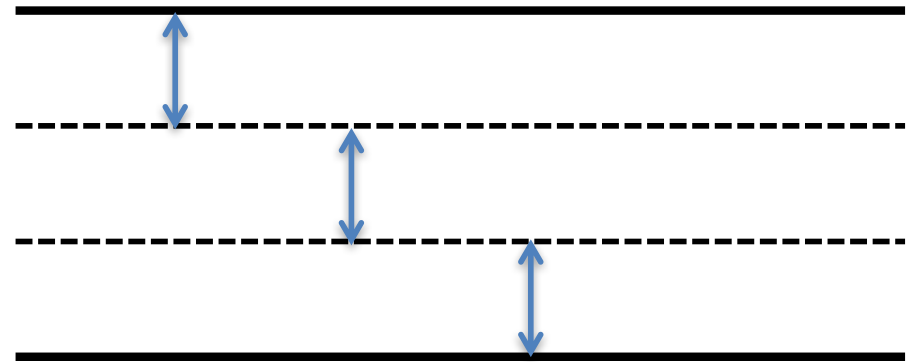
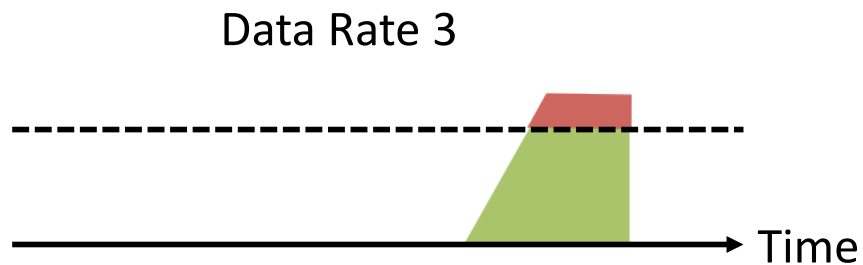
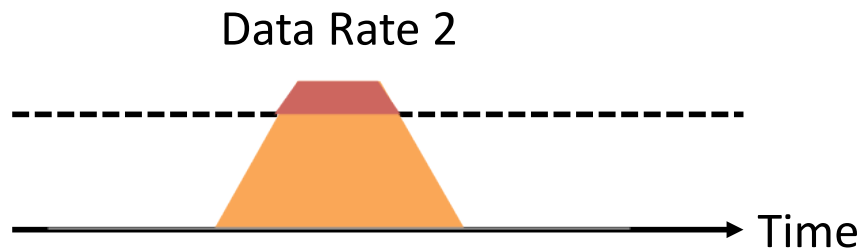
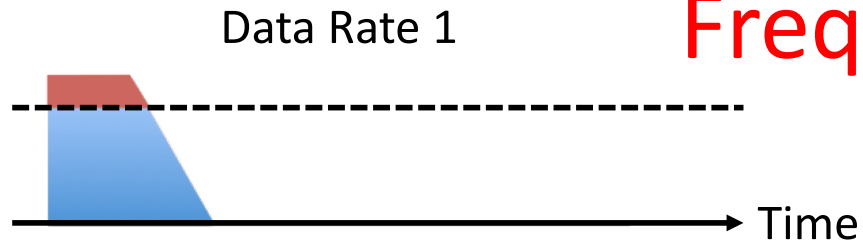


# Three Flows with Bursty Traffic



# When Each Flow Gets 1/3<sup>rd</sup> of Capacity

Frequent Overloading



# When Flows Share Total Capacity



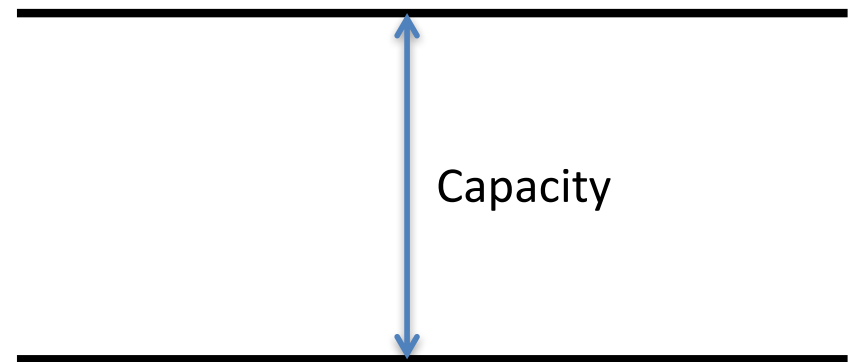
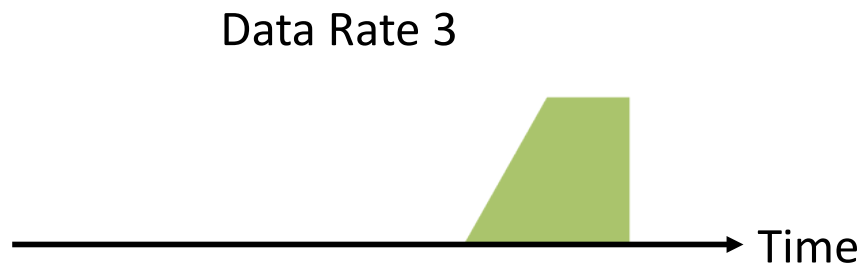
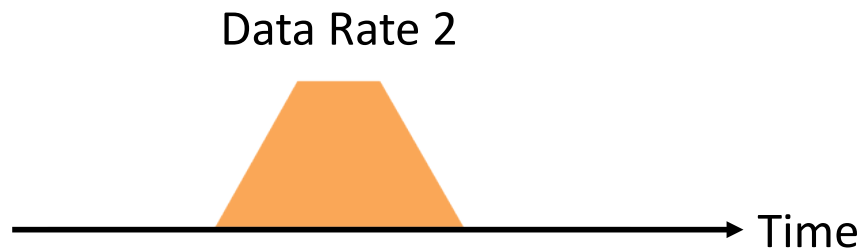
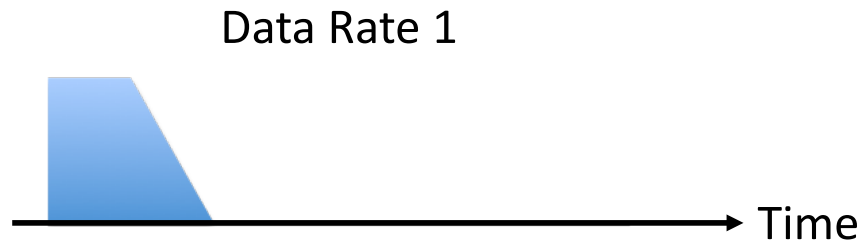
No Overloading



Statistical multiplexing relies on the assumption that not all flows burst at the same time.

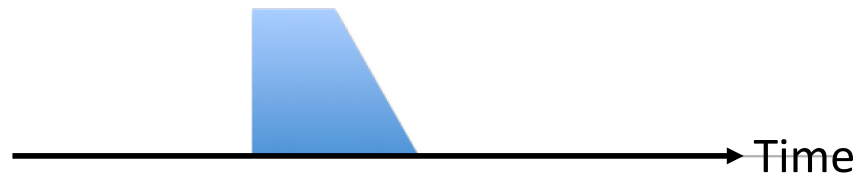
Very similar to insurance, and has same failure case

# Three Flows with Bursty Traffic

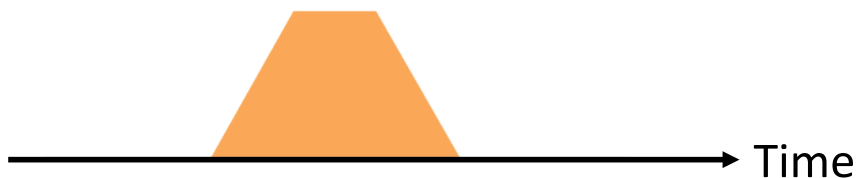


# Three Flows with Bursty Traffic

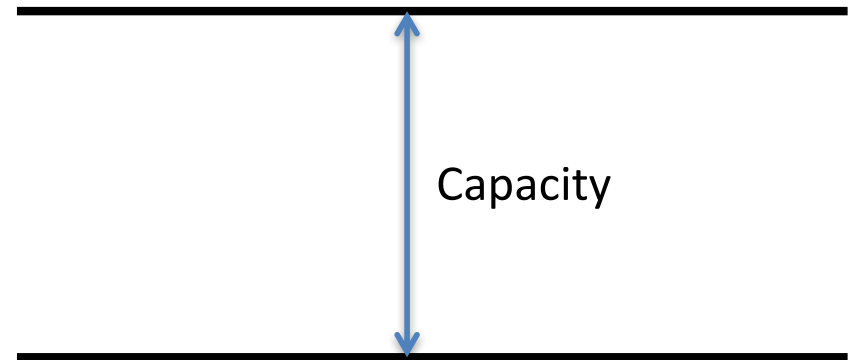
Data Rate 1



Data Rate 2

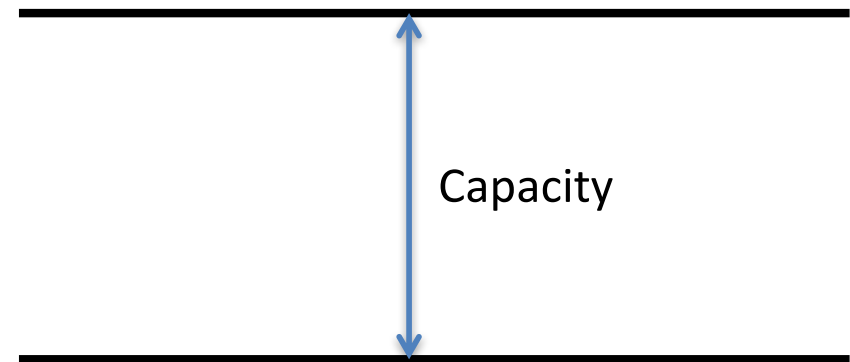
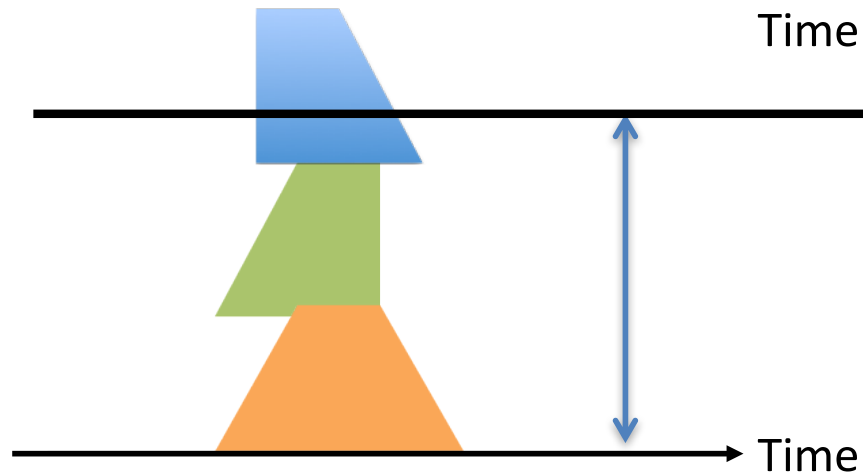


Data Rate 3



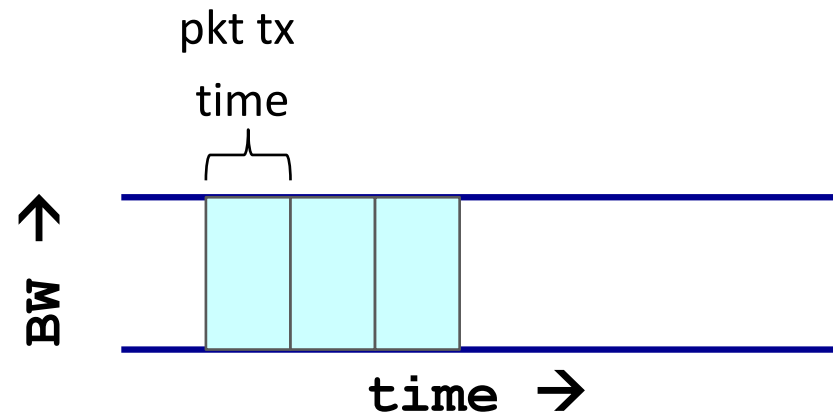
# Three Flows with Bursty Traffic

Data Rate  $1+2+3 \gg$  Capacity

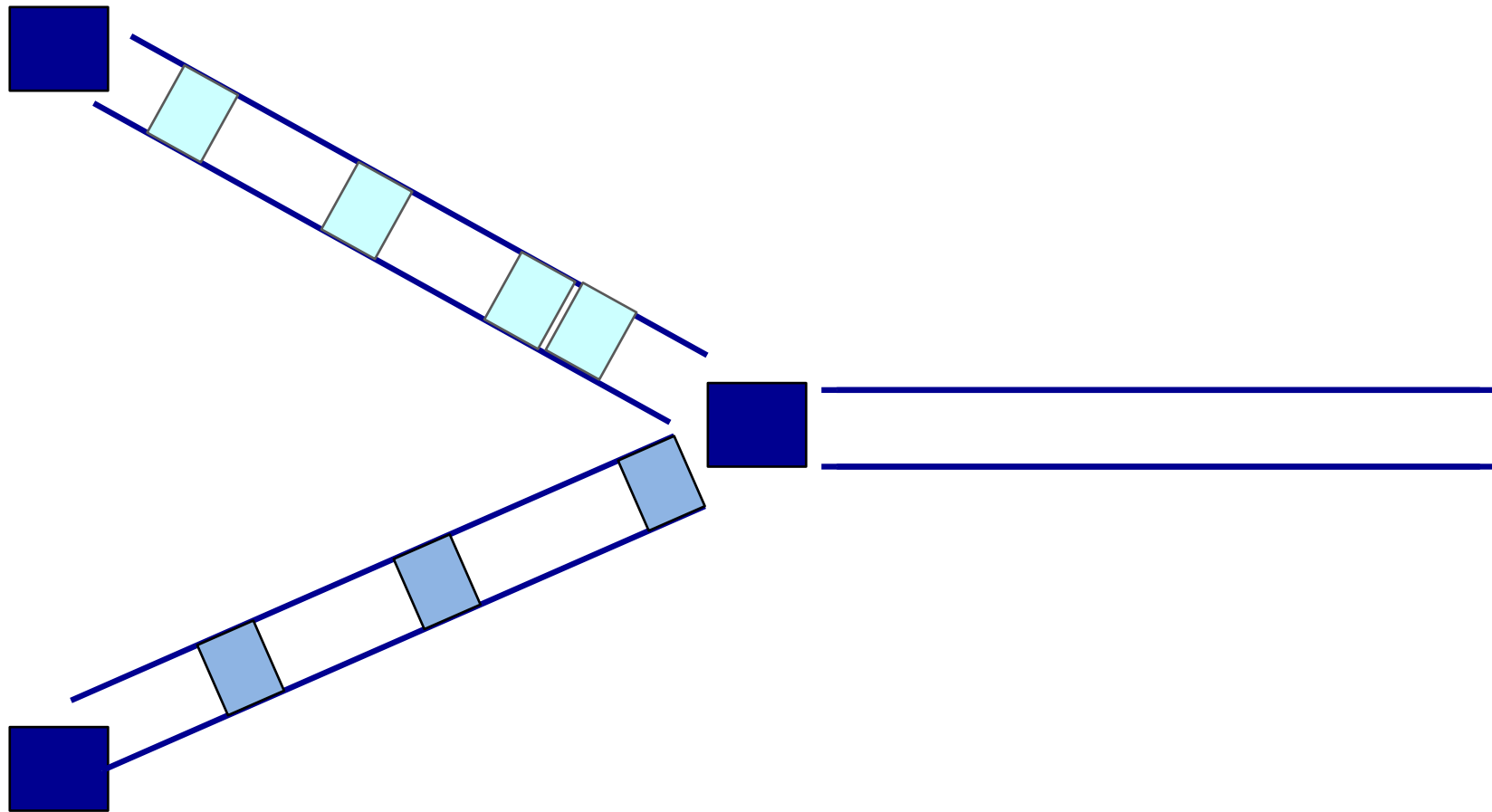


What do we do under overload?

# Statistical multiplexing: pipe view

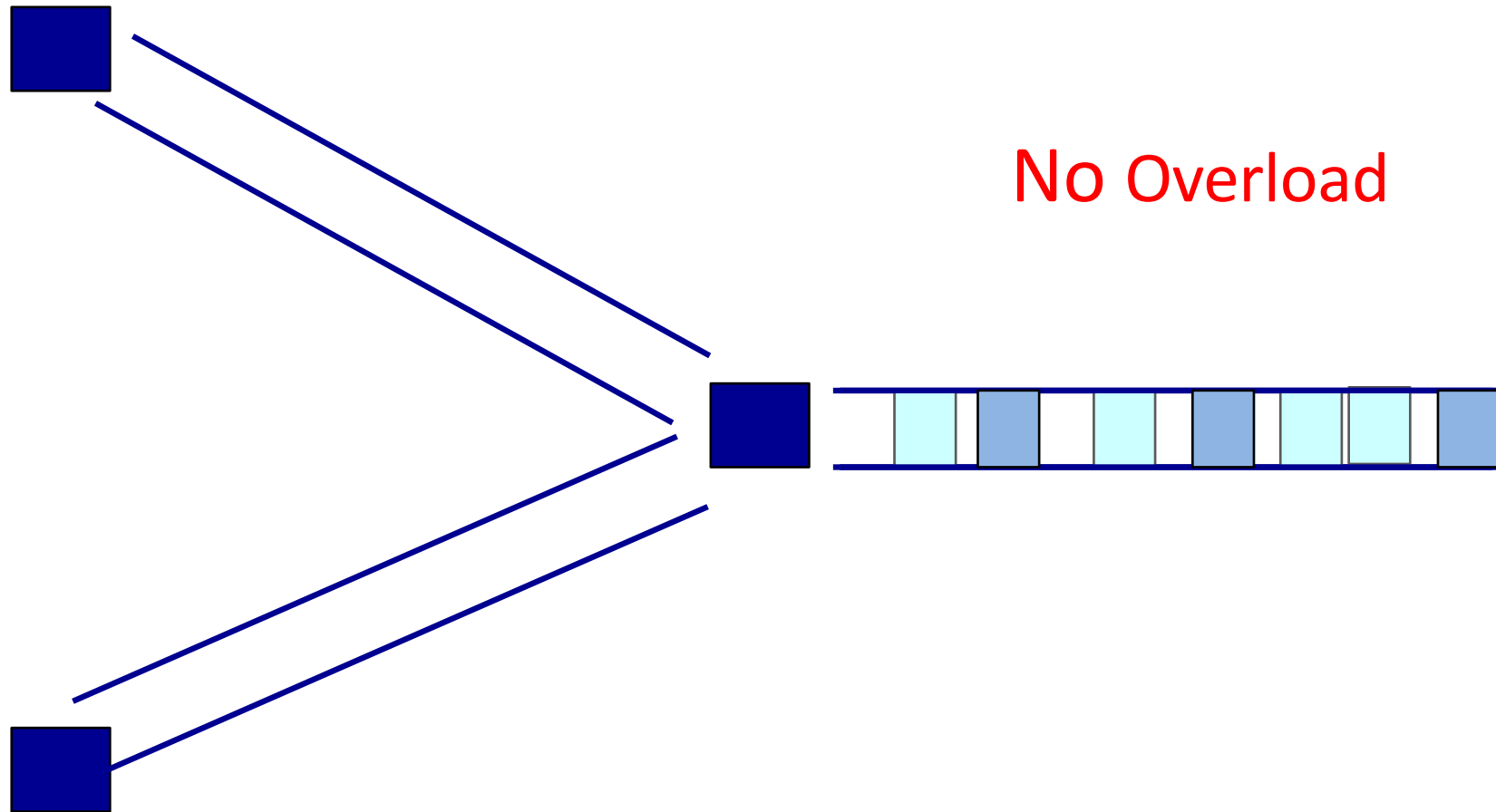


# Statistical multiplexing: pipe view

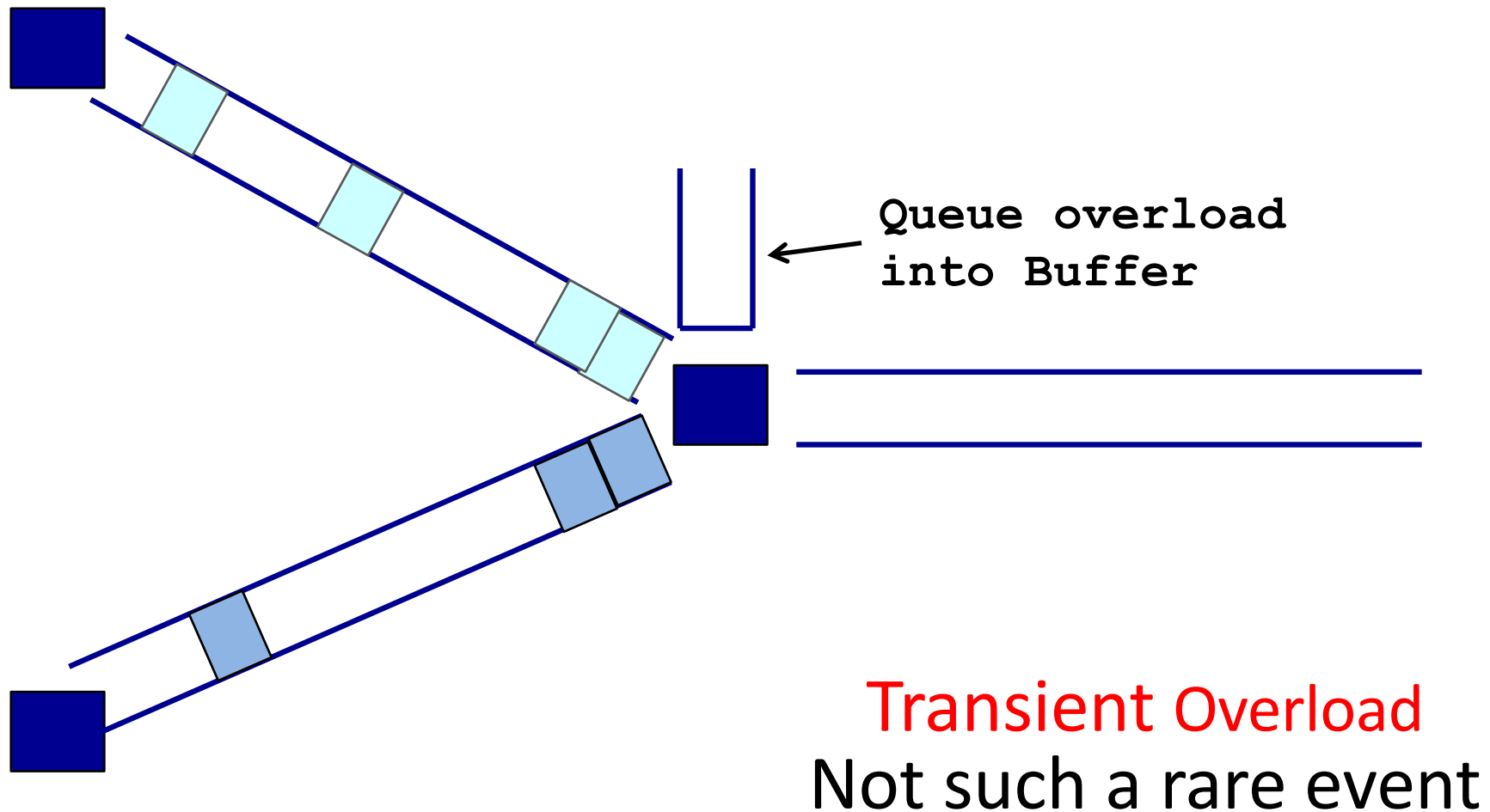




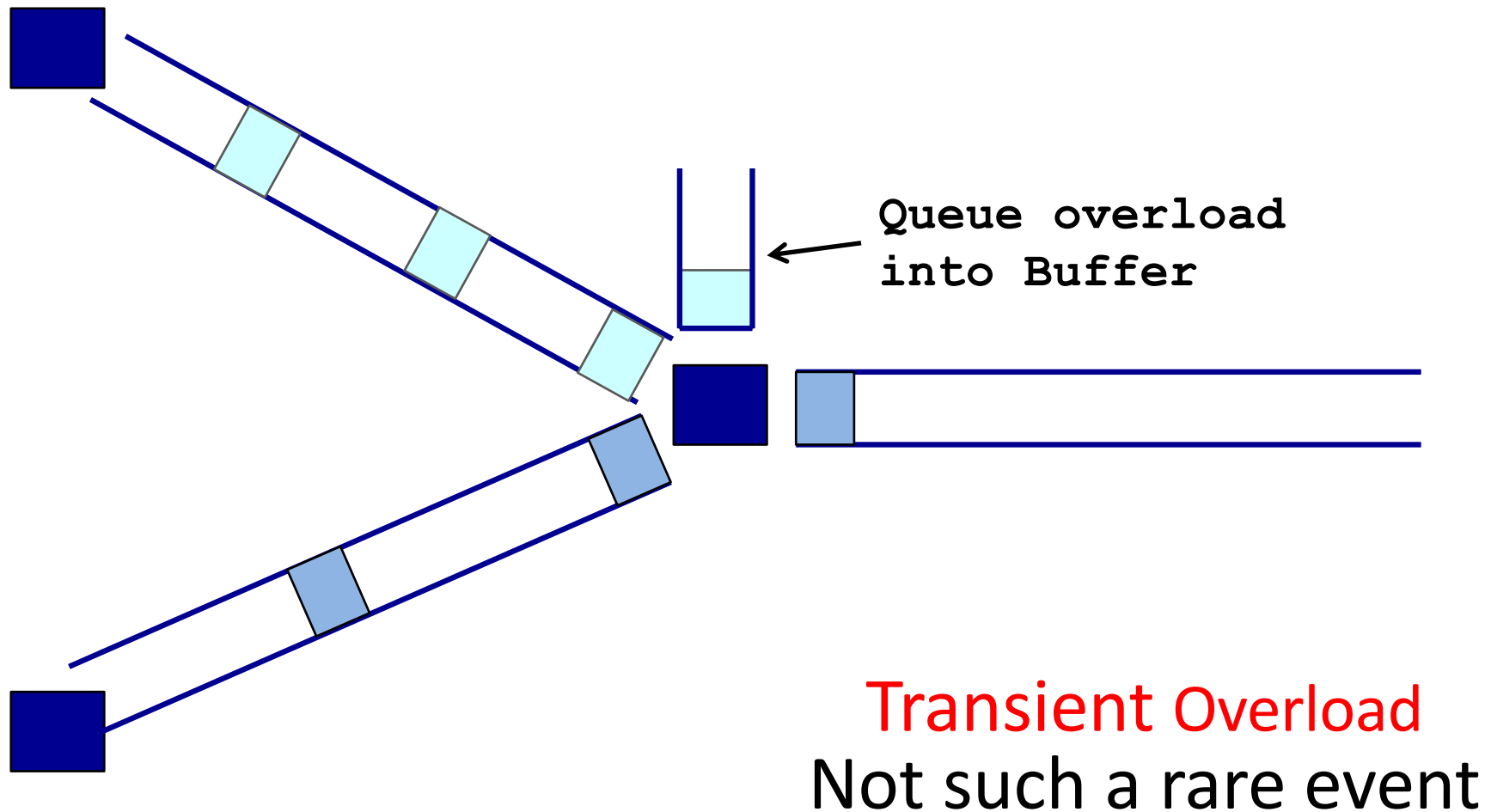
# Statistical multiplexing: pipe view



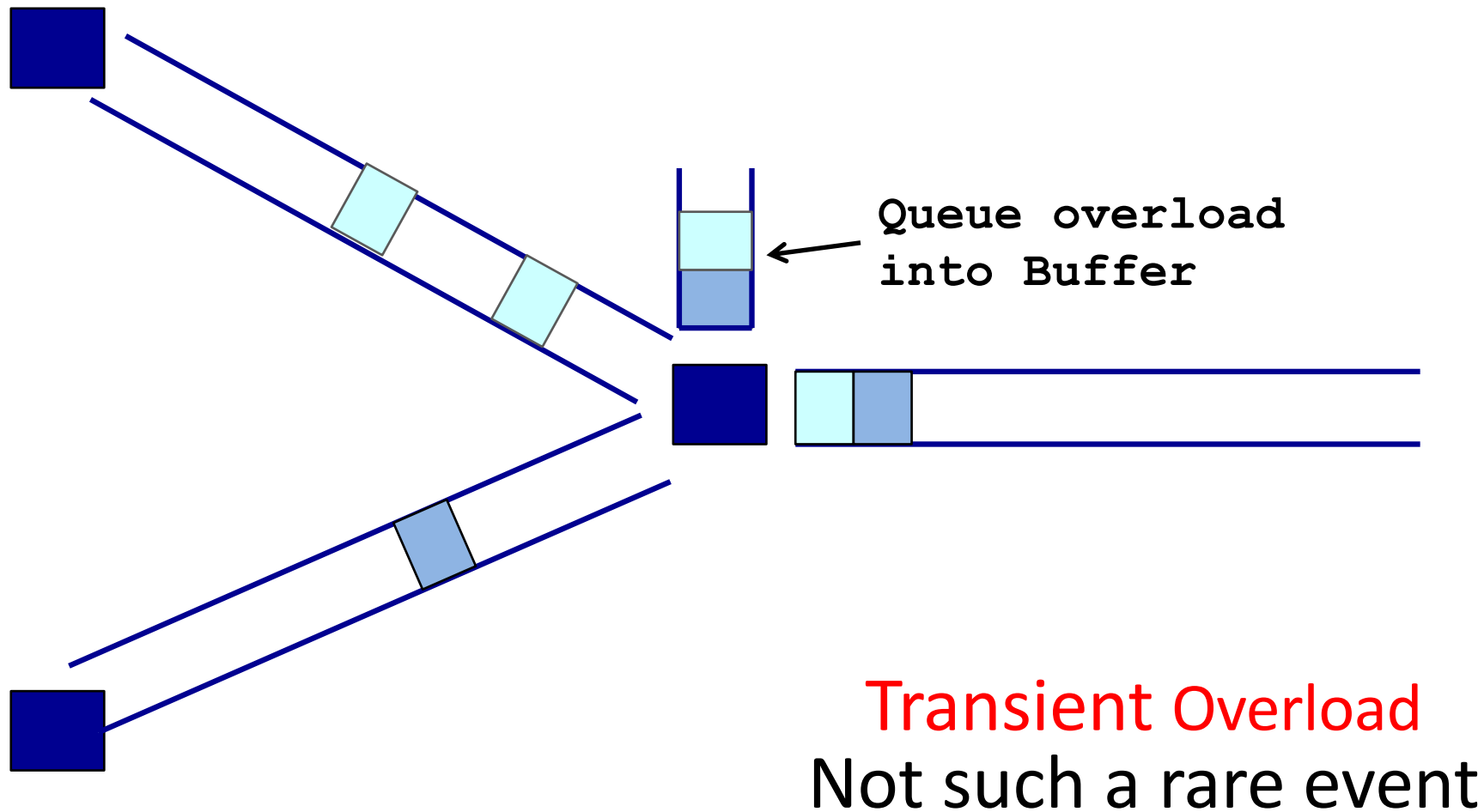
# Statistical multiplexing: pipe view



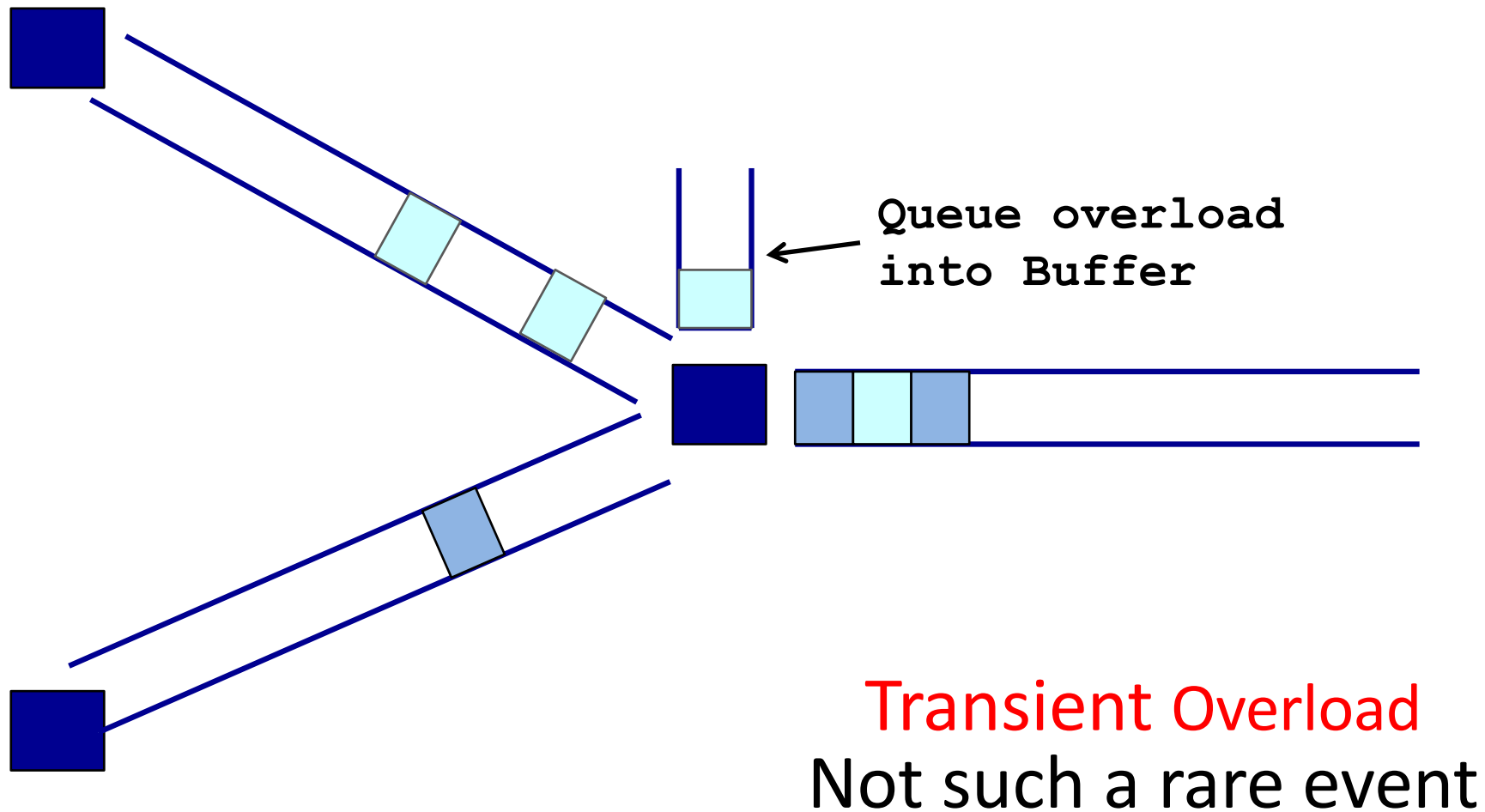
# Statistical multiplexing: pipe view



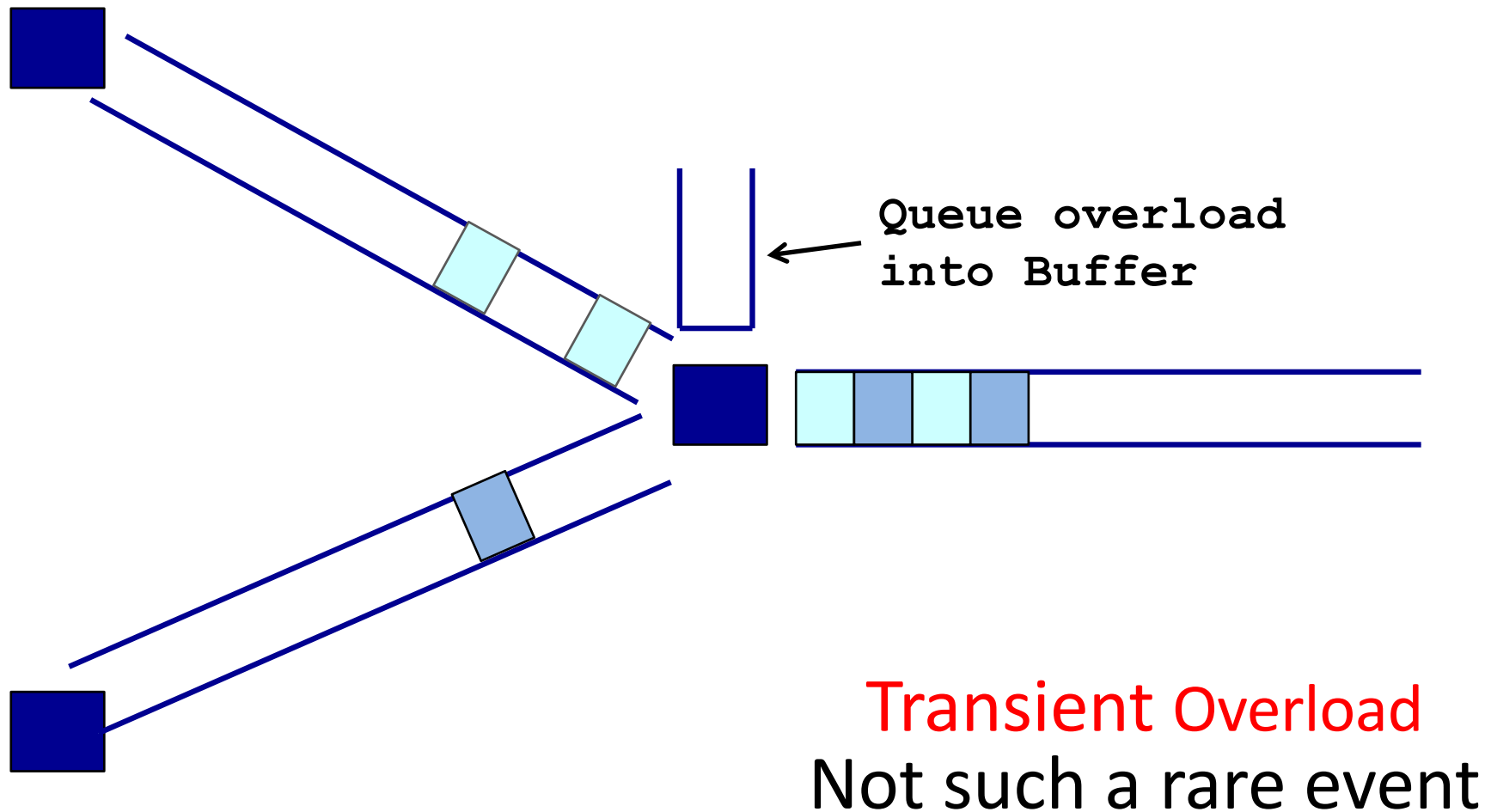
# Statistical multiplexing: pipe view



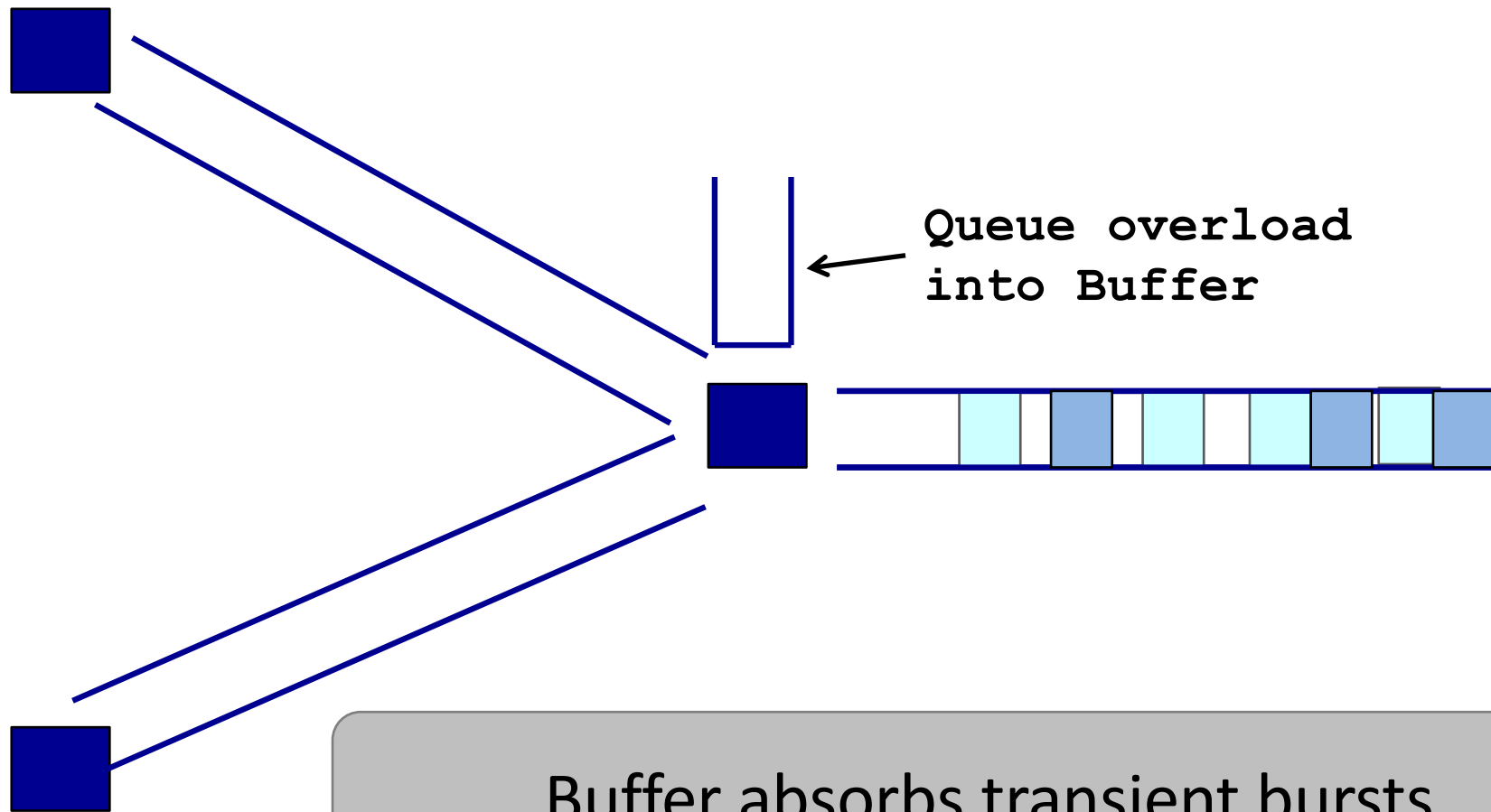
# Statistical multiplexing: pipe view



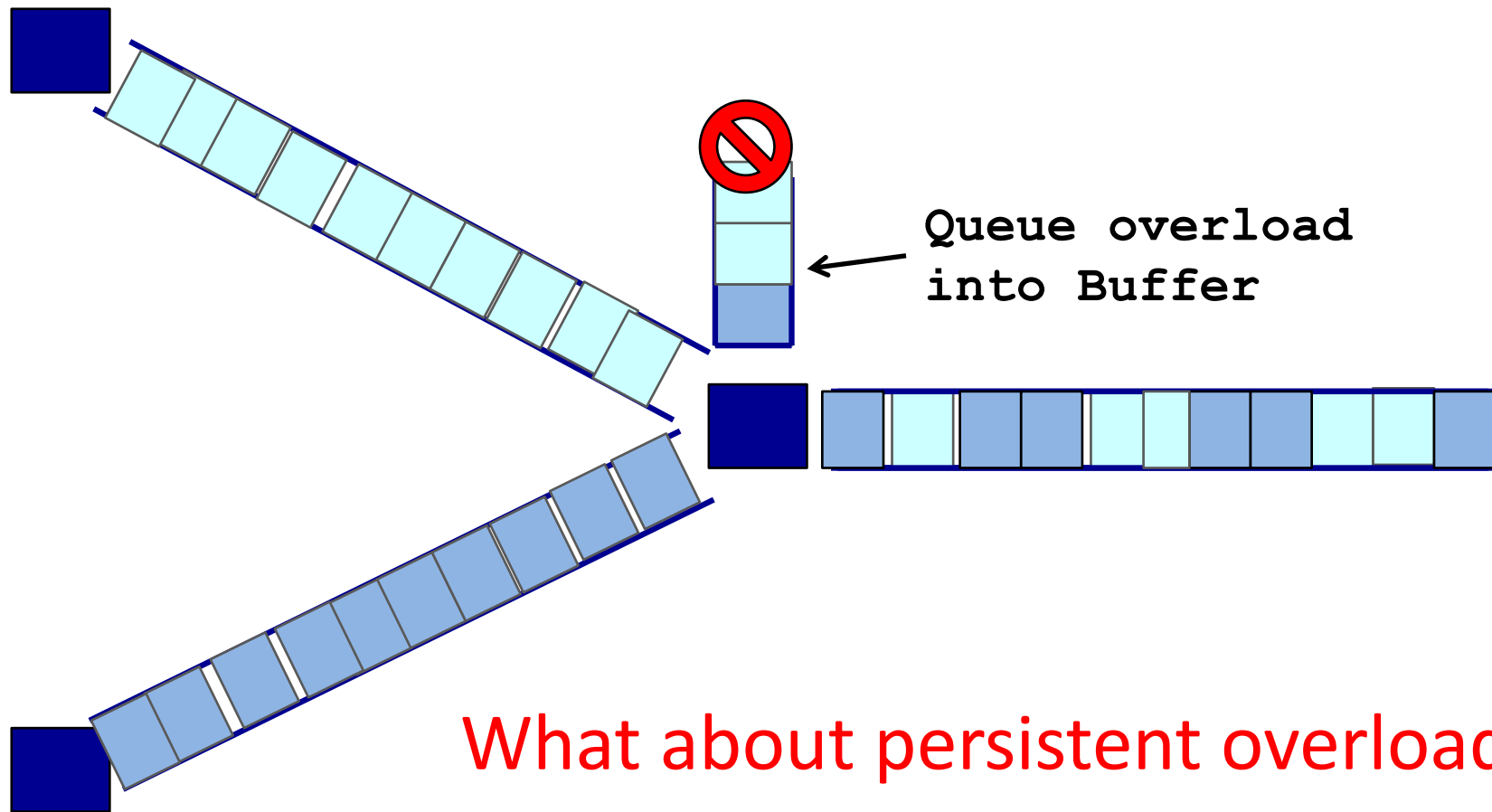
# Statistical multiplexing: pipe view



# Statistical multiplexing: pipe view



# Statistical multiplexing: pipe view



What about persistent overload?

Will eventually drop packets



# Queues introduce queuing delays

- Recall,

packet delay = transmission delay + propagation delay (\*)

- With queues (statistical multiplexing)

packet delay = transmission delay + propagation delay + queuing delay (\*)

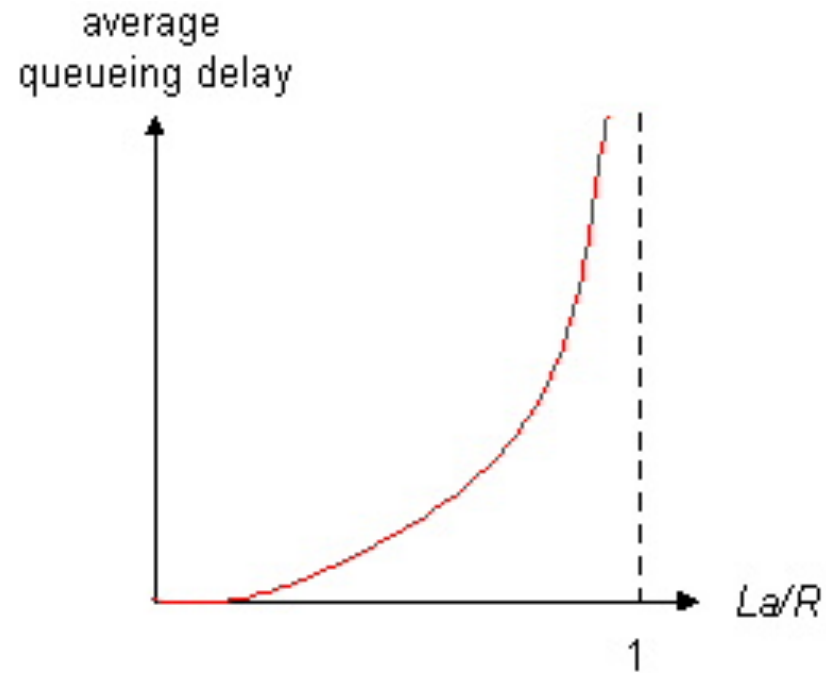
- Queuing delay caused by “packet interference”
- Made worse at high load
  - less “idle time” to absorb bursts
  - think about traffic jams at rush hour  
or rail network failure

(\* plus per-hop *processing* delay that we define as negligible)

# Queuing delay

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

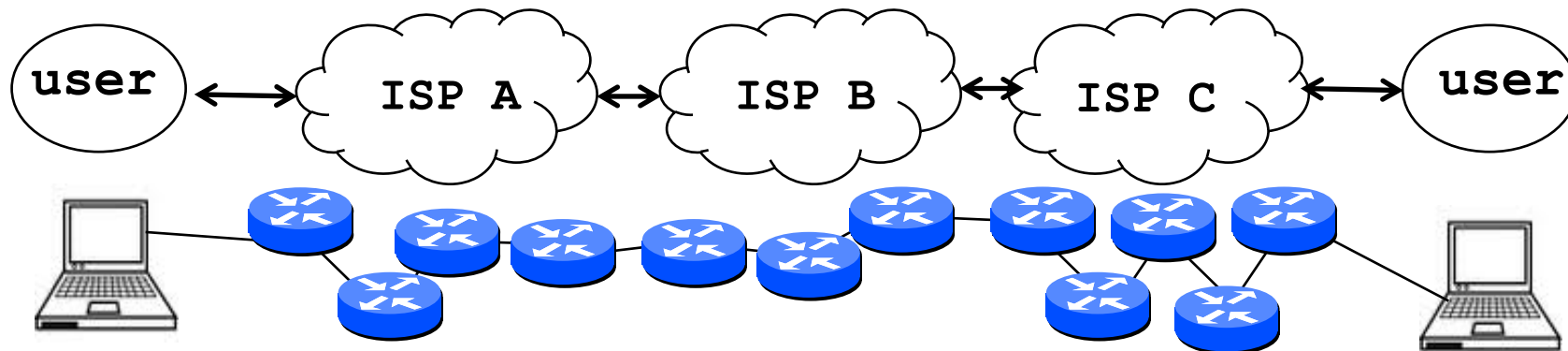
traffic intensity =  $\frac{La}{R}$



- ❑  $\frac{La}{R} \sim 0$ : average queuing delay small
- ❑  $\frac{La}{R} \rightarrow 1$ : delays become large
- ❑  $\frac{La}{R} > 1$ : more “work” arriving than can be serviced, average delay infinite – or data is lost (*dropped*).

# Recall the Internet *federation*

- The Internet ties together different networks
  - >18,000 ISP networks



We can see (hints) of the nodes and links using traceroute...

# “Real” Internet delays and routes

traceroute: rio.cl.cam.ac.uk to munnari.oz.au

(tracepath on pwf is similar)

Three delay measurements from  
rio.cl.cam.ac.uk to gatwick.net.cl.cam.ac.uk

traceroute munnari.oz.au

traceroute to munnari.oz.au (202.29.151.3), 30 hops max, 60 byte packets

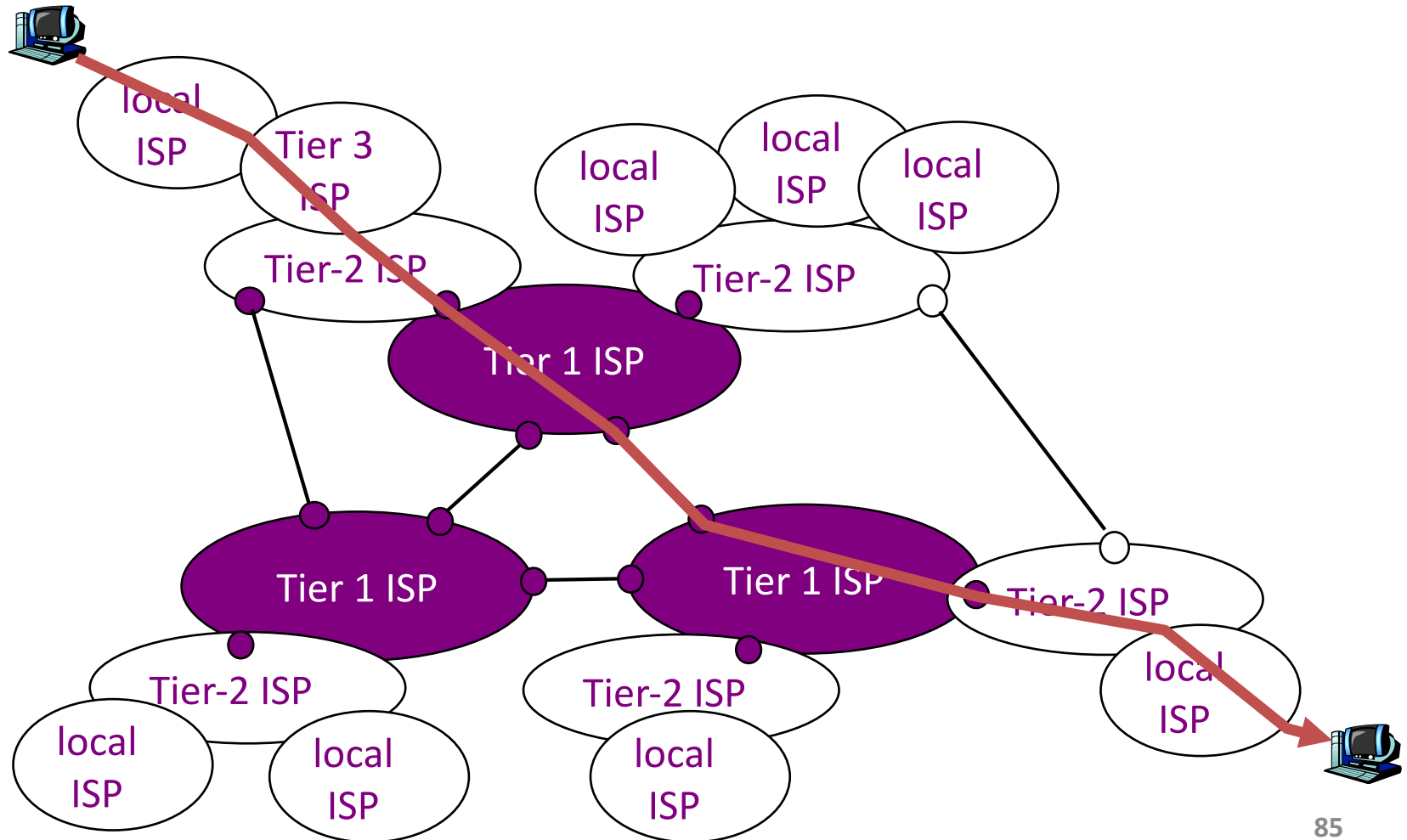
```
1  gatwick.net.cl.cam.ac.uk (128.232.32.2) 0.416 ms 0.384 ms 0.427 ms
2  cl-sby.route-nwest.net.cam.ac.uk (193.60.89.9) 0.393 ms 0.440 ms 0.494 ms
3  route-nwest.route-mill.net.cam.ac.uk (192.84.5.137) 0.407 ms 0.448 ms 0.501 ms
4  route-mill.route-enet.net.cam.ac.uk (192.84.5.94) 1.006 ms 1.091 ms 1.163 ms
5  xe-11-3-0.camb-rbr1.eastern.ja.net (146.97.130.1) 0.300 ms 0.313 ms 0.350 ms
6  ae24.lowdss-sbr1.ja.net (146.97.37.185) 2.679 ms 2.664 ms 2.712 ms
7  ae28.londhx-sbr1.ja.net (146.97.33.17) 5.955 ms 5.953 ms 5.901 ms
8  janet.mx1.lon.uk.geant.net (62.40.124.197) 6.059 ms 6.066 ms 6.052 ms
9  ae0.mx1.par.fr.geant.net (62.40.98.77) 11.742 ms 11.779 ms 11.724 ms
10 ae1.mx1.mad.es.geant.net (62.40.98.64) 27.751 ms 27.734 ms 27.704 ms
11 mb-so-02-v4.bb.tein3.net (202.179.249.117) 138.296 ms 138.314 ms 138.282 ms
12 sg-so-04-v4.bb.tein3.net (202.179.249.53) 196.303 ms 196.293 ms 196.264 ms
13 th-pr-v4.bb.tein3.net (202.179.249.66) 225.153 ms 225.178 ms 225.196 ms
14 pyt-thairen-to-02-bdr-pyt.uni.net.th (202.29.12.10) 225.163 ms 223.343 ms 223.363 ms
15 202.28.227.126 (202.28.227.126) 241.038 ms 240.941 ms 240.834 ms
16 202.28.221.46 (202.28.221.46) 287.252 ms 287.306 ms 287.282 ms
17 * * *
18 * * *
19 * * *
20 coe-gw.psu.ac.th (202.29.149.70) 241.681 ms 241.715 ms 241.680 ms
21 munnari.OZ.AU (202.29.151.3) 241.610 ms 241.636 ms 241.537 ms
```

trans-continent  
link

\* means no response (probe lost, router not replying)

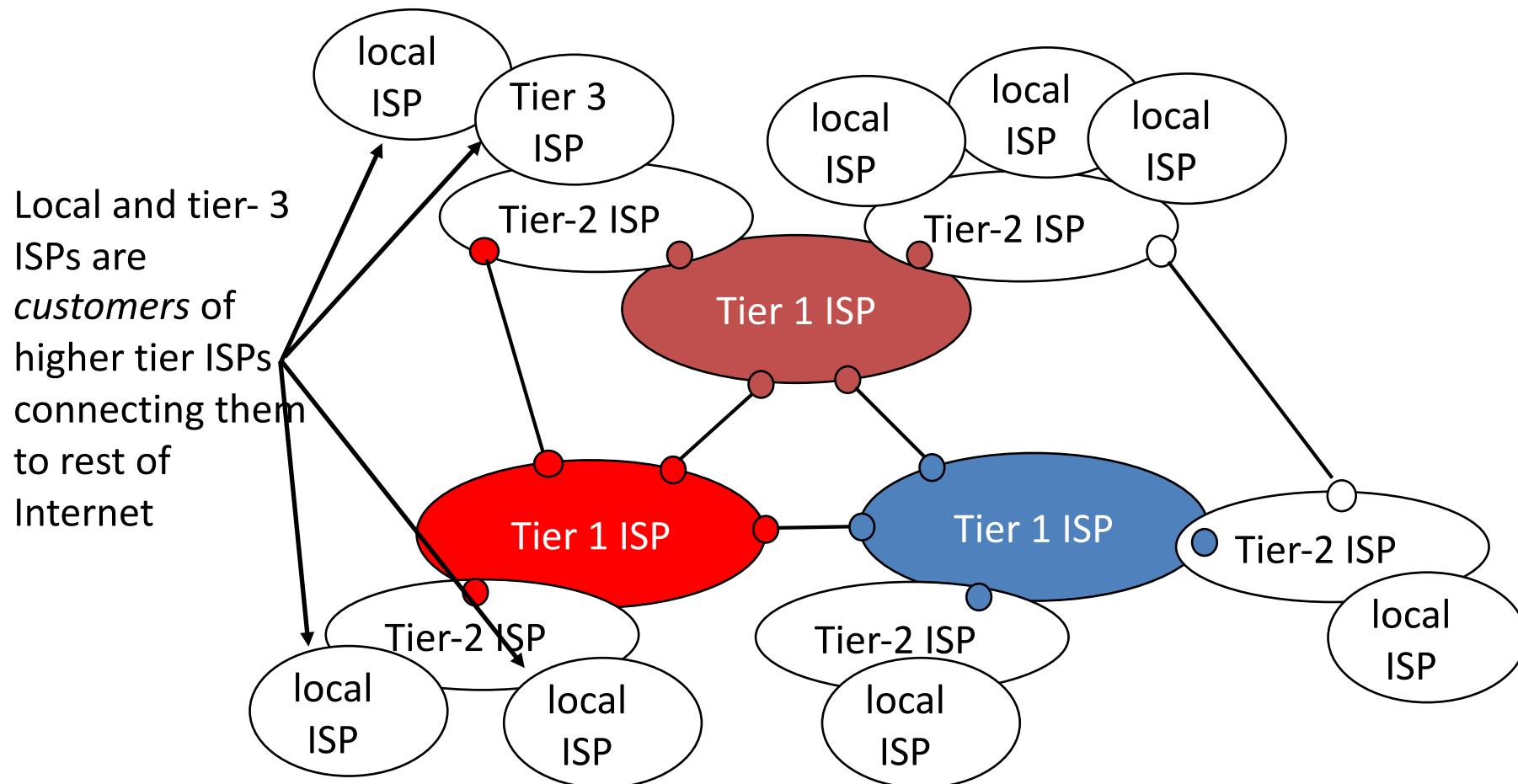
# Internet structure: network of networks

- a packet passes through many networks!



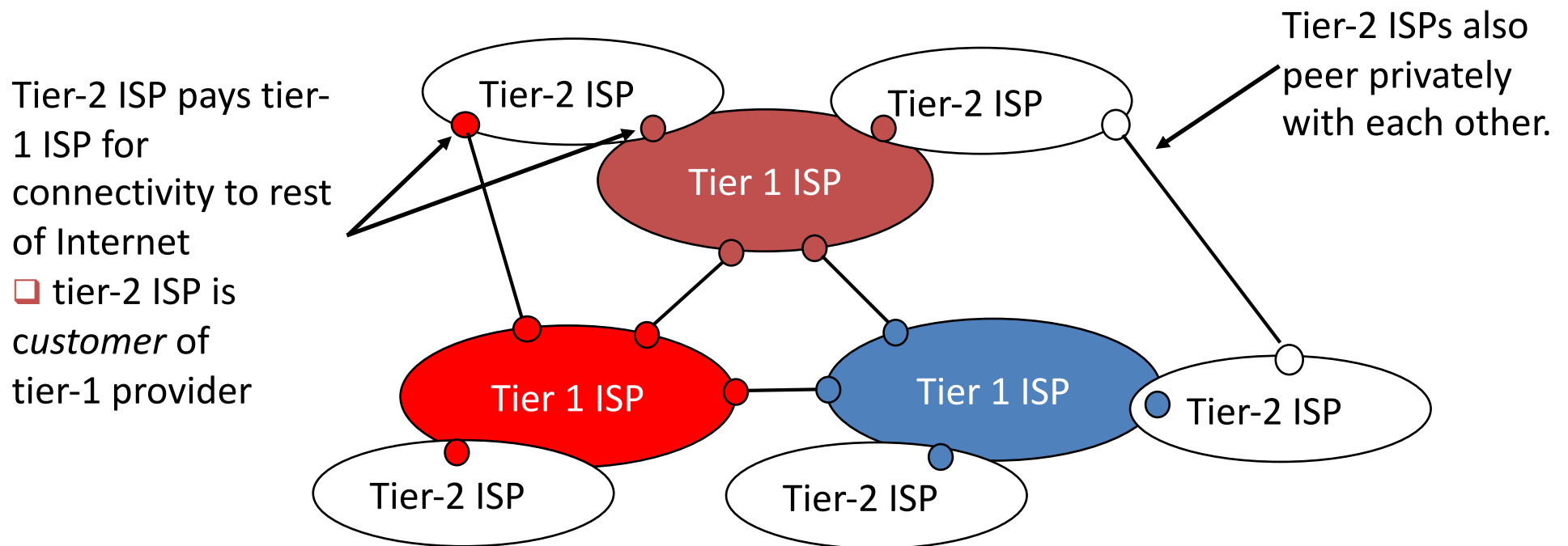
# Internet structure: network of networks

- “Tier-3” ISPs and local ISPs
  - last hop (“access”) network (closest to end systems)



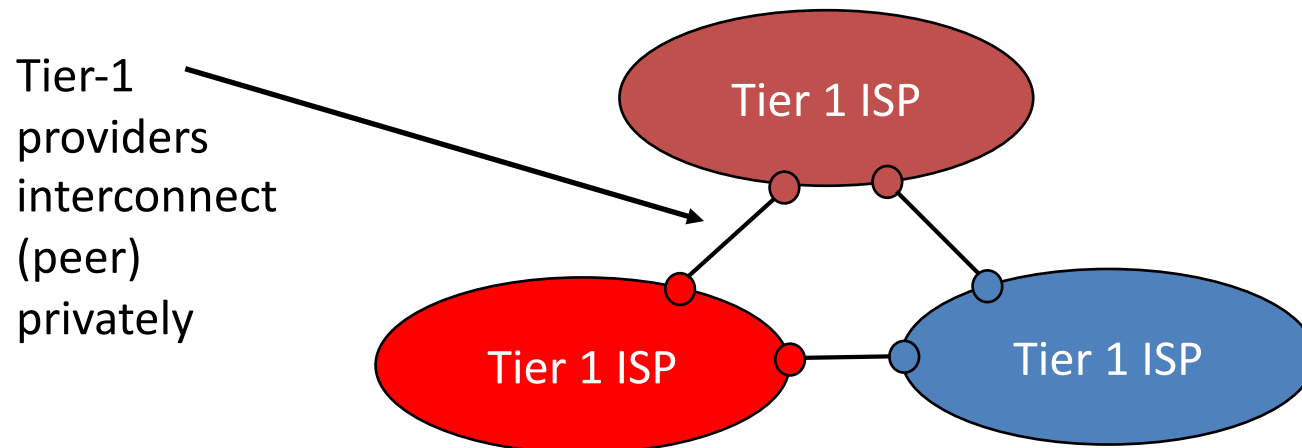
# Internet structure: network of networks

- “Tier-2” ISPs: smaller (often regional) ISPs
  - Connect to one or more tier-1 ISPs, possibly other tier-2 ISPs



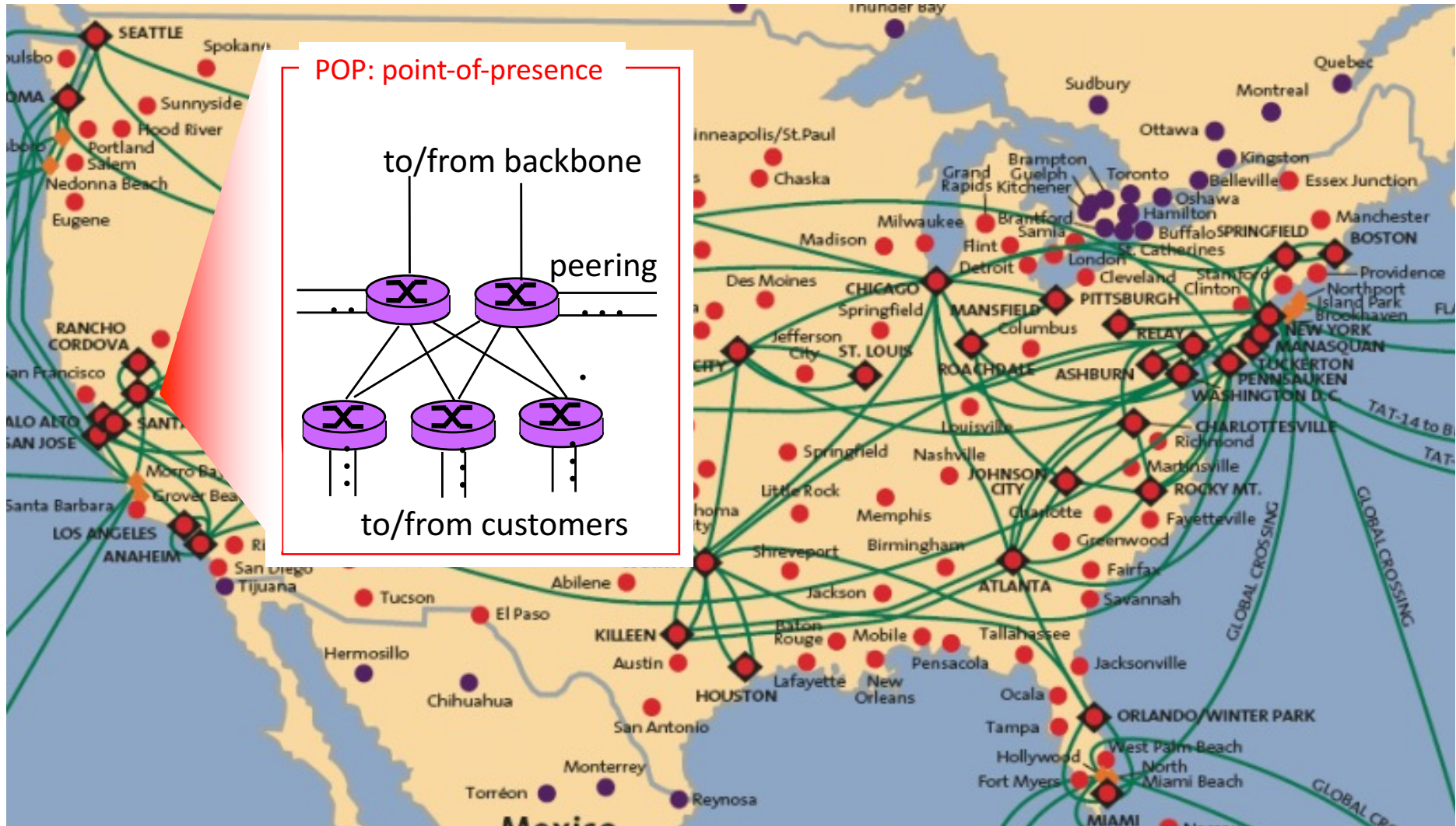
# Internet structure: network of networks

- roughly hierarchical
- **at center: “tier-1” ISPs** (e.g., Verizon, Sprint, AT&T, Cable and Wireless), national/international coverage
  - treat each other as equals





# Tier-1 ISP: e.g., Sprint



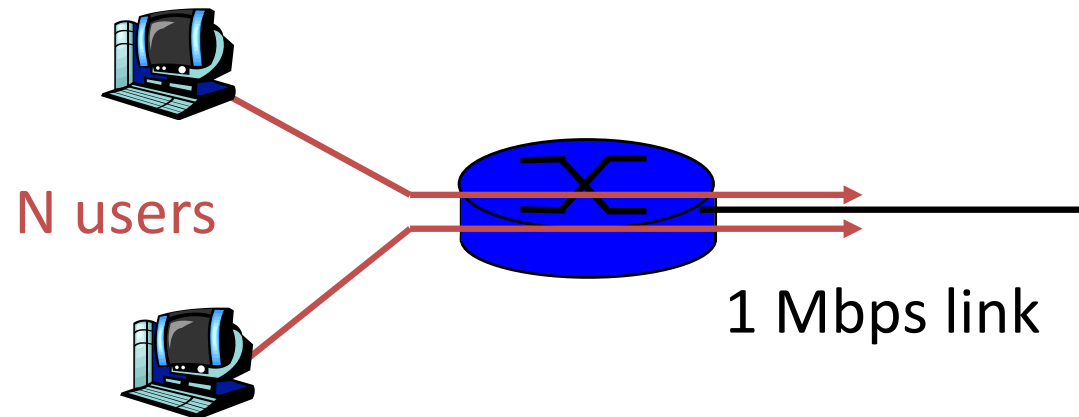
# Packet Switching

- Data is sent as chunks of formatted bits (Packets)
- Packets consist of a “header” and “payload”
- Switches “forward” packets based on their headers
- Each packet travels independently
- No link resources are reserved in advance. Instead packet switching leverages **statistical multiplexing**
  - allows efficient use of resources
  - but introduces queues and queuing delays

# Packet switching versus circuit switching

*Packet switching may (does!) allow more users to use network*

- 1 Mb/s link
- each user:
  - 100 kb/s when “active”
  - active 10% of time
- *circuit-switching:*
  - 10 users
- *packet switching:*
  - with 35 users, probability > 10 active at same time is less than .0004



Q: how did we get value 0.0004?

# Packet switching versus circuit switching

Q: how did we get value 0.0004?

- 1 Mb/s link
- each user:
  - 100 kb/s when “active”
  - active 10% of time
- *circuit-switching:*
  - 10 users
- *packet switching:*
  - with 35 users, probability > 10 active at same time is less than .0004

**HINT:** Binomial Distribution

# Circuit switching: pros and cons

- Pros
  - guaranteed performance
  - fast transfers (once circuit is established)
- Cons
  - wastes bandwidth if traffic is “bursty”
  - connection setup adds delay
  - recovery from failure is slow

# Packet switching: pros and cons

- Cons

- no guaranteed performance
- header overhead per packet
- queues and queuing delays

- Pros

- efficient use of bandwidth (stat. muxing)
- no overhead due to connection setup
- resilient -- can `route around trouble`

# Summary

- A sense of how the basic `plumbing' works
  - links and switches
  - packet delays = transmission + propagation + queuing + (negligible) per-switch processing
  - statistical multiplexing and queues
  - circuit vs. packet switching