Memory
Lecture Content

- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

Memory
Learning Objectives

- Understand the concepts of semiconductor memory
- Understand the operations of error detecting and correcting code
- Familiarize yourself with modern DRAM
- Understand the different concepts and organizations of external memory

Internal Memory

- In the past, main memory employed an array of doughnut-shaped ferromagnetic loops referred to as cores
- Hence, main memory was an is often referred to as core

- Today, the use of semiconductor chips for main memory is almost universal
- The basic element is the memory cell:
  - A variety of electronic technologies are used
  - They exhibit two stable (or semistable) states; binary 1 and 0.
  - They are capable of being written into (at least once), to set the state.
  - They are capable of being read to sense the state.

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)				
Erasable PROM (EPROM)	Read-mostly memory	UV light, chip-level	Electrically	
Electrically Erasable PROM (EEPROM)		Electrically, byte-level		
Flash memory		Electrically, block-level		

Two Different RAM Technologies

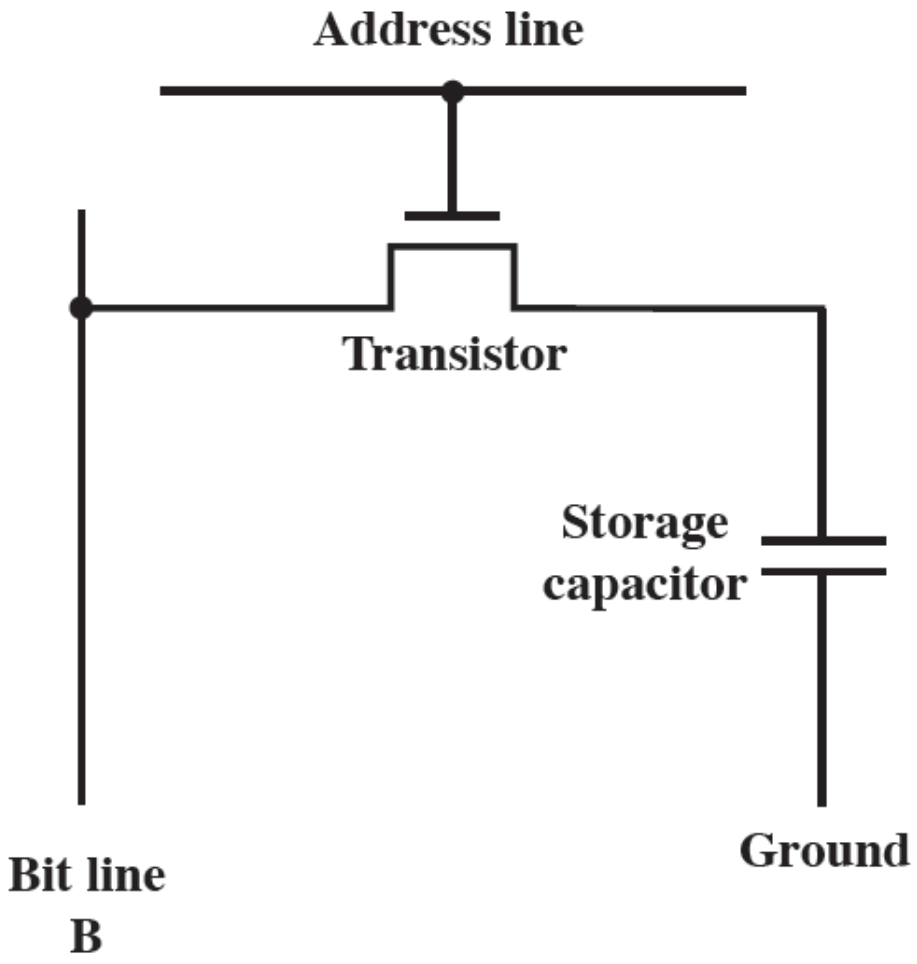
- **DRAM**
  - Dynamic RAM (DRAM)
  - Made with cells that store data as charge on capacitors
  - Presence or absence of charge in a capacitor is interpreted as a binary 1 or 0
  - Requires periodic charge refreshing to maintain data storage
  - The term dynamic refers to tendency of the stored charge to leak away, even with power continuously applied

Two Different RAM Technologies

- **DRAM**
  - Dynamic RAM (DRAM)
  - Made with cells that store data as charge on capacitors
  - Presence or absence of charge in a capacitor is interpreted as a binary 1 or 0
  - Requires periodic charge refreshing to maintain data storage
  - The term dynamic refers to tendency of the stored charge to leak away, even with power continuously applied

- **SRAM**
  - Static RAM (SRAM)
  - Digital device that uses the same logic elements used in the processor
  - Binary values are stored using traditional flip-flop logic gate configurations
  - Will hold its data as long as power is supplied to it

DRAM


Address line
Transistor
Storage
capacitor
Ground
Bit line
B

SRAM vs. DRAM

- Both are volatile
  - Power must be continuously supplied to preserve the bit values
- Dynamic cell
  - Simpler to build, smaller
  - Denser packed and less expensive
  - Requires the supporting refresh circuitry
  - Tend to be favored for large memory requirements
  - Used for main memory
- Static
  - Faster
  - Used for cache memory (both on and off chip)

Memory
Lecture Content

- **Internal Memory**
  - General
  - **Error Correction Codes**
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

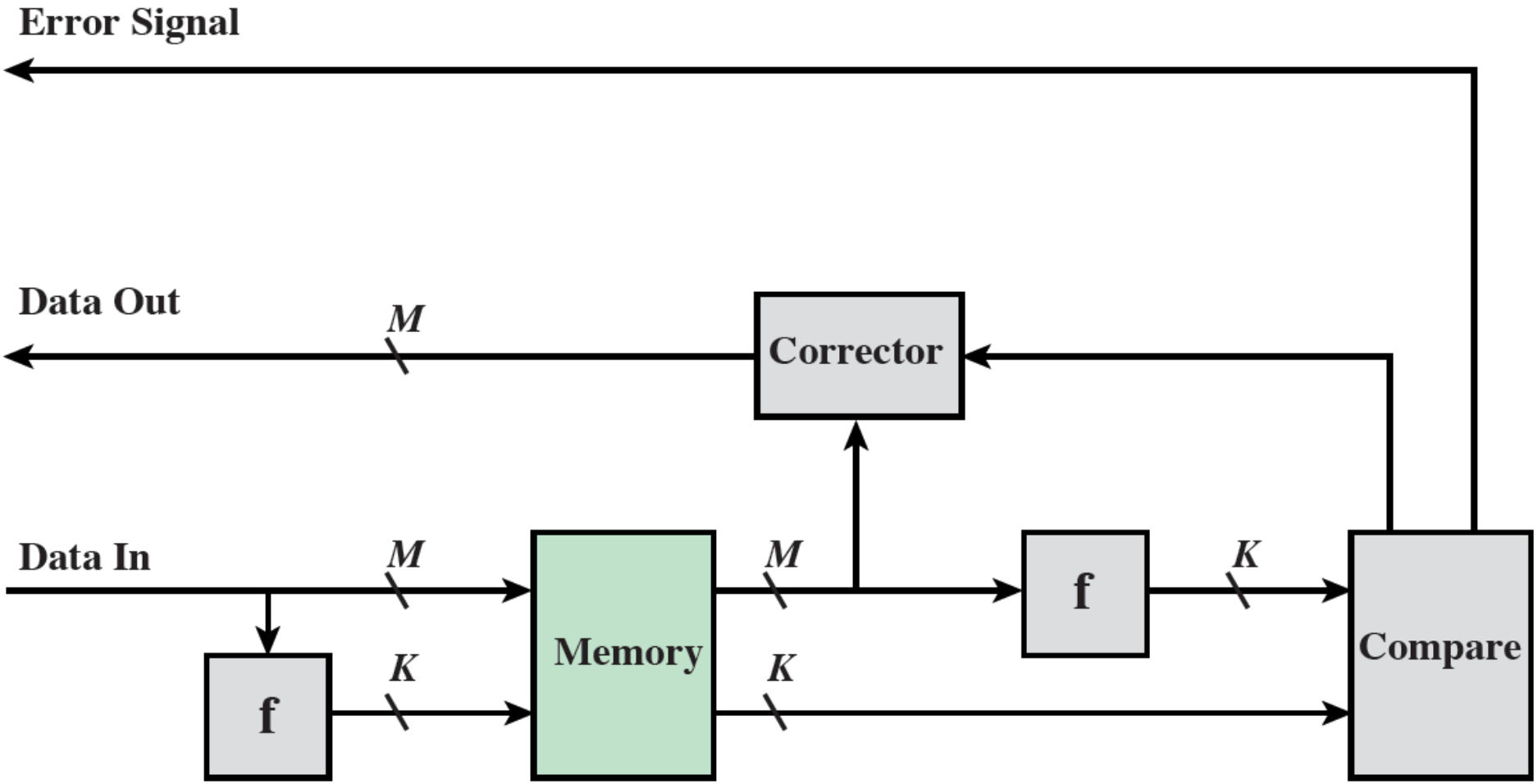
Memory is Subject to Errors

- **Hard Failure**
  - Permanent physical defect
  - Memory cell or cells affected cannot reliably store data but become stuck at 0 or 1 or switch erratically between 0 and 1
  - Can be caused by:
    - Harsh environmental abuse
    - Manufacturing defects
    - Wear

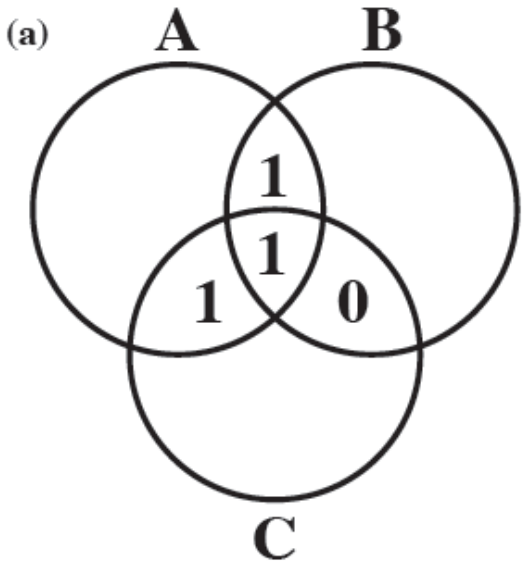
Memory is Subject to Errors

- **Hard Failure**
  - Permanent physical defect
  - Memory cell or cells affected cannot reliably store data but become stuck at 0 or 1 or switch erratically between 0 and 1
  - Can be caused by:
    - Harsh environmental abuse
    - Manufacturing defects
    - Wear
- **Soft Error**
  - Random, non-destructive event that alters the contents of one or more memory cells
  - No permanent damage to memory
  - Can be caused by:
    - Power supply problems
    - Alpha particles

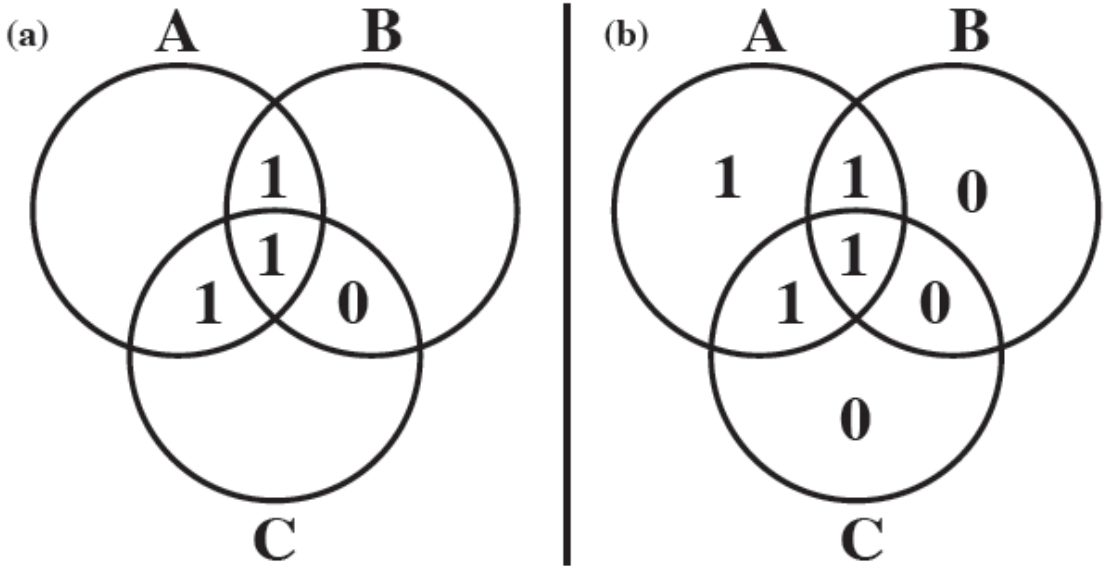
Error Correcting Codes



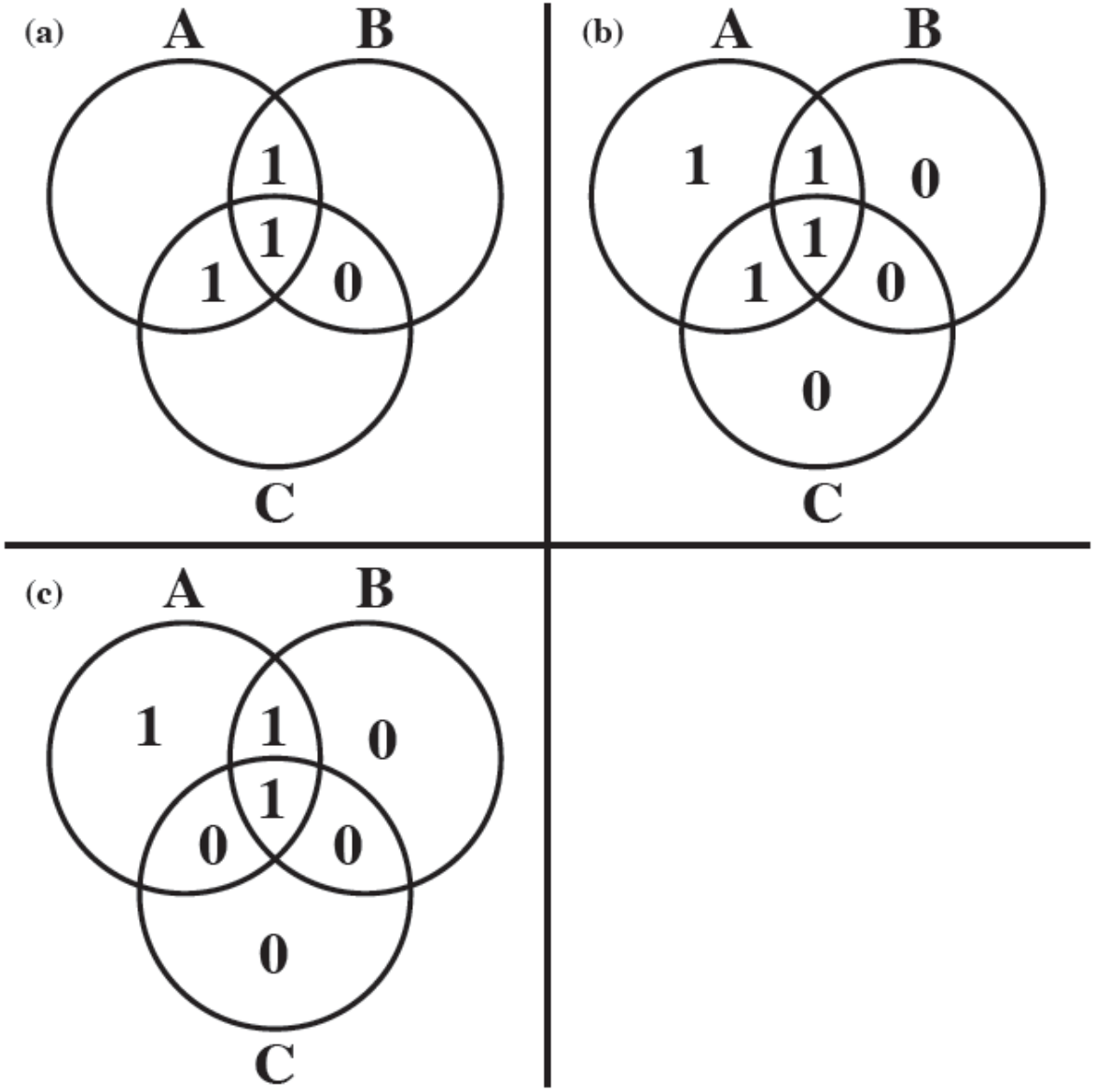
Hamming Error Correction Code



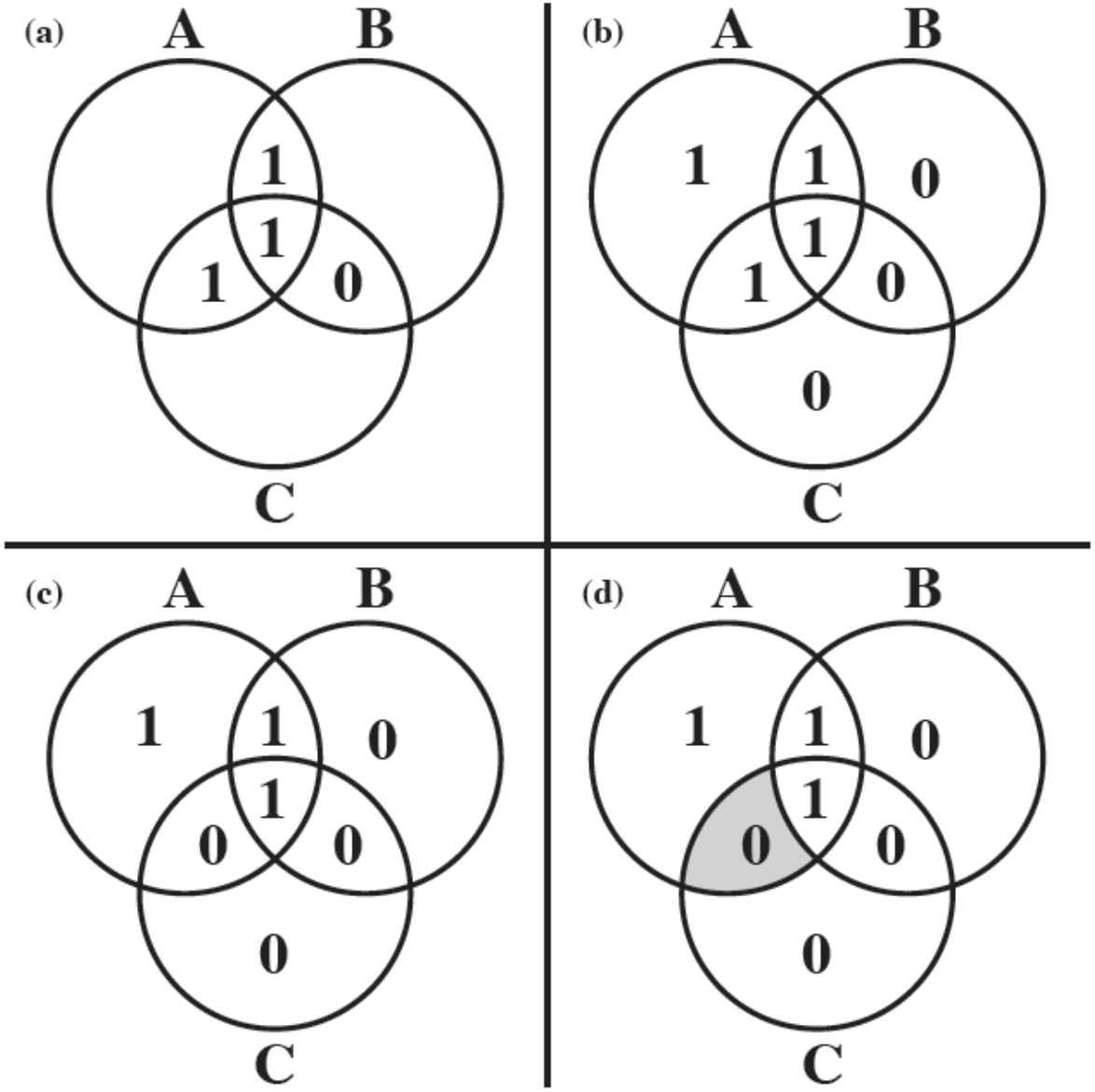
Hamming Error Correction Code



Hamming Error Correction Code



Hamming Error Correction Code



Hamming Error Correction Code

- The general Procedure is as follows:
  - A check bit is included on every 2^N position, e.g.:
    - …0001
    - …0010
    - …0100
    - …1000
  - For the parity calculation, all successive bits which index in binary representation has a one on that position, are included.
  - For 10, that would be:
    - …011
    - …110
    - …111
    - …

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Calculating the check bits:
  - XOR all indicated bits in a row (remember XOR is 1 if there is an odd number of 1s)

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Calculating the check bits:
  - XOR all indicated bits in a row (remember XOR is 1 if there is an odd number of 1s)
  - $c_0 = d_0 \oplus d_1 \oplus d_3 \oplus d_4 \oplus d_6 \oplus d_8 \oplus d_{10}$

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Calculating the check bits:
  - XOR all indicated bits in a row (remember XOR is 1 if there is an odd number of 1s)
  - $c_0 = d_0 \oplus d_1 \oplus d_3 \oplus d_4 \oplus d_6 \oplus d_8 \oplus d_{10}$
  - $c_1 = d_0 \oplus d_2 \oplus d_3 \oplus d_5 \oplus d_6 \oplus d_9 \oplus d_{10}$

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Calculating the check bits:
  - XOR all indicated bits in a row (remember XOR is 1 if there is an odd number of 1s)
  - $c_0 = d_0 \oplus d_1 \oplus d_3 \oplus d_4 \oplus d_6 \oplus d_8 \oplus d_{10}$
  - $c_1 = d_0 \oplus d_2 \oplus d_3 \oplus d_5 \oplus d_6 \oplus d_9 \oplus d_{10}$
  - $c_2 = d_1 \oplus d_2 \oplus d_3 \oplus d_7 \oplus d_8 \oplus d_9 \oplus d_{10}$
  - $c_3 = d_4 \oplus d_5 \oplus d_6 \oplus d_7 \oplus d_8 \oplus d_9 \oplus d_{10}$

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Calculating the check bits:
  - XOR all indicated bits in a row
  - No parity bit is part of another parity bit's calculation
  - How does this help us?

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Lets assume we have an error in bit 9

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- Lets assume we have an error in bit 9
- That means the parity pits 0 and 3 are erroneous

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- As the parity bits are set to the positions 2^N, they work as binary code to encode the erroneous bit

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- As the parity bits are set to the positions 2^N, they work as binary code to encode the erroneous bit
- Assume, the parity bits c_1, c_2, c_3 indicate an error
  - That means in binary encoding 1110
  - => We have an error in position 14

Hamming Error Correction Code

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		d_0		d_1	d_2	d_3		d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
c_0		x		x		x		x		x		x		x
	c_1	x			x	x			x	x			x	x
			c_2	x	x	x					x	x	x	x
							c_3	x	x	x	x	x	x	x

- As the parity bits are set to the positions 2^N, they work as binary code to encode the erroneous bit
- Assume, the parity bits c_1, c_2, c_3 indicate an error
  - That means in binary encoding 1110
  - => We have an error in position 14

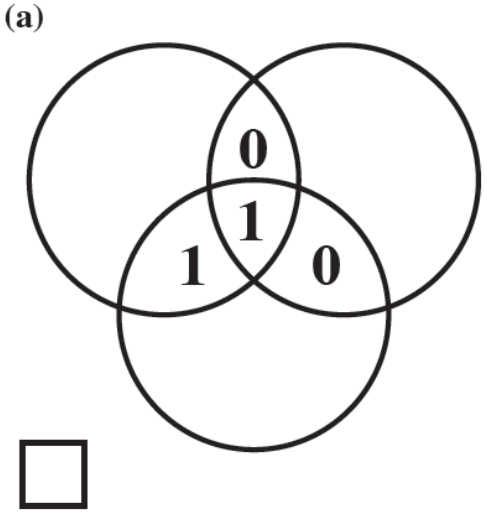
Summary

- The Hamming Code:
  - **Single-Error-Correcting (SEC)** code
    - Error detecting and correction for one faulty bit
  - Easy to calculate

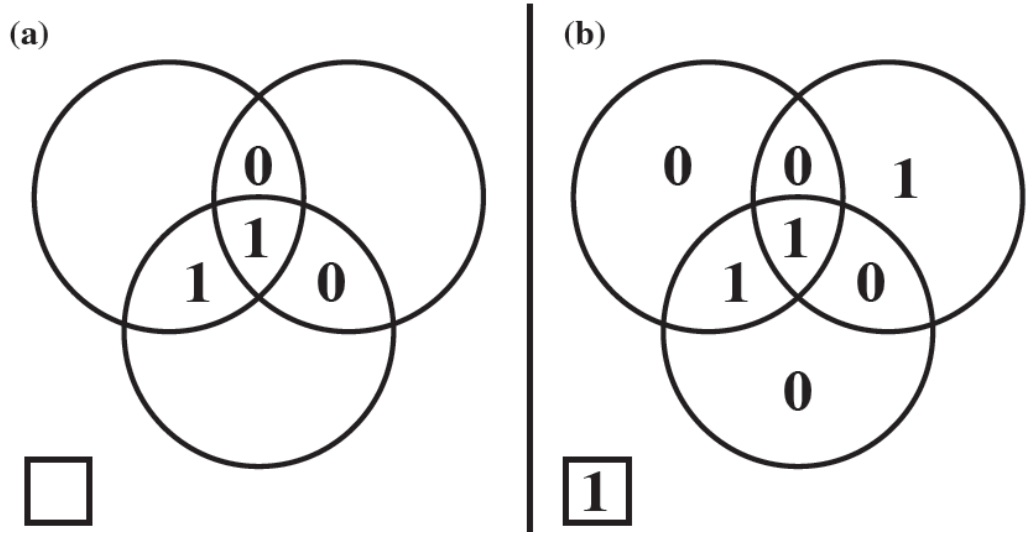
- What happens if more than one error occurs
  - Basically two possibilities
  - The error goes undetected
  - The error detection goes berserk and introduces an third error

- **Single-Error-Correcting, Double-Error-Detecting (SEC-DED)** code
  - Introduces one additional bit
  - Is considered okay in terms overhead vs. additional security

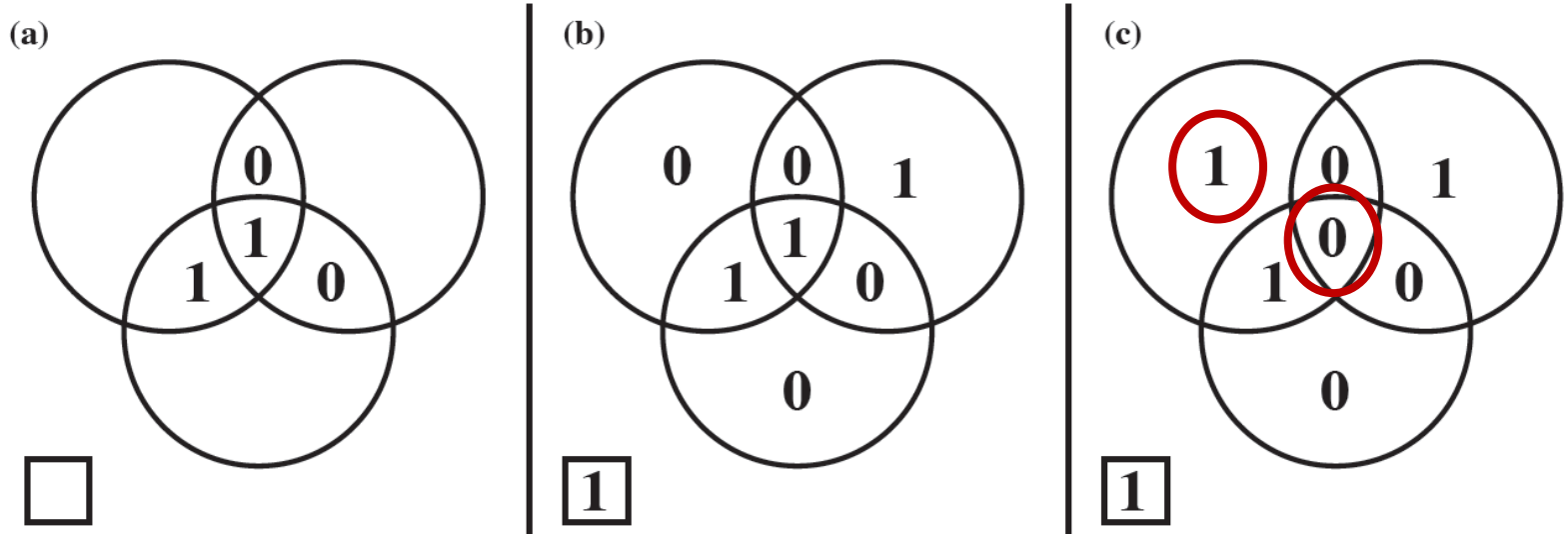
Hamming SEC-DED Code



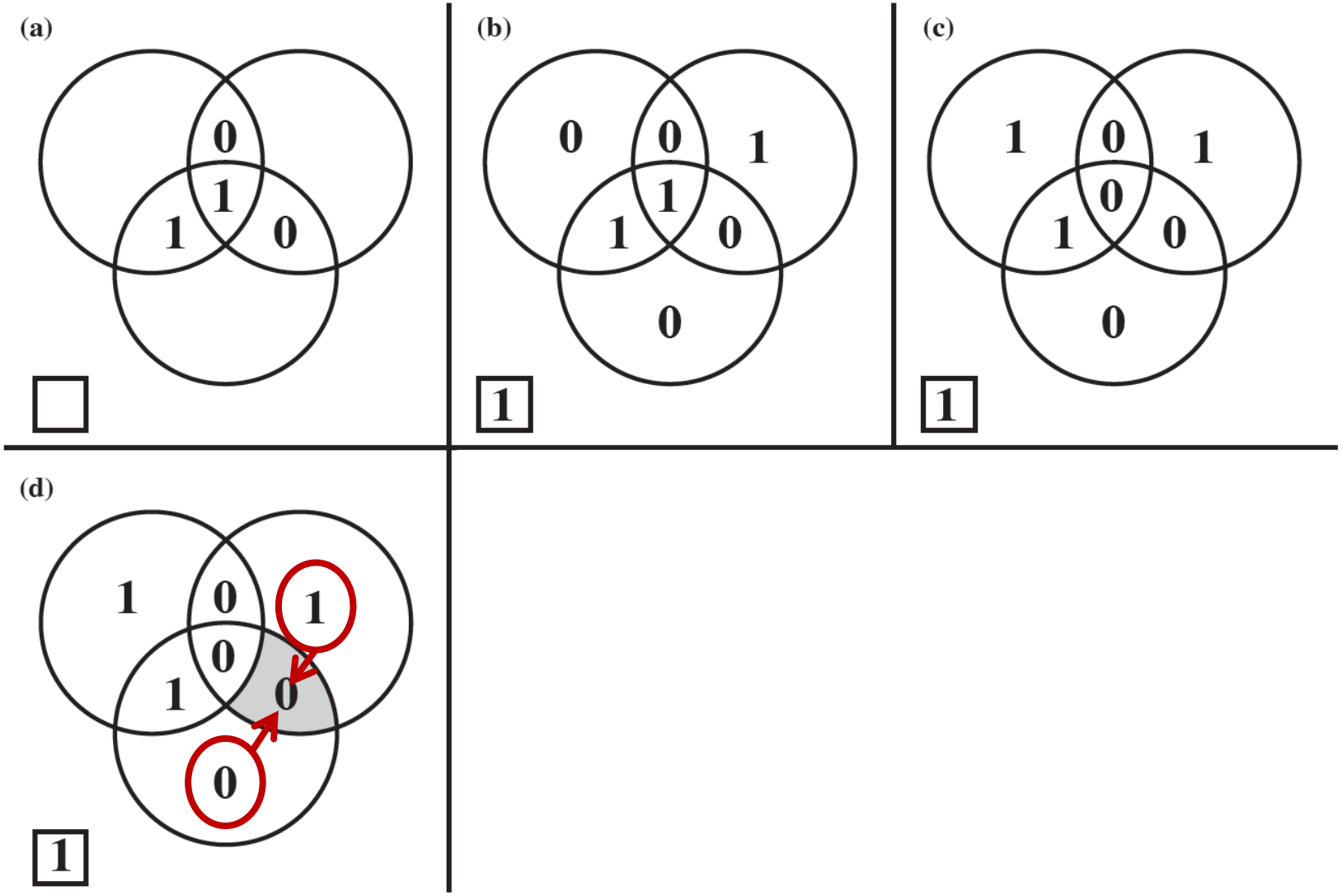
Hamming SEC-DED Code



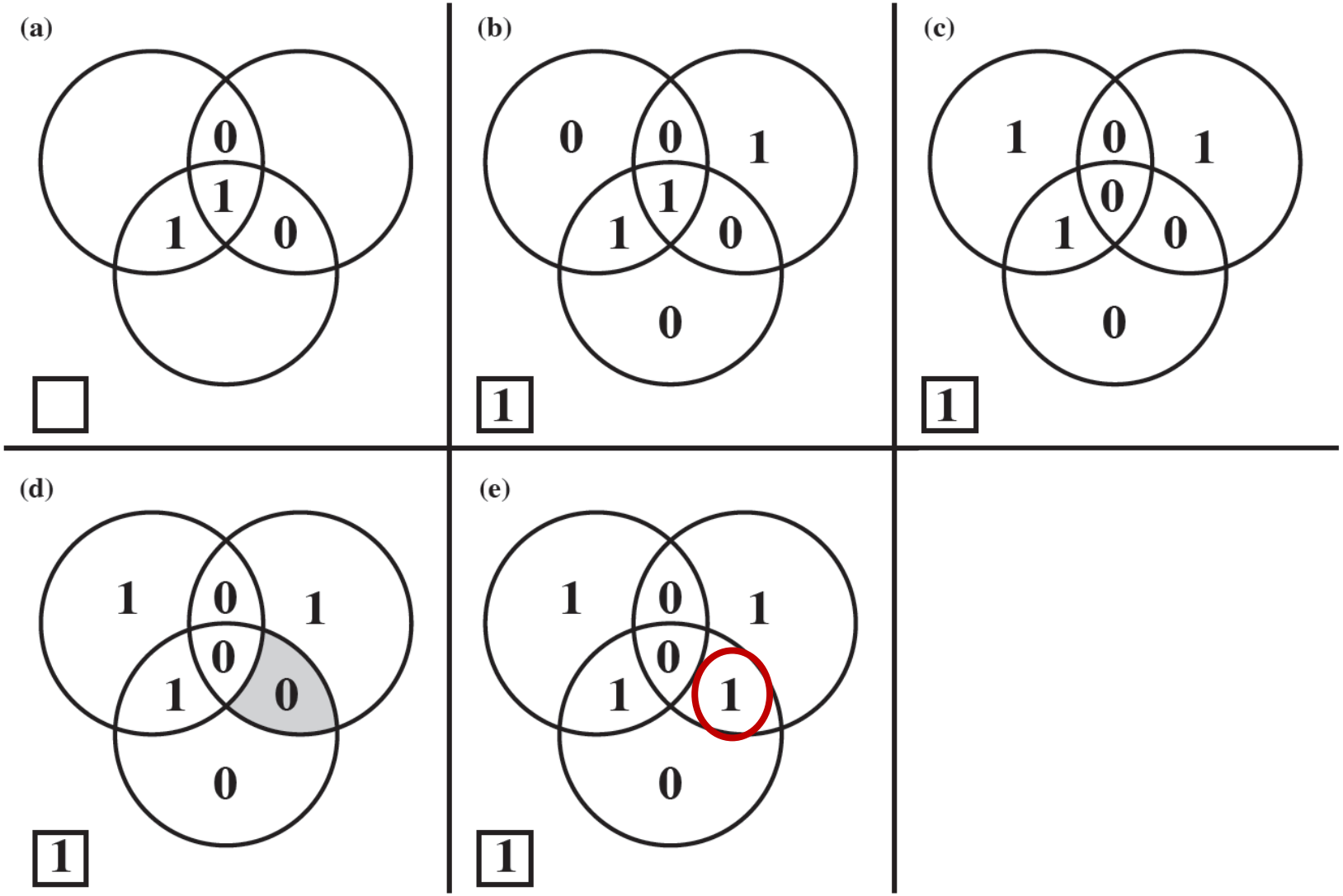
Hamming SEC-DED Code



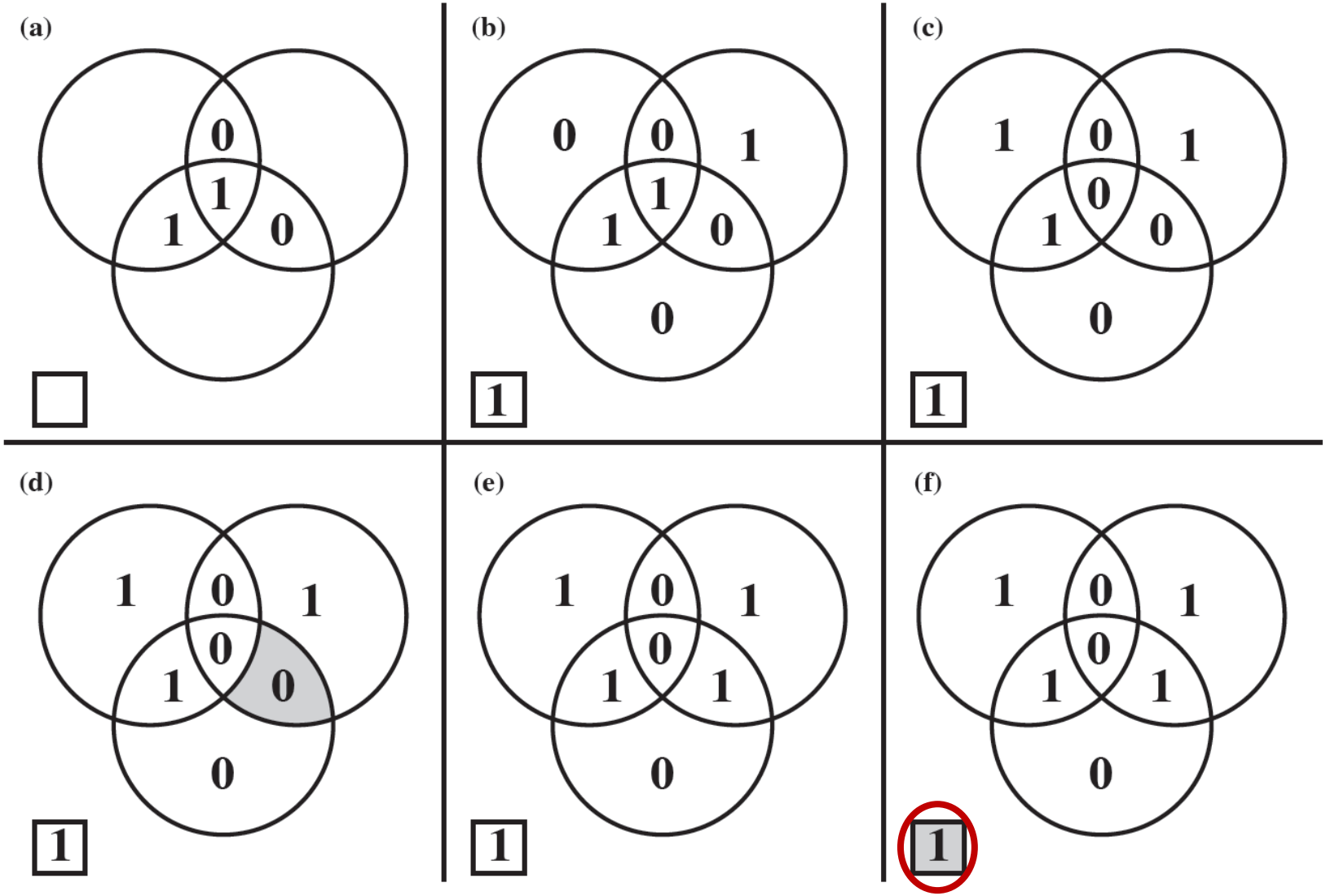
Hamming SEC-DED Code



Hamming SEC-DED Code



Hamming SEC-DED Code



Increase in Word Length with Error Correction

	Single-Error Correction		Single-Error Correction/ Double-Error Detection	
Data Bits	Check Bits	% Increase	Check Bits	% Increase
8	4	50	5	62.5
16	5	31.25	6	37.5
32	6	18.75	7	21.875
64	7	10.94	8	12.5
128	8	6.25	9	7.03
256	9	3.52	10	3.91

Memory
Lecture Content

- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

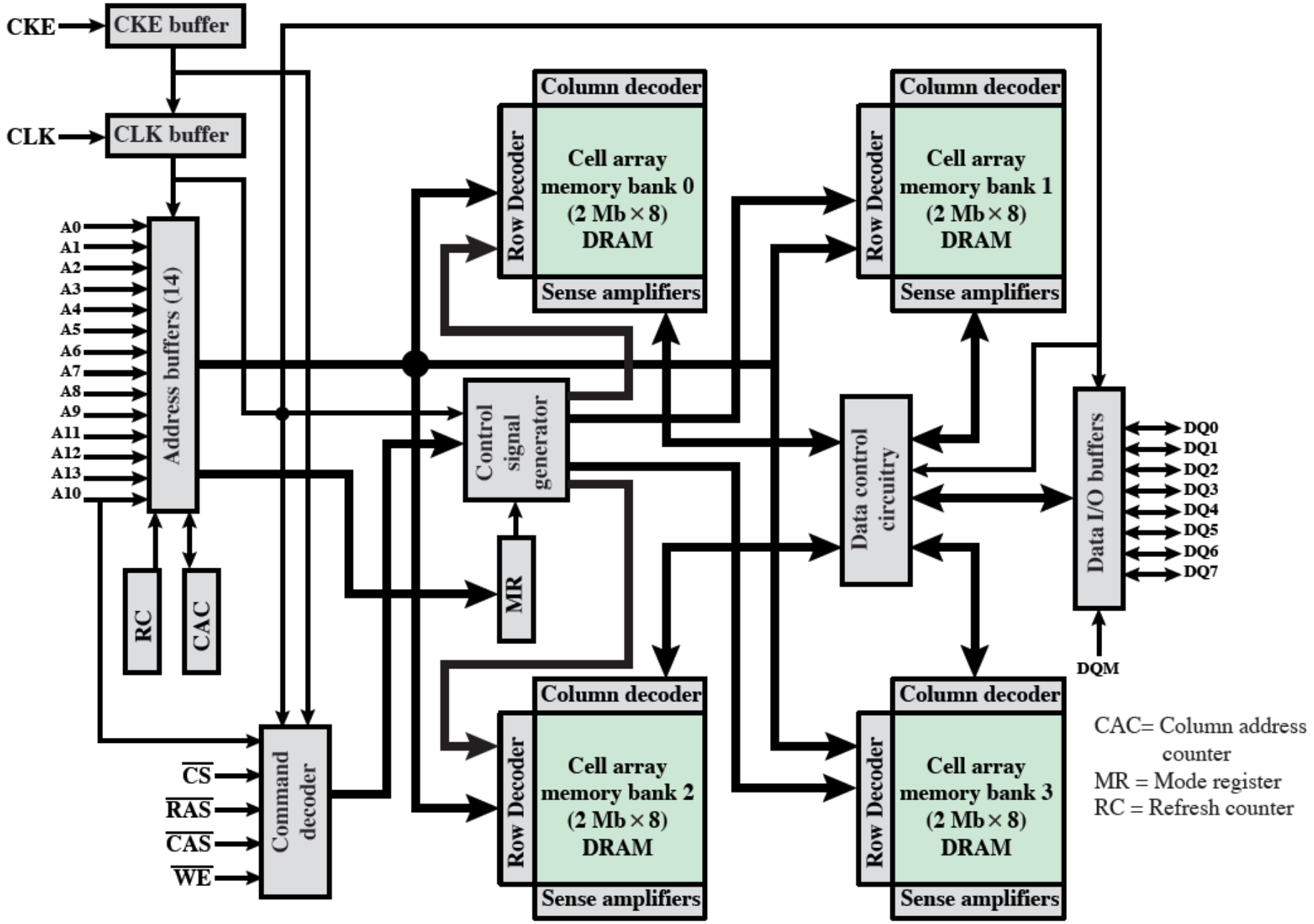
Advanced DRAM Organization

- One of the most critical system bottlenecks when using high-performance processors is the interface to main internal memory
- The traditional DRAM chip is constrained both by its internal architecture and by its interface to the processor's memory bus
- A number of enhancements to the basic DRAM architecture have been explored:

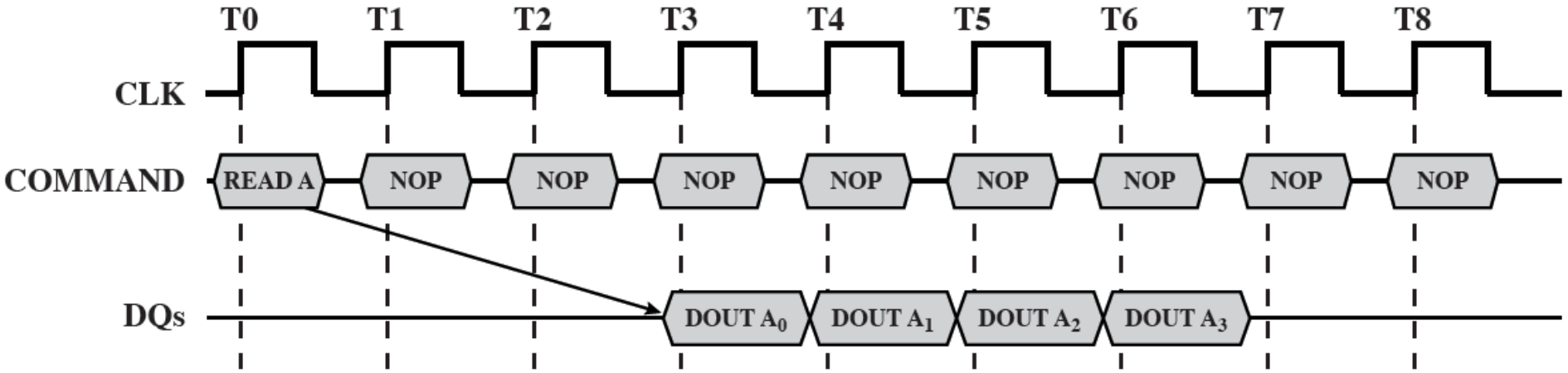
	Clock Frequency (MHz)	Transfer Rate (GB/s)	Access Time (ns)	Pin Count
SDRAM	166	1.3	18	168
DDR	100/200	3.2	12.5	184
DDR4-1600	200/800/1600	12.8	10	288
DDR4-3200	400/1600/3200	25.6	10	288
RDRAM	300/600	4.8	12	162

SDRAM

- One of the most widely used forms of DRAM
- Exchanges data with the processor synchronized
  - Running at the full speed of the processor/memory bus without imposing wait states
- With synchronous access the DRAM moves data in and out under control of the system clock
  - The processor or other master issues the instruction and address information which is latched by the DRAM
  - The DRAM then responds after a set number of clock cycles
  - Meanwhile the master can safely do other tasks while the SDRAM is processing


CKE
CKE buffer
CLK
CLK buffer
A0
A1
A2
A3
A4
A5
A6
A7
A8
A9
A11
A12
A13
A10
Address buffers (14)
RC
CAC
Column decoder
Row Decoder
Cell array memory bank 0 (2 Mb × 8) DRAM
Sense amplifiers
Column decoder
Row Decoder
Cell array memory bank 1 (2 Mb × 8) DRAM
Sense amplifiers
Control signal generator
MR
Data control circuitry
Data I/O buffers
DQ0
DQ1
DQ2
DQ3
DQ4
DQ5
DQ6
DQ7
DQM
Column decoder
Row Decoder
Cell array memory bank 2 (2 Mb × 8) DRAM
Sense amplifiers
Column decoder
Row Decoder
Cell array memory bank 3 (2 Mb × 8) DRAM
Sense amplifiers
CS
RAS
CAS
WE
Command decoder
CAC= Column address counter
MR = Mode register
RC = Refresh counter

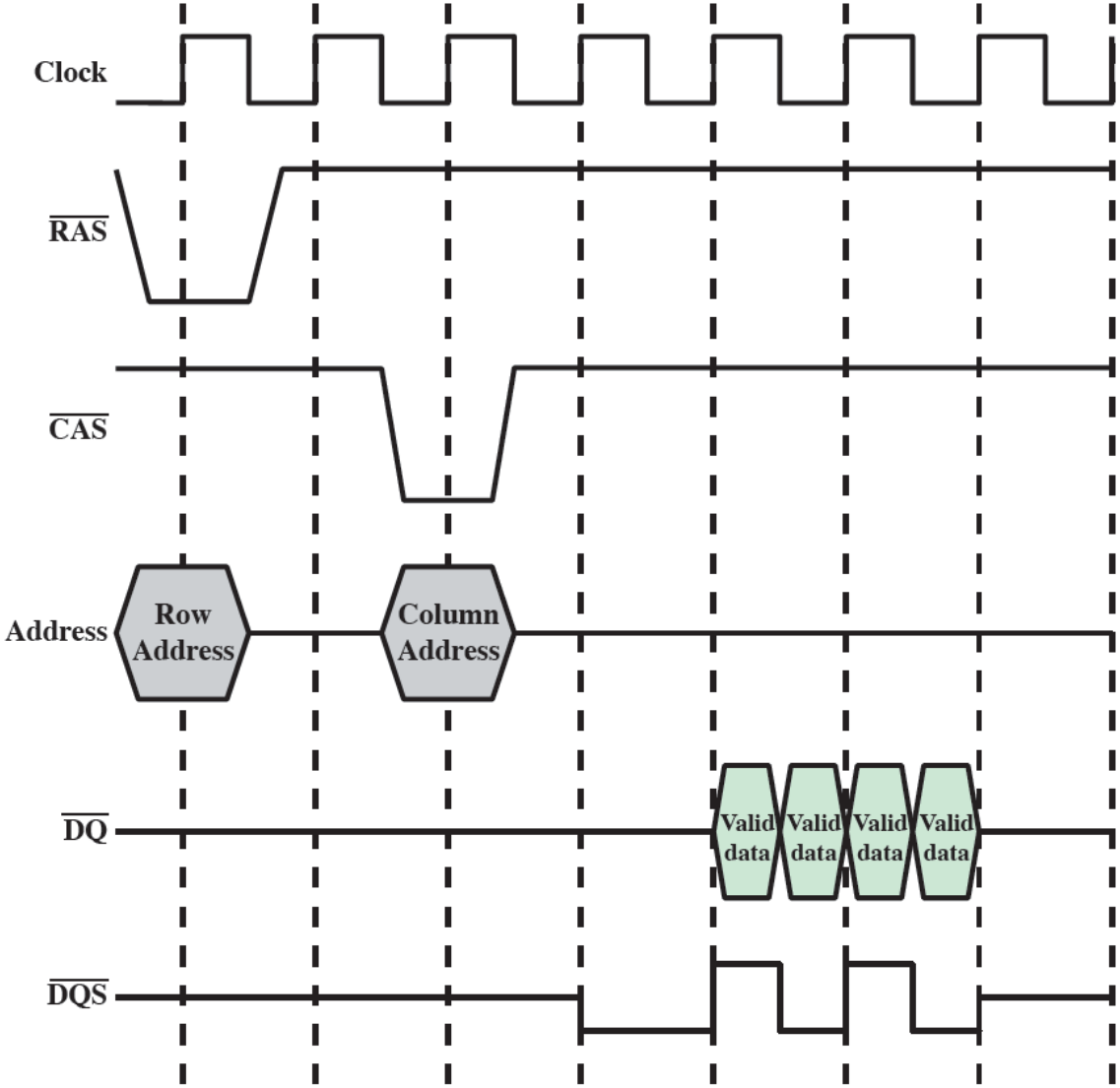
SDRAM Read Timing
(Burst Length = 4, CAS latency =2)



Double Data Rate SDRAM (DDR SDRAM)

- SDRAM can only send data once per bus clock cycle
- Double-data-rate SDRAM can send data twice per clock cycle, once on the rising edge of the clock pulse and once on the falling edge
- Developed by the JEDEC Solid State Technology Association (Electronic Industries Alliance's semiconductor-engineering-standardization body)

DDR-RAM Read Timing



Judging The Performance of RAM: Common Measures

- Clock Frequency
  - Memory clock
  - Bus clock
  - Effective clock
- Timing
  - CL (CAS Latency)
  - TRCD (Row Address to Column Address Delay)
  - TRP (Row Precharge Time)
  - TRAS (Row Active Time)

Memory
Lecture Content

- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

Magnetic Disk

- Constructed of nonmagnetic material, called the substrate, coated with a magnetizable material
- Substrate typically is aluminum or aluminum alloy, sometimes glass

- Benefits of the glass substrate:
  - Improvement in the uniformity of the magnetic film surface to increase disk reliability
  - A significant reduction in overall surface defects to help reduce read-write errors
  - Ability to support lower fly heights
  - Better stiffness to reduce disk dynamics
  - Greater ability to withstand shock and damage

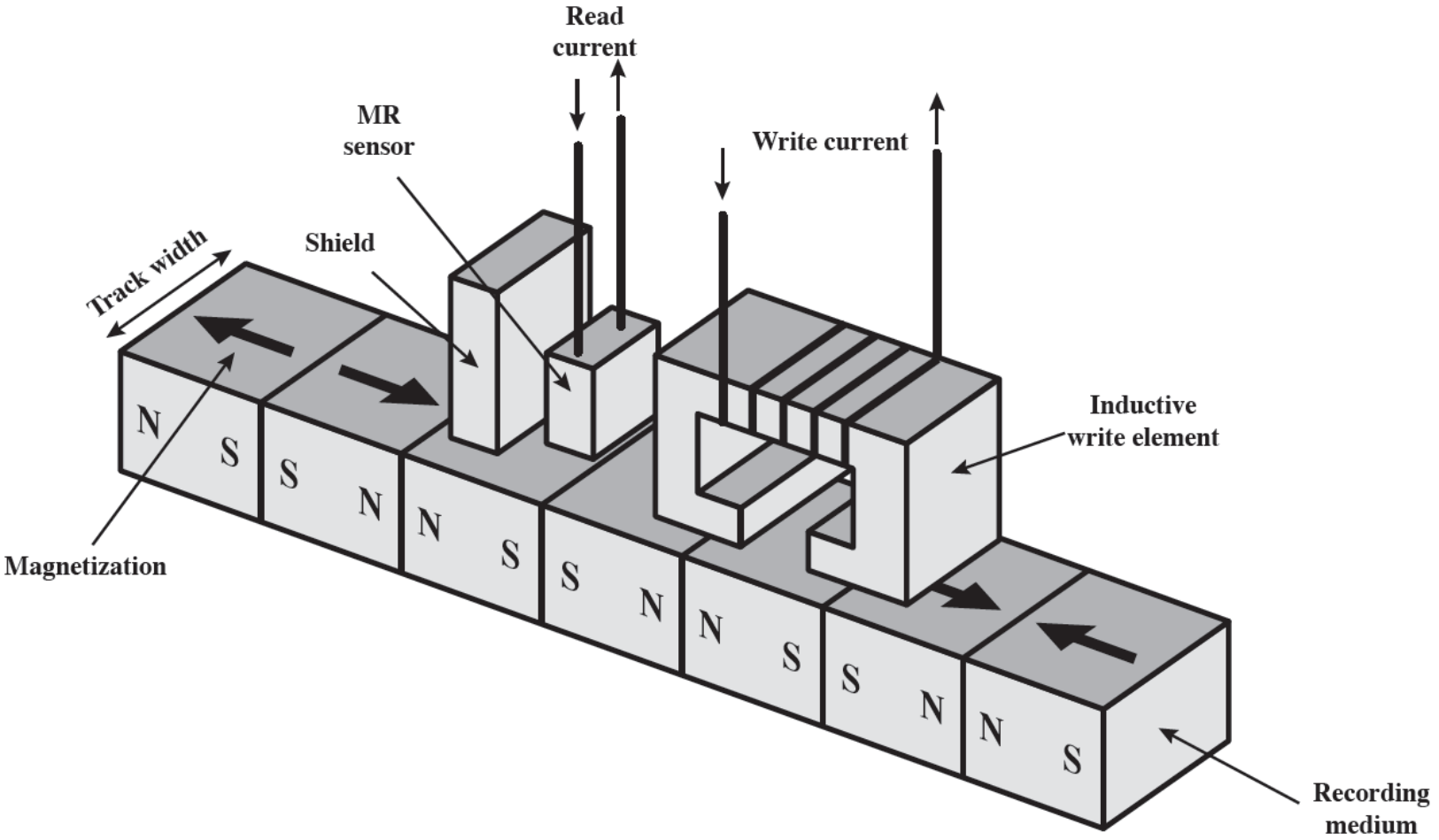
The Read/Write Head

- Data are recorded on and later retrieved from the disk via a conducting coil named the head
  - In many systems there are two heads, a read head and a write head
  - During a read or write operation the head is stationary while the platter rotates beneath it
- The write mechanism exploits the fact that electricity flowing through a coil produces a magnetic field

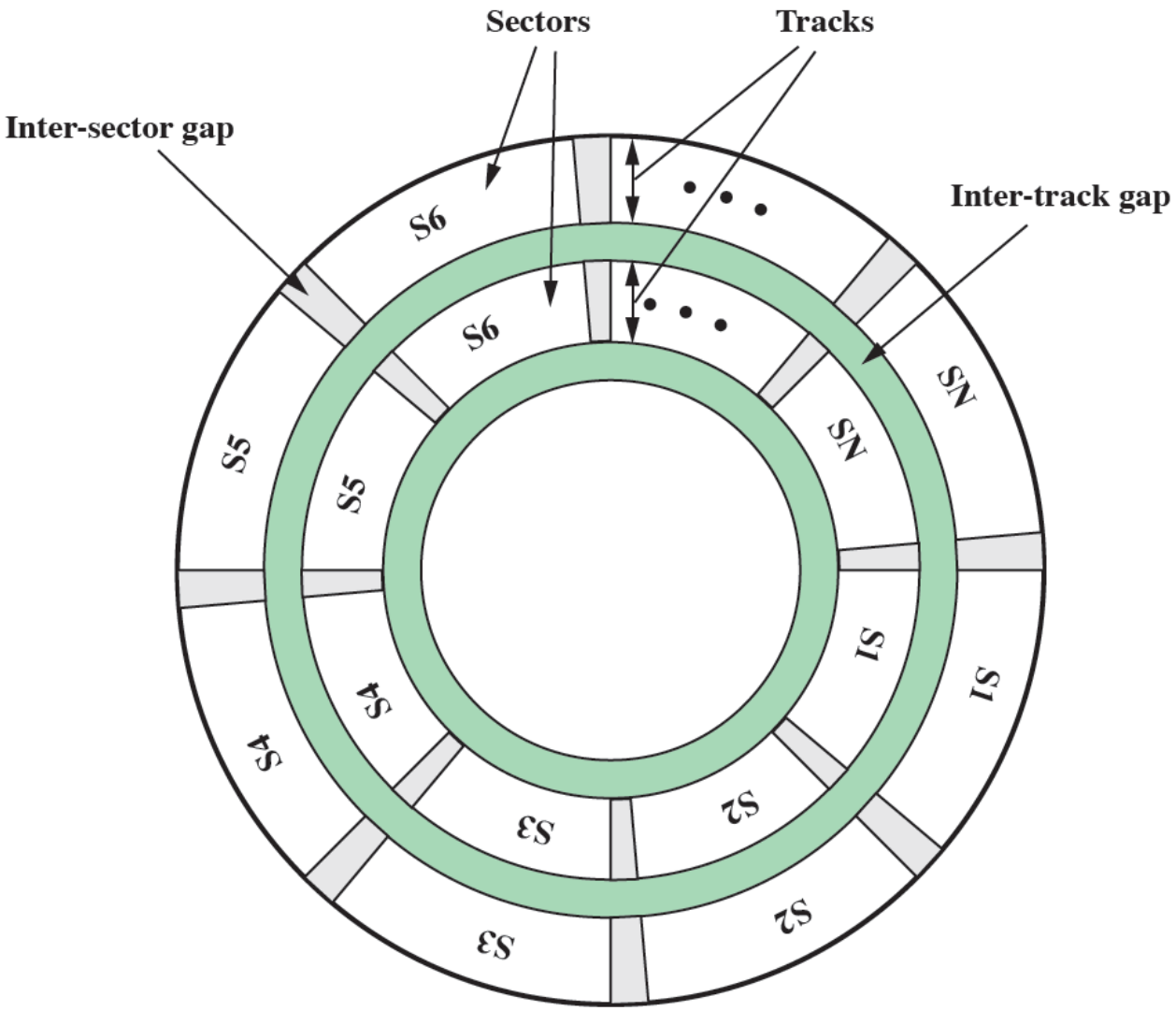
Writing on a Disk

- Electric pulses are sent to the write head
  - The resulting magnetic patterns are recorded on the surface below
  - Different patterns for positive and negative currents
- The write head itself is made of easily magnetizable material
  - Shaped like an U with the gap facing the surface
- An electric current induces a magnetic field across the gap, which in turn magnetizes a small area of the recording medium
- Reversing the direction of the current reverses the direction of the magnetization on the recording medium

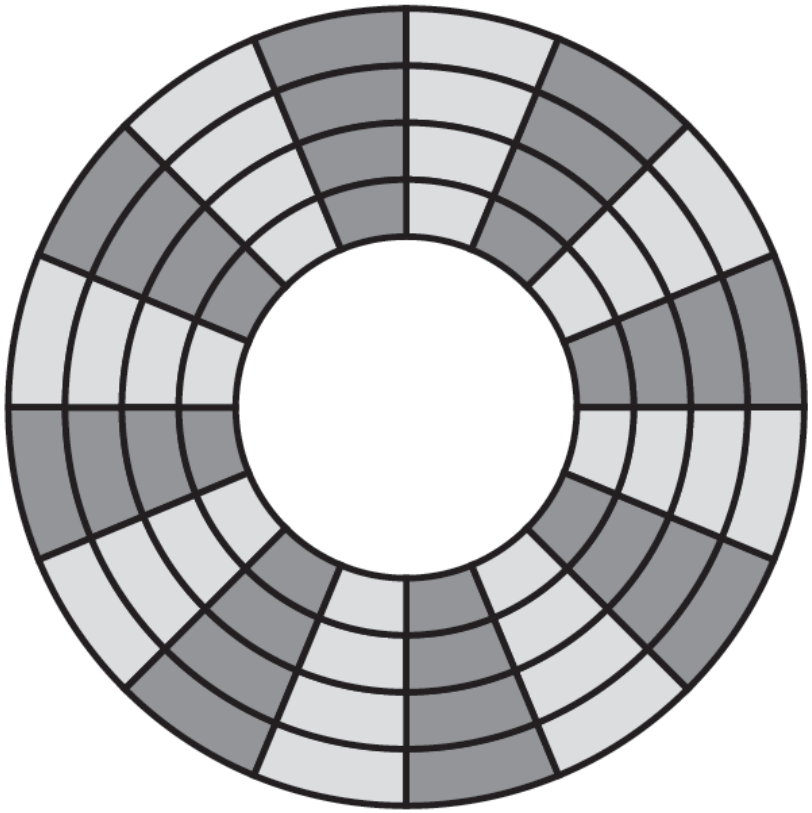
The Read/Write Head



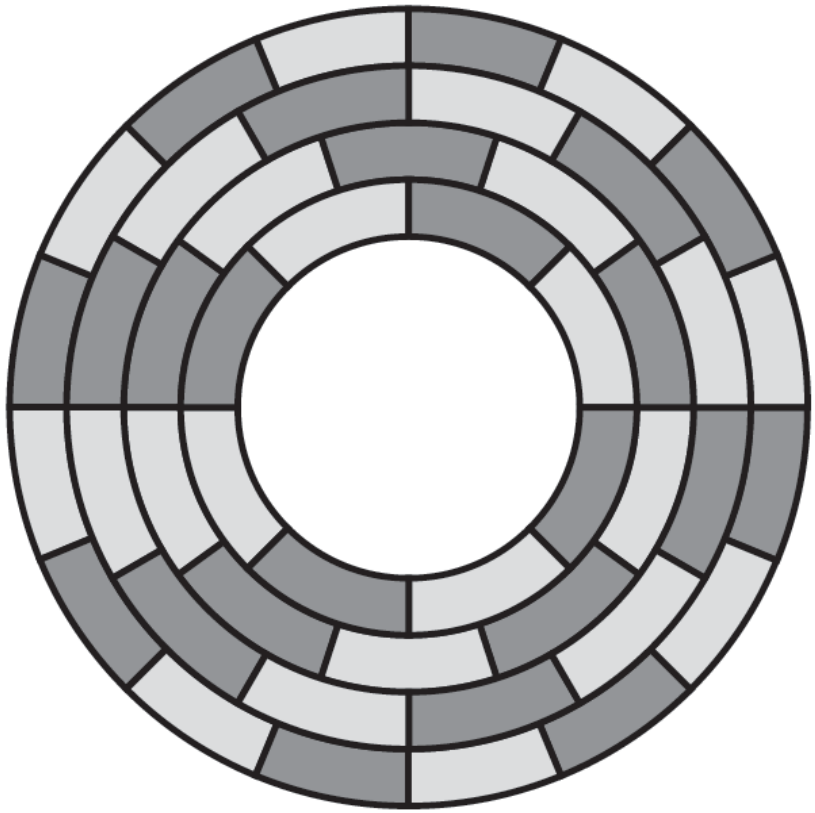
General Disk Layout



Disk Layout Methods

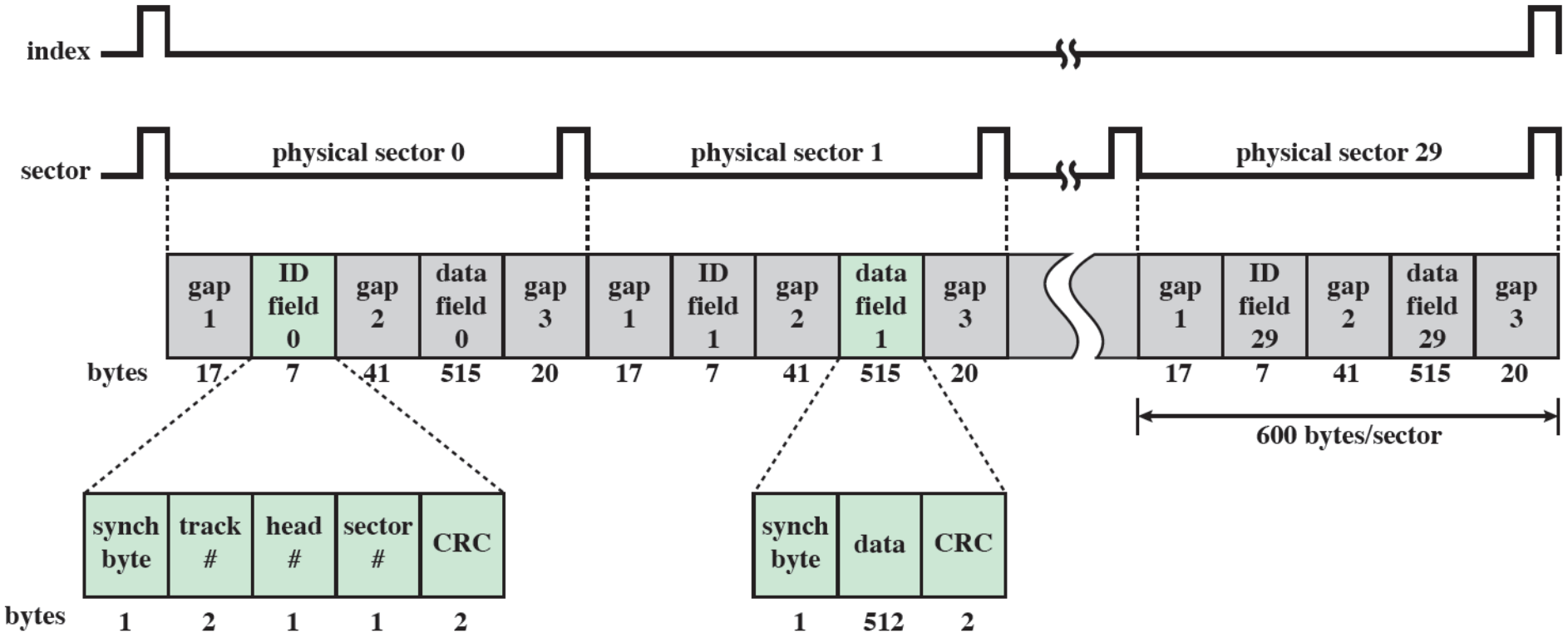


(a) Constant angular velocity



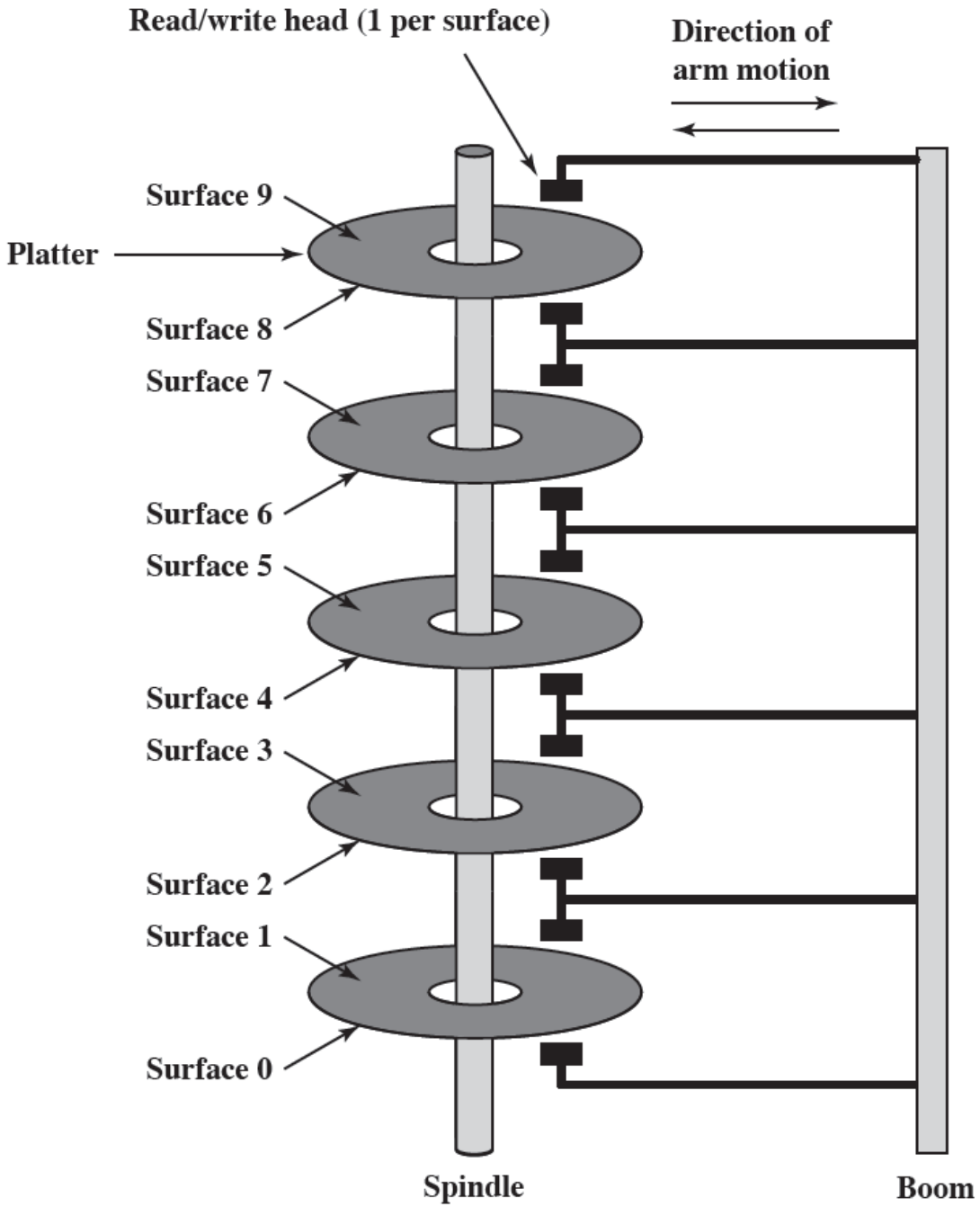
(b) Multiple zoned recording

Winchester Disk Format

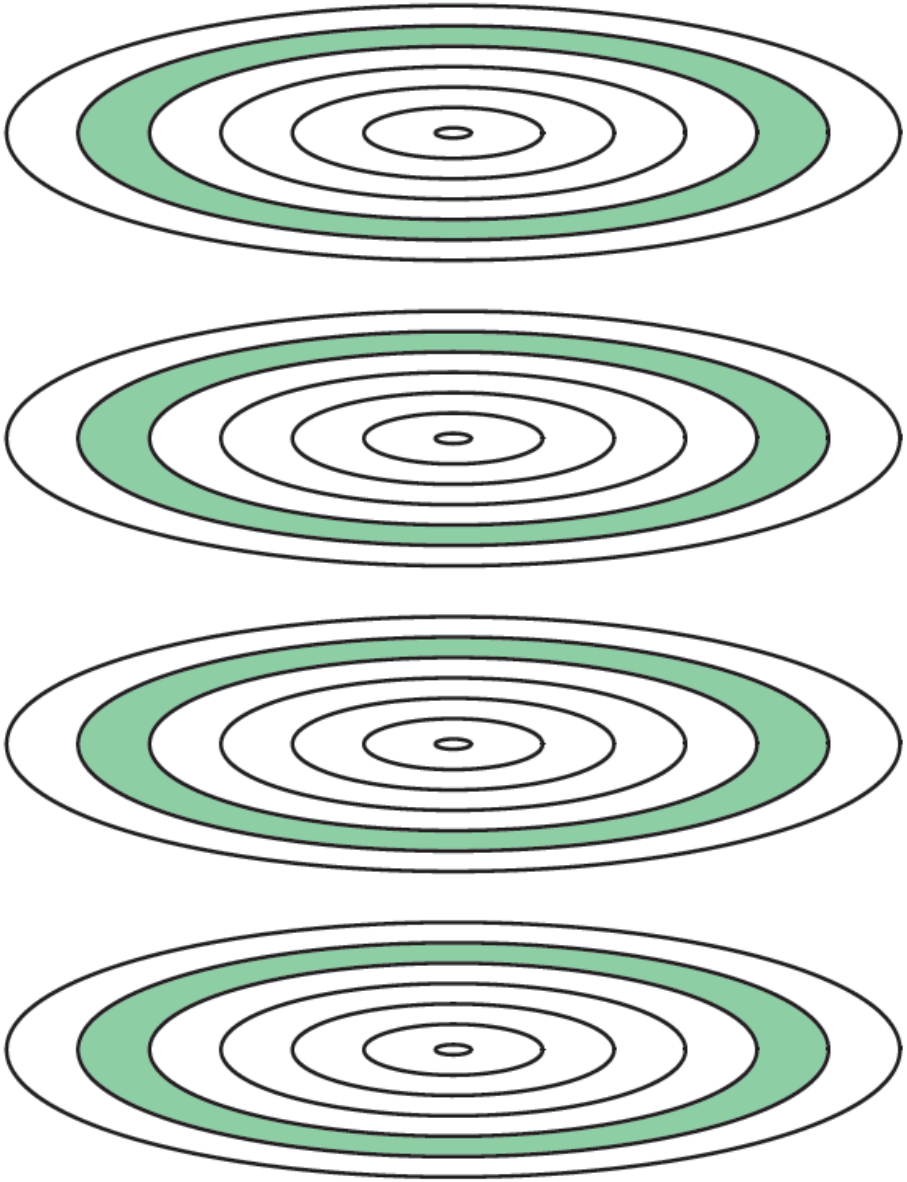


Physical Characteristics of Disk Systems

- **Head Motion**
  - Fixed head (one per track)
  - Movable head (one per surface)
- **Disk Portability**
  - Nonremovable disk
  - Removable disk
- **Sides**
  - Single sided
  - Double sided
- **Platters**
  - Single platter
  - Multiple platter
- **Head Mechanism**
  - Contact (floppy)
  - Fixed gap
  - Aerodynamic gap (Winchester)


Read/write head (1 per surface)
Direction of arm motion
Surface 9
Platter
Surface 8
Surface 7
Surface 6
Surface 5
Surface 4
Surface 3
Surface 2
Surface 1
Surface 0
Spindle
Boom

Cylinders



Characteristics: Head Mechanisms

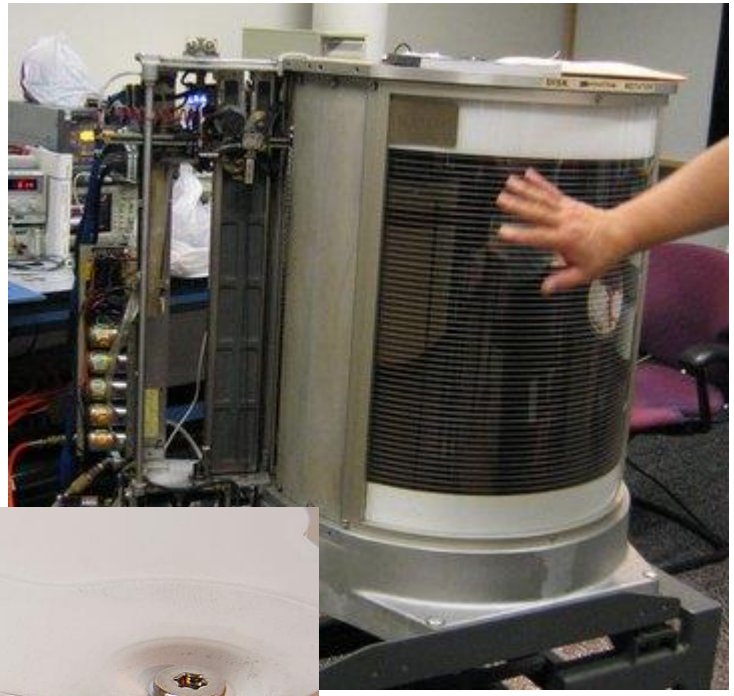
- The head must generate or sense an electromagnetic field of sufficient magnitude to write and read properly
- The narrower the head, the closer it must be to the platter surface to function
- A narrower head means narrower tracks and therefore greater data density
- The closer the head is to the disk the greater the risk of error from impurities or imperfections

Characteristics: Winchester Heads

- Used in sealed drive assemblies that are almost free of contaminants
- Designed to operate closer to the disk's surface than conventional rigid disk heads, thus allowing greater data density
- Is actually an aerodynamic foil that rests lightly on the platter's surface when the disk is motionless
  - The air pressure generated by a spinning disk is enough to make the foil rise above the surface

Head Flighthight

- First Harddisk: IBM 305 RAMAC
  - About 51 µm
  - Used forced air to maintain flight hight
- First Winchester Heads
  - About 6 µm
- Todays HDDs
  - About 3 nm



➢ Images: en.wikipedia.org

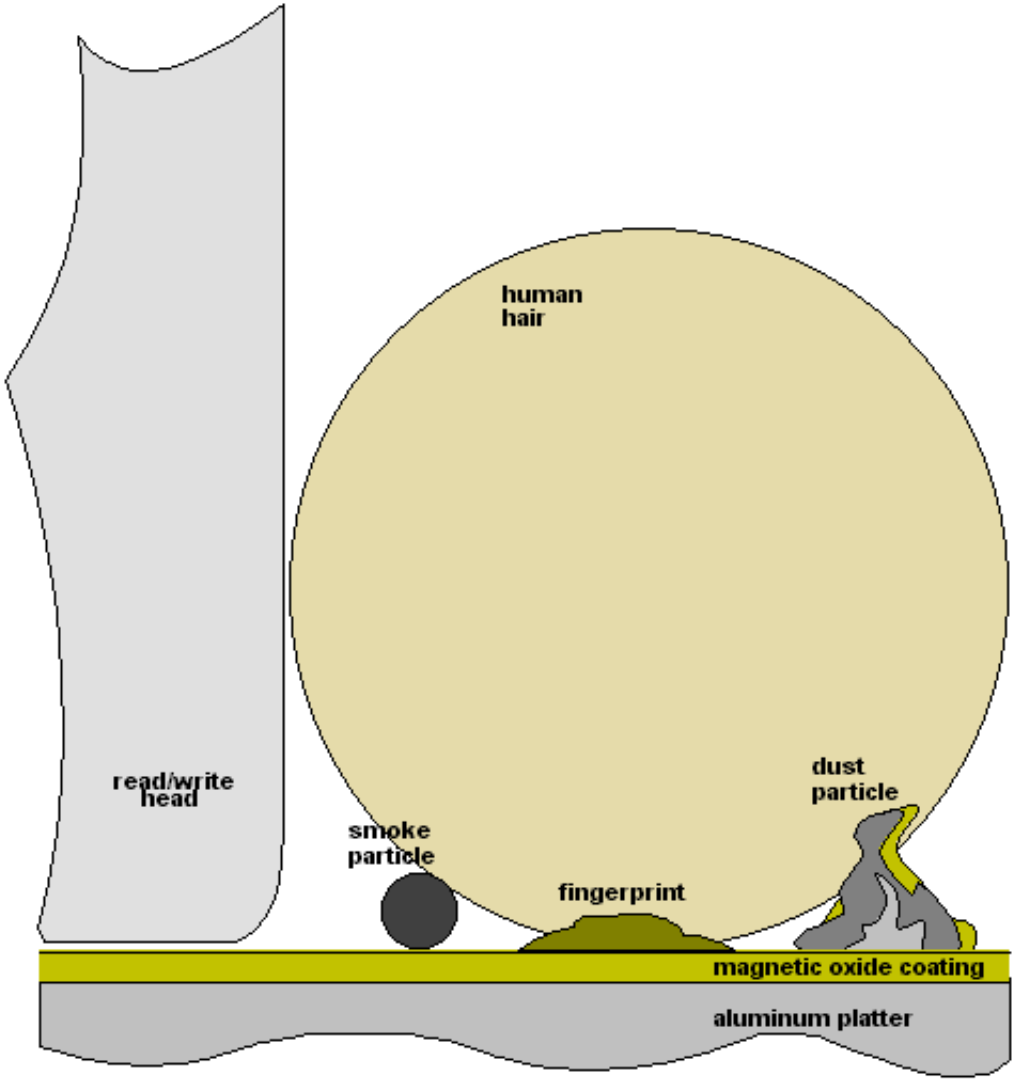
Bring it to Scale

- With a width of less than a hundred nanometers and a thickness of about ten, it flies above the platter at a speed of up to 15,000 RPM, at a height that's the equivalent of 40 atoms.

- Consider the head would be a Boing 747 and we are reading the earth's surface:
  - The head would fly at Mach 800 (Mach 10 is the record by X-43 scramjet)
  - At less than one centimeter from the ground
  - And count every blade of grass
  - Making fewer than 10 unrecoverable counting errors in an area equivalent to all of Ireland

➢ Source: http://www.tomshardware.com/picturestory/476-6-seagate-hard-drive.html

Head-to-Surface Distance Size Comparison



Development of Harddrives

Parameter	Started with	Developed to	Improvement
Capacity (formatted)	3.75 megabytes	> four terabytes	> million
Physical volume	1.9 m^3	34 cm^3	57,000
Weight	910 kg	62 g	15,000
Average access time	~ 600 ms	a few milliseconds	200
Price	US$9,200 per megabyte	< $0.05 per gigabyte by 2013	180
Areal density	2,000 bits per square inch	826 gigabits per square inch	> 400-million

- Source: http://en.wikipedia.org/wiki/Hard_disk_drive

Disk Performance Parameters

- The disks are rotating at constant speed

- To read or write the head must be positioned at the desired track and at the beginning of the desired sector on the track
  - Track selection involves moving the head in a movable-head system or electronically selecting one head on a fixed-head system
  - Once the track is selected, the disk controller waits until the appropriate sector rotates to line up with the head

Disk Performance Parameters

- **Seek Time**
  - The time it takes to position the head at the track
- **Rotational Delay (rotational latency)**
  - The time it takes for the beginning of the sector to reach the head
- **Access Time**
  - The time it takes to get into position (sum of seek and rotational delay)
- **Transfer Time**
  - Once the head is in position, the read or write operation is then performed as the sector moves under the head
  - This is the data transfer portion of the operation

Memory
Lecture Content

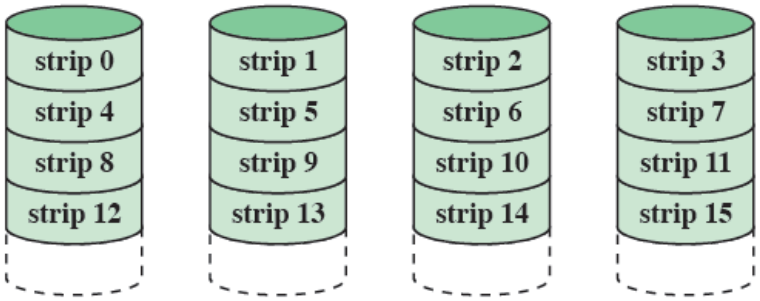
- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

Overview

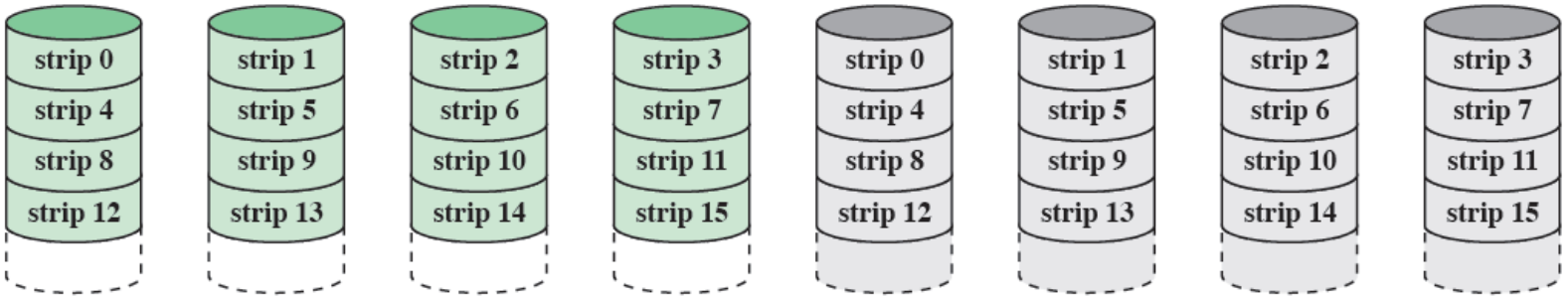
- There are 7 different levels of RAID

- Levels do not imply a hierarchical relationship but designate different design architectures that share three common characteristics:
  - Set of physical disk drives viewed by the operating system as a single logical drive
  - Data are distributed across the physical drives of an array in a scheme known as striping
  - Redundant disk capacity is used to store parity information, which guarantees data recoverability in case of a disk failure

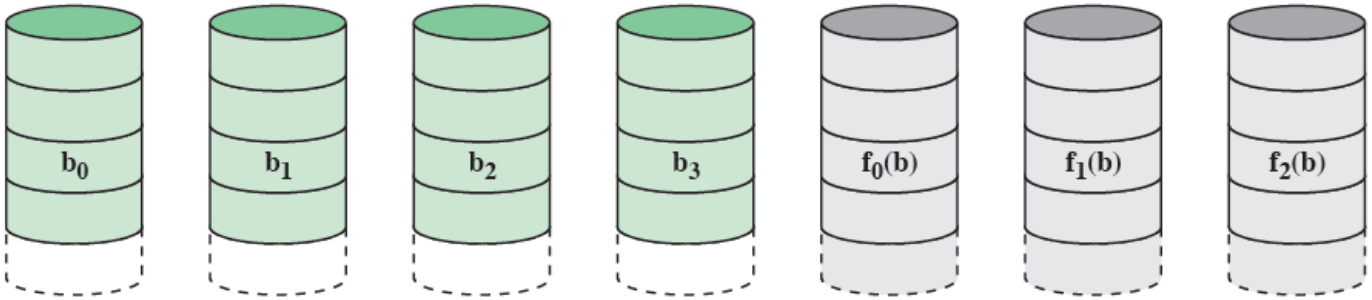
RAID 0,1,2



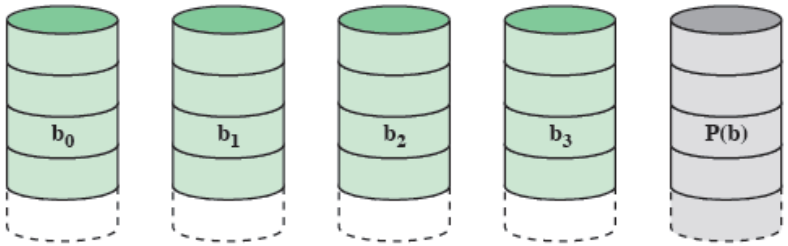
(a) RAID 0 (non-redundant)



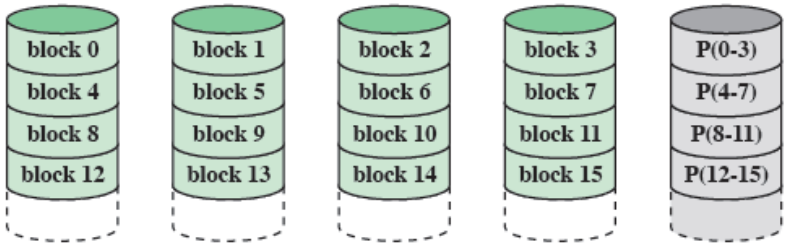
(b) RAID 1 (mirrored)



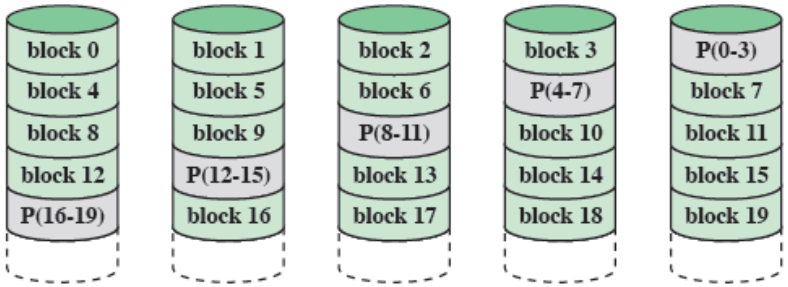
(c) RAID 2 (redundancy through Hamming code)



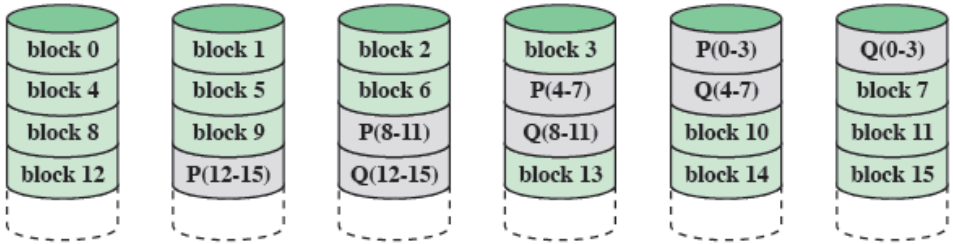
(d) RAID 3 (bit-interleaved parity)



(e) RAID 4 (block-level parity)



(f) RAID 5 (block-level distributed parity)



(g) RAID 6 (dual redundancy)

Raid Overview 1/2

Level	Advantages	Disadvantages	Applications
0	• I/O performance is greatly improved by spreading the I/O load across many channels and drives • No parity calculation overhead is involved • Very simple design • Easy to implement	The failure of just one drive will result in all data in an array being lost	• Video production and Editing • Image editing • Pre-press applications • Any application requiring high bandwidth
1	• 100% redundancy of data means no rebuild is necessary in case of a disk failure, just a copy to the replacement disk • Under certain circumstances, RAID 1 can sustain multiple simultaneous drive failures • Simplest RAID storage subsystem design	Highest disk overhead of all RAID types (100%) - inefficient	• Accounting • Payroll • Financial • Any application requiring very high availability
2	• Extremely high data transfer rates possible • The higher the data transfer rate required, the better the ratio of data disks to ECC disks • Relatively simple controller design compared to RAID levels 3,4 & 5	Very high ratio of ECC disks to data disks with smaller word sizes - inefficient Entry level cost very high - requires very high transfer rate requirement to justify	No commercial implementations exist / not commercially viable

Raid Overview 2/2

3	• Very high read data transfer rate • Very high write data transfer rate • Disk failure has an insignificant impact on throughput • Low ratio of ECC (parity) disks to data disks means high efficiency	Transaction rate equal to that of a single disk drive at best (if spindles are synchronized) Controller design is fairly complex	• Video production and live-streaming • Image editing • Video editing • Any application requiring high throughput
4	• Very high Read data transaction rate • Low ratio of ECC (parity) disks to data disks means high efficiency	• Quite complex controller design • Worst write transaction rate and Write aggregate transfer rate • Difficult and inefficient data rebuild in the event of disk failure	No commercial implementations exist / not commercially avialable
5	• Highest Read data transaction rate • Low ratio of ECC (parity) disks to data disks means high efficiency • Good aggregate transfer rate	Most complex controller design Difficult to rebuild in the event of a disk failure (as compared to RAID level 1)	• File and application servers • Database servers • Web, e-mail, and news servers • Intranet servers • Most versatile RAID level
6	• Provides for an extremely high data fault tolerance and can sustain multiple simultaneous drive failures	• More complex controller design • Controller overhead to compute parity addresses is extremely high	Perfect solution for mission critical applications

Memory
Lecture Content

- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

Flash Memory and SSDs

- A memory device made with solid state components that can be used as a replacement to a hard disk drive (HDD)
  - The term solid state refers to electronic circuitry built with semiconductors

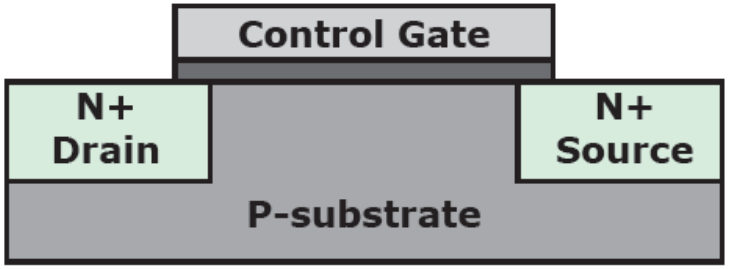
- Flash memory
  - A type of semiconductor memory used in many consumer electronic products including smart phones, GPS devices, MP3 players, digital cameras, and USB devices
  - Cost and performance has evolved to the point where it is feasible to use to replace HDDs

Types of Flash Memory

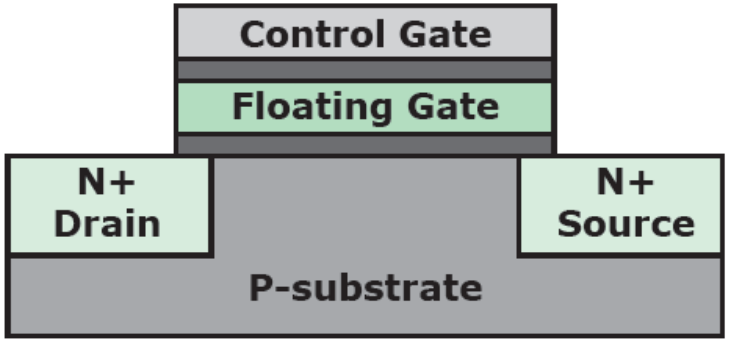
- Two distinctive types of flash memory:
- NOR
  - The basic unit of access is a bit
  - Provides high-speed random access
  - Used for BIOS program that runs at start-up

- NAND
  - The basic unit is 16 or 32 bits
  - Reads and writes in small blocks
  - Used in USB flash drives, memory cards, and in SSDs
  - Does not provide a random-access external address bus so the data must be read on a block-wise basis

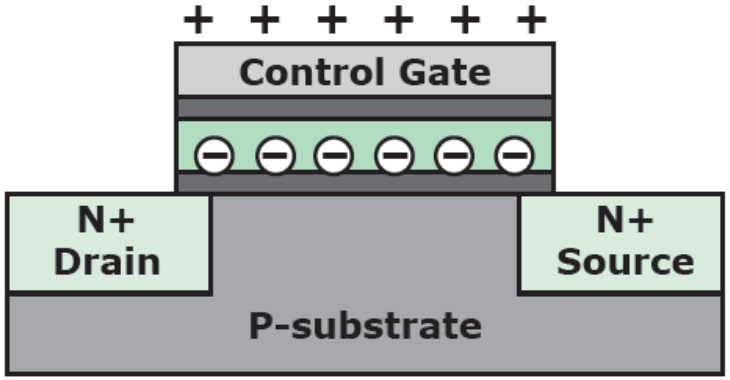
Flash Memory



(a) Transistor structure



(b) Flash memory cell in one state



(c) Flash memory cell in zero state

SSDs vs. HDDs

- High-performance input/output operations per second (IOPS)
- Durability
- Longer lifespan (usually)
- Lower power consumption
- Quieter and cooler running capabilities
- Lower access times and latency rates

	NAND Flash Drives	Disk Drives
I/O per second (sustained)	Read: 45,000 / Write: 15,000	300
Throughput (MB/s)	Read: 400+ / Write: 200+	up to 150
Random access time (ms)	0.1	4–10
Storage capacity	up to 1 TB	up to 8 TB

Practical Issues

- **SDD performance has a tendency to slow down over time**
  - Files are stored on disk as a set of pages, typically 4 KB in length.
  - These are typically not stored as a contiguous set of pages on the disk.
  - However, flash memory is accessed in blocks, typically of 512 KB

  - Now consider what must be done to write a page onto a flash memory:
    - The entire block must be read from the flash memory and placed in a RAM buffer
    - Before the block can be written back to flash memory, the entire block of flash memory must be erased
    - The entire block from the buffer is now written back to the flash memory

  - When the disks gets full, files tend to be separated over more and more blocks

Practical Issues

- **SDD performance has a tendency to slow down over time**
- Counter Measures:

- **Over Provisioning**
  - a substantial portion of the SSD as extra space for write operations
  - erase inactive pages during idle time used to defragment the disk

- **TRIM command**
  - allows an operating system to inform a solid state drive (SSD) which blocks of data are no longer considered in use and can be wiped internally

Practical Issues

- **Flash memory becomes unusable over time (writes)**
  - Techniques for prolonging life:
    - Front-ending the flash with a cache to delay and group write operations
    - Using wear-leveling algorithms that evenly distribute writes across block of cells
  - Bad-block management techniques
  - Most flash devices estimate their own remaining lifetimes so systems can anticipate failure and take preemptive action

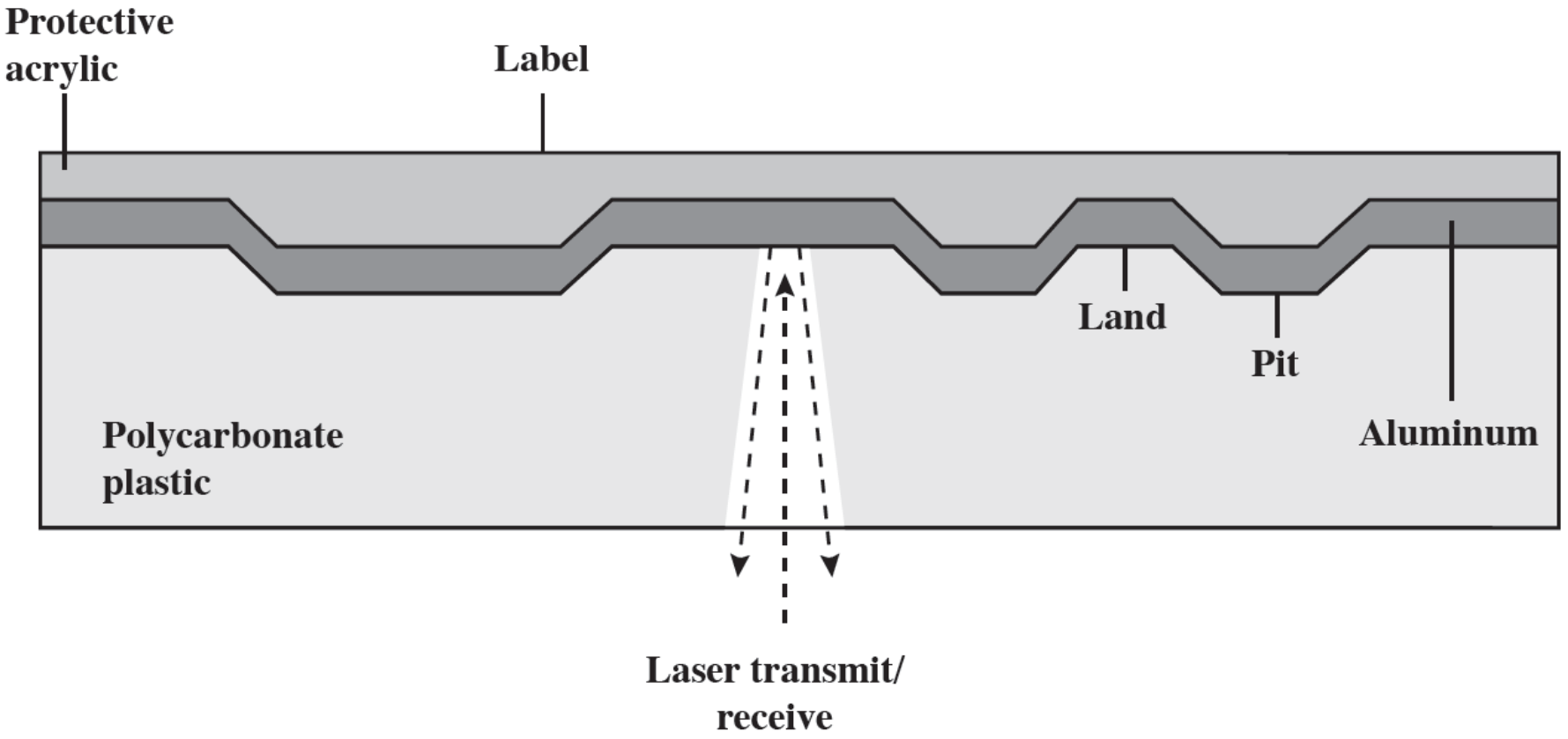
Memory
Lecture Content

- Internal Memory
  - General
  - Error Correction Codes
  - Current DRAMs
- External Memory
  - Hard Drives
  - Raid
  - Flash & SSDs
  - Optical Discs

Optical Discs Overview

- CD-ROM / R / RW
  - Compact Disc Read Only Memory / Recordable / Rewriteable
  - Invented 1983 by Philipps and Sony
  - Stores up to 800 MB, 12cm diameter
- DVD R / RW
  - Digital Versatile Disk (not video disc)
  - doubled sided capacity up to 17 GB
- Blue-Ray Disc
  - Using 405-nm laser (blue-violet)
  - A single layer can store up to 25 GB

Principle of Optical Disks



Principle of Optical Disks

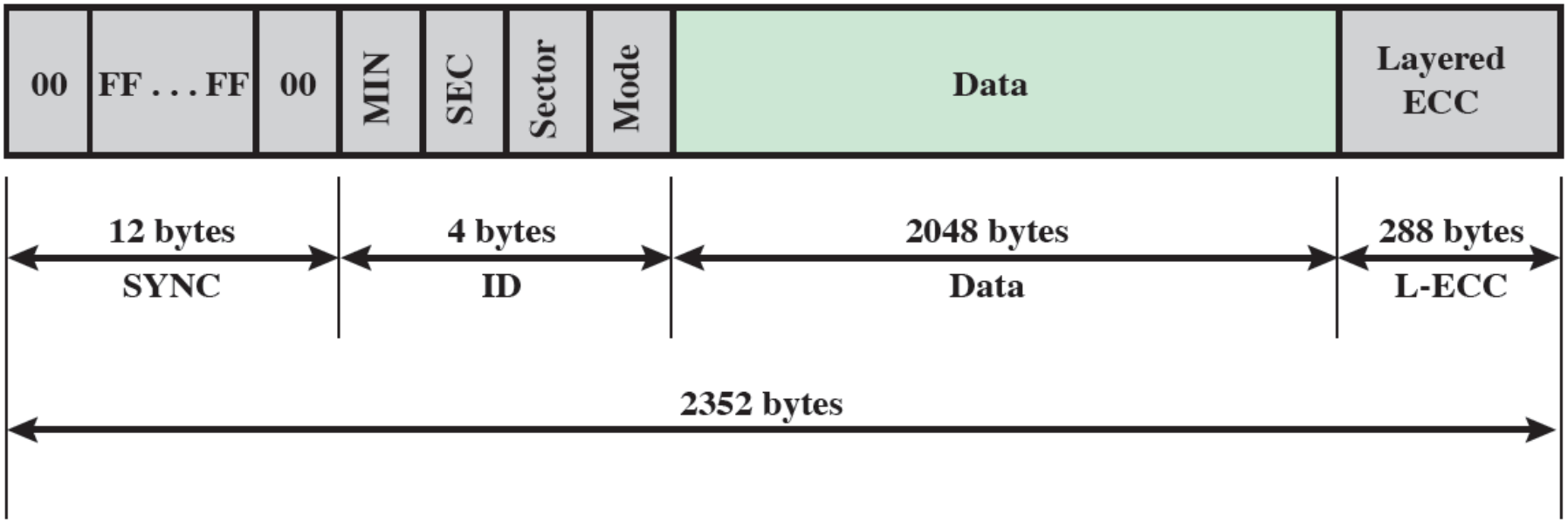
- The intensity of the reflected laser changes as it encounters a pit.
  - Specifically, if the laser beam falls on a pit, which has a somewhat rough surface, the light scatters and a low intensity is reflected back to the source.
  - A land is a smooth surface, which reflects back at higher intensity.

- The beginning or end of a pit represents a 1
- When no change in elevation occurs between intervals, a 0 is recorded

Principle of Optical Disks

- On a magnetic disk, information is recorded in concentric tracks.
  - Simplest Form is constant angular velocity (CAV) system
  - An increase in density is achieved with multiple zoned recording
- CDs contain a single spiral track, beginning near the center
- Thus, all sectors are the same length
  - Information is packed evenly across the disk in segments of the same size
  - Scanned at the same rate by rotating the disk at a variable speed
  - The pits are then read by the laser at a constant linear velocity (CLV)

A CD-ROM Sector



- **Sync**: Identifies the beginning of a block
- **Header**: The header contains the block address and the mode
- **Mode**:
  - 0 specifies a blank data field
  - 1 specifies the use of an error-correcting code and 2048 bytes of data
  - 2 specifies 2336 bytes of user data with no error-correcting code