Basic Concepts (part 2): packet delays and layering

CS168, Fall 2014

Sylvia Ratnasamy

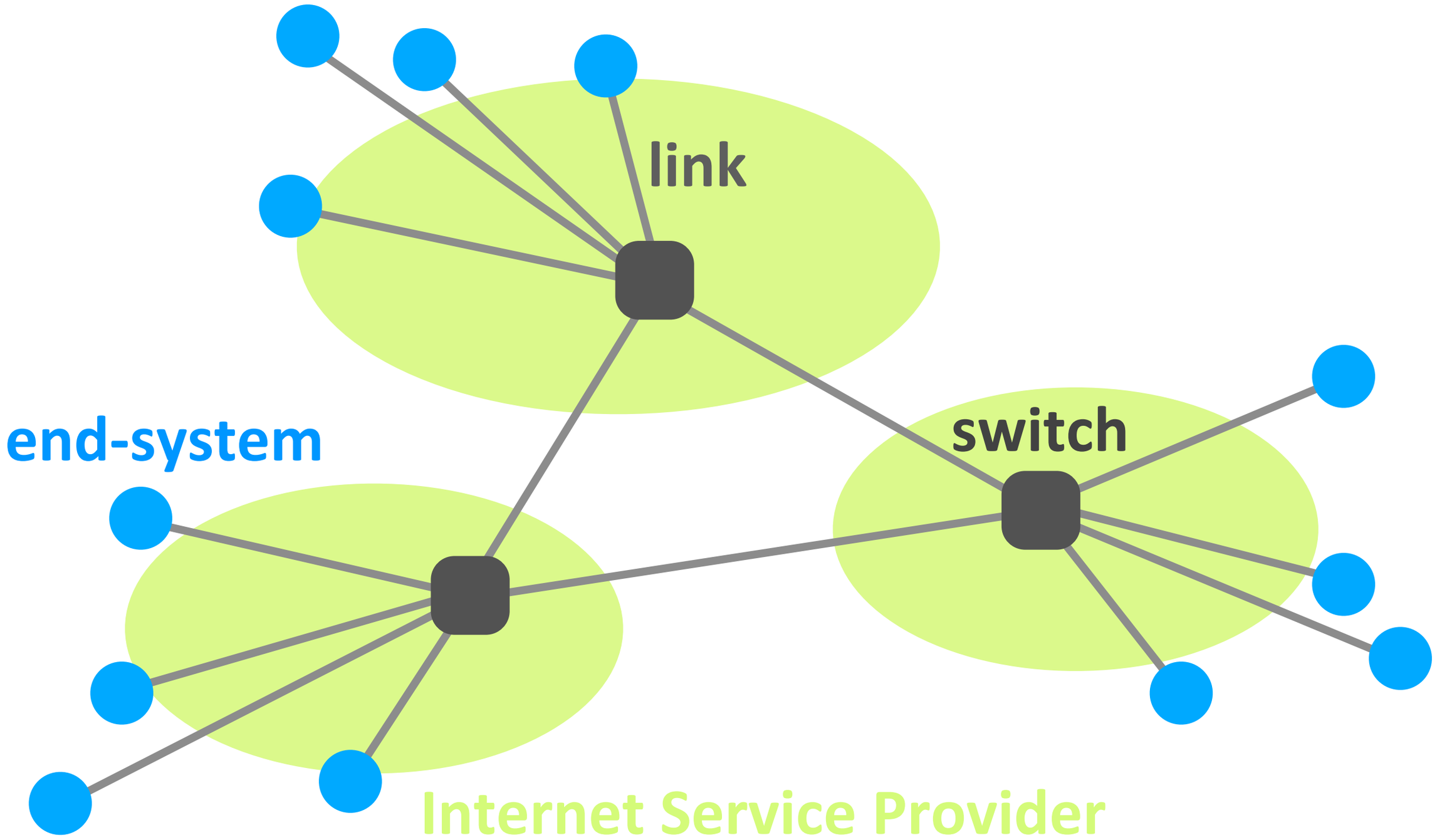
http://inst.eecs.berkeley.edu/~cs168/fa14/

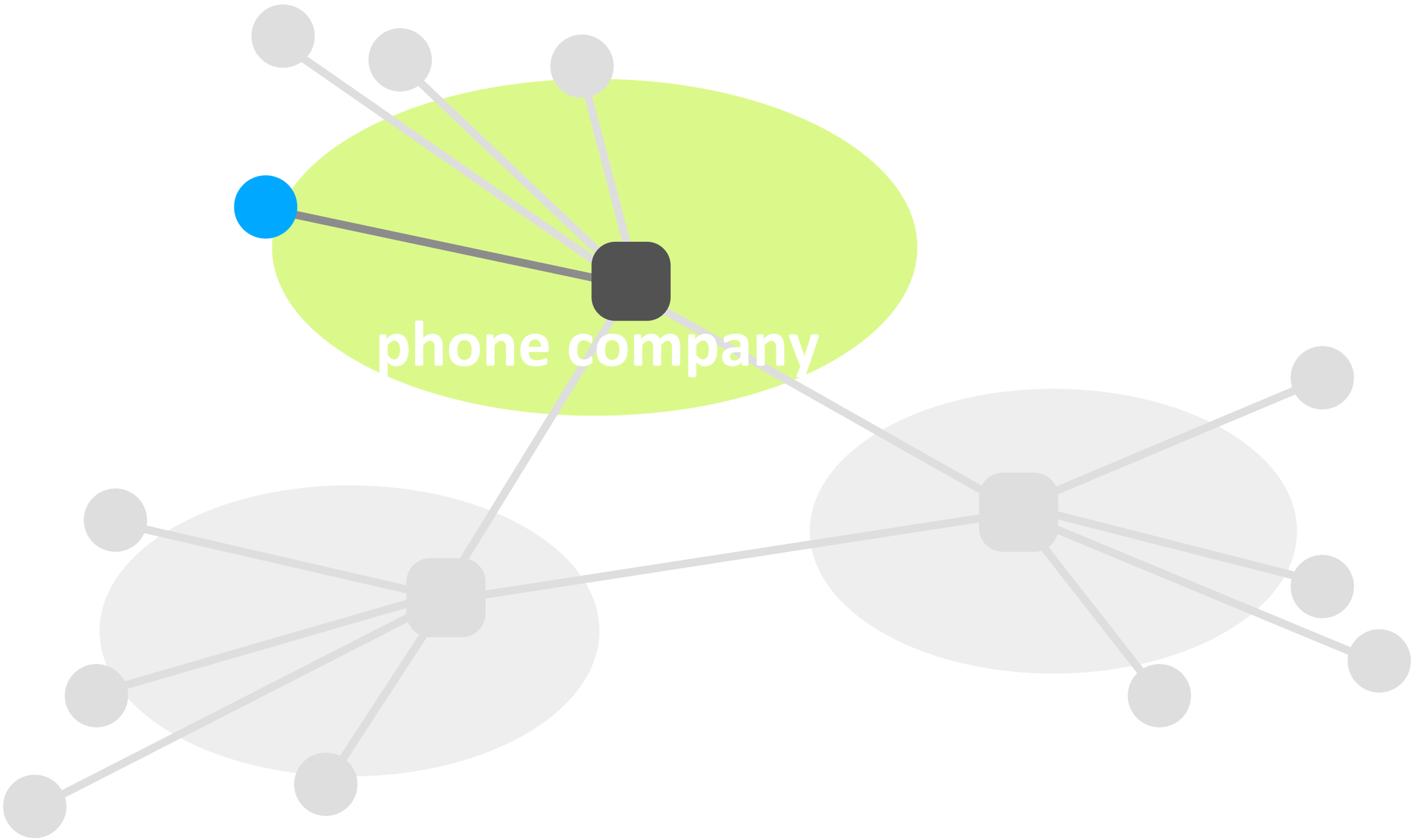
Administrivia

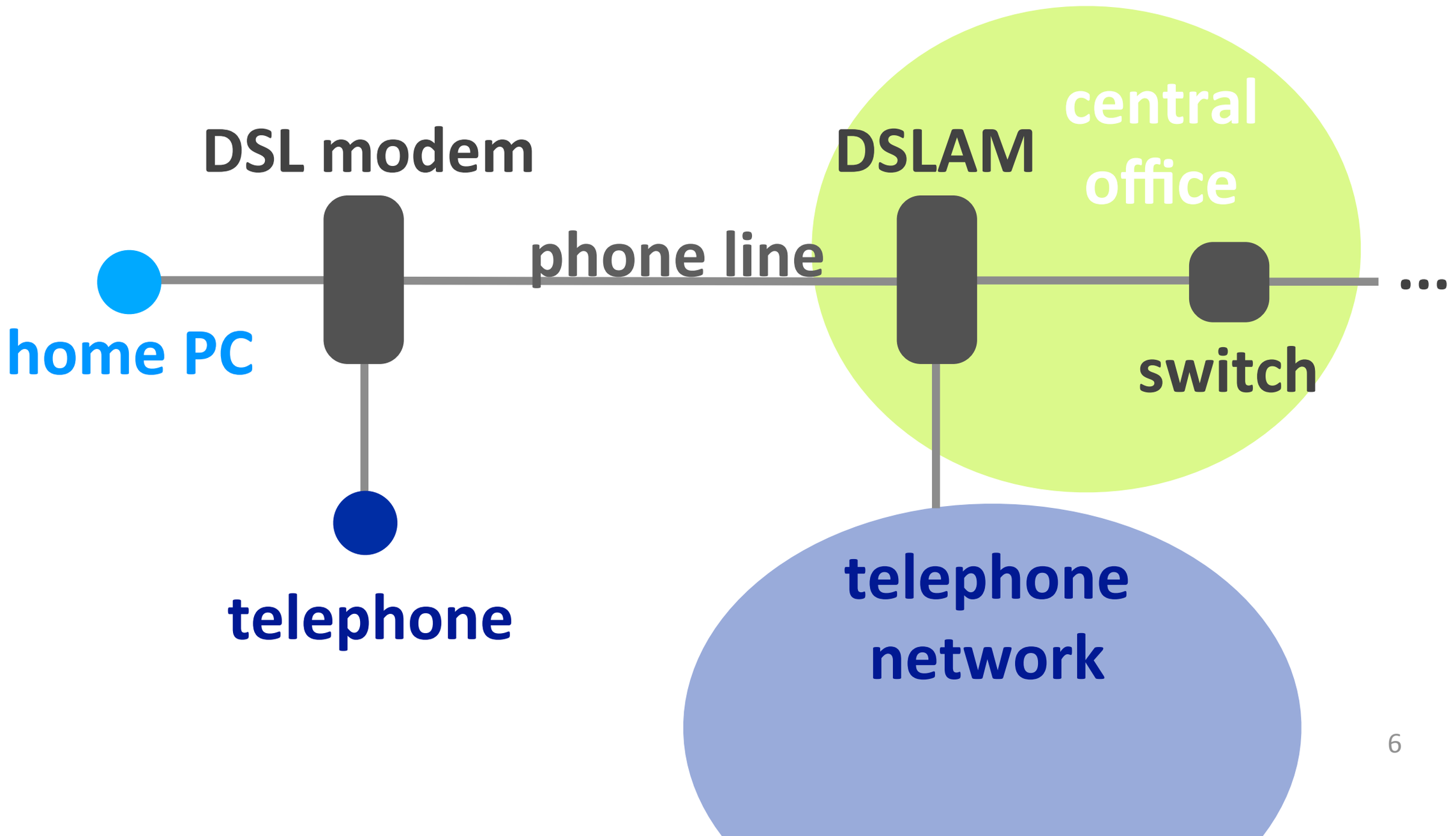
- Discussion sections start Monday, Sep 15
- Homework#1 out tonight
  - *announce homework logistics*

Plan of attack

- What is a network made of?
- How is it shared?
- How do we evaluate a network?
- How is communication organized?


link
end-system
switch
Internet Service Provider


phone company


DSL modem
DSLAM
central office
phone line
...
home PC
switch
telephone
telephone network

Two approaches to sharing

- Reservations → circuit switching
- On demand → packet switching

Circuit vs. Packets

- Circuits
  - *predictable performance*
  - *inefficient use of network resources*
  - *complex (state in the network)*

- Packets
  - *unpredictable performance*
  - *efficient use of network resources*
  - *simple (no state in the network)*

- *What physical infrastructure is already available?*
- *Reserve or on-demand?*

Today

- What is a network made of?
- How is it shared?
- How do we evaluate a network?
- How is communication organized?

Performance Metrics

- Delay
- Loss
- Throughput

Delay

- *How long does it take to send a packet from its source to destination?*

Delay

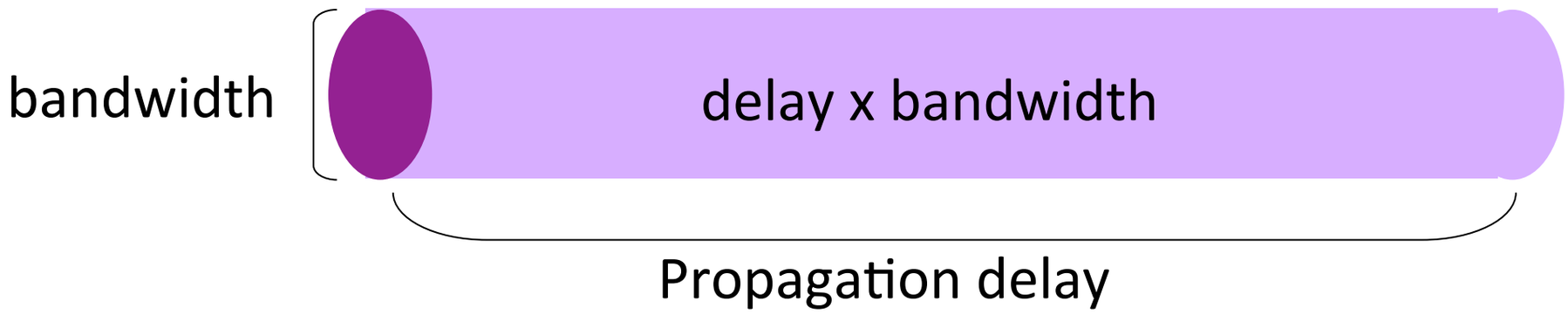
- Consists of four components
  - *transmission delay*
  - *propagation delay*

  } *due to link properties*

  - *queuing delay*
  - *processing delay*

  } *due to traffic mix and switch internals*

A network link



- Link bandwidth
  - number of bits sent/received per unit time (bits/sec or bps)
- Propagation delay
  - time for one bit to move through the link (seconds)
- Bandwidth-Delay Product (BDP)
  - number of bits “in flight” at any time
  - BDP = bandwidth × propagation delay

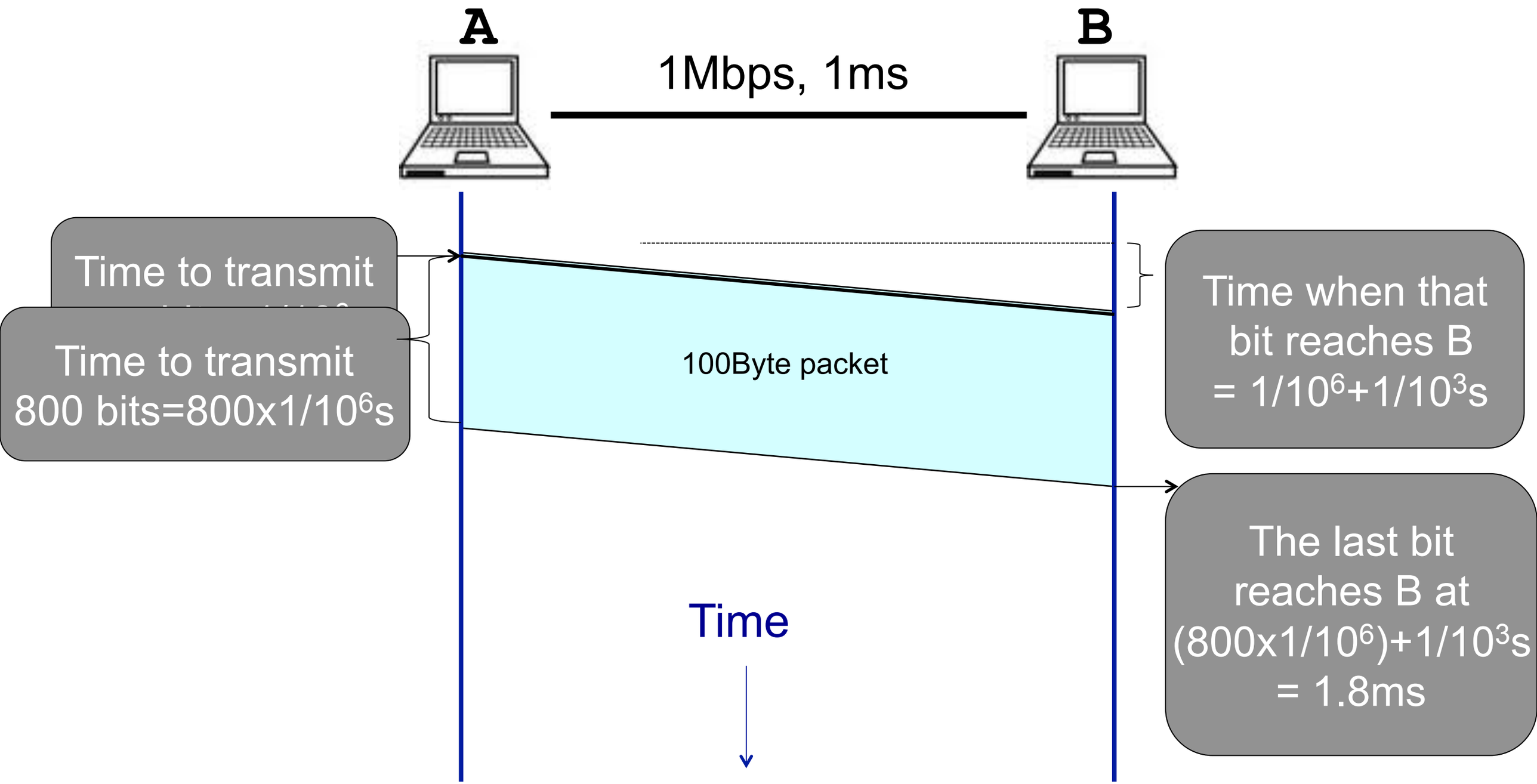
Examples

- Same city over a slow link:
  - bandwidth: ~100Mbps
  - propagation delay: ~0.1msec
  - BDP: 10,000bits (1.25KBytes)

- Cross-country over fast link:
  - bandwidth: ~10Gbps
  - propagation delay: ~10msec
  - BDP: 10^8bits (12.5MBytes)

Packet Delay

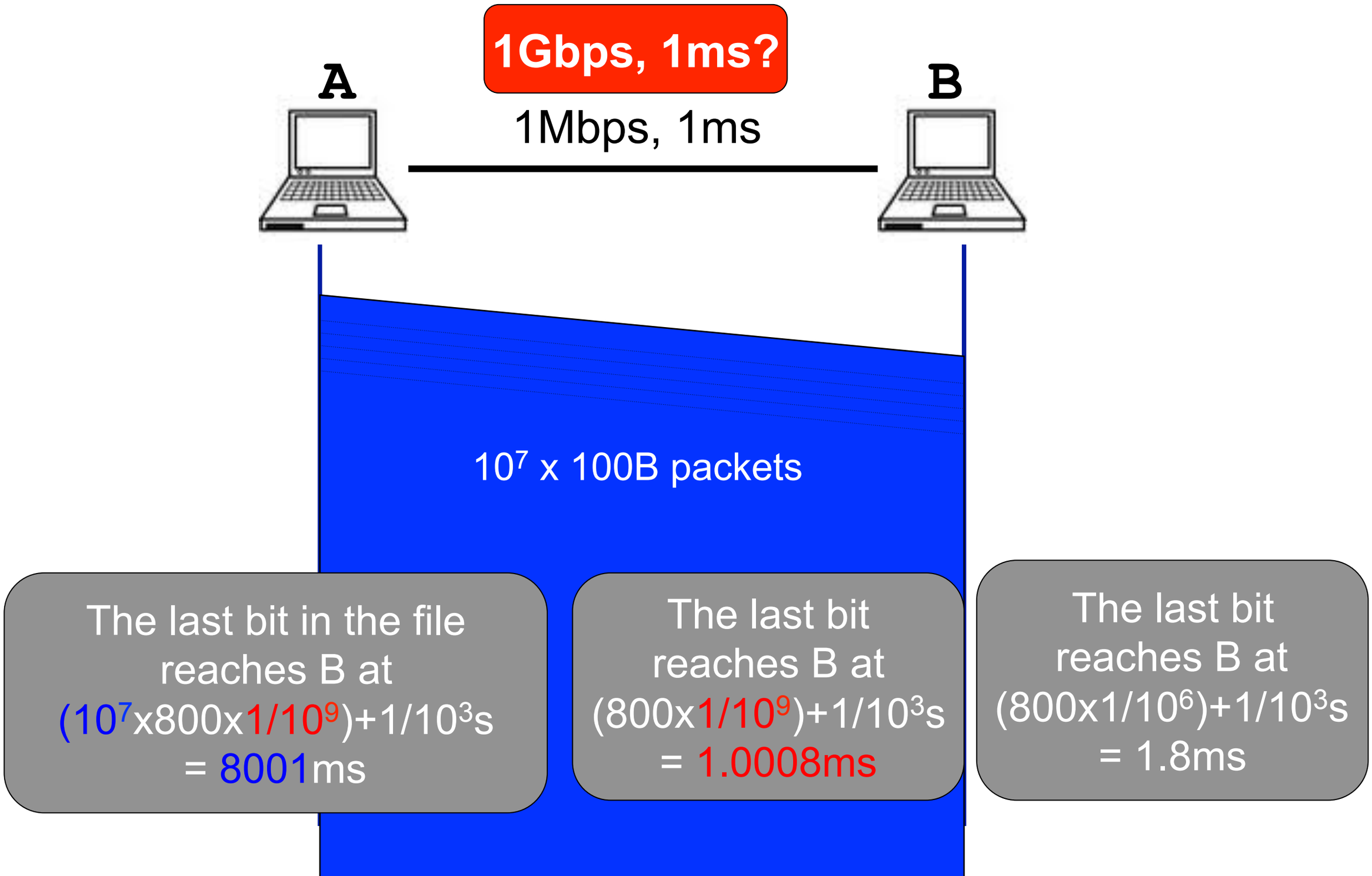
Sending 100B packets from A to B?



1GB file in 100B packets

ay

Sending 100B packets from A to B?



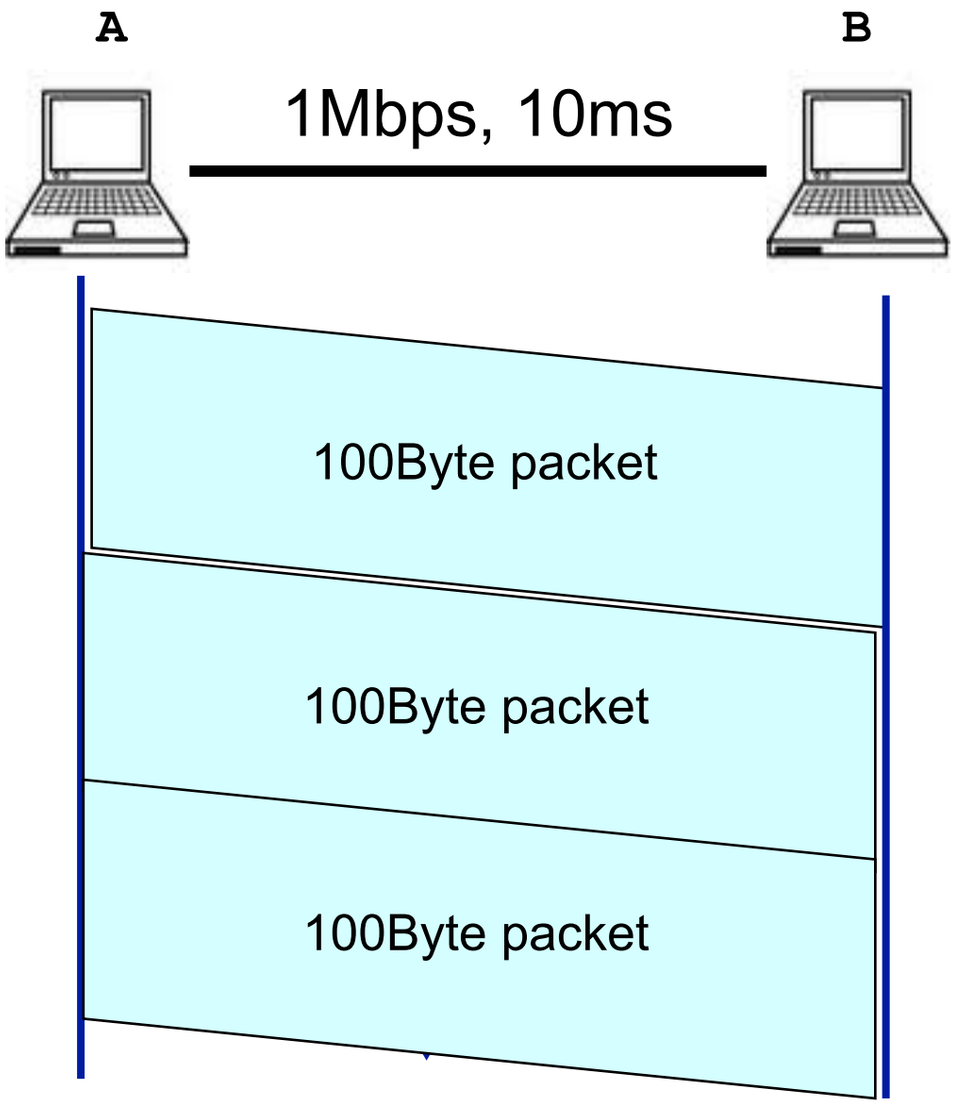
The last bit in the file reaches B at $(10^7 \times 800 \times 1/10^9)+1/10^3$s = 8001ms

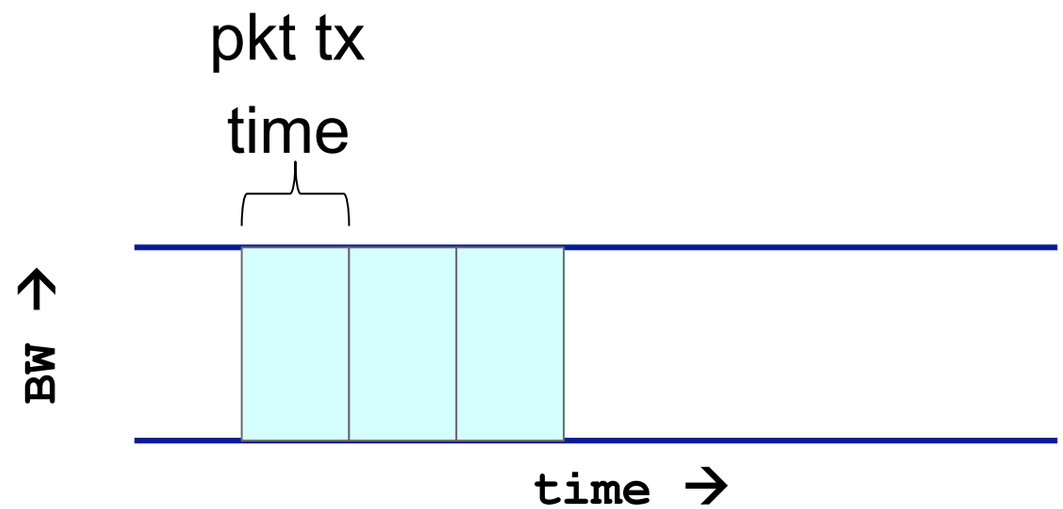
The last bit reaches B at $(800 \times 1/10^9)+1/10^3$s = 1.0008ms

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

Packet Delay: The "pipe" view

Sending 100B packets from A to B?

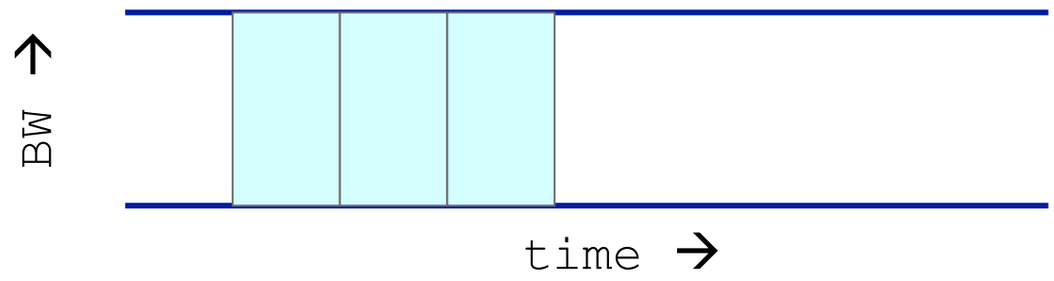




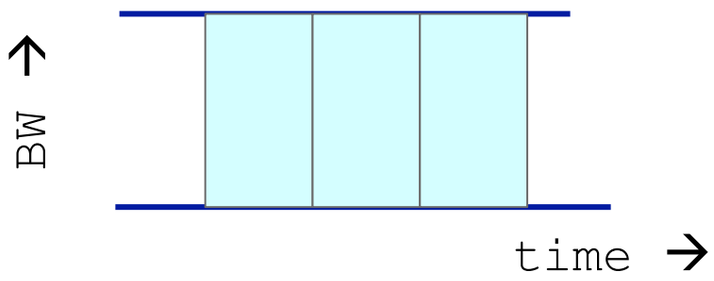
Packet Delay: The "pipe" view

Sending 100B packets from A to B?

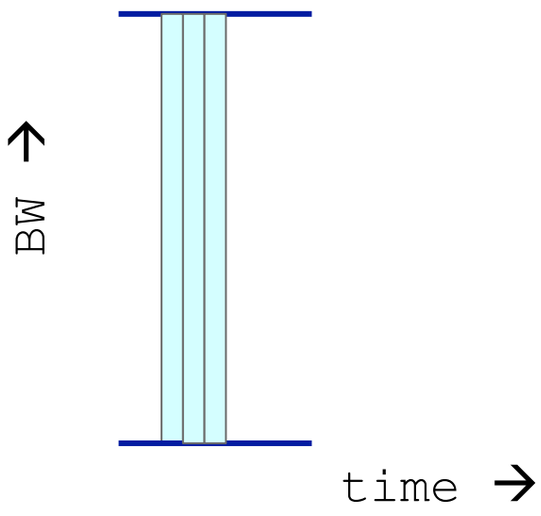
1Mbps, 10ms (BDP=10,000)



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

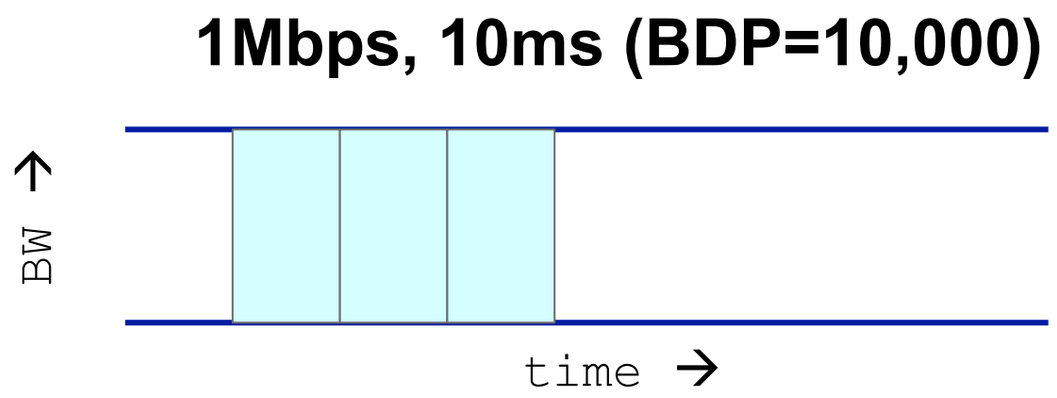


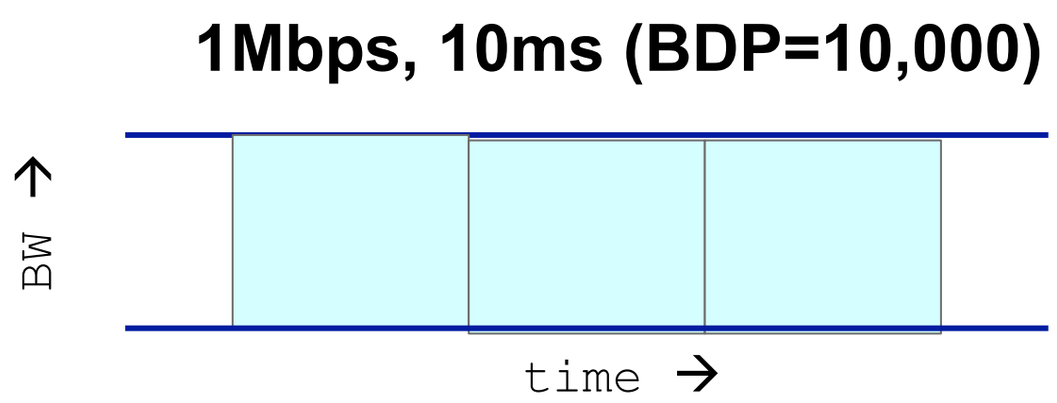
10Mbps, 1ms (BDP=10,000)



Packe 200B?: The “pipe” view

Sending 100B packets from A to B?





1. Transmission delay

- How long does it take to push all the bits of a packet into a link?
- Packet size / Link bandwidth
  - *e.g. 1000 bits / 100 Mbits per sec = 10^{-5} sec*

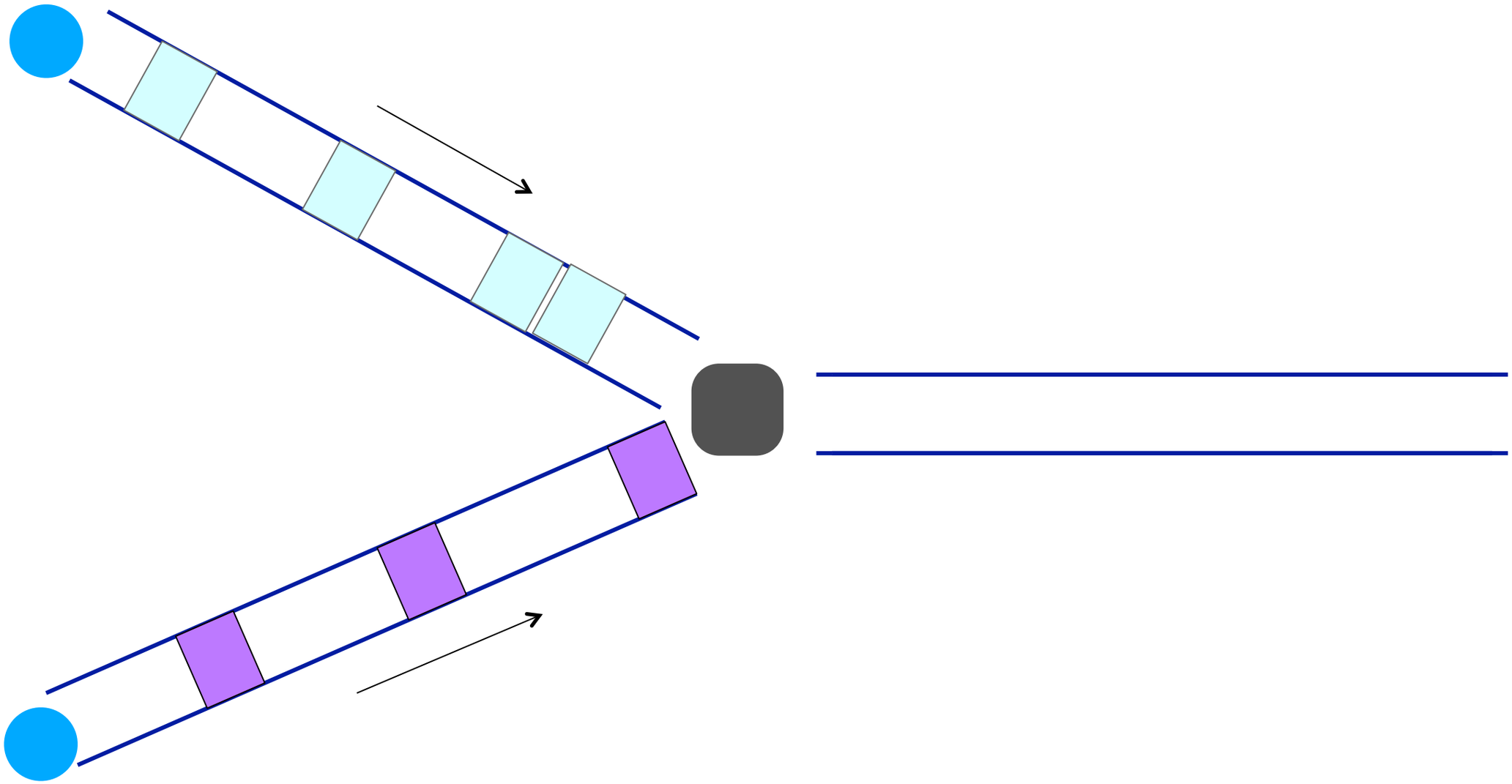
2. Propagation delay

- How long does it take to move one bit from one end of a link to the other?
- Link length / Link propagation delay
  - *E.g. 30 kilometers / 3 10^8 meters per sec = 10^{-4} sec*

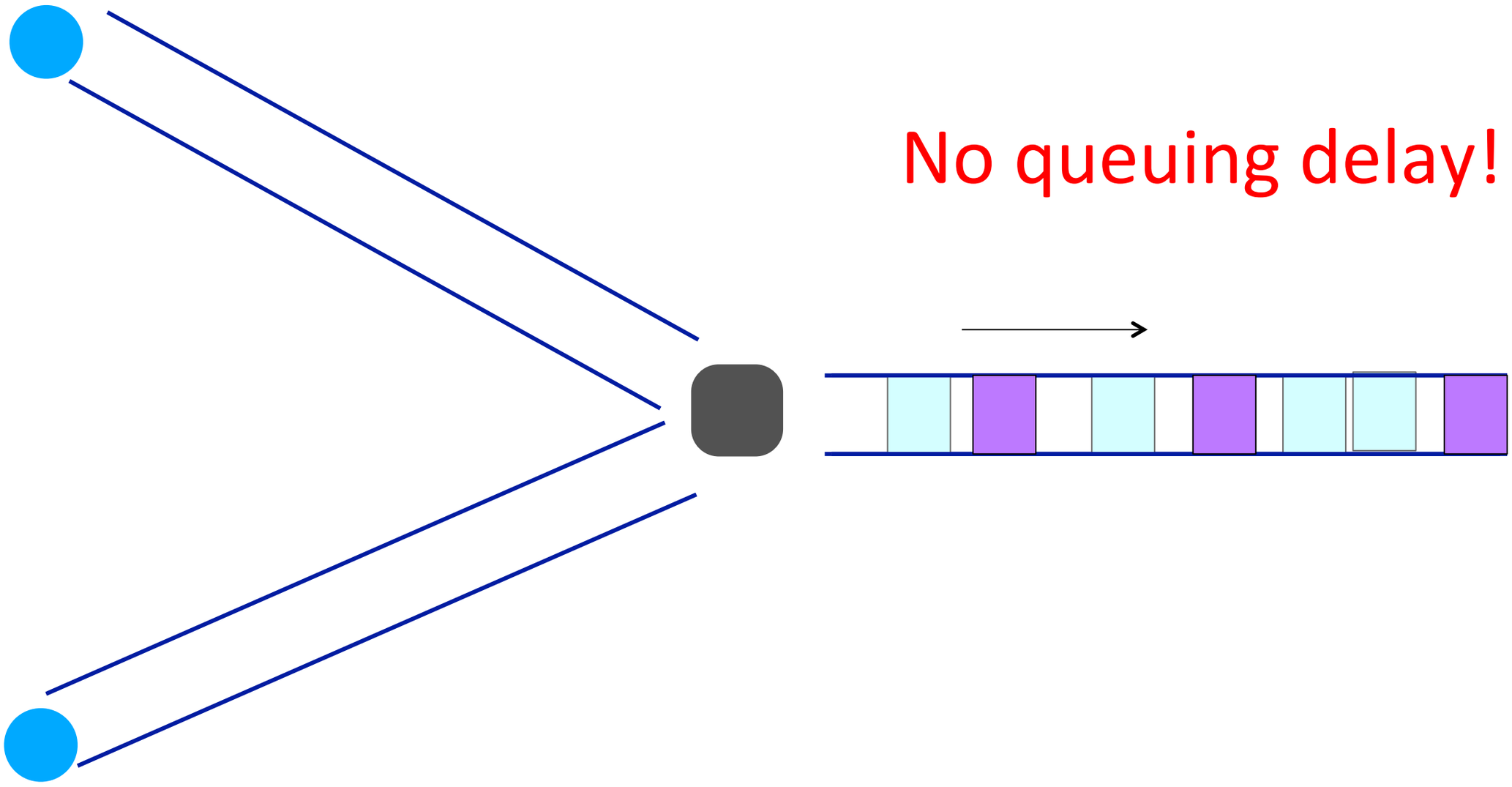
3. Queuing delay

- *How long does a packet have to sit in a buffer before it is processed?*

Queuing delay: “pipe” view

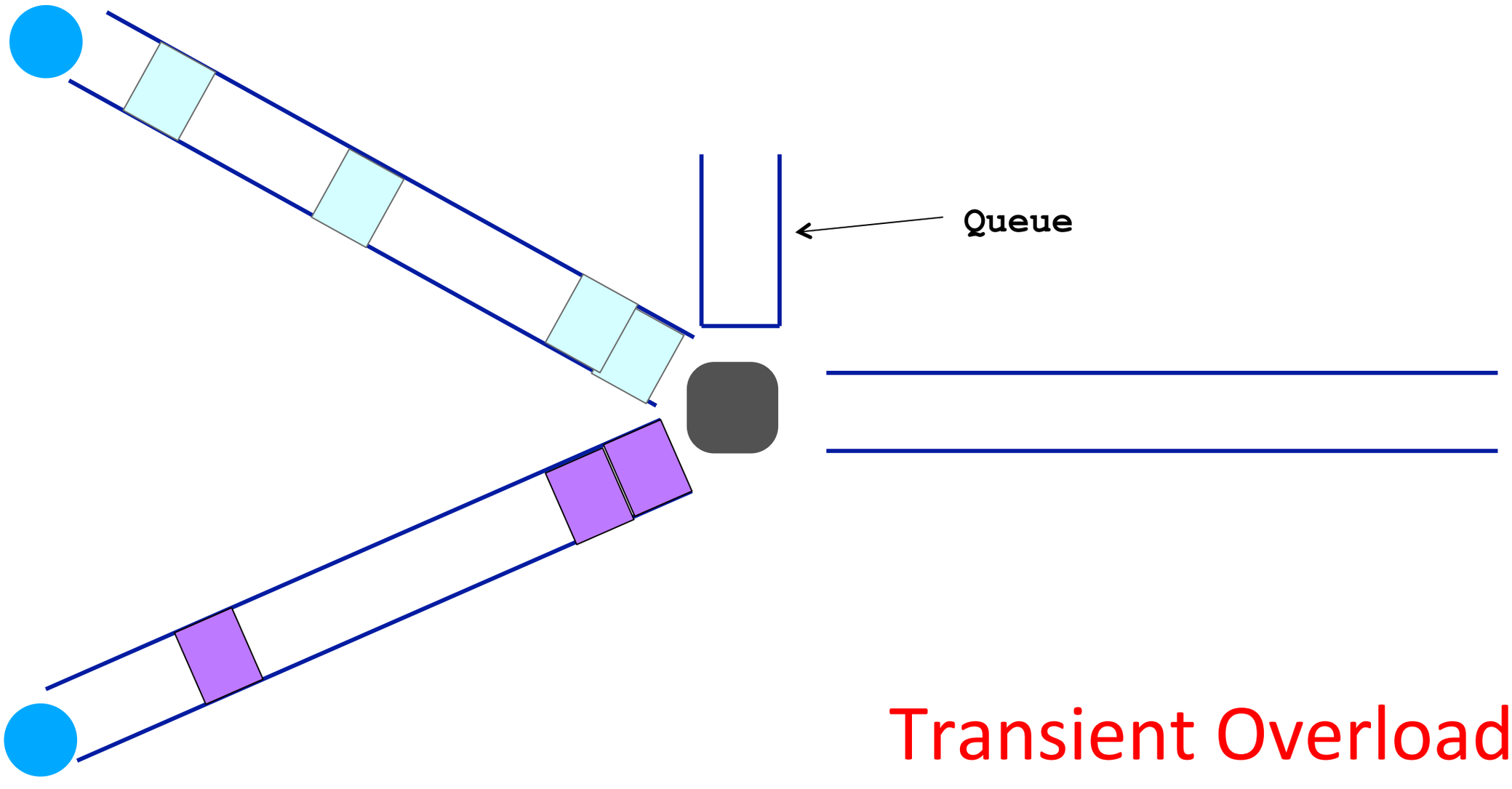


Queuing delay: “pipe” view



No queuing delay!

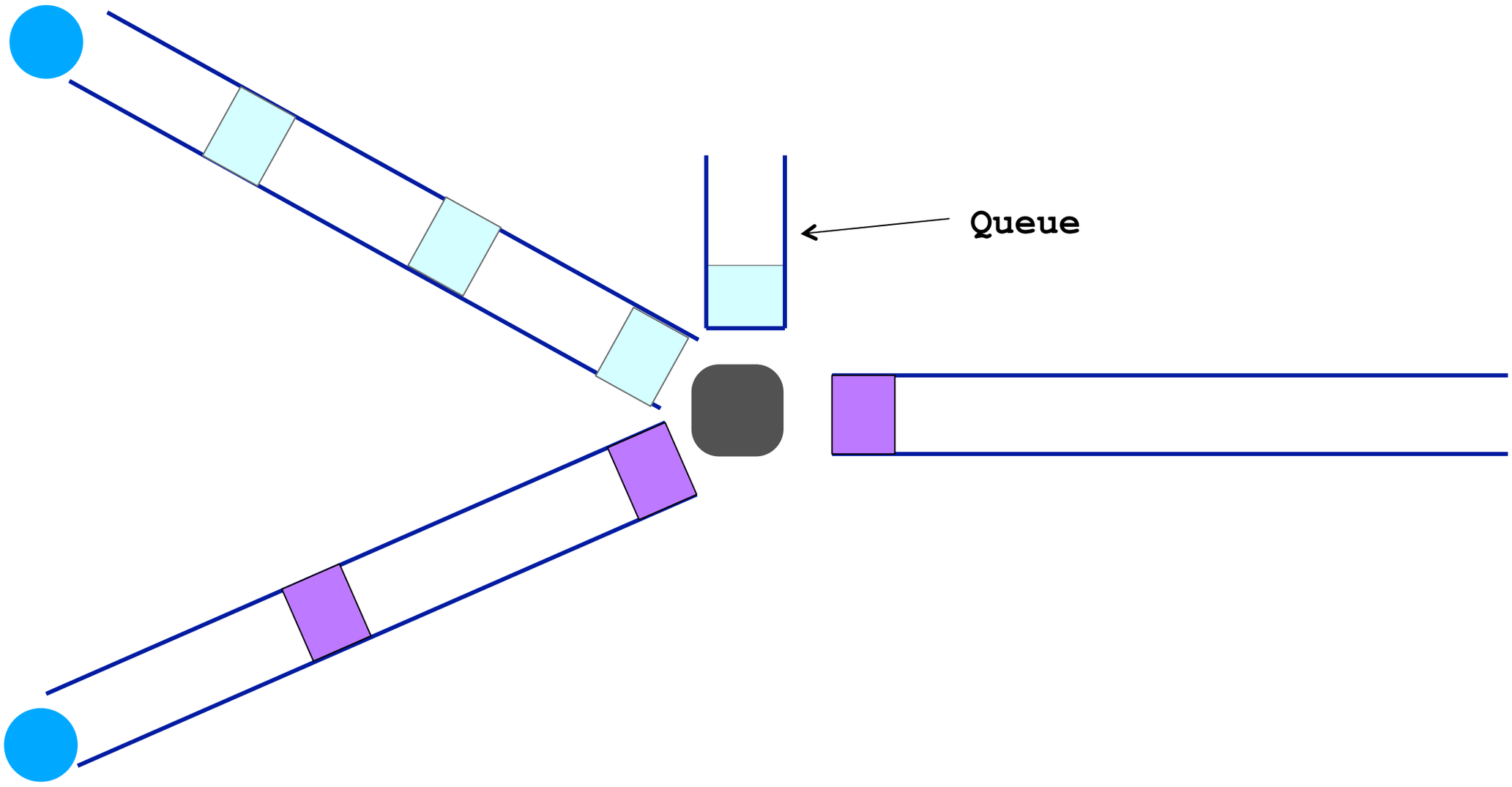
Queuing delay: “pipe” view



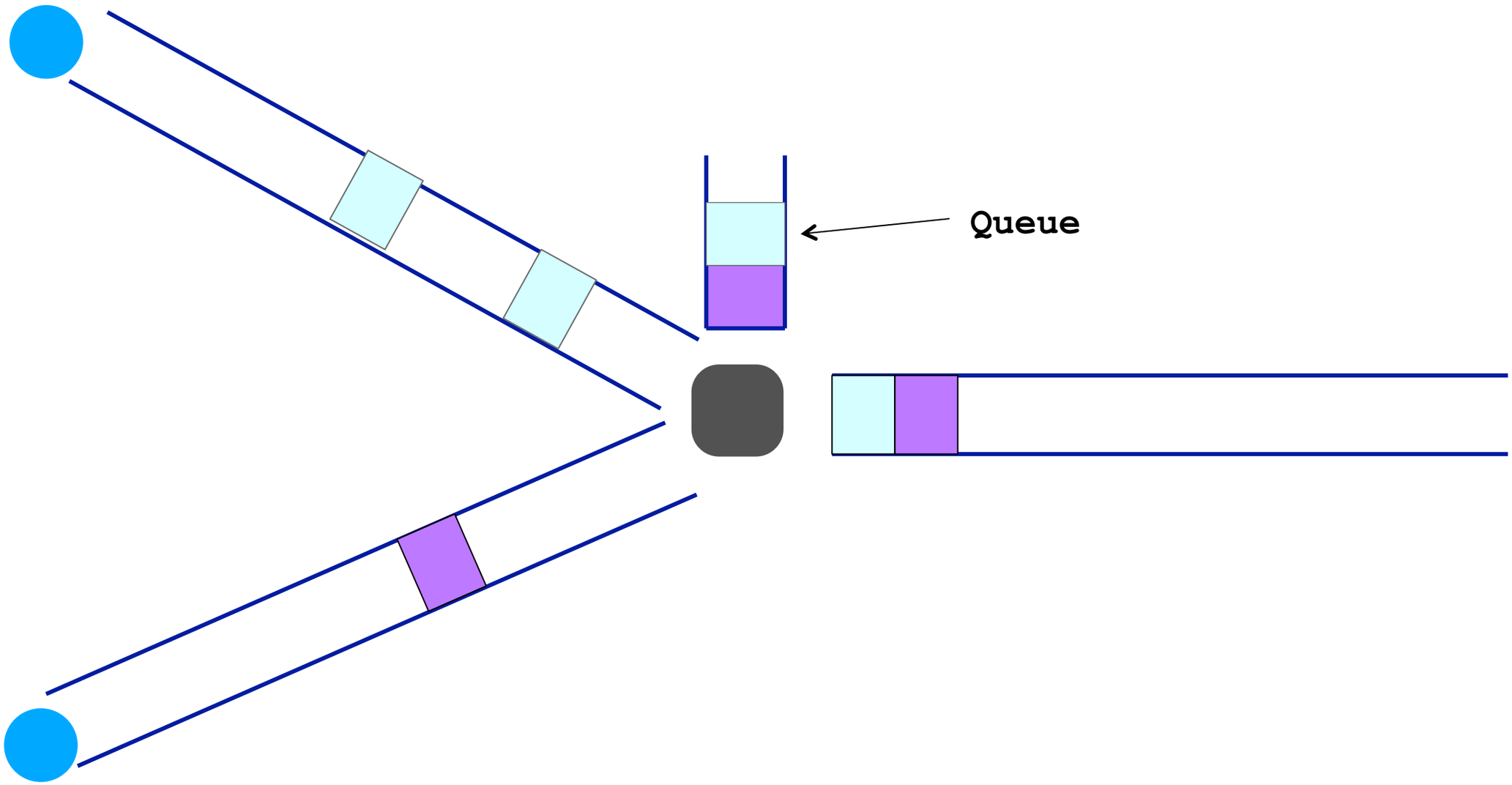
Transient Overload

Not a rare event!

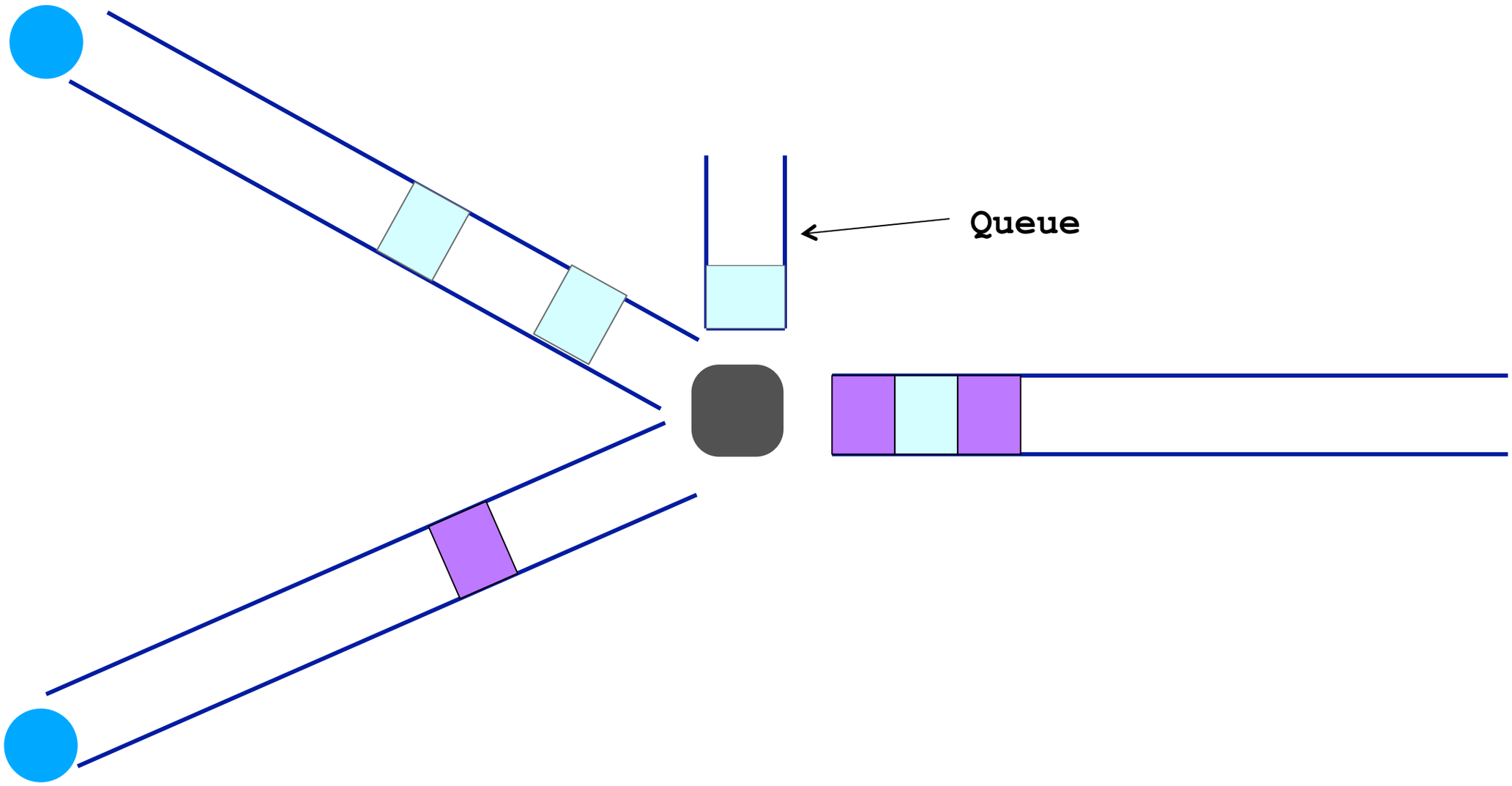
Queuing delay: “pipe” view



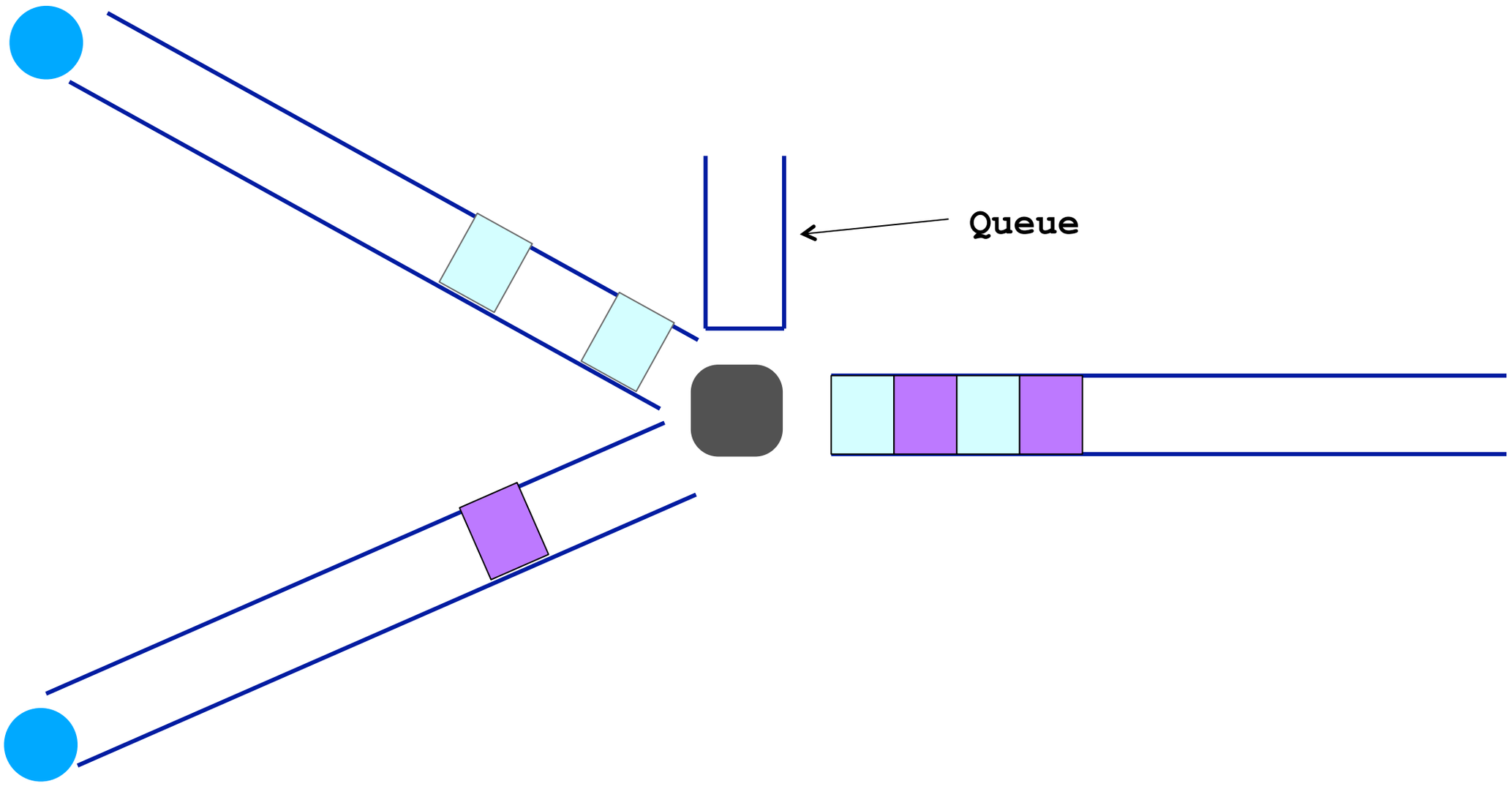
Queuing delay: “pipe” view



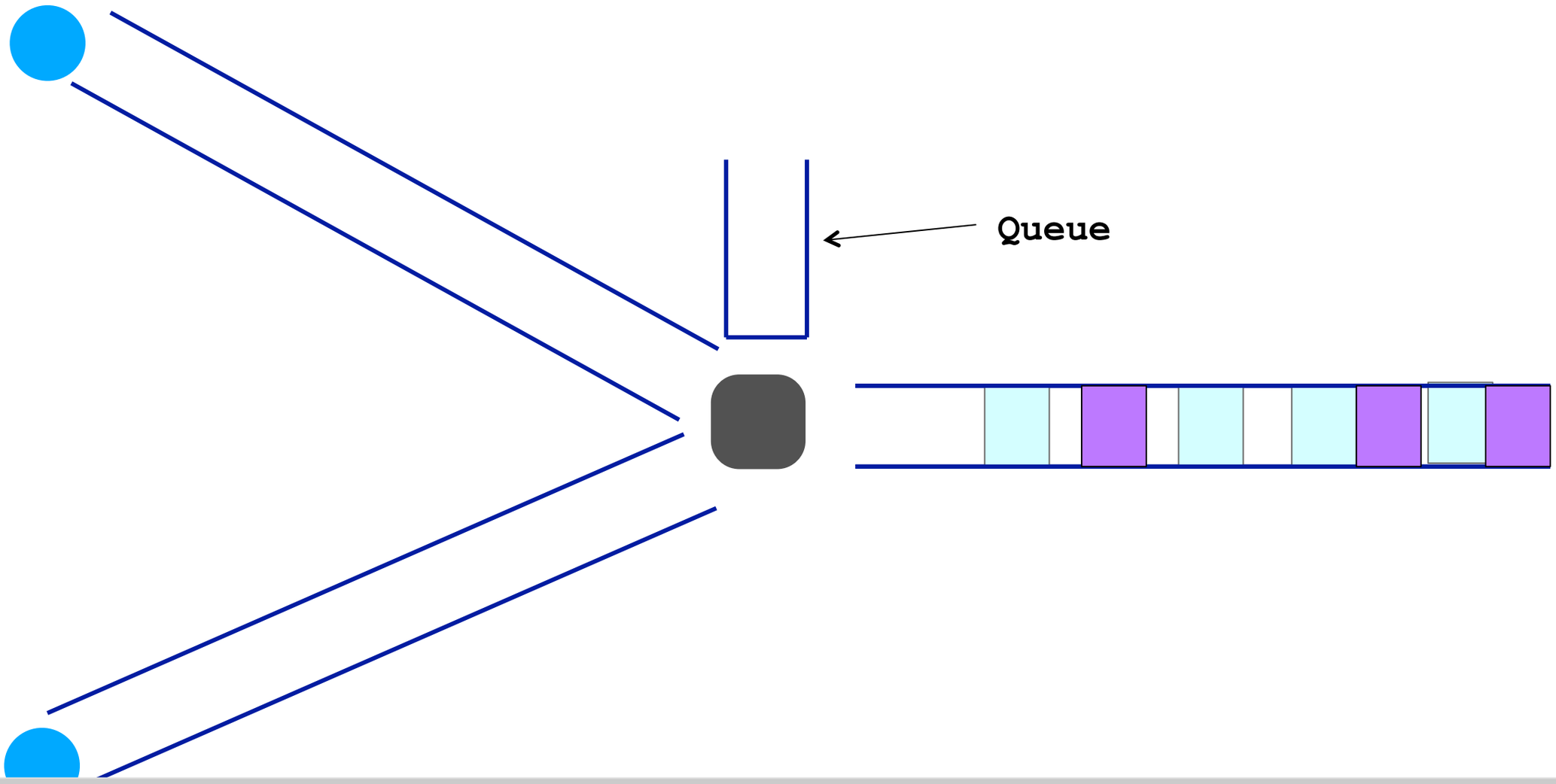
Queuing delay: “pipe” view



Queuing delay: “pipe” view

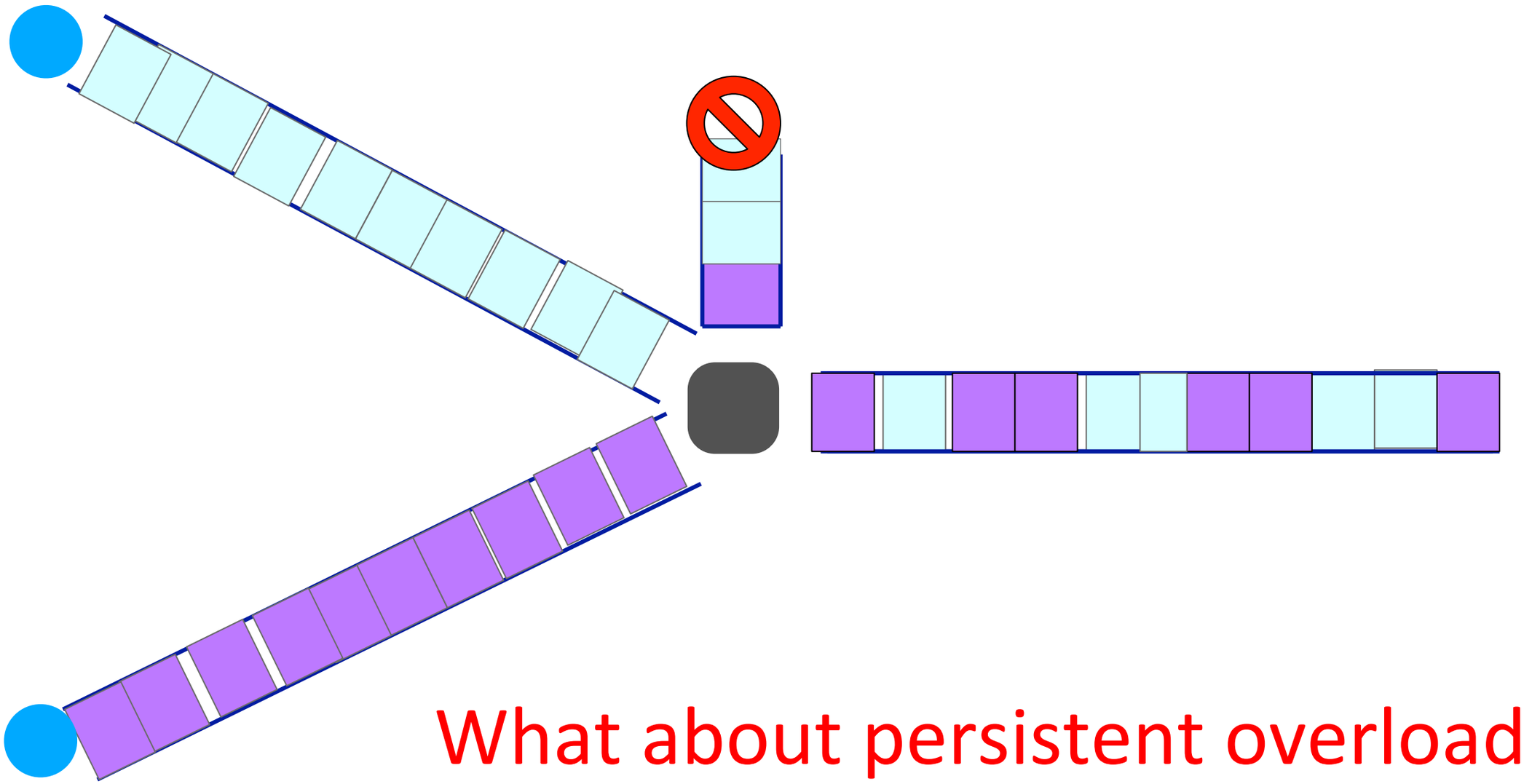


Queuing delay: “pipe” view



Queues absorb transient bursts but introduce queuing delay

Queuing delay: “pipe” view

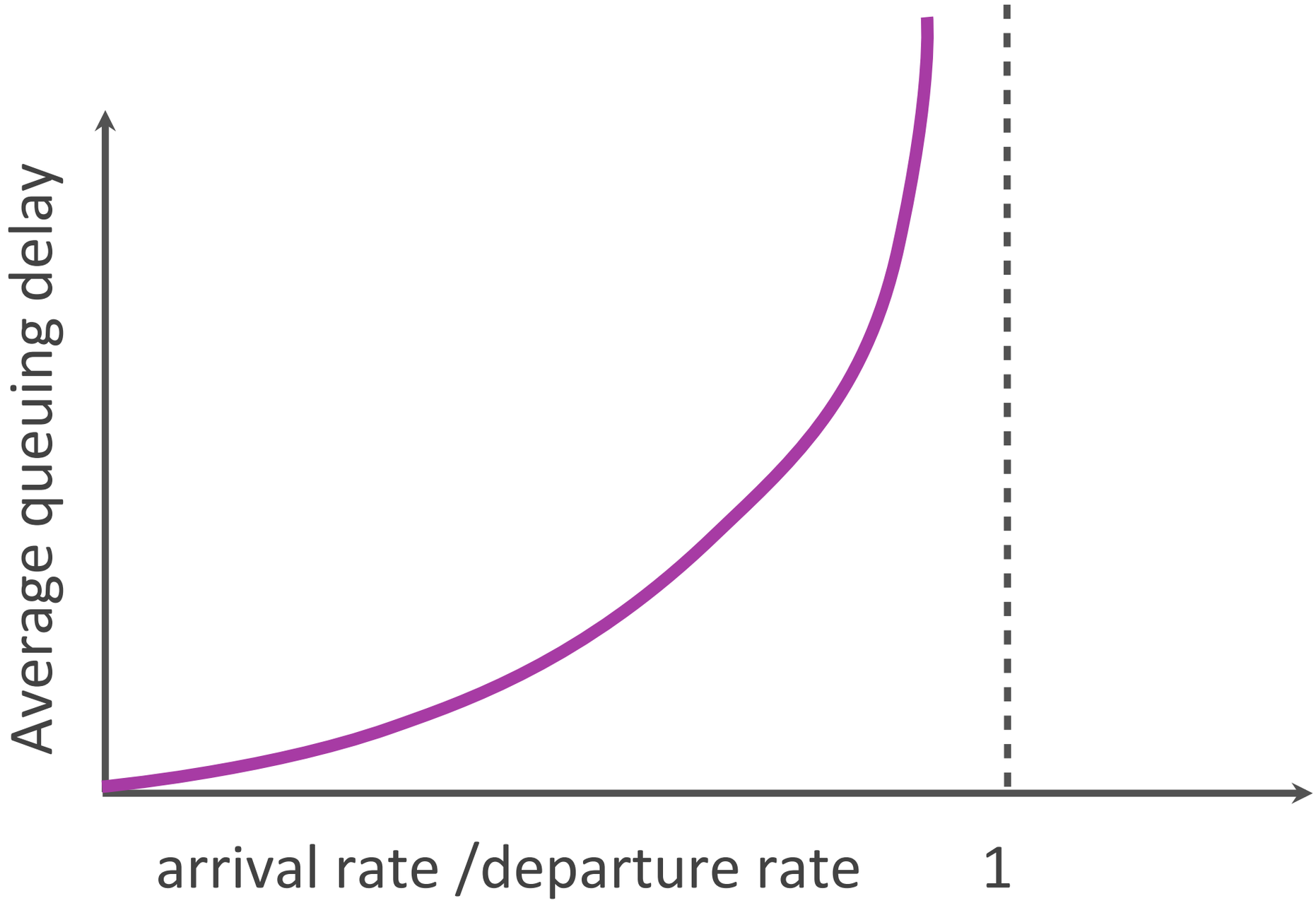


What about persistent overload?

Will eventually drop packets (“loss”)

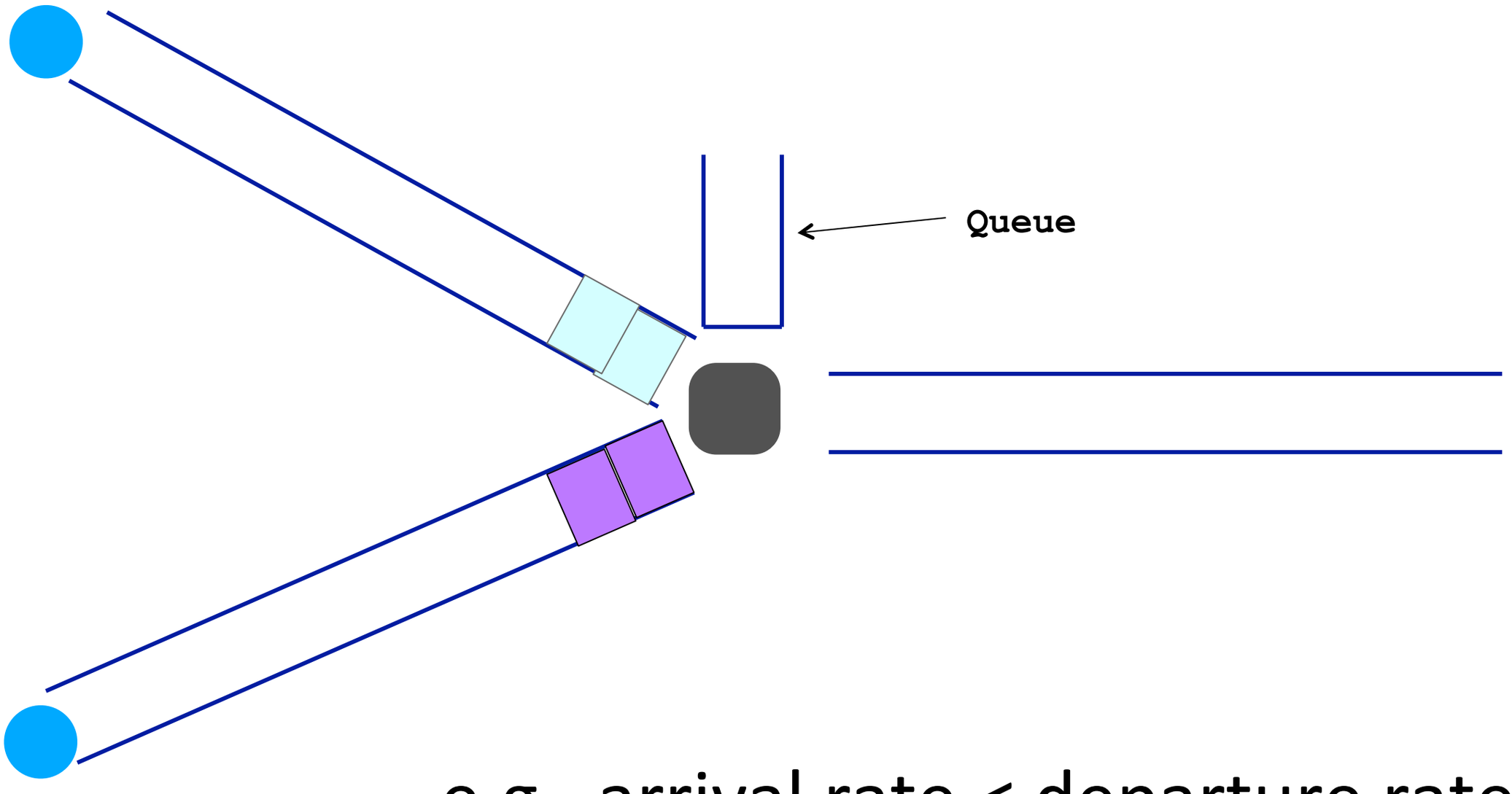
Queuing delay

- If arrival rate > departure rate
  - *approaches infinity (assuming an infinite buffer)*


Average queuing delay
arrival rate /departure rate
1

Queuing delay

- If arrival rate > departure rate
  - *approaches infinity (assuming an infinite buffer)*
  - *in practice, finite buffer → loss*
- If arrival rate < departure rate



e.g., arrival rate < departure rate

Queuing delay

- If arrival rate > departure rate
  - *approaches infinity (assuming an infinite buffer)*
  - *in practice, finite buffer → loss*
- If arrival rate < departure rate
  - *depends on burst size*

Queuing Delay

- *How long does a packet have to sit in a buffer before it is processed?*
- Depends on traffic pattern

Queuing Delay

- *How long does a packet have to sit in a buffer before it is processed?*

- Depends on traffic pattern

- Characterized with statistical measures
  - average queuing delay
  - average arrival rate
  - average departure rate

Basic Queuing Theory Terminology

- Arrival process: how packets arrive
  - Average rate A
- W: average time packets wait in the queue
  - W for “waiting time”
- L: average number of packets waiting in the queue
  - L for “length of queue”

Little's Law (1961)

$$L = A \times W$$

- Compute L: count packets in queue every second
  - How often does a single packet get counted? W times

- Why do you care?
  - Easy to compute L, harder to compute W

4. Processing Delay

- *How long does the switch take to process a packet?*
  - typically assume this is negligible

Delay

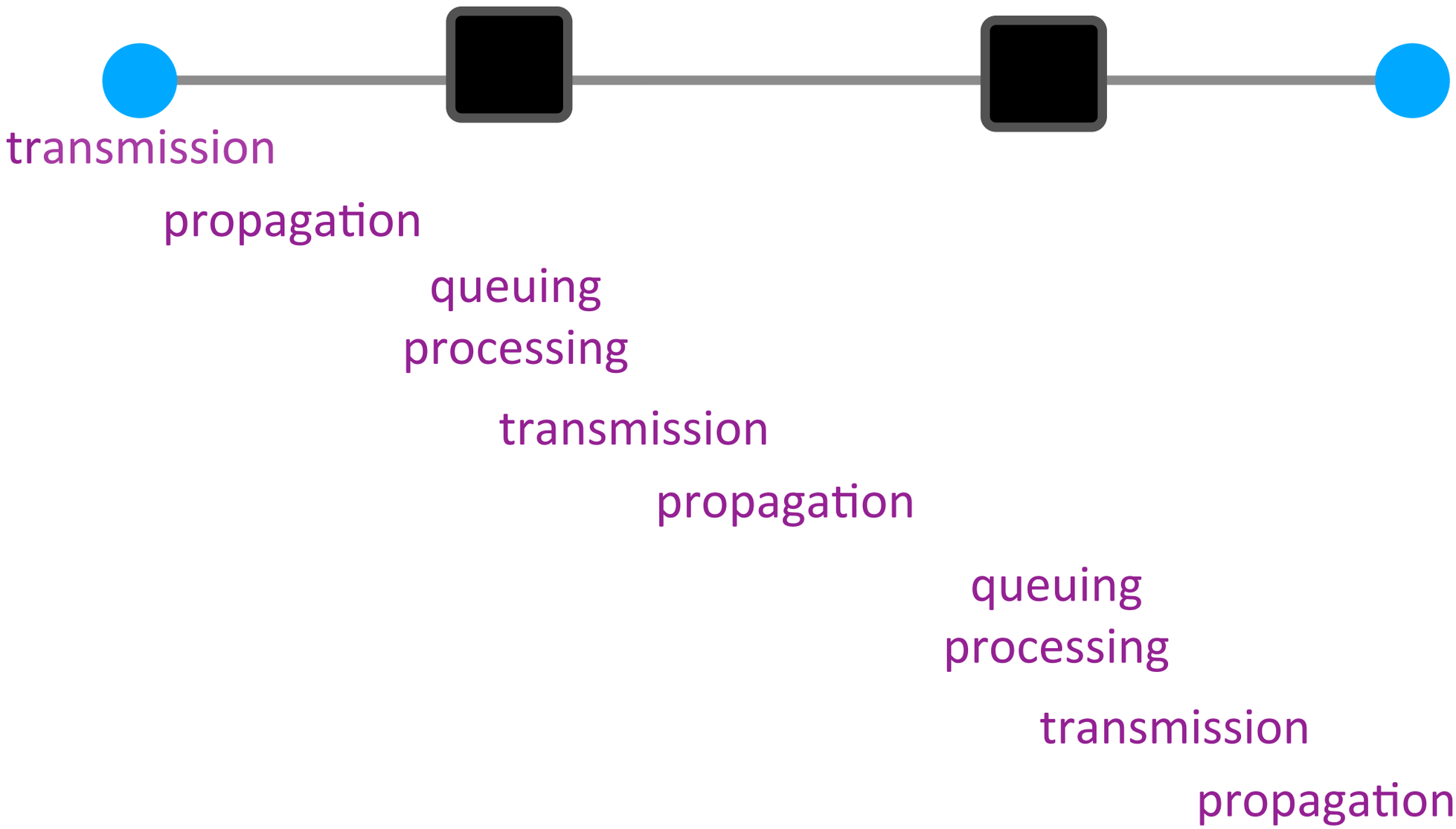
- Consists of four components
  - *transmission delay*
  - *propagation delay*

  } *due to link properties*

  - *queuing delay*
  - *processing delay*

  } *due to traffic mix and switch internals*

End-to-end delay

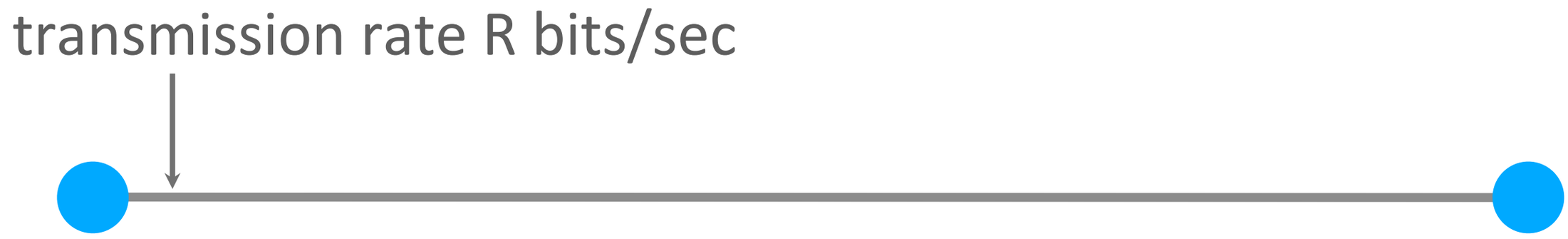


Loss

- *What fraction of the packets sent to a destination are dropped?*

Throughput

- *At what rate is the destination receiving data from the source*
  - Data size / transfer time

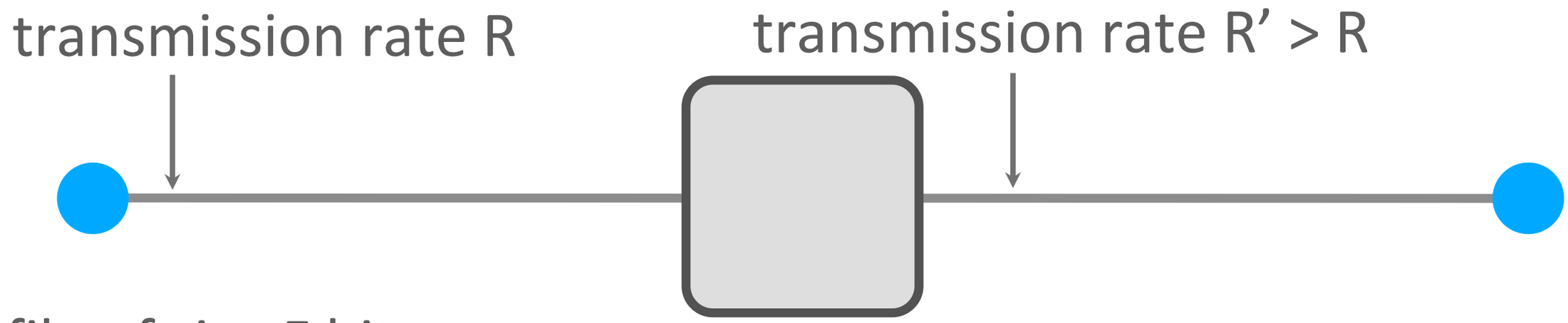


file of size F bits

packets of size L bits

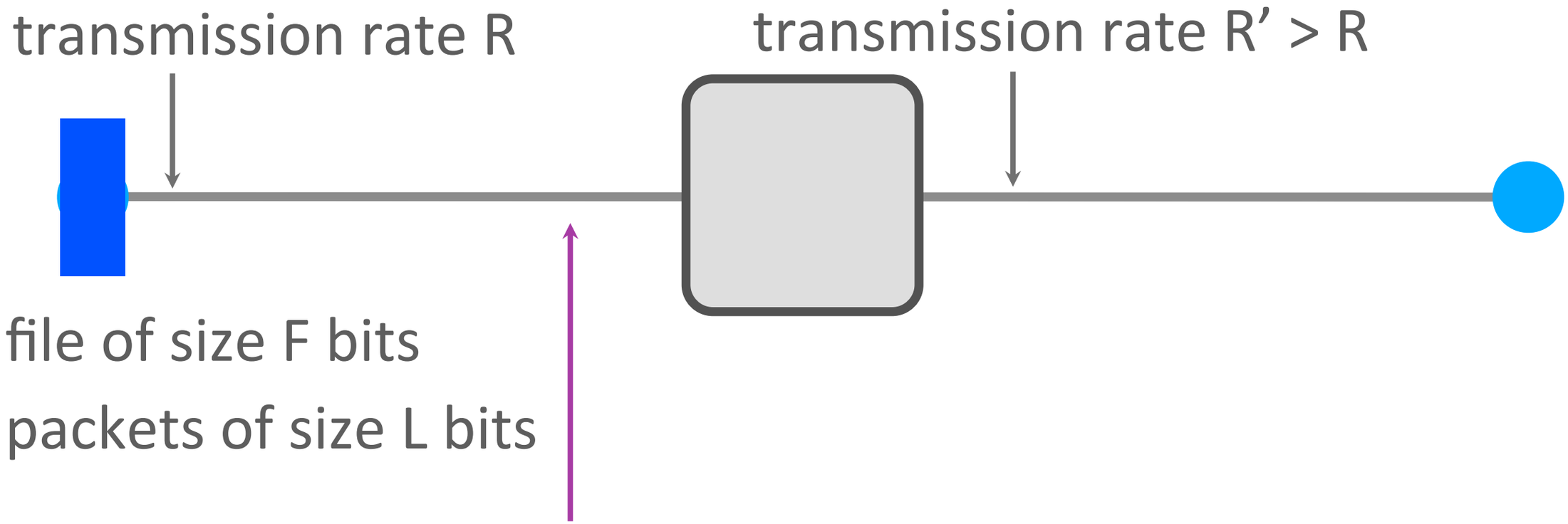
Transfer time = F/R + propagation delay

Average throughput = R



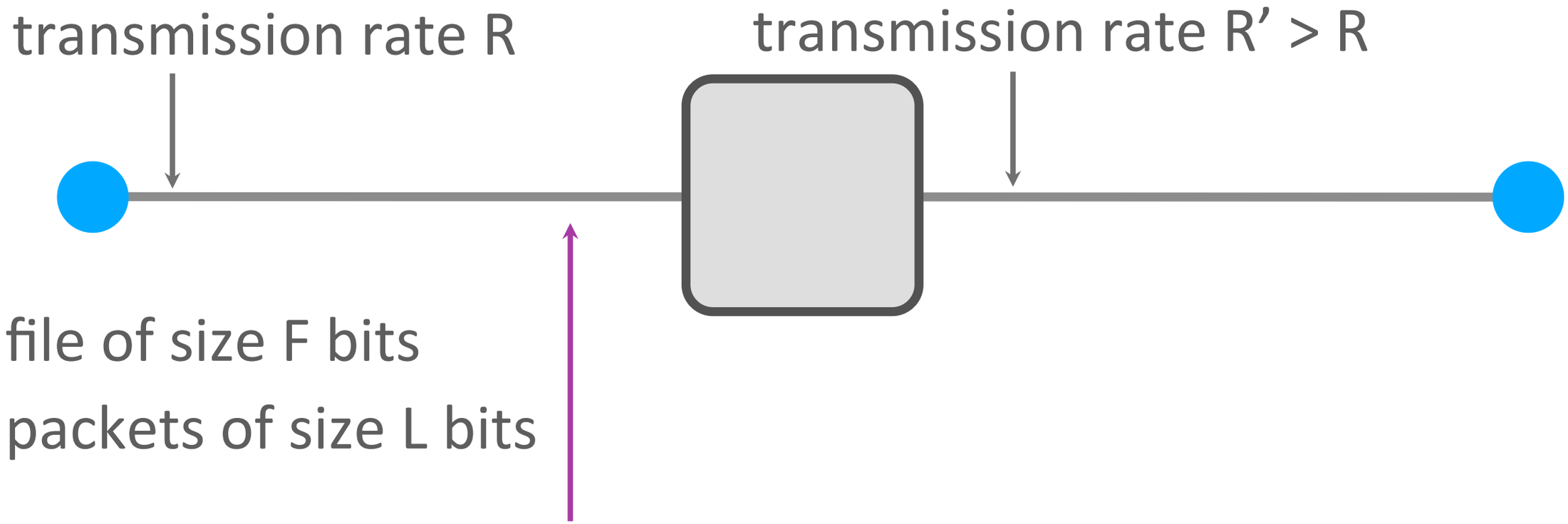
file of size F bits

packets of size L bits

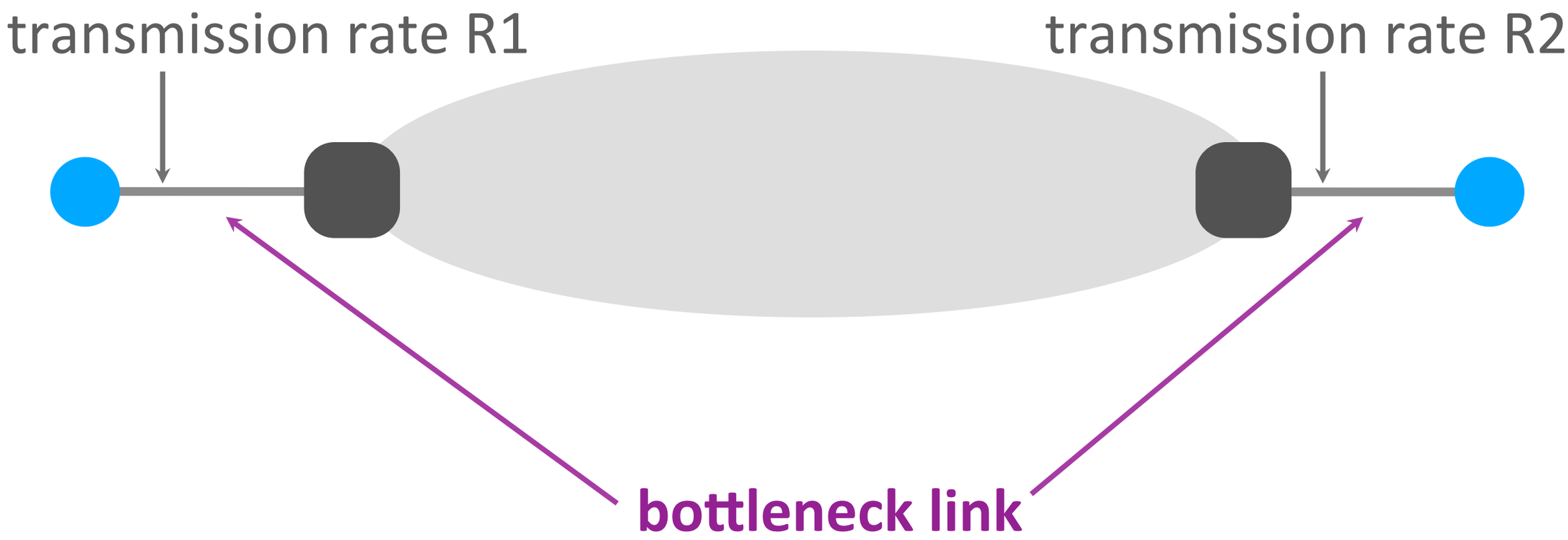


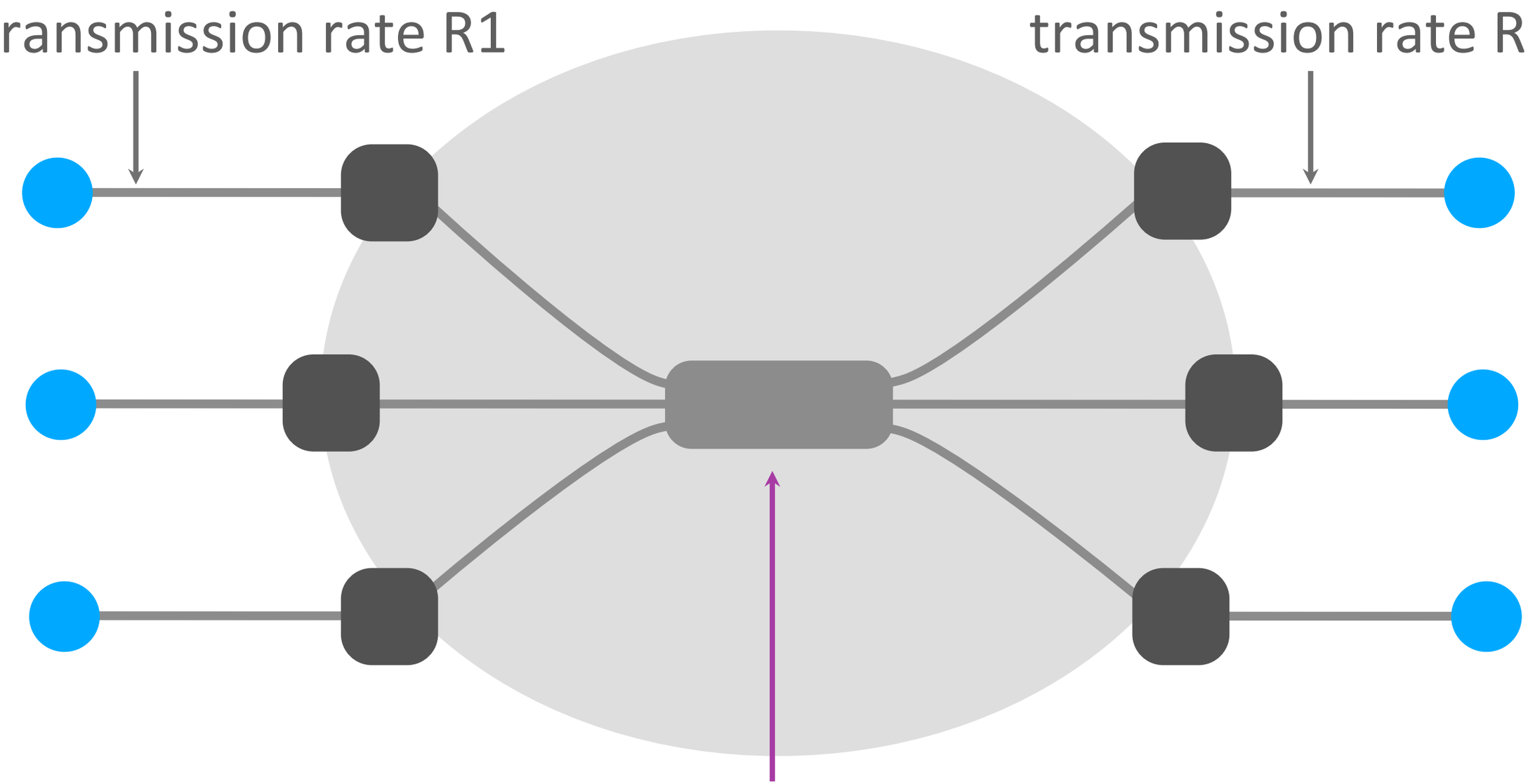
Transfer time = F/R + propagation delay + L/R’

Average throughput = min { R, R’ } = R



Average throughput = min { R, R’ } = R


transmission rate R1
transmission rate R2
bottleneck link


transmission rate R1
transmission rate R2
bottleneck link

Throughput

- At what rate is the destination receiving data from the source?

- Later in the semester
  - *TCP throughput, application-level throughput, etc.*
  - *throughput vs. "goodput"*

- *What physical infrastructure is already available?*
- *Reserve or on-demand?*
- *Where's my delay coming from?*