Performance metrics

Technology convergence

Originally:

- Telephone networks for telephone calls
- Internet for file transfer & email
- Cable TV networks for broadcast video
- Cell phone networks for telephone calls

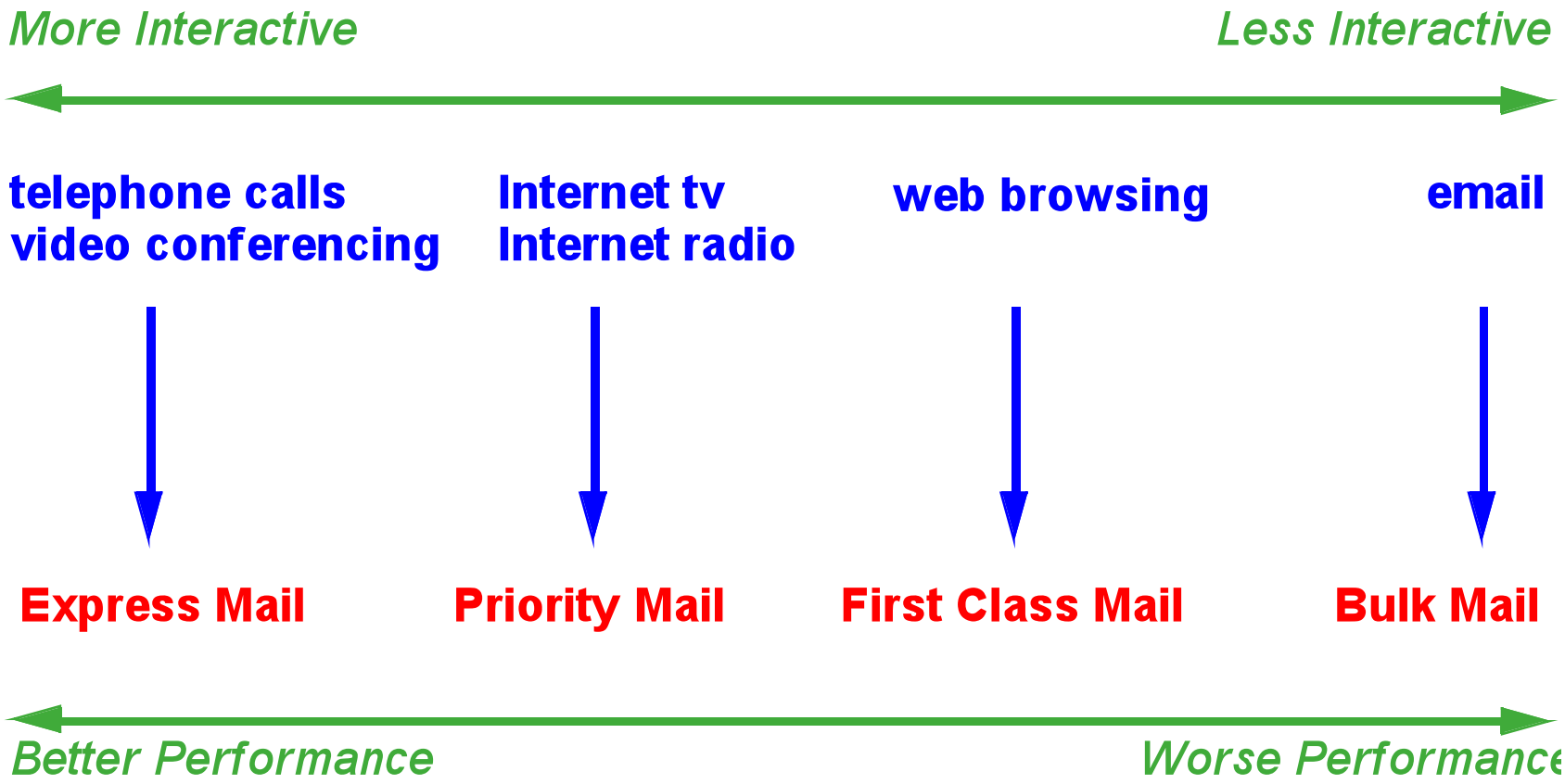
Now:

- telephone networks: telephone + Internet + video
- cable networks: video + Internet + telephone
- cellular networks: telephone + Internet + video

Convergence:

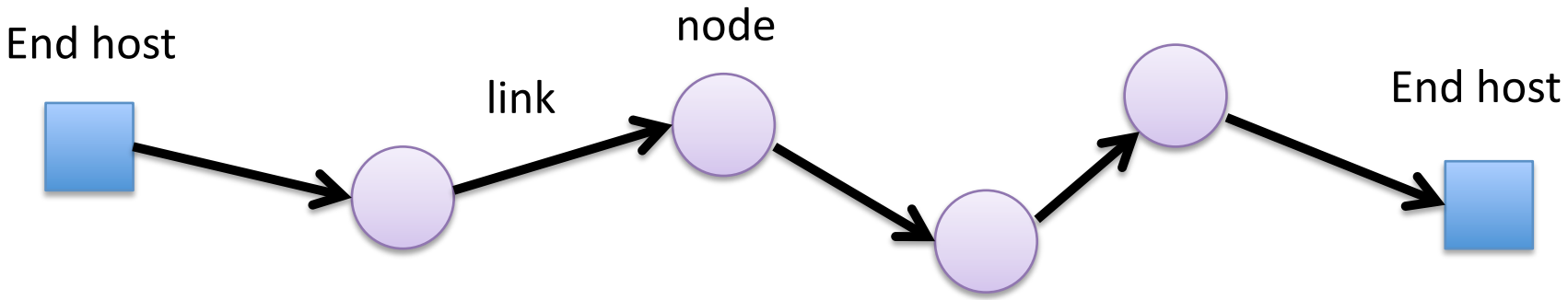
- email, web browsing, streaming, telephone calls, video chat, gaming, …
- over all networks

Technology convergence



Different applications may have different Quality of Service (QoS) constraints

Packet Switching Performance metrics



End-to-end performance:

- Multi-hop
- Link (capacity, physical length, etc.)
- Node (processing speed, buffer)
- Topology
- Active end hosts

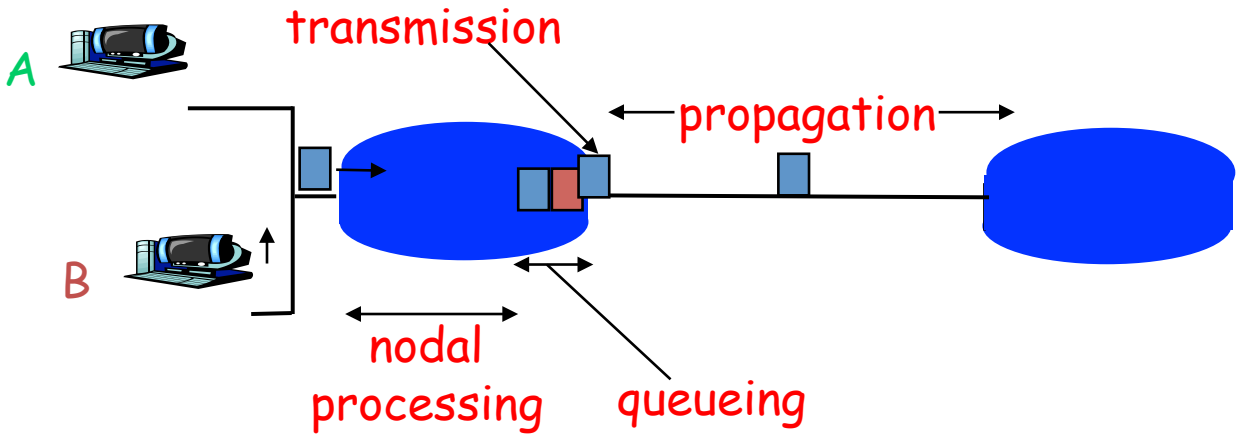
Packet Switching Performance metrics

- Packet Delay
- Packet Loss
- Throughput
- (Efficiency)

Performance metrics are interdependent

e.g.: no delay constraint = packet loss probability close to zero

Delay (node)



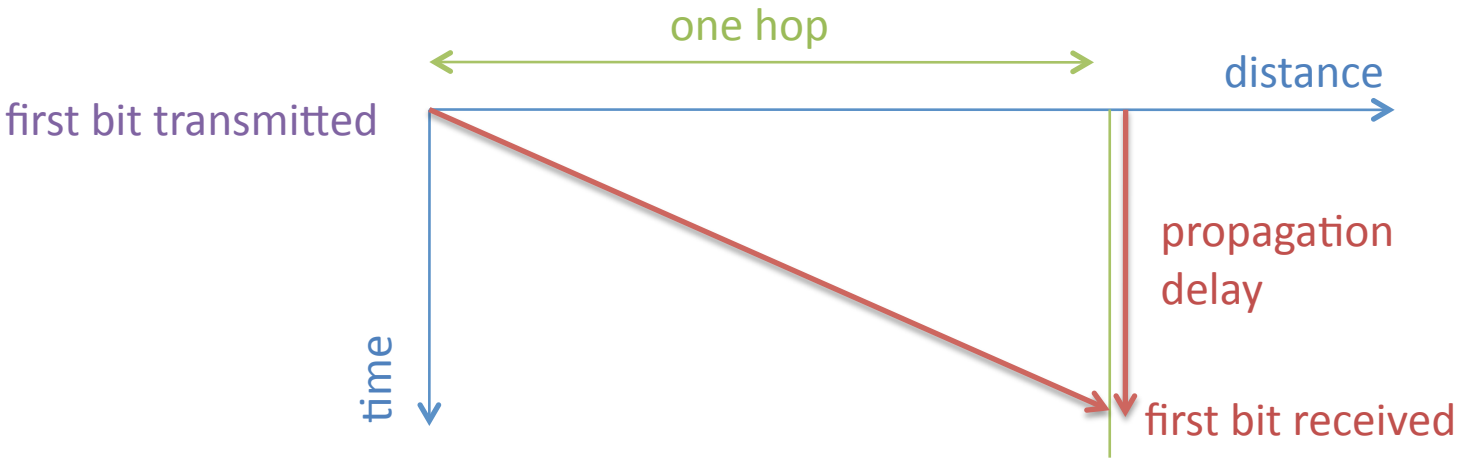
$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Kurose fig. 1.16

Processing delay

Defined by the hardware

Propagation delay

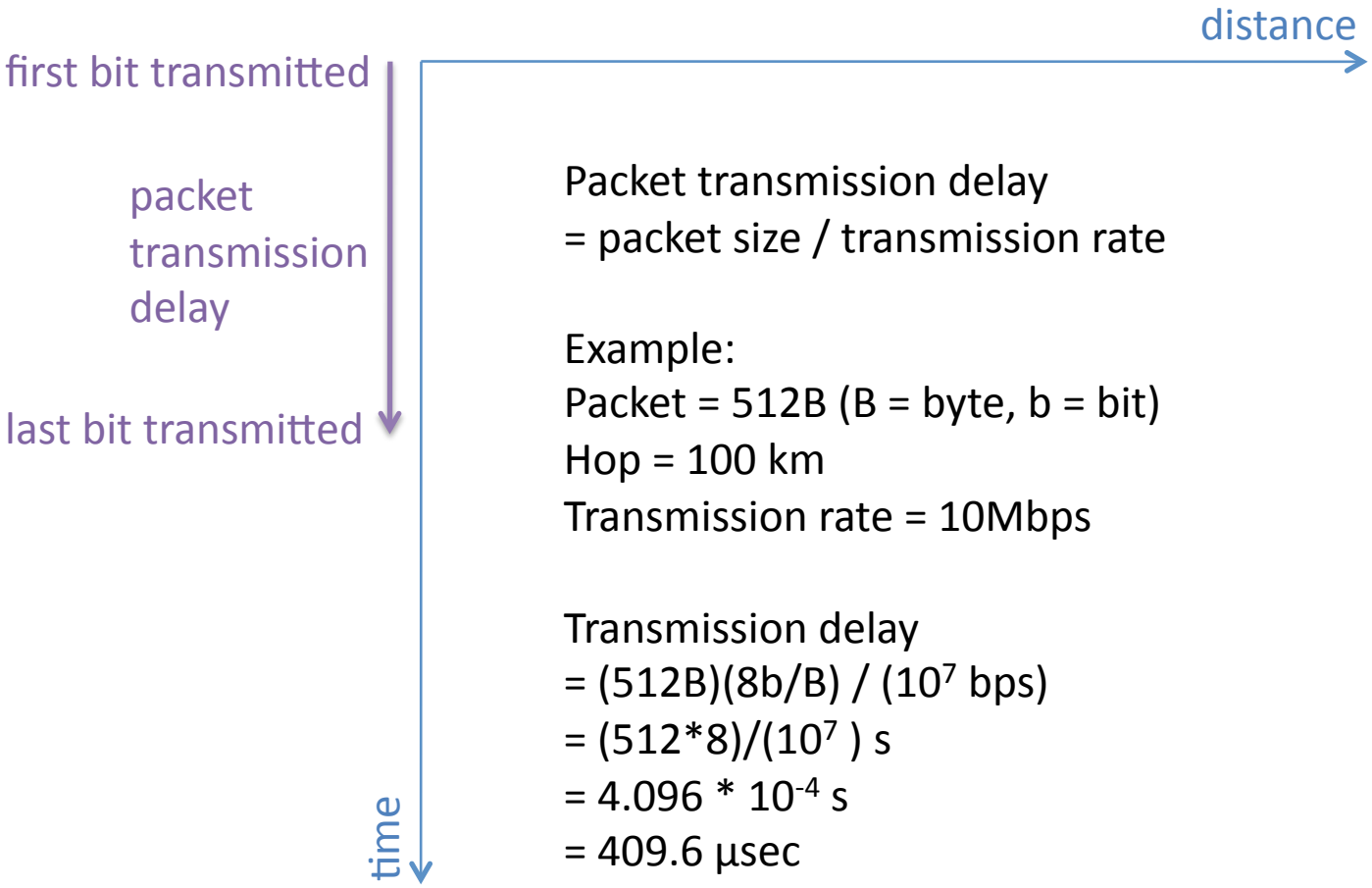


Propagation delay
= distance / propagation speed

Example:
Packet = 512B (B = byte, b = bit)
Hop = 100 km
Transmission rate = 10Mbps

Propagation delay
= 100 km / (3 * 10^8 m/s)
= 333 µsec

Transmission delay

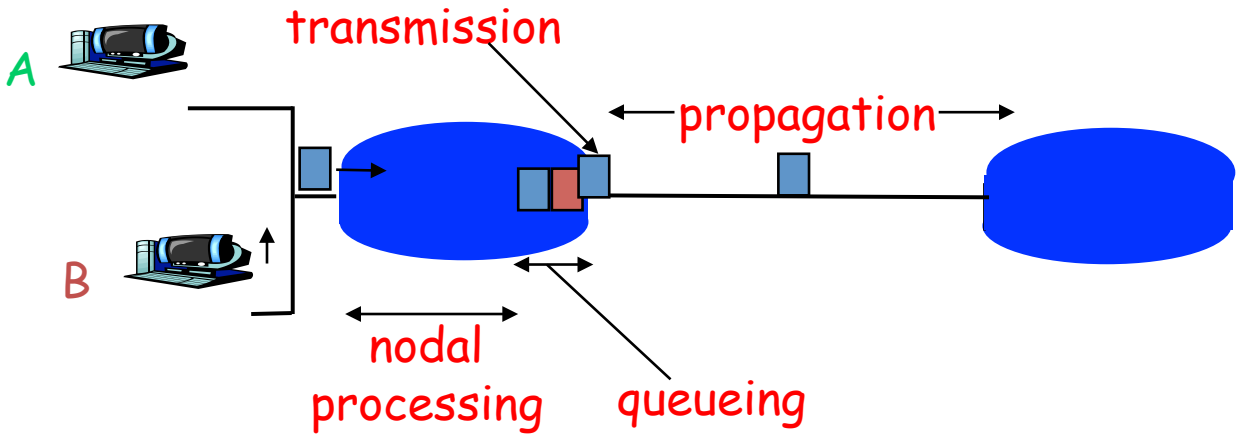


Packet transmission delay
= packet size / transmission rate

Example:
Packet = 512B (B = byte, b = bit)
Hop = 100 km
Transmission rate = 10Mbps

Transmission delay
= (512B)(8b/B) / (10^7 bps)
= (512*8)/(10^7) s
= 4.096 * 10^{-4} s
= 409.6 µsec

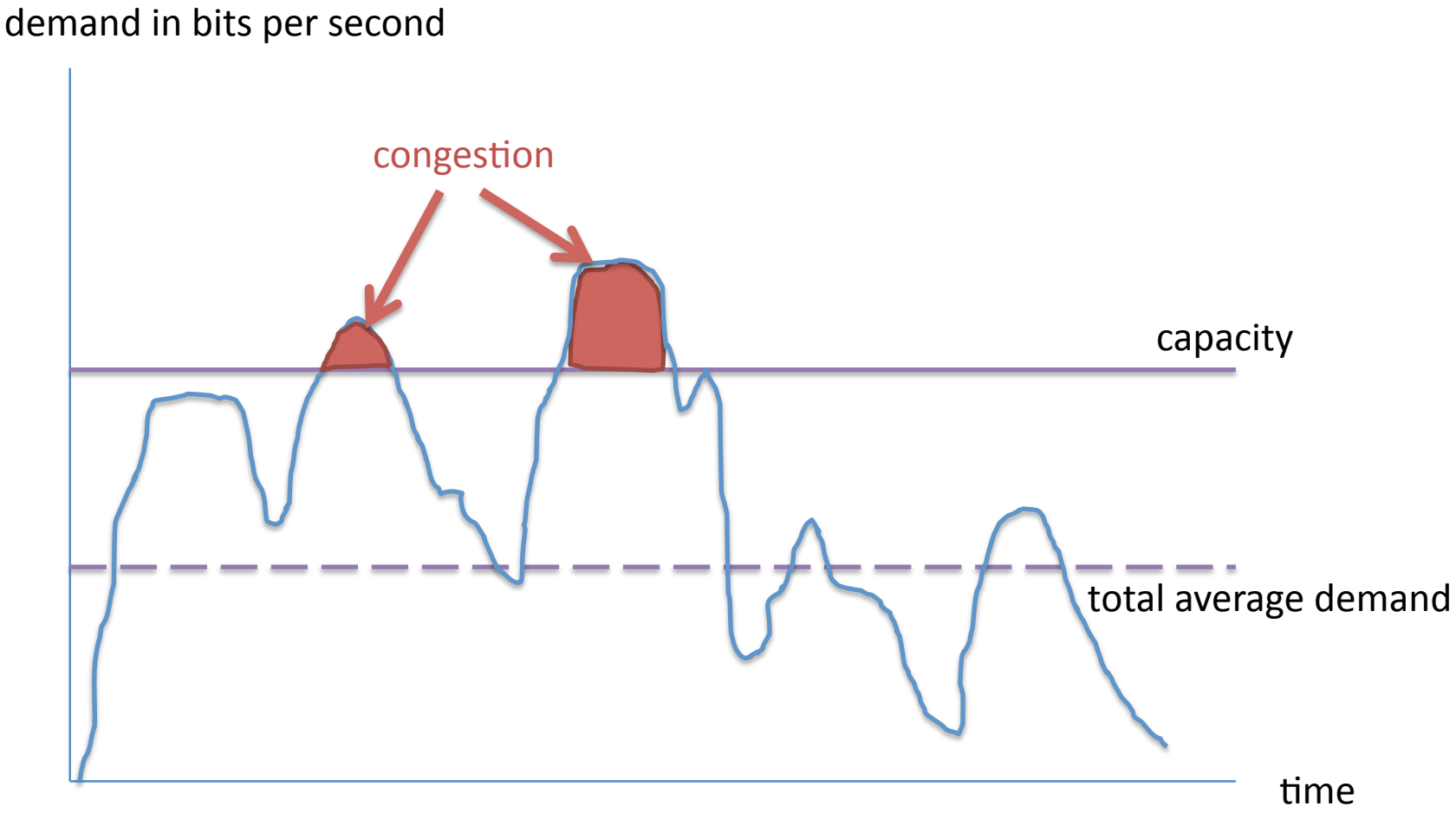
Delay (node)



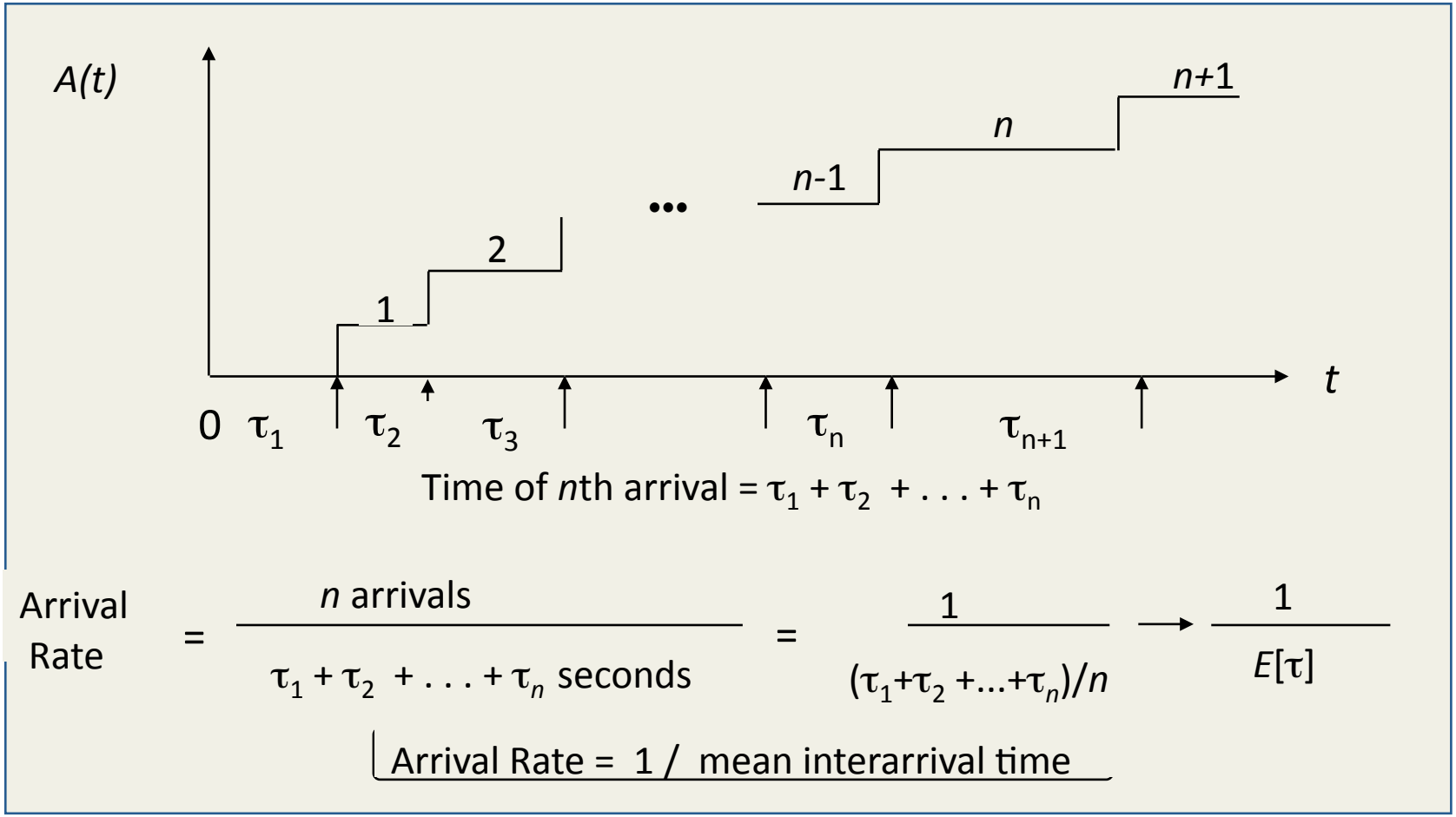
$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Kurose fig. 1.16

Congestion


demand in bits per second
congestion
capacity
total average demand
time

Arrivals to a queue



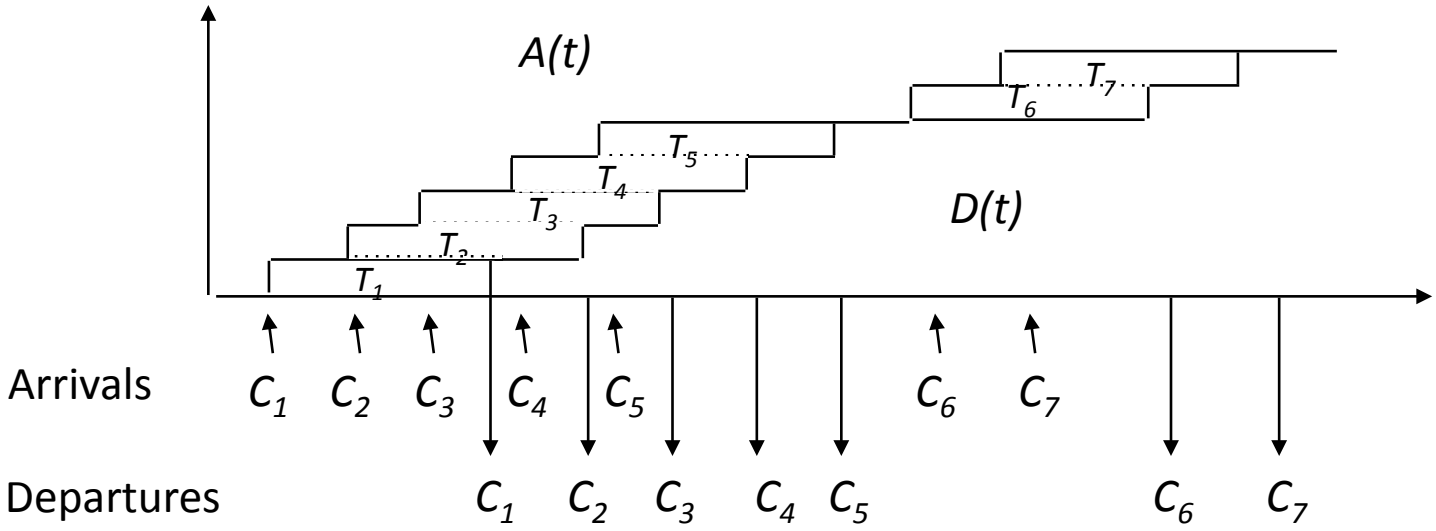
Time of *n*th arrival = $\tau_1 + \tau_2 + \ldots + \tau_n$

$$\text{Arrival Rate} = \frac{n \text{ arrivals}}{\tau_1 + \tau_2 + \ldots + \tau_n \text{ seconds}} = \frac{1}{(\tau_1+\tau_2+\ldots+\tau_n)/n} \rightarrow \frac{1}{E[\tau]}$$

Arrival Rate = 1 / mean interarrival time

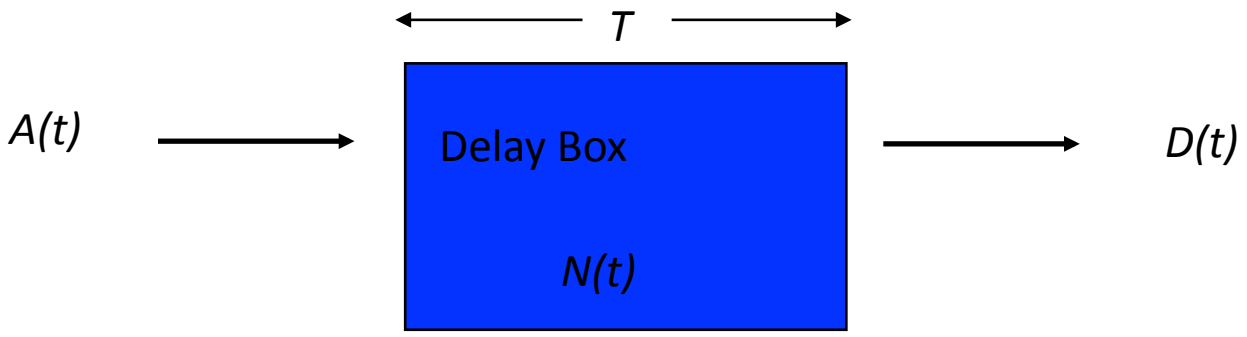
Leon-Garcia fig. A.2

Arrivals and Departures



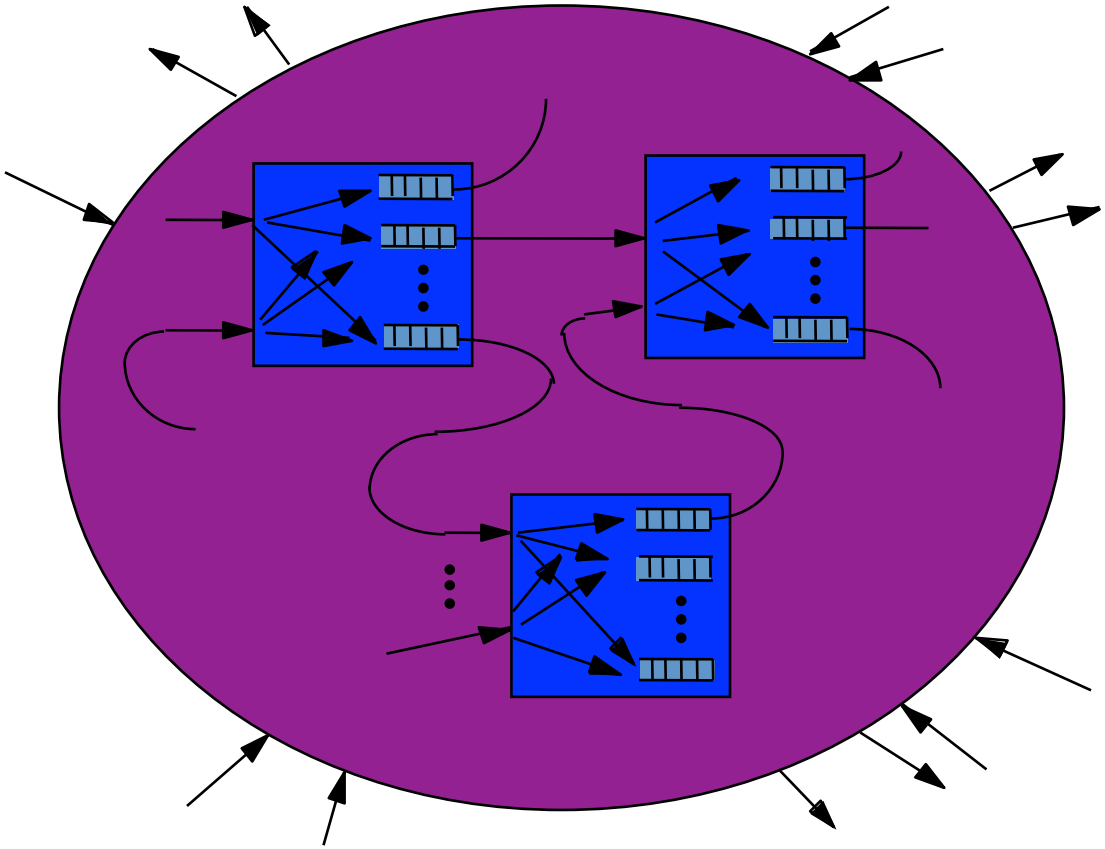
Leon-Garcia fig. A.4

Queue Delay



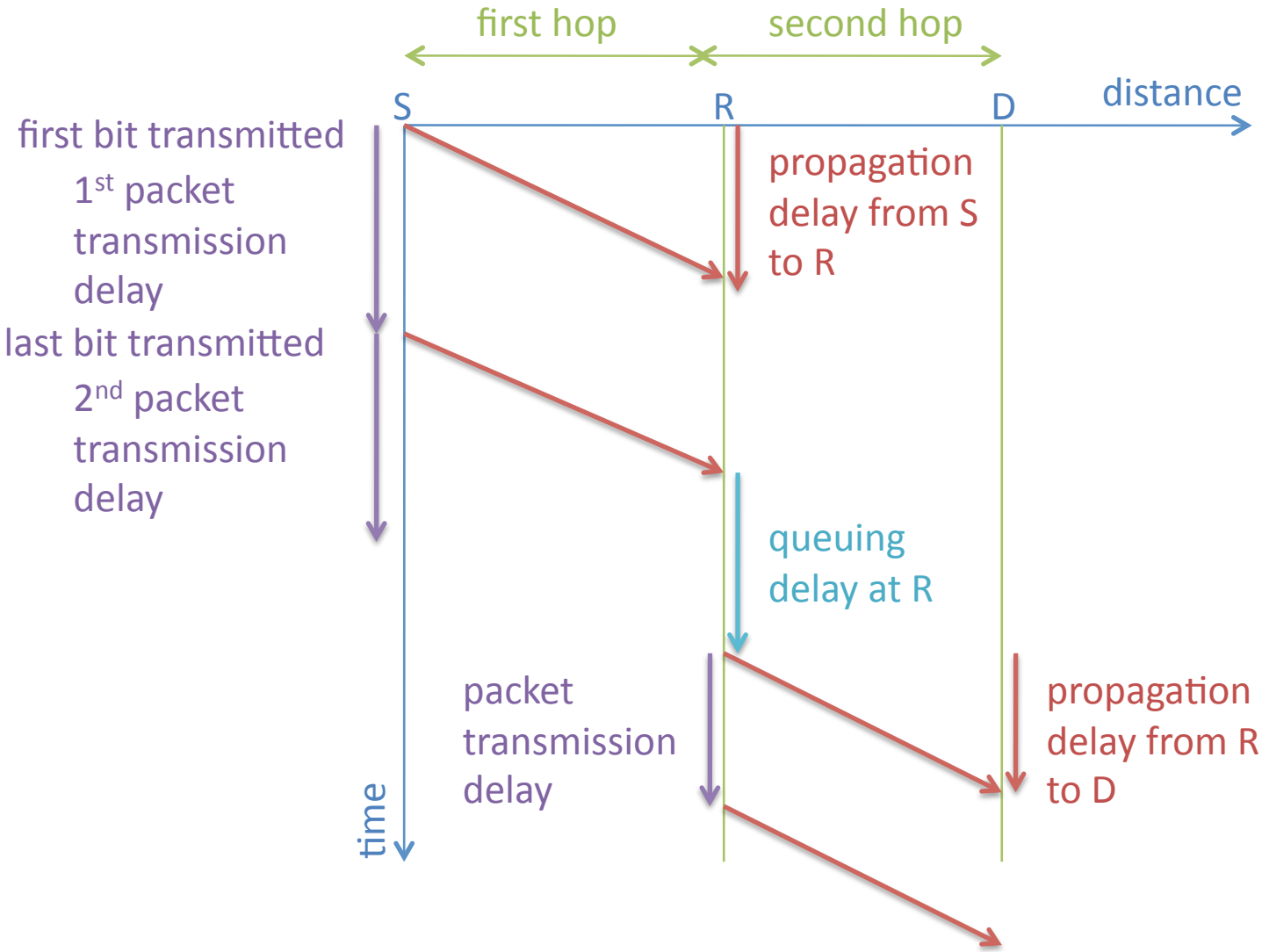
Leon-Garcia fig. A.3

Network of queues

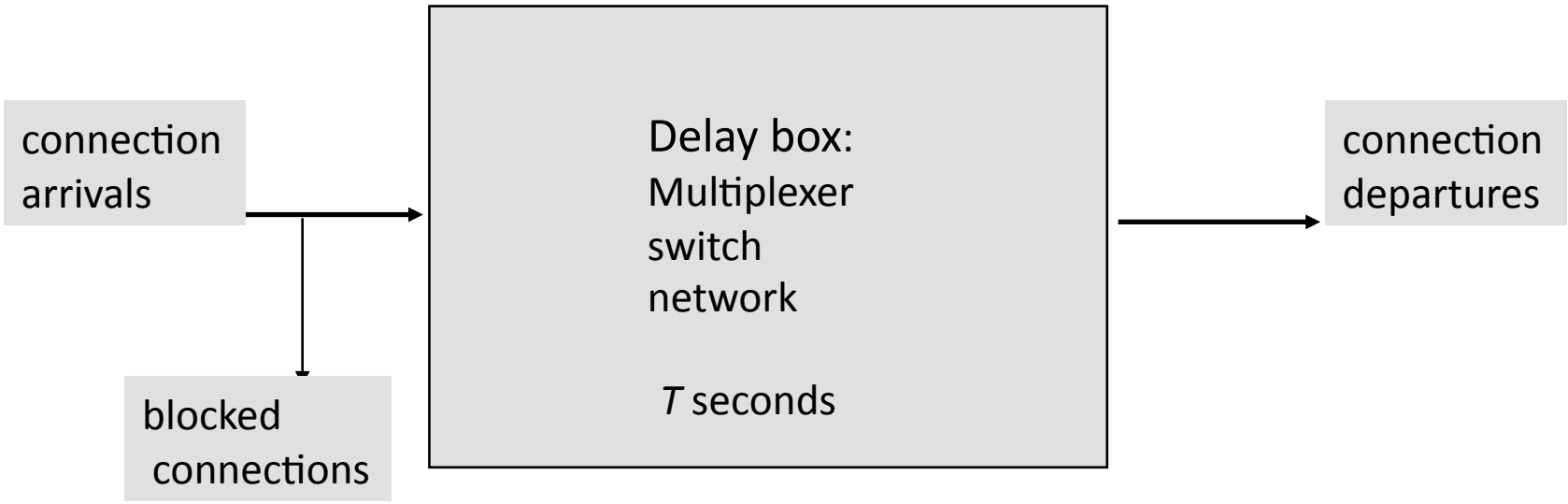


Leon-Garcia fig. A.5

Multiple hops



Queue



Leon-Garcia fig. A.1

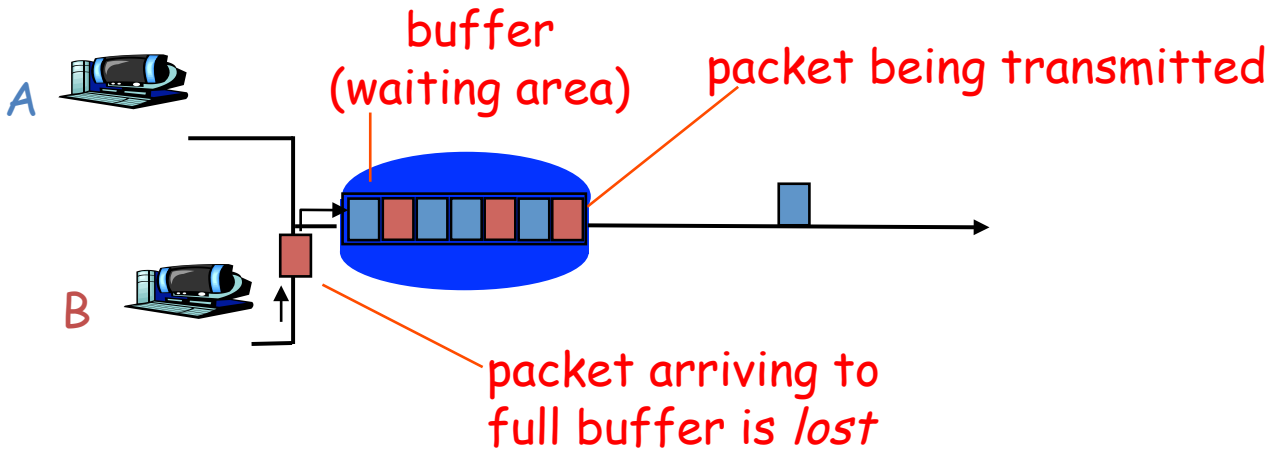
Delay

What is the difference between the queue delay and the processing, transmission and propagation delay?

Queue delay is random

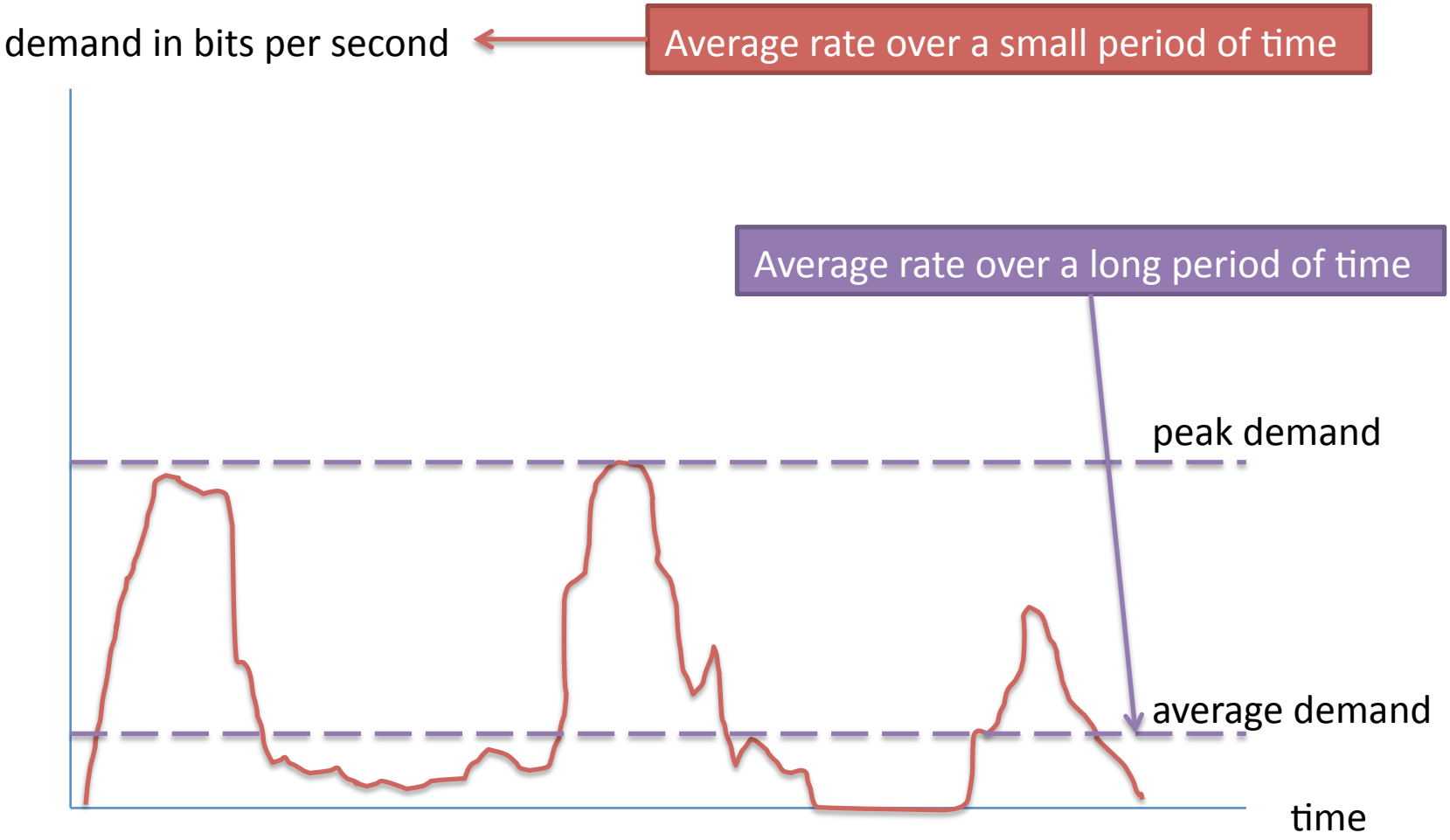
Processing, transmission and propagation delay are deterministic

Packet Loss

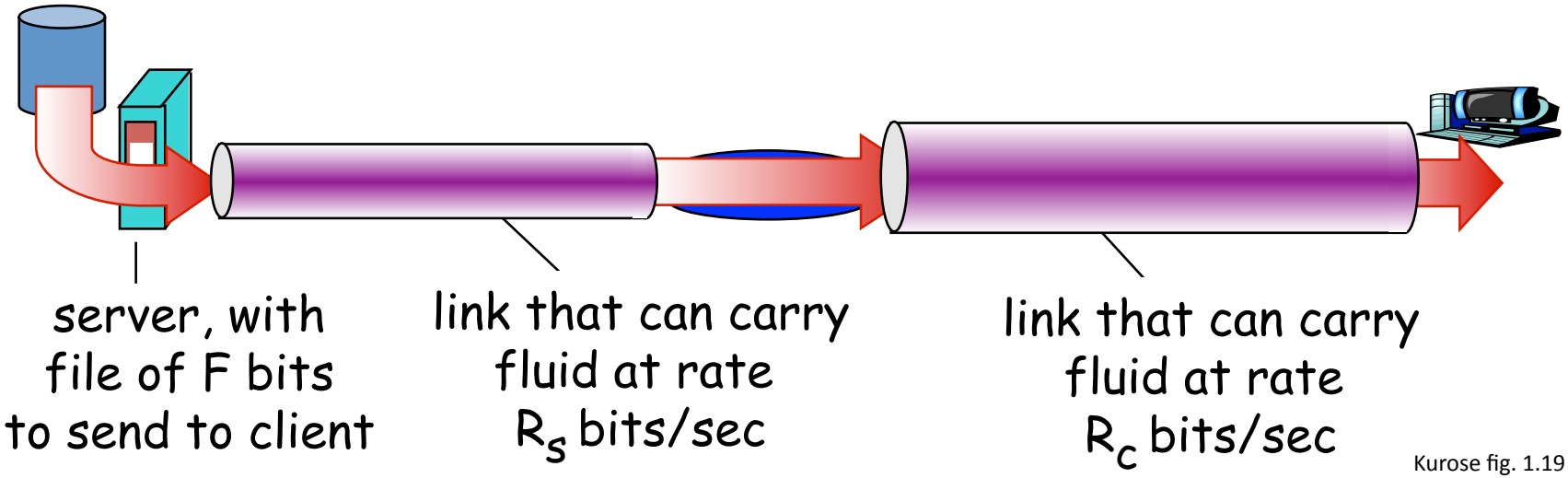


Kurose slide 1-54

Throughput: fluid model



Throughput

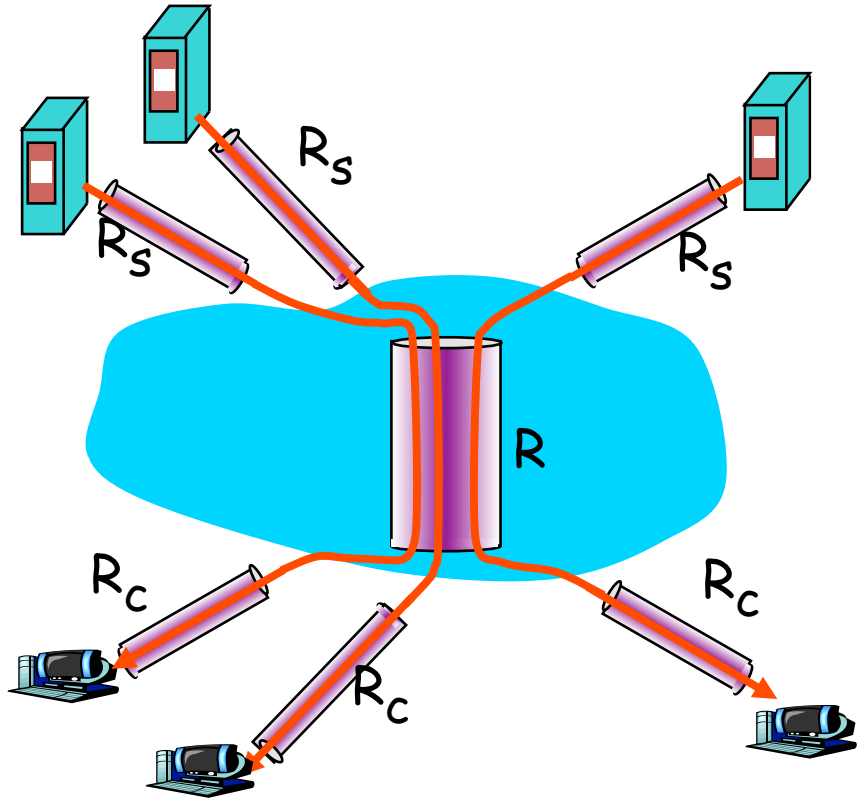


Kurose fig. 1.19

- Throughput = bits per second, calculated over a time window of many seconds
- What throughput will the user see in this web page download?
  - Min(R_s, R_c)

Throughput

- What is the throughput on a single connection?
  - Min(R_s, R/3, R_c)



Kurose fig. 1.20

Efficiency

$$\frac{\text{Data successfully delivered to the destination}}{\text{Overall resource used}}$$

Includes overhead, idle time/frequency, lost packets, etc.

Resource sharing!