

CS 107 Fall 2006

Lecture 5: Disks and Data Storage

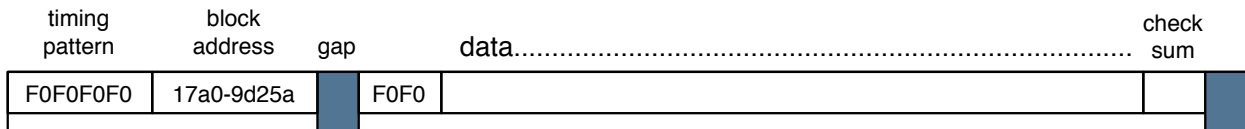
1 Hardware Components of a Disk

- Controller.
- Platter(s)
- Read arm and seek time.
- Read head and rotational delay
- Data channel

2 How the Hardware Finds the Data

- Platters and surfaces.
- Tracks and sectors.
- Addresses
- Timing bytes
- Data and check sums.

What happens when you format a disk?



February 2004:

Hitachi GST 300 gigabyte 3.5" drive

61 billion bits per square inch, 10,025 rpm, 2Gb per second fiber optic channel.

A 300 Gigabyte disk with 2 recording surfaces and 512 bytes per sector, would need 300 million sectors per platter. That could be 100,000 tracks with 30,000 sectors per track.

Figure 1: The form of one disk sector.

3 Data and Metadata

- Partition table
- The block map says which blocks are in use and which are free. A block may have 1, 2, 3, 8, or 16 sectors. A block is the smallest amount of data that can be read or stored.

- Volume directory is the top level directory listing the folders and files on the disk. For each file, it provides a way to find the blocks that belong to that file.
- Some files are very short and do not fill a block. The rest of the space in the block is wasted. This is called “internal fragmentation”. When you archive a set of files, this wasted space is squeezed out and many file are packed together in one block. This saves space, but costs time to unpack the archive the next time you need to use one of the files
- Many files are more than one block long. When you write a file, it is stored in the first free blocks on the disk, and these are often not contiguous.
- After many files are written and deleted, there may be many small areas of free disk space, and the recently written files may be broken up into many small pieces. This is **external fragmentation**. To improve performance when your disk space becomes fragmented, you should copy all files that you care about to a backup medium, then run a de-fragmentation program. If all goes well, you will not need that backup copy. However, if the power fails or the defragmentation process gets interrupted for any reason, you will need your backup copy.

4 Ways a Disk can Fail

- Physical failure: Heat, power burst, dirt. Symptoms.
- Corruption: a bad program can destroy a crucial part of the disk, such as the block table or the volume directory. Symptoms.

In either case, all or most of your data will still be on the disk but you will not be able to access it. There are commercial businesses that specialize in recovering data after disk failures. They charge substantial fees.

5 Using reasonable caution.

- What will happen when your disk fails? Will you be a big loser?
- How can you protect yourself?
 - Avoid overheating. Use a fan on hot humid days, if necessary.
 - Learn to recognize trouble in the early stages.
 - Backup any files that matter to a removable medium. Keep it in a different room or a different site.
- RAID: (Redundant Array of Independent Devices) is a solution for commercial systems and servers. The same data is written almost simultaneously to two disks. If one fails, the other still has the data.
- Caution: when you erase a file, traces remain. There are tools to read these faint traces. When you discard a disk, protect your privacy by using your system’s “secure erase” command.