Disk Controller

• One or more micro-controllers
  – Firmware
  – Positioning system controller (critical to performance and capacity)
  – Spindle motor controller

• A built-in cache (64KB-4MB)
  – Important to the performance of some workloads
  – Speed-matching buffer and disk block caching

• Some buffers (such as a track buffer)

• An Interface (SCSI, EIDE, Fibre Channel, etc)
  – Transfer data between host and controller buffer
  – External transfer rate: 20-200 MB/sec

• A read channel
  – Internal transfer rate: 15-60 MB/sec
  – Transfer data between controller buffer and disk platter
  – Disk bandwidth is the total number of bytes transferred, divided by the total time between the first request for service and the completion of the last transfer.
Disk Controller-1

- Bus time
  - SCSI bus contention/arbitration
  - Data/command transfer time
  - Command read time
  - Firmware overhead (0.3-1ms)
Disk Controller-2-caching

- Cache management algorithms
  - Segment (cache line, same size)
  - Segment size matches # sectors per track
  - Dynamically modify # of segments
  - Overlap?

- **Read-ahead/track buffer**
  - A read that hits in the cache can be satisfied “immediately”
  - Zero-latency read for whole-track transfers

- **Write-back/write-through**
  - Immediate reporting for latency
  - Throughput benefits:
    - Data in a write buffer are often overwritten in place
    - Large # of stored writes provides scheduling chances for near-optimal fashion
Disk Controller -3

- Command queuing
  - Support multiple outstanding requests
  - Can perform disk requests scheduling (FCFS, SSTF…) by firmware, rather than external scheduler (device driver)
    - More accurate info about head position and various positioning delays
  - Exploit inter-request concurrency among command processing, bus transfer, media access components
Time Overlap

**Read**
- Host sends command
- Controller decodes it
- SCSI bus data transfers to host
- Status message to host
- Controller disconnects from bus & starts seek
- Disk mechanism

**Write**
- Host sends command
- Controller decodes it
- SCSI bus data transfer from host
- Status message to host
- Controller starts seek
- Disk mechanism
State of the Art: Barracuda 180

- 181.6 GB, 3.5 inch disk
- 12 platters, 24 surfaces
- 24,247 cylinders
- 7,200 RPM; (4.2 ms avg. latency)
- 7.4/8.2 ms avg. seek (r/w)
- 64 to 35 MB/s (internal)
- 0.1 ms controller time
- 10.3 watts (idle)

source: www.seagate.com

Latency = Queuing Time + Controller time + Seek Time + Rotation Time + Size / Bandwidth

per access +

source: www.seagate.com