### Types of External Memory

- **Magnetic Disk**
  - RAID
  - Removable
- **Optical**
  - CD-ROM
  - CD-Recordable (CD-R)
  - CD-R/W
  - DVD
- **Magnetic Tape**
- Flash memories are often used as a “solid-state drives”

### Magnetic Disk

- Disk substrate coated with magnetizable material (iron oxide...rust)
- Substrate used to be aluminium
- Now glass or ceramic
  - Improved surface uniformity
    - Increases reliability
  - Reduction in surface defects
    - Reduced read/write errors
  - Lower flight heights (See later)
  - Better stiffness
  - Better shock/damage resistance

### Read and Write Mechanisms

- Recording & retrieval via conductive coil called a head
- May be single read/write head or separate ones
- During read/write, head is stationary, platter rotates
- Write
  - Current through coil produces magnetic field
  - Pulses sent to head
  - Magnetic pattern recorded on surface below
- Read (traditional)
  - Magnetic field moving relative to coil produces current
  - Coil is the same for read and write
- Read (contemporary)
  - Separate read head, close to write head
  - Partially shielded magneto resistive (MR) sensor
  - Electrical resistance depends on direction of magnetic field
  - High frequency operation
    - Higher storage density and speed

### Data Organization and Formatting

- Concentric rings or tracks
  - Gaps between tracks
  - Reduce gap to increase capacity
  - Same number of bits per track (variable packing density)
  - Constant angular velocity (rotational speed is constant so linear velocity varies)
- Tracks divided into sectors
- Minimum block size is one sector
- May have more than one sector per block
Disk Data Layout
- Bit near center of rotating disk passes fixed point slower than bit on outside of disk
- Increased spacing between bits in outer tracks compared to inner
- Rotate disk at constant angular velocity (CAV)
  - Gives pie shaped sectors and concentric tracks
  - Individual tracks and sectors addressable
  - Move head to given track and wait for given sector
  - Waste of space on outer tracks
    - Lower data density
- Can use zones to increase capacity (typical 16)
  - Each zone has fixed bits per track
  - More complex circuitry
  - Common on modern disks

Disk Layout Methods
- Must be able to identify start of track and sector
- Format disk
  - Additional information not available to user
  - Marks tracks and sectors

Winchester Disk Format
Seagate ST506
- Fixed (rare) or movable head
- Disk is removable or fixed
- Single or double (usually) sided
- Single or multiple platter
- Head mechanism
  - Contact (Floppy)
  - Fixed gap
  - Flying (Winchester)
Fixed/Movable Head Disk

- Fixed head
  - One read write head per track
  - Heads mounted on fixed ridged arm
- Movable head
  - One read write head per side
  - Mounted on a movable arm

Removable or Not

- Removable disk
  - Can be removed from drive and replaced with another disk
  - Provides unlimited storage capacity
  - Easy data transfer between systems
- Nonremovable disk
  - Permanently mounted in the drive

Multiple Platters

- One head per side
- Heads are joined and aligned
- Aligned tracks on each platter form cylinders
- Data is striped by cylinder
  - reduces head movement
  - Increases speed (transfer rate)

Floppy Disk

- 8” (175kb), 5.25” (360kb/720kb), 3.5” (1440kb, 2880kb)
- Small capacity
  - Aside: 3.5” are almost always described as 1.44Mb but they are really 1.40625Mb
- Slow, cheap, universal, almost obsolete
- With Windows 2000 Microsoft modified the floppy driver to be much faster (and also much less reliable)
- Omitted from most current computers
- To access old data floppy drives can be attached through USB ports
Winchester Hard Disk (1)
• Developed by IBM in Winchester (USA)
• Sealed unit
• One or more platters (disks)
• Heads fly on boundary layer of air as disk spins, rest on disk when power off
• Very small head to disk gap
• Getting more robust
  – Automatic correction of errors, remapping of bad sectors

Winchester Hard Disk (2)
• Universal
• Cheap
• Fastest external storage
• Getting larger all the time
  – Hundreds of gigabytes now easily available

Speed
• Seek time
  – Time to move head to correct track
  – Typically < 10ms today
• (Rotational) latency
  – Waiting for data to rotate under head
  – Speeds of 3,600 to 15,000 rpm common
  – Floppies 300-600 rpm
• Access time = Seek + Latency
• Transfer rate
• Operating system queuing adds additional time

Transfer Time and Rates
• Transfer time $T = \frac{b}{rN}$
  $b =$ bytes to transfer, $r =$ revolutions/sec, $N =$ number of bytes/track
• Average access time $T_a$
  $T_a = T_s + \frac{1}{2r} + \frac{b}{rN}$
  Where $T_s$ is average seek time

Operating system queuing
• Operating system queuing adds additional time
• But queues increase overall I/O throughput

Disk scheduling algorithms
• Improve performance with efficient scheduling
• Covered in Operating Systems
• Examples:
  – Shortest Seek Time First
  – SCAN head continually moves innermost to outermost and back to center
Rotational Position Sensing

- Used in servers
- Channel released when seek command issued
- Device attempts to reestablish communication when seek is complete

For more detailed information...

- See http://www.pcguide.com/ref/hdd/index.htm

Typical Specs (Now a bit old...)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Seagate Barracuda 9SE</th>
<th>Seagate Barracuda 7KB.8</th>
<th>Seagate Barracuda 75KB.9</th>
<th>Seagate Barracuda 95KB.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>High-density server</td>
<td>High-performance desktop</td>
<td>Entry-level desktop</td>
<td>Entry-level desktop</td>
</tr>
<tr>
<td>Capacity</td>
<td>134 GB</td>
<td>160 GB</td>
<td>75 GB</td>
<td>95 GB</td>
</tr>
<tr>
<td>Recording speeds/track</td>
<td>0.14 sec</td>
<td>0.18 sec</td>
<td>0.8 sec</td>
<td>0.8 sec</td>
</tr>
<tr>
<td>Write Precompensation</td>
<td>10 sec</td>
<td>12 sec</td>
<td>12 sec</td>
<td>12 sec</td>
</tr>
<tr>
<td>Spin-up speed</td>
<td>7200 rpm</td>
<td>7200 rpm</td>
<td>5400 rpm</td>
<td>5400 rpm</td>
</tr>
<tr>
<td>Average rotational delay</td>
<td>3.2 sec</td>
<td>3.2 sec</td>
<td>2.7 sec</td>
<td>2.7 sec</td>
</tr>
<tr>
<td>Maximum transfer rate</td>
<td>15 MB/s</td>
<td>14 MB/s</td>
<td>14 MB/s</td>
<td>14 MB/s</td>
</tr>
<tr>
<td>Drive power</td>
<td>3.5 W</td>
<td>3.5 W</td>
<td>3.5 W</td>
<td>3.5 W</td>
</tr>
<tr>
<td>Tracks per cylinder</td>
<td>802</td>
<td>922</td>
<td>952</td>
<td>952</td>
</tr>
</tbody>
</table>

RAID

- Two interpretations of acronym:
  - Redundant Array of Independent Disks
  - Redundant Array of Inexpensive Disks
  - 2nd interpretation probably more common
- 7 levels defined
  - NOT a hierarchy
  - Levels 2 and 4 not implemented in practice
- 3 principles in common:
  - Set of physical disks viewed as single logical drive by O/S
  - Data distributed across physical drives
  - Can use redundant capacity to store parity information for error detection and correction

Why RAID?

- Improve I/O throughput through parallelism
- Improve recoverability
  - Note that multiple devices actually decreases the reliability of a system
  - Disks are independent, so probability of failure is the sum of the individual probabilities

Overview of RAID Levels

<table>
<thead>
<tr>
<th>Category</th>
<th>Level</th>
<th>Description</th>
<th>Disk configuration</th>
<th>Data availability</th>
<th>Large I/O drive support</th>
<th>Small I/O request rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stripping</td>
<td>0</td>
<td>Non-striped</td>
<td>0</td>
<td>Low</td>
<td>High</td>
<td>Very high</td>
</tr>
<tr>
<td>Mirroring</td>
<td>1</td>
<td>Mirroring</td>
<td>1 stripe</td>
<td>Higher than RAID 3, 5, 6, 7</td>
<td>Higher than striping for RAID 3, 5, 6, 7</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reconstruction via Mirroring</td>
<td></td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
<tr>
<td>Parity Services</td>
<td>2</td>
<td>Parity Services</td>
<td>4 parity</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
<tr>
<td>Banked mirroring</td>
<td>3</td>
<td>Banked Mirroring</td>
<td>6 mirror</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
<tr>
<td>Independent services</td>
<td>4</td>
<td>Independent services</td>
<td>8 mirror</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
<tr>
<td>Banked mirroring</td>
<td>5</td>
<td>Banked Mirroring</td>
<td>12 mirror</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
<tr>
<td>Banked striped</td>
<td>6</td>
<td>Banked Striped</td>
<td>16 mirror</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>Data distributed across disks for RAID 6, 7</td>
<td>High</td>
</tr>
</tbody>
</table>
I/O Transfer Rate and I/O Request Rate

- Note the last two columns in table
  - Transfer Rate and I/O Request Rate are not the same!
  - High transfer rate useful when large blocks of data have to read (or written); e.g., large database
  - Ability to satisfy high I/O request rate useful when many small independent requests have to be satisfied; e.g., web server, mail server, database server, multi-user computing, other transaction-oriented environments

RAID 0

- No redundancy
- Appears as logical disk on which all data is stored: an abstraction of the real disk(s)
- Data is placed in segments called strips
  - A *stripe* is all of the strips at the same location on all of the disks
  - Strips are placed on disks in round-robin:

RAID 0 Logical to Physical Mapping

RAID 0 Performance

- Depends on
  - High transfer capacity on entire path to memory
  - Application data requests need to drive I/O efficiently, either:
    - Large requests for logically contiguous data that can be satisfied by parallel access to different disks, or
    - Many small requests, each of which requires access to a single strip of a disk
  - Strip size has to be balanced with typical I/O patterns
  - Depends on whether you want large transfer capacity or high I/O request rate

Raid 0 Request Patterns

- Transfer capacity increases:
  - When the strip size is smaller than a single request, and requests are made for contiguous data, so that a single request is handled by multiple disks (in parallel)

- Request Rate
  - Multiple data requests probably not on same disk so provides increased request rate
  - Strip size should be larger than average request size

Redundancy

- RAID 0 does not provide redundant storage
- RAID 1 provides redundant storage in the simplest form: data is duplicated on mirrored disks
- RAID 2-6 provide redundancy through parity calculations
**RAID 1**
- Mirrored Disks
- Data is (usually) striped across disks
  - Each logical stripe mapped to 2 disks
- Read from either
- Write to both
- Recovery is simple
  - Swap faulty disk & re-mirror
  - No down time
- Relatively expensive: double disks

**RAID 1 Pros and Cons**
- Expensive (2 disks per logical disk)
  - Usually used for system or other highly critical data only
- Can achieve very high transfer rates (2 x RAID0) in transaction-oriented environment but only if most requests are Reads
- Write are limited to the slower of the two drives

**RAID 2**
- Parallel access across all disks for each I/O request
  - Requires synchronized disks and specialized controllers
- Very small strips, single byte/word
  - Data split at bit level across disks
- Error correction calculated across corresponding bits on disks
  - Multiple parity disks store Hamming code error correction in corresponding positions
- Lots of redundancy
  - Expensive
  - No longer used; most disks incorporate ECC already

**RAID 2 Layout**

**RAID 3**
- Similar to RAID 2, but only one redundant disk, no matter how large the array
- Data is striped at the byte level; a block of data has parts written to each drive
- Simple parity bit for each set of corresponding bits
- Data on failed drive can be reconstructed from surviving data and parity info
- Very high transfer rates, but low I/O request rates (1 at a time, because all disks are involved in each I/O)
- RAID 3 is not common; poor I/O request performance
RAID 3 Layout
- Each disk operates independently
- Good for high I/O request rate
- Large strips (e.g., 16k or 32k)
- Bit by bit parity calculated across stripes on each disk
- Parity stored on parity disk
- Poor write performance on small requests
- The single parity drive is a bottleneck

RAID 4 Layout
- One of the most widely used RAID schemes
- Similar to RAID 4, but parity striped across ALL disks
- Round robin allocation for parity stripe
- Avoids RAID 4 bottleneck at parity disk
- Commonly used in network servers

RAID 5 Layout
- Two parity calculations
- Stored in separate blocks on different disks
- User requirement of N disks needs N+2
- High data availability
  - Three disks need to fail for data loss
  - Significant write penalty
- Rarely used because possibility of multiple simultaneous disk failure is very slim
- Catastrophic events would normally destroy most or all disks

RAID 6 Layout
- Each disk operates independently
- Good for high I/O request rate
- Large strips (e.g., 16k or 32k)
- Bit by bit parity calculated across stripes on each disk
- Parity stored on parity disk
- Poor write performance on small requests
- The single parity drive is a bottleneck
RAID 6 Layout

RAID Comparison

<table>
<thead>
<tr>
<th>Level</th>
<th>Advantage</th>
<th>Disadvantages</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>High availability, no parity required</td>
<td>None</td>
<td>High write performance, good for transactional applications</td>
</tr>
<tr>
<td>1</td>
<td>No parity, high performance</td>
<td>None</td>
<td>Good for transactional applications</td>
</tr>
<tr>
<td>5</td>
<td>Parity and block striping</td>
<td>High parity overhead</td>
<td>Good for large data sets</td>
</tr>
<tr>
<td>6</td>
<td>Parity and block striping</td>
<td>None</td>
<td>Good for large data sets</td>
</tr>
</tbody>
</table>

Other RAID systems

- The list above is not all-inclusive
- Some systems combine levels, for example RAID 0+1
- Some are proprietary, such as RAID 7

Optical Storage

- Primarily CD and DVD
- Other technologies exist:
  - Magneto-optical
  - Floptical

Optical Storage Products
**Optical Storage CD-ROM**
- Originally designed for audio
- 650-700 MB stores over 70 minutes audio
- Polycarbonate plastic coated with highly reflective coat, usually aluminium
- Data stored as pits and lands
- Read by reflecting laser

**Constant Linear Velocity**
- Unlike a magnetic disk, CD has only a single track - a very long spiral
- Constant packing density
  - Same numbers of bits/inch on outer edge and inner edge
  - Requires drive to change velocity
  - Higher speeds on inner part of disc
- Complicates seek operation; very slow on CDs

**CD Operation**
- Protective acrylic
- Label
- Polycarbonate plastic
- Land
- Hg
- Aluminum
- Laser transmit/receive

**CD-ROM Drive Speeds**
- Audio is “single” speed
  - Track (spiral) is 5.27km long
  - Gives 4391 seconds = 73.2 minutes
- Other speeds are quoted as multiples
  - e.g. 24x
- Quoted figure is maximum drive can achieve

**CD-ROM Block Format**
- Mode 0=blank data field
- Mode 1=2048 byte data+error correction
- Mode 2=2336 byte data
- CD-ROM Error Correction

**CD-ROM Error Correction**
- CD-ROMs use Reed-Solomon error correction, developed in the late 1960’s
  - Theory is based on abstract algebra and the details are beyond the scope of this course
  - Widely used for storage media including hard disks
  - Specific technique is called “Cross-Interleave Reed-Solomon Coding” (CIRC)
- Reed-Solomon error correction is particularly useful for “error bursts”
- For CD-ROMs error bursts as long as 450 bytes can be completely corrected (about 2.5mm on disk surface)
**2nd Level Error Correction**

- CD Drives in computers have a 2nd level of error correction that is not typically used in audio drives
- 4 byte cyclic redundancy check code in aux header allows detection of errors
- Additional 276 bytes of CIRC is used for correction
- Requires additional hardware on drive (embedded microprocessor) to handle error correction efficiently.

**Random Access on CD-ROM**

- Difficult to implement because of constant linear velocity
  1. Move head to rough position
  2. Set correct speed
  3. Read address
  4. Adjust to required location
  5. (Yawn!)

**CD-ROM pros and cons**

- Large capacity (compared to floppy disks)
- Easy to mass produce
- Removable
- Robust (but not permanent)
- Expensive for small runs
- Slow
- Read only

**CD-ROM Variants**

- CD-Recordable (CD-R)
  - WORM (Write once, read many)
  - Very inexpensive
  - Compatible with CD-ROM drives
- CD-Rewritable (CD-RW)
  - Erasable
  - Fairly expensive compared to CD-R
  - Mostly CD-ROM drive compatible
  - Phase change
    - Material has two different reflectivities in different phase states
    - Good for 500,000 + erase cycles

**DVD - what’s in a name?**

- Officially - nothing
- Digital Video Disk
  - Used to indicate a player for movies
  - Only plays video disks
- Digital Versatile Disk
  - Used to indicate a computer drive
  - Will read computer disks and play video disks
- Same size as CD but stores 6 times as much data

**DVD - technology**

- Multi-layers possible
  - DVDs with 2 layers and two sides can store 17GB
- Very high capacity (4.7G per layer)
- Full length movie on single disk
  - Using MPEG compression
- Finally standardized
- Movies carry regional coding
  - Players only play correct region films
CD and DVD

- Many different physical format standards
  - DVD-R
  - DVD+R
  - DVD+RW
  - DVD-RW
  - DVD-RAM
  - ... and more
- These differences are largely irrelevant to most people today
- Many modern drives can handle most formats but not necessarily with the cheapest media
- For more information:
  http://www.digitalfaq.com/guides/media/dvd-formats.htm

DVD -Writable

- Many different physical format standards
  - DVD-R
  - DVD+R
  - DVD+RW
  - DVD-RW
  - DVD-RAM
  - ... and more
- These differences are largely irrelevant to most people today
- Many modern drives can handle most formats but not necessarily with the cheapest media
- For more information:
  http://www.digitalfaq.com/guides/media/dvd-formats.htm

High Definition Optical Disks

- Designed for high-def video with greater storage capacity than DVDs.
- Higher density achieved with by using blue-violet laser with shorter wavelength than red lasers
- Two technologies (HD-DVD and Blu-Ray) competed for the market
- 2008 saw the withdrawal of HD-DVD and the dominance of Blu-Ray
- Blu-ray allows 25GB/50GB disks
- 1920×1080 pixel resolution at up to 60 frames per second interlaced or 24 frames per second progressive

Magnetic Tape

- Serial access
- Slow (very!)
- Very cheap
- Backup and archive
- Low cost of backup but high cost to restore
  - 2006 failure of 2-terabyte RAID 0 at Fogler Library required 30 days to restore from tape