Outline

- Network Performance
- Evaluating the performance of computer networks
  - Key metrics
    - Latency (Delay)
    - Bandwidth (Data Rate)
  - Key parameters
Bandwidth

Definitions (use)
- Width of the frequency band (Hz) [exact, EE]
- Number of bits per second that can be transmitted over a communication link (b/s) [liberally…]

Examples
- 1 Mbps: $1 \times 10^6$ bits/second $\sim 1 \times 2^{20}$ bits/sec
  - $1 \times 10^{-6}$ seconds to transmit each bit
  - imagine a timeline: each bit occupies 1 micro second (length)
  - on a 2 Mbps link: each bit occupies 0.5 micro second
- The smaller the width, more bits will be transmitted per unit time
Bits transmitted at a particular bandwidth can be regarded as having some width:
(a) bits transmitted at 1Mbps (each bit 1 µs wide);
(b) bits transmitted at 2Mbps (each bit 0.5 µs wide).
Latency (Delay)

- Latency = propagation + transmission + queueing (delay)

- Propagation delay = distance/speed of light (propagation)
- Transmission delay = packet size / data rate
  - (the data rate is often called bandwidth)

- Important for the
  - transmission of…
    - 1 bit => propagation delay
    - (very) many bits => bandwidth
Delay-Bandwidth Product

- Think of the channel (between a pair of processes) as a hollow pipe
  - **Latency** (delay) is the length of the pipe
  - **Bandwidth** is the width of the pipe
- Delay of 50 ms and bandwidth of 45 Mbps
  - $50 \times 10^{-3} \text{ s} \times 45 \times 10^6 \text{ b/s}$
  - $2.25 \times 10^6 \text{ bits} = 280 \text{ KB data}$

The **delay-bandwidth product** shows how much information is ‘stored’ in the pipe (channel)
Delay x Bandwidth

- How many bits the sender must transmit before the first bit arrives at the receiver if the sender keeps the pipe full
- It takes another one-way latency to receive a response from the receiver...
- If the sender does not fill the pipe—send a whole delay \(\times\) bandwidth product’s worth of data before it stops to wait for a signal—the sender will not fully utilize the channel (network)

- Infinite bandwidth
  - Round-Trip Time (RTT) dominates
  - Throughput = Transfer Size / Transfer Time
  - Transfer Time = RTT + 1/Bandwidth \(\times\) Transfer Size

- It’s all relative
  - 1-MB file to 1-Gbps link looks like a 1-KB packet to 1-Mbps link
A 1-MB file would fill the 1-Mbps link 80 times, but only fill the 1-Gbps link 1/12.5 of one time

Note: assuming cross-country propagation delay = 100 ms
Relative importance of bandwidth and latency (delay) depends on application
- For large file transfer, bandwidth is critical
- For small messages (HTTP, NFS, etc.), latency is critical
- Variance in latency (jitter) can also affect some applications (e.g., audio/video conferencing)
Summary

- We have discussed two performance metrics using which we can analyze the performance of computer networks:
  - latency (delay)
  - bandwidth (data rate)