Lorenzo Scaturchio

Homework 5

1. Consider the following:

- The user program continuously performs reads of 64KB blocks, and requires 2 million cycles to process each block.
- The clock rate is 3GHz
- The maximum sustained transfer rate of the memory bus is 640MB/sec
- The read/write bandwidth of the disk controller and the disk drives is 64MB/sec, disk average seek plus rotational latency is 9 ms

   a. How much time is needed for CPU to process each 64 KB block?

1MHz = 1 million cycles per second therefore ->
1GHz = 1 billion cycles per second
Tcpu = (2,000,000 bytes (per 64 KB blocks) / 3,000,000,000 cycles per second) * 1000 ->
Tcpu = 0.667 ms

   b. How much time is spent in memory transfer for a 64 KB block?

Tmem = -> Tread * 2 * 10 / Ttransfer_rate * 2 * 20
[(64 KB * 2 * 10) / (640 MB/sec * 2 * 20)] * 1000 ->
Tmem = 0.097 ms

   c. How much time is spent in I/O transfer for a 64 KB block?

Ti/o = Tseek + Trotation + Ttransfer ->
Tseek + Trotation = 9 ms ->
Ttransfer = 4096 B / 64 MB/s * 1000 ms/sec / 2^20 MB/B = 0.98 ms
Ti/o = 9.98 ms

   d. Based on the answers of a) – c), what is the bottleneck (slowest component) in this computer system setup, the CPU, memory bus, or the disk set?

The biggest bottleneck in the computer system setup would be the disk set based off of the calculations. We could clearly see that I/O transfer took 9 ms while the other components took less than 1 ms to execute.

2. Answer the following:

a. What is the average time to read or write a 512-byte sector for a typical disk rotating at 7200 RPM? The advertised average seek time is 8ms, the transfer rate is 20MB/sec, and the controller overhead is 2ms. Assume that the disk is idle so that there is no waiting time.

Tavg = Trotation + Tseek + Ttransfer + Toverhead
Tavg = 8 ms + [(0.5 / (7200 RPM / 60 sec)) * 1000 ms] + (0.5 KB / 20.0 MB/sec) + 2 ms ->
Tavg = 8 ms + 4.15 ms + 0.025 ms + 2 ms ->
Tavg = 14.18 ms

b. A program repeatedly performs a three-step process: It reads in a 4-KB block of data from disk, does some processing on that data, and then writes out the result as another 4-KB block elsewhere on the disk. Each block is contiguous and randomly located on a single track on the disk. The disk drive rotates at 7200RPM, has an average seek time of 8ms, and has a transfer rate of 20MB/sec. No other program is using the disk or processor, and there is no overlapping of disk operation with processing. The processing step takes 20 million clock cycles, and the clock rate is 400MHz. What is the overall speed of the system in blocks processed per second assuming no other overhead?

Trotation = 0.5 / 7200RPM = 4.17 ms ->
4 KB / 20 MB/sec = 0.2 ms ->

Ti/o = Trotation + Tseek + Ttransfer + Toverhead ->
Ti/o = 4.17 ms + 8 ms + 0.2 ms ->
Ti/o = 14.36 ms

20 million cycles/(400MHz) = 50ms
Therefore Speed = 50 ms + (2 * 14.362) = 78.7 ms / block ->
Speed = 13.38 blocks/sec

c. If a system contains 1,000 disk drives, and each of them has a 500,000 hour MTBF (Mean Time Between Failures), how often a drive failure will occur in that disk system?

P(Failure) = 500,000 hour MTBF / 1000 disk drives ->
P(Failure) = 500 hours

3. Consider a hard disk with 3 platters. Each platter has two recording surfaces (top and bottom). Each surface has 100,000 tracks. Each track holds 2000 sectors of 4K bytes

each. The disk rotates at 10,000 RPM. Assume the disk has a bandwidth of 100MB/sec. Note that a block is the unit of data transfer between disk and main memory. Each block may contain a number of sectors. Here we assume there is only one sector per block. Note also that traditionally the sector size is 512 bytes, although 4KB is now more commonly used. In addition, we assume that the time it takes the head to move n tracks is 1 + 0.0001n milliseconds

a. What is the capacity of the disk?

C = platters * surfaces * [tracks (per surface) (sectors * 4 KB) ->
C = 3 * 2 * [100,000 * (2000 * 4 KB)] ->
C = 6 * [100,000 * 8 MB] ->
C = 6 * 800 GB ->
C = 4.8 TB

b. What is the maximum and average seek time? (assume that the average seek time is approximately 1/3 of the maximum seek time.)

1 seek time= 100,000 tracks * 1.00025 ms = 11 ms ->
Max seek time = 11 ms ->
Avg seek time = 11 / 3 = 3.67 ms

c. What is the maximum and average rotational latency?

1 rotation = 60000 ms / 10000 = 6 ms ->
Max rotation time = 6 ms ->
Avg rotation time = 6 / 2 = 3 ms

d. How long does it take to read an 8 Mbyte file located randomly on the drive, i.e., 2000 reads of size 4KB each? You may assume that each 4KB occupies a sector.

Ti/o = 2000 * (Tseek(avg) + Trotation(avg) + Ttransfer ->
Tseek(avg) = 3.67 ms->
Trotation(avg) = 3 ms ->
Ttransfer = 8 MB * (1 sec / 100 MB) * (1000 ms / sec) = 80 ms ->

Ti/o = 2000 * (3.667 ms + 3 ms) + 80 ms ->
Ti/o = 13.42 sec

e. What if all the 2000 sectors of the above file to be read are located next to each other on the same track, i.e., sequential read?

Ti/o = Tseek + Trotation + Ttransfer ->
Ti/o = 3.67 ms + 3 ms + [(4 KB / (100 MB / sec) * 2000] ->
Ti/o = 3.67 ms + 3 ms + 80 ms ->
Ti/o = 86.67 ms

f. How long does it take to read an 8 Mbyte file when the size of block is 4MB and blocks are randomly located on the drive? Assume that each 4MB occupies a group of sequential sectors on the same track which can be read without extra seek or rotational latency.

Ti/o = 2 * (Tseek + Trotation) + Transfer ->
Ti/o = 2 * (3.667 ms + 3 ms) + 80 ms ->
Ti/o = 93.34 ms

4. What are the purposes of two master nodes in Hadoop? How does the master achieve high reliability to avoid the single point of failure?

The purposes of two master nodes in Hadoop are for utilization of different purposes which are named: NameNode, and Resource manager. The NameNode's purpose is to manage the HDFS storage. The Resource manager on the other hand is responsible for overseeing the scheduling of application tasks. The high reliability comes from the parallelization of multiple slave nodes on top of the master node such that all the tasks that are being executed, if there is one point of failure then the tasks are re-allocated accordingly.

5. What is speculative execution (also called backup tasks) in Hadoop? What problem does it solve?

Speculative execution is a backup task that is created by Hadoop when a particular task is slower to fully execute relative to other tasks. If the task happens to be completed then any duplicate tasks made by the process of speculative execution are terminated. This process solves the issue of having severe bottlenecks from tasks that aren't supposed to be running so slow.

6. You have a data file (480MB) and want to store it in HDFS which block size is 128 MB. Your Hadoop cluster has 10 data nodes. Explain (draw a figure if needed) how this file is divided and stored in this system in details (e.g., with specific locations of data blocks). And explain the steps how this file can be accessed later.

Allocation toward 10 blocks of 128 MB in size would consist of the client first contacting NameNode, which would then inform the client of the closest DataNodes storing blocks of the file when reading, and selects DataNodes for holding its replicas when writing as a basis for the logistics of the HDFS. One block, which is 128 MB splits the data with a replication factor of 2 (because 3 exceeds the total amount of block storage allocated) consisting of 96 MB of the 128 MB per block used. The replication is to ensure recovery and there is still storage to write additional files within the HDFS.