Homework 1

Software Systems **Allen B. Downey**
Spring 2005

Due: Tuesday 1 February

The reading for this assignment is Patterson, "Latency Lags Bandwidth." Note: in the table on page 72, in the Local Area Network section, the line marked Latency is in units of microseconds (μs), not milliseconds (ms) as labelled.

1.1 Cheat Sheet

b	bit	
B	Byte	8 bits
s	second	
G	giga	10^9 or 2^{30}
M	mega	10^6 or 2^{20}
K	kilo	10^3 or 2^{10}
m	milli	10^{-3}
μ	micro	10^{-6}
n	nano	10^{-9}

Annoyingly, network speeds are usually measured in bits (Mb/s for example); most other measurements are in bytes.

1.2 Distance limits latency

For wireless media and optical fiber, the lower bound on latency is determined by the speed of light, which is roughly 3.0×10^8 m/s. Electrical signals in wire propagate about 30% more slowly; a good rule of thumb is 2×10^8 m/s.

1. How far does light travel in one nanosecond? one microsecond? one millisecond?

2. How long (in ns) does it take an electrical signal to cross a chip that is 15 mm wide?

3. Latencies for current DRAMs are roughly 50 ns. Do you think electrical propagation time is a significant part of this latency?

4. Latencies for current disk drives are roughly 6 ms. Do you think electrical propagation time is a significant part of this latency?

5. If a disk spins at 15,000 rev/minute, how long (in ms) does it take to make one revolution?

6. What was the point of the last two questions?

1.3 Bandwidth is for big things

1. If the capacity of your hard drive is 100 GB, and it's full, how long would it take to transmit the entire contents over the two T1 lines that connect the Olin campus to the rest of the world, which provide aggregate bandwidth of roughly 3 Mb/s? What does this mean for remote backup systems?

2. If your monitor displays 1280 x 1024 pixels, and the "depth" is 32 bits per pixel, and the refresh rate is 85 Hz, what is the bandwidth from your video card to your monitor? How does that compare with the bandwidth of local area networks? What does this mean for remote display protocols?

3. According to a press release (ha!), Netflix has the capacity to ship 1 million DVDs per day by USPS. The contents of the DVDs vary, of course, but I would guess that the average size of the contents is 7 GB. Assuming that my guess is at least in the ballpark, what is the total data delivery capacity of the Netflix distribution centers, in Mb/s?

4. If Netflix switched from mailing DVDs to sending movies by 10 Gigabit Ethernet, how many Ethernet lines would they need? What effect would this change have on latency?

1.4 Latency is for little things

1. Use `ping` to estimate the network latency from your machine to four sites: one on campus, one on the East Coast, one on the West Coast, and one in Asia or Australia. Let it run for 10-20 iterations. When you hit Control-C, ping prints some summary statistics.

   For each site, what is the minimum latency, and what is the range between the minimum and maximum?

2. In general, humans cannot detect delays shorter than 1/10th of a second, or 100 ms. If you use `ssh` to connect to a machine in California, every time you press a key, it takes one round-trip latency before the key appears on your screen. How often would you see a noticeable delay?

1.5 Size amortizes cost

1. If the time to get one byte from disk is 6 ms and the bandwidth from the disk to memory is 90 MB/s, how long does it take to transfer a 512 B block from disk to memory? How about an 8 KB block? How about an entire track, which is 369 KB?

2. How long would it take to write 1 GB of data from memory onto disk, assuming that the limiting factor is bandwidth?

3. How long would it take to write 1 GB of data if the average time to write each 512 B block is 3 ms?

4. What is the point of the previous two questions?

1.6 Total latency, bottleneck bandwidth

1. Imagine a network path that contains three links, a Fast Ethernet, a Gigabit Ethernet, and a 10 Gigabit Ethernet. What is the total time it would take a small packet to traverse all three links and return?

2. At what bandwidth could we transfer a large file along this path?

3. A cheap way to build a "supercomputer" is to connect up a rack of off-the-shelf servers with a Gigabit Ethernet. In a setup like this, how long does it take for a given processor to access the memory module on another processor? How does this compare with the time to access local memory? How about a local disk?

4. In the same "supercomputer," how fast (bandwidth) could we move data from one processor's memory to another's?

5. What was the point of the last two questions?