Microprocessor Design & Organisation HCA2102

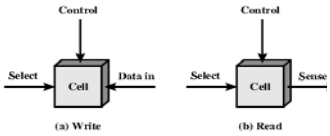
Internal & External Memory

Semiconductor Memory Types

Memory Type	Category	Erasure	Write Mechanism	Volatility
Random-access memory (RAM)	Read-write memory	Electrically, byte-level	Electrically	Volatile
Read-only memory (ROM)	Read-only memory	Not possible	Masks	Nonvolatile
Programmable ROM (PROM)			Electrically	
Erasable PROM (EPROM)	Read-mostly memory	UV light, chip-level		
Electrically Erasable PROM (EEPROM)		Electrically, byte-level		
Flash memory		Electrically, block-level		

Semiconductor Memory

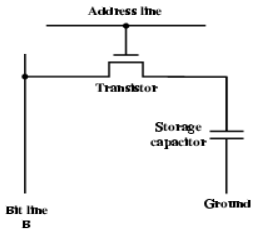
- RAM
  - Misnamed as all semiconductor memory is random access
  - Read/Write
  - Volatile
  - Temporary storage
  - Static or dynamic



Memory Cell Operation

Dynamic RAM

- Bits stored as charge in capacitors
- Charges leak
- Need refreshing even when powered
- Simpler construction
- Smaller per bit - Less expensive
- Need refresh circuits
- Slower
- Main memory
- Essentially analogue
  - Level of charge determines value

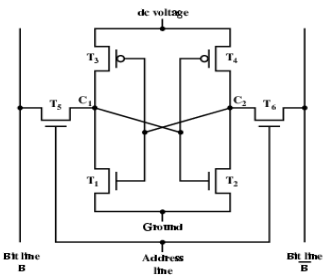


DRAM Operation

- Address line active when bit read or written
  - Transistor switch closed (current flows)
- Write
  - Voltage to bit line
    - High for 1 low for 0
  - Then signal address line
    - Transfers charge to capacitor
- Read
  - Address line selected
    - transistor turns on
  - Charge from capacitor fed via bit line to sense amplifier
    - Compares with reference value to determine 0 or 1
  - Capacitor charge must be restored

Static RAM

- Bits stored as on/off switches
- No charges to leak
- No refreshing needed when powered
- More complex construction
- Larger per bit - More expensiv
- Does not need refresh circuits
- Faster
- Cache
- Digital
  - Uses flip-flops



Static RAM Operation

- Transistor arrangement gives stable logic state
- State 1
  - C_1 high, C_2 low
  - T_1 T_4 off, T_2 T_3 on
- State 0
  - C_2 high, C_1 low
  - T_2 T_3 off, T_1 T_4 on
- Address line transistors T_5 T_6 is switch
- Write – apply value to B & complement to $\overline{B}$
- Read – value is on line B

SRAM v DRAM

- Both volatile
  - Power needed to preserve data
- Dynamic cell
  - Simpler to build, smaller
  - More dense
  - Less expensive
  - Needs refresh
  - Larger memory units
- Static
  - Faster
  - Cache

Read Only Memory (ROM)

- Permanent storage
  - Non-volatile
- Microprogramming (see later)
- Library subroutines
- Systems programs (BIOS)
- Function tables

Types of ROM

- Written during manufacture
  - Very expensive for small runs
- Programmable (once)
  - PROM
  - Needs special equipment to program
- Read “mostly”
  - Erasable Programmable (EPROM)
    - Erased by UV
  - Electrically Erasable (EEPROM)
    - Takes much longer to write than read
  - Flash memory
    - Erase whole memory electrically

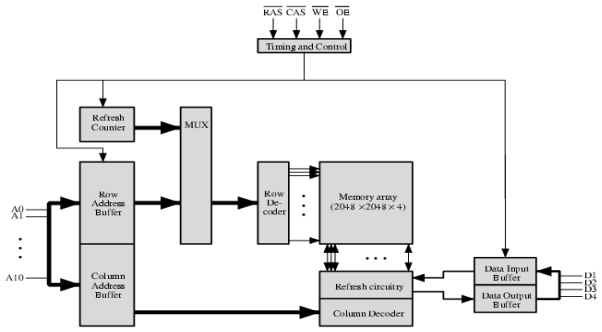
Organisation in detail

- A 16Mbit chip can be organised as 1M of 16 bit words
- A bit per chip system has 16 lots of 1Mbit chip with bit 1 of each word in chip 1 and so on
- A 16Mbit chip can be organised as a 2048 x 2048 x 4bit array
  - Reduces number of address pins
    - Multiplex row address and column address
    - 11 pins to address (2^{11}=2048)
    - Adding one more pin doubles range of values so x4 capacity

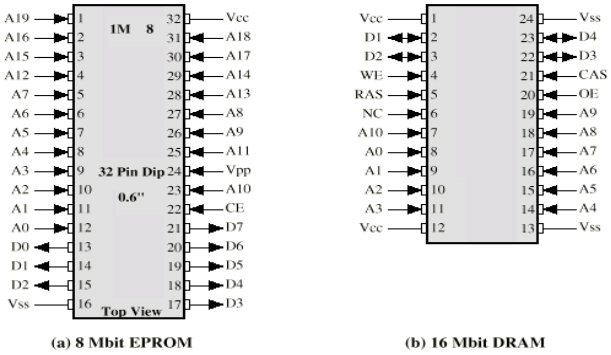
Refreshing

- Refresh circuit included on chip
- Disable chip
- Count through rows
- Read & Write back
- Takes time
- Slows down apparent performance

Typical 16 Mb DRAM (4M x 4)



Packaging



Error Correction

- Hard Failure
  - Permanent defect
- Soft Error
  - Random, non-destructive
  - No permanent damage to memory
- Detected using Hamming error correcting code

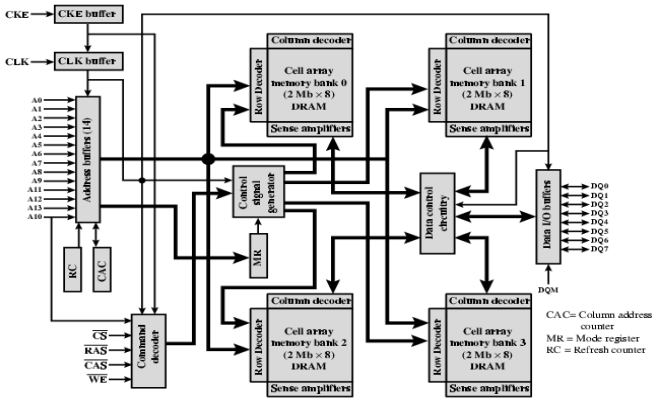
Advanced DRAM Organization

- Basic DRAM same since first RAM chips
- Enhanced DRAM
  - Contains small SRAM as well
  - SRAM holds last line read (c.f. Cache!)
- Cache DRAM
  - Larger SRAM component
  - Use as cache or serial buffer

Synchronous DRAM (SDRAM)

- Access is synchronized with an external clock
- Address is presented to RAM
- RAM finds data (CPU waits in conventional DRAM)
- Since SDRAM moves data in time with system clock, CPU knows when data will be ready
- CPU does not have to wait, it can do something else
- Burst mode allows SDRAM to set up stream of data and fire it out in block
- DDR-SDRAM sends data twice per clock cycle (leading & trailing edge)

IBM 64Mb SDRAM



Types of External Memory

- Magnetic Disk
  - RAID
  - Removable
- Optical
  - CD-ROM
  - CD-Recordable (CD-R)
  - CD-R/W
  - DVD+-R/RW
- Magnetic Tape

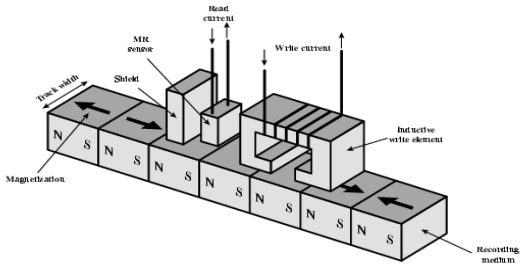
Magnetic Disk

- Disk substrate coated with magnetizable material (iron oxide…rust)
- Substrate used to be aluminium
- Now glass
  - 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 and 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

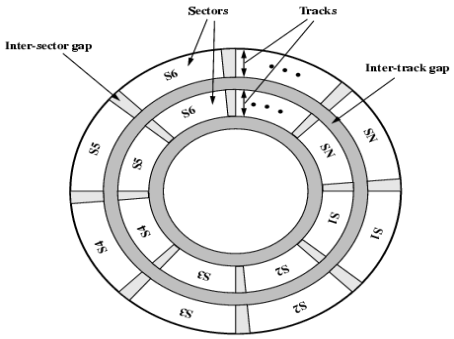
Inductive Write MR Read



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
- Tracks divided into sectors
- Minimum block size is one sector
- May have more than one sector per block

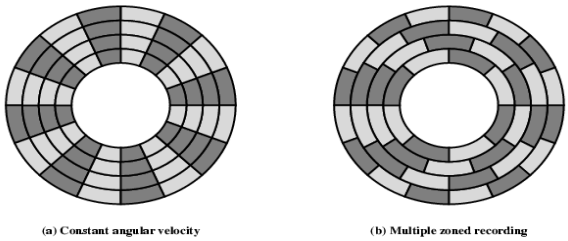
Disk Data Layout



Disk Velocity

- Bit near centre of rotating disk passes fixed point slower than bit on outside of disk
- Increase spacing between bits in different tracks
- 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
  - Each zone has fixed bits per track
  - More complex circuitry

Disk Layout Methods Diagram



(a) Constant angular velocity (b) Multiple zoned recording

Characteristics

- Fixed (rare) or movable head
- Removable or fixed
- Single or double (usually) sided
- Single or multiple platter
- Head mechanism
  - Contact (Floppy)
  - Fixed gap
  - Flying (Winchester)

Fixed/Movable Head Disk

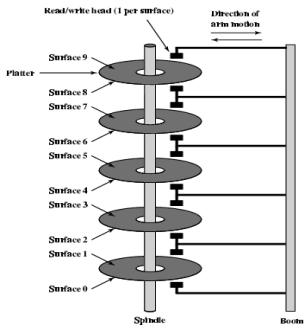
- Fixed head
  - One read write head per track
  - Heads mounted on fixed ridged arm
- Movable head
  - One read write head per side
  - Mounted on a movable arm

Removable or Not

- Removable disk
  - Can be removed from drive and replaced with another disk
  - Provides unlimited storage capacity - Easy data transfer between systems
- Non-removable 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)



Winchester Hard Disk

- Developed by IBM in Winchester (USA)
- Sealed unit
- One or more platters (disks)
- Heads fly on boundary layer of air as disk spins
- Very small head to disk gap
- Getting more robust
- Universal -Cheap
- Fastest external storage
- Getting larger all the time - 100s of Gigabyte now usual

Floppy Disk

- 8", 5.25", 3.5"
- Small capacity
  - Up to 1.44Mbyte (2.88M never popular)
- Slow
- Universal
- Cheap
- Obsolete?

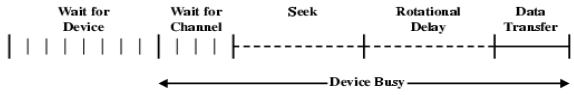
Removable Hard Disk

- ZIP
  - Cheap - Very common
  - Only 100M
- JAZZ
  - Not cheap
  - 1GB
- L-120 (floppy drive)
  - Also reads 3.5" floppy
  - Becoming more popular?
- All obsoleted by CD-R and CD-R/W?

Speed

- Seek time
  - Moving head to correct track
- (Rotational) latency
  - Waiting for data to rotate under head
- Access time = Seek + Latency
- Transfer rate

Timing of Disk I/O Transfer



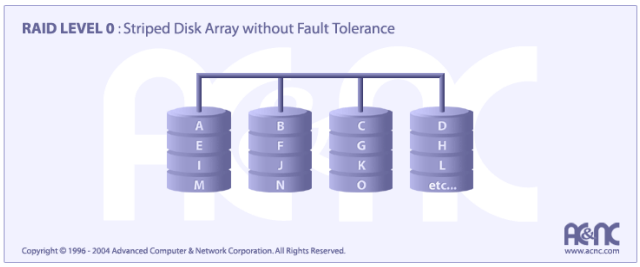
RAID

- Redundant Array of Independent Disks
- Redundant Array of Inexpensive Disks
- 10 levels in common use
- Not a hierarchy
- 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.

RAID 0 – Disk Striping

- No redundancy – Min. 2 disks
- Data striped across all disks
- Round Robin striping
- Increase I/O speed
  - Multiple data requests probably not on same disk
  - Disks seek in parallel
  - Can read/write blocks on one drive while seeking on another
  - A set of data is likely to be striped across multiple disks

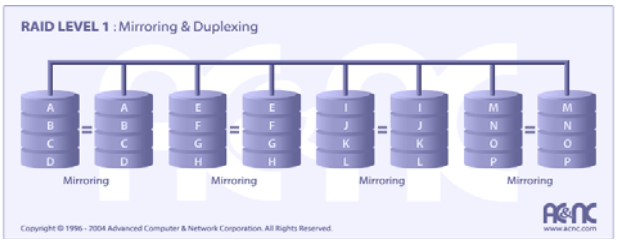
Raid 0 at work



RAID 1 – Disk Mirroring

- Provides Redundancy – Min. 2 **equal** size disks
- Same data is written to each disk in array.
- Read from either
- Write to all
- High Data Reliability:
  - Recovery is simple
    - Swap faulty disk & re-mirror
    - No down time
- Expensive

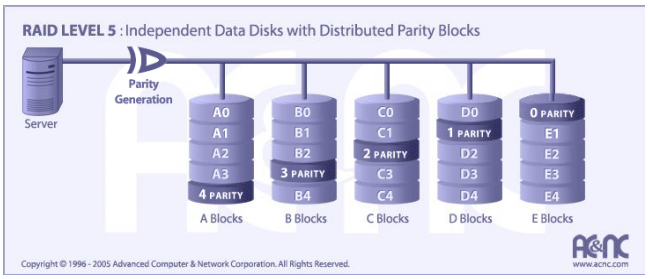
RAID 1 at work



RAID 5 – Disk Striping with parity

- Provides redundancy – Min 3 disks required
- Parity striped across all disks
- Round robin allocation for parity stripe
- High read but medium write performance
- Commonly used in network servers
- Most popular: balance redundancy & cost

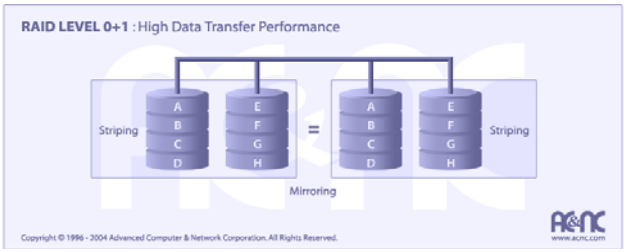
RAID 5 at work



RAID 0+1 – Disk Stripping with Mirroring

- Provides redundancy – Min. 2 similar disks pairs (in effect min. 4 disks)
- High Data transfer Performance
- Limited data reliability
- Failure of any disk will cause whole array to become a RAID 0 array.
- Very expensive

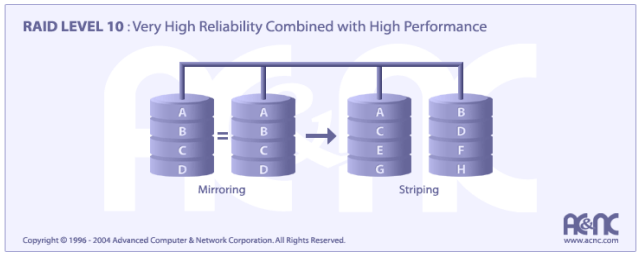
RAID 0+1 at work



RAID 10 – Disk Mirroring with Striping

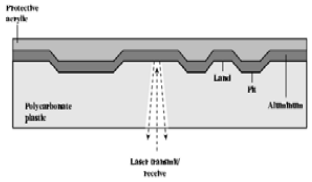
- Provides redundancy – Min. 2 similar disks pairs (in effect min. 4 disks)
- High Data transfer Performance
- High Data reliability
- Can sustain multiple simultaneous disk failures under certain conditions.
  - More expensive than RAID 0+1

RAID 10 at work



Optical Storage CD-ROM

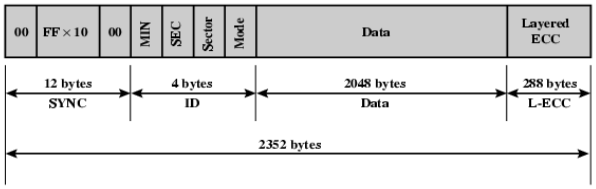
- Originally for audio
- 700 Megabytes giving over 80 minutes audio
- Polycarbonate coated with highly reflective coat, usually aluminium
- Data stored as pits
- Read by reflecting laser
- Constant packing densit
- Constant linear velocity



CD-ROM Drive Speeds

- Audio is single speed
  - Constant linear velocity
  - 1.2 ms^{-1}
  - Track (spiral) is 5.27 km long !
  - Gives 4391 seconds = 73.2 minutes
- Other speeds are quoted as multiples
- e.g. 24x (each x is 150 kB/s)
- Quoted figure is maximum speed the drive can achieve.

CD-ROM Format



- Mode 0=blank data field
- Mode 1=2048 byte data+error correction
- Mode 2=2336 byte data

Random Access on CD-ROM

- Difficult
- Move head to rough position
- Set correct speed
- Read address
- Adjust to required location

CD-ROM pros & cons

- Large capacity - Removable - Robust
- Expensive for small runs – Slow - Read only

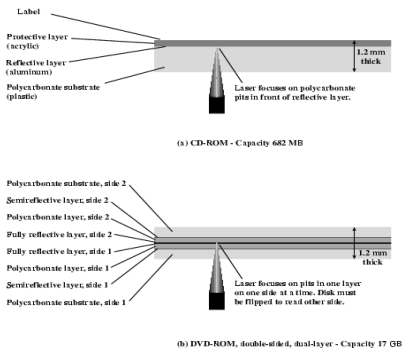
Other Optical Storage

- CD-Recordable (CD-R)
  - WORM
  - Now affordable
  - Compatible with CD-ROM drives
- CD-RW
  - Erasable
  - Getting cheaper
  - Mostly CD-ROM drive compatible
  - Phase change
    - Material has two different reflectivities in different phase states

DVD - Technology

- Digital Video Disk or Digital Versatile Disk
- Single or Dual sided and layered
- Very high capacity (4.7 GB per layer per side)
- Full length movie on single disk
  - Using MPEG-2 compression
- Finally standardized
- Movies carry regional coding
- Players only play correct region films - Can be “fixed”

CD vs. DVD



Magnetic Tape

- Serial access
- Slow
- Very cheap
- Backup and archive

Digital Audio Tape (DAT)

- Uses rotating head (like video)
- High capacity on small tape
  - 4 Gigabyte uncompressed
  - 8 Gigabyte compressed
- Backup of PC/network servers