Network Performance

Packet Delay

- Packet travels from source to destination via intermediate nodes
- Packet suffers from several types of delays at *each node* along the path
    - Delay: waiting time (rough understanding)

Types of delays:

1. Processing Delay
2. Queuing Delay
3. Transmission Delay
4. Propagation Delay

Processing Delay (Typically Microseconds (10^{-6}s))

Time required for the router to:

- Examing the packet header
- Detemrine where to forward the packet
- Check for errors in the packet

Queueing Delay (Typically Microseconds (10^{-6}s) to Milliseconds (10^{-3}s))

Time a pakcet waits to be transmitted onto the outbound link.

A cause for this can be other packets at the same node.

- Outbound link may be busy sending other packets
- Empty Queue: 0 queuing delay
- Long Queue: large queueing delay

Cases

Suppose that in every second:

- x new packets arrive at a node; max y packets forwarded through outbound links

Roughly 3 cases:

- $x << y$: average queuing delay, small
- $x \approx y$: average queuing delay can be small or large
- $x > y$: queuing delay is indefinite

Transmission Delay (Typically Microseconds (10^{-6}s) to Milliseconds (10^{-3}s))

Time required to push all bits of the packet into a link

- From transmission of the first bit to transmission of the last bit
- Equal to packet length (bits) / transmission rate (bits/s)

Propagation Delay

Time required for each bit to propagate to the next node

- Equal to distance (length of link) / propagation speed
- Propagation link depends on the physical medium of the link

---

Packet Loss

A router has limited buffer (queueing capacity) so when that buffer is full, a router will proceed to drop any new arriving packets making them considered as lost.

Certain protocols may include mechanisms for the retransmission of lost packets.

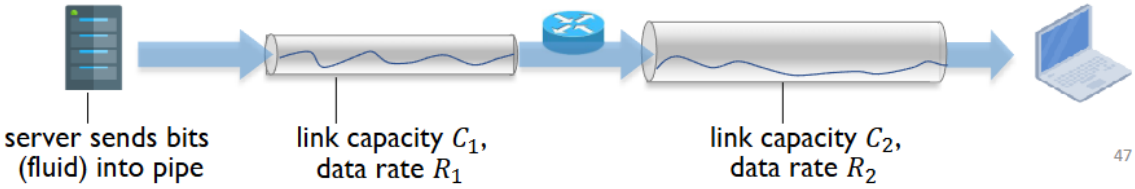
Throughput

Relates to the rate of successful data transmission over a link.

- Unit: bits (unless stated otherwise)
- Instantaneous throughput: rate at a given point in time
- Average throughput: rate over a longer period of time.

Throughput: the Pipe Analogy

- For now: consider throughput as the end-to-end successful data transmission rate (unit: bits/s)
- Link capacity (bandwidth): maximum data rate supported by a link
- Question: what is the end-to-end throughput
  - When $R_1 < R_2$? when $R_1 > R_2$?
  - What if there are N links with throughput R_1 , R_2 ... R_N
- Bottleneck link: the link on an end-to-end path that constraints the end-to-end throughput



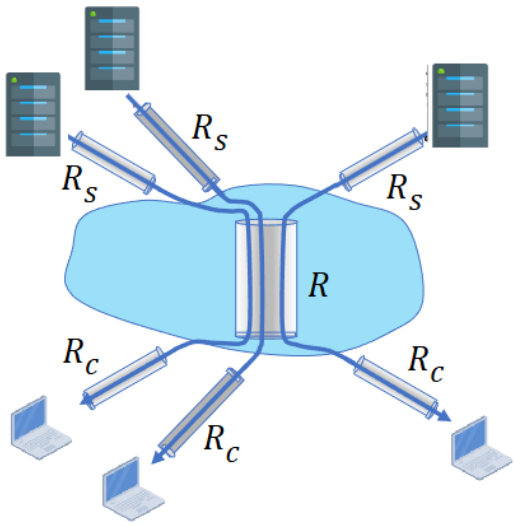
47

Throughput: Network Scenario

- What if there are multiple connections (data flows)?
- Consider the following scenario:
  - 10 server-client connections fairly share a backbone link in the network core
  - Server-side data rate: R_s
  - Client-side data rate: R_c
  - Backbone link data rate: R
- Per-connection end-end throughput:

$$\min\left(R_c,\ R_s,\ \frac{R}{10}\right)$$

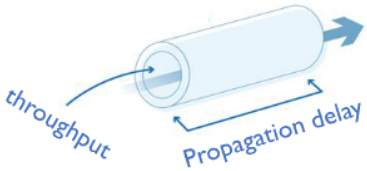
- In practice: R_c or R_s is often the bottleneck

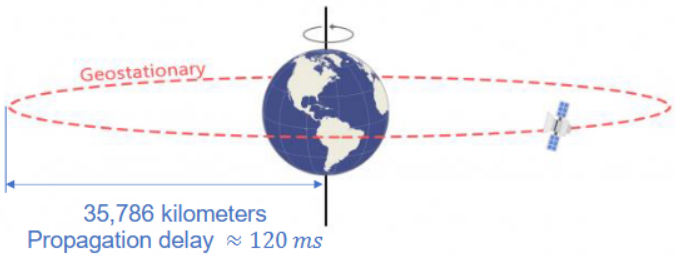


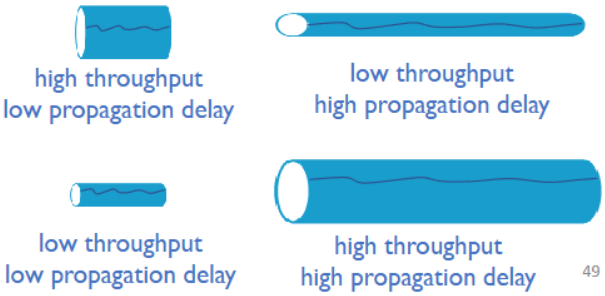
48

Throughput vs Delay

- Suppose that the size of a data packet is given
- Question:
  - Are throughput and delay related (both end-to-end)? If yes, which type of delay?
  - Does a high end-to-end throughput mean a low end-to-end delay? Why?







49