[Chap.6-1] The Memory Hierarchy

Young Ik Eom (yieom@skku.edu, 031-290-7120)
Distributing Computing Laboratory
Sungkyunkwan University
http://dclab.skku.ac.kr

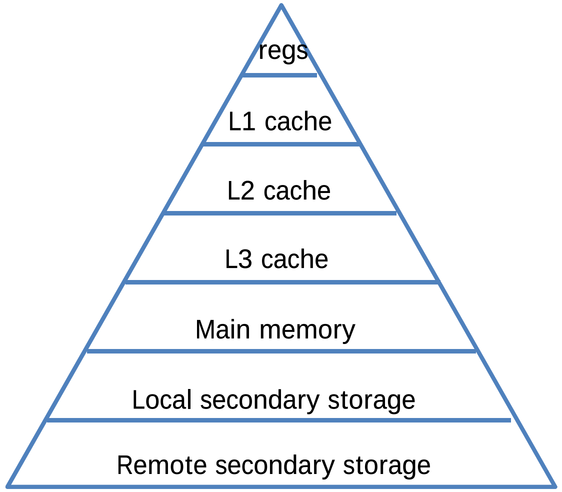
Contents

- Introduction
- Storage technologies
- Locality
- The memory hierarchy
- Cache memories
- Writing cache-friendly code
- Impact of caches on program performance

Introduction

Memory system

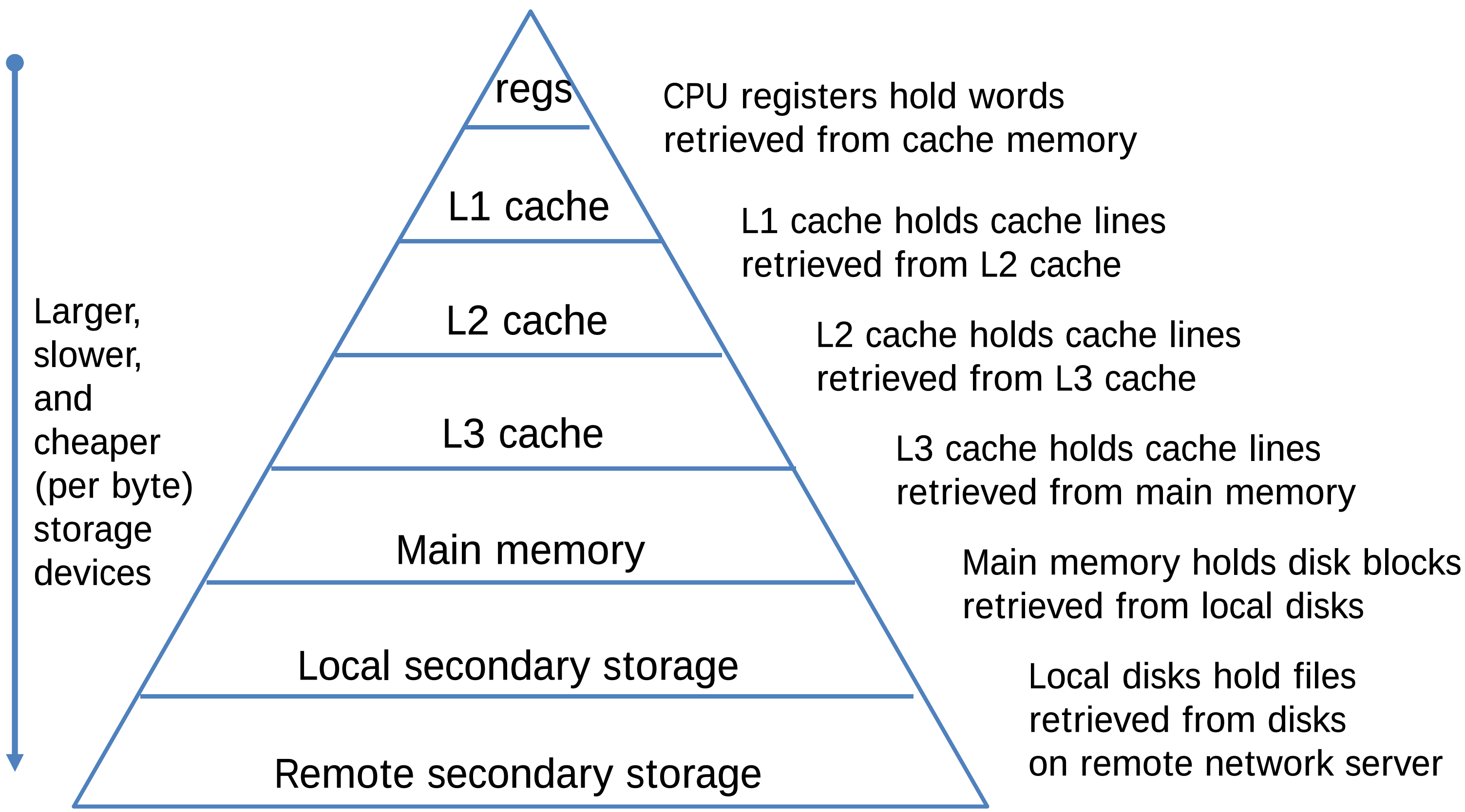
- Hierarchy of storage devices
  with different capacities, costs, and access times
  - CPU registers (can be accessed in 0 cycles)
  - Cache memories (can be accessed in 4~75 cycles)
  - Main memory (can be accessed in hundreds of cycles)
  - Hard disk (can be accessed in tens of millions of cycles)
  - Etc



Introduction

■ Memory system

- Hierarchy of storage devices



Introduction

Memory system

- Locality
  - Programs with good locality
    - ✓ Tend to access the same set of instructions and data items over and over again
    - ✓ Tend to access more instructions and data items from the upper levels of memory hierarchy than programs with poor locality
    - ✓ Run faster than programs with poor locality

Introduction

- **Memory system**
  - Cache memory
    - Have the most impact on application program performance
    - Focuses on …
      - ✓ Analyzing programs for locality
      - ✓ Improving locality

Storage Technologies

- **RAM (Random Access Memory)**
  - SRAM (Static RAM)
    - Faster and more expensive
    - Used for cache memories
  - DRAM (Dynamic RAM)
    - Used for main memory and frame buffer of a graphic system
  - Both are volatile
    - Lose information if the supply voltage is turned off

Storage Technologies

SRAM and DRAM

	Trans. per bit	Access time	Needs refresh?	Sensitive?	Cost	Applications
SRAM	6	1×	No	No	1,000×	Cache M
DRAM	1	10×	Yes	Yes	1×	Main M, Frame buffers

- SRAM
  - Persistent
    - ✓ No refresh is necessary
  - Faster than DRAM
  - Not sensitive to disturbances such as light and electrical noise
  - Lower density
  - Consumes more power
  - More expensive

Storage Technologies

■ Conventional DRAMs

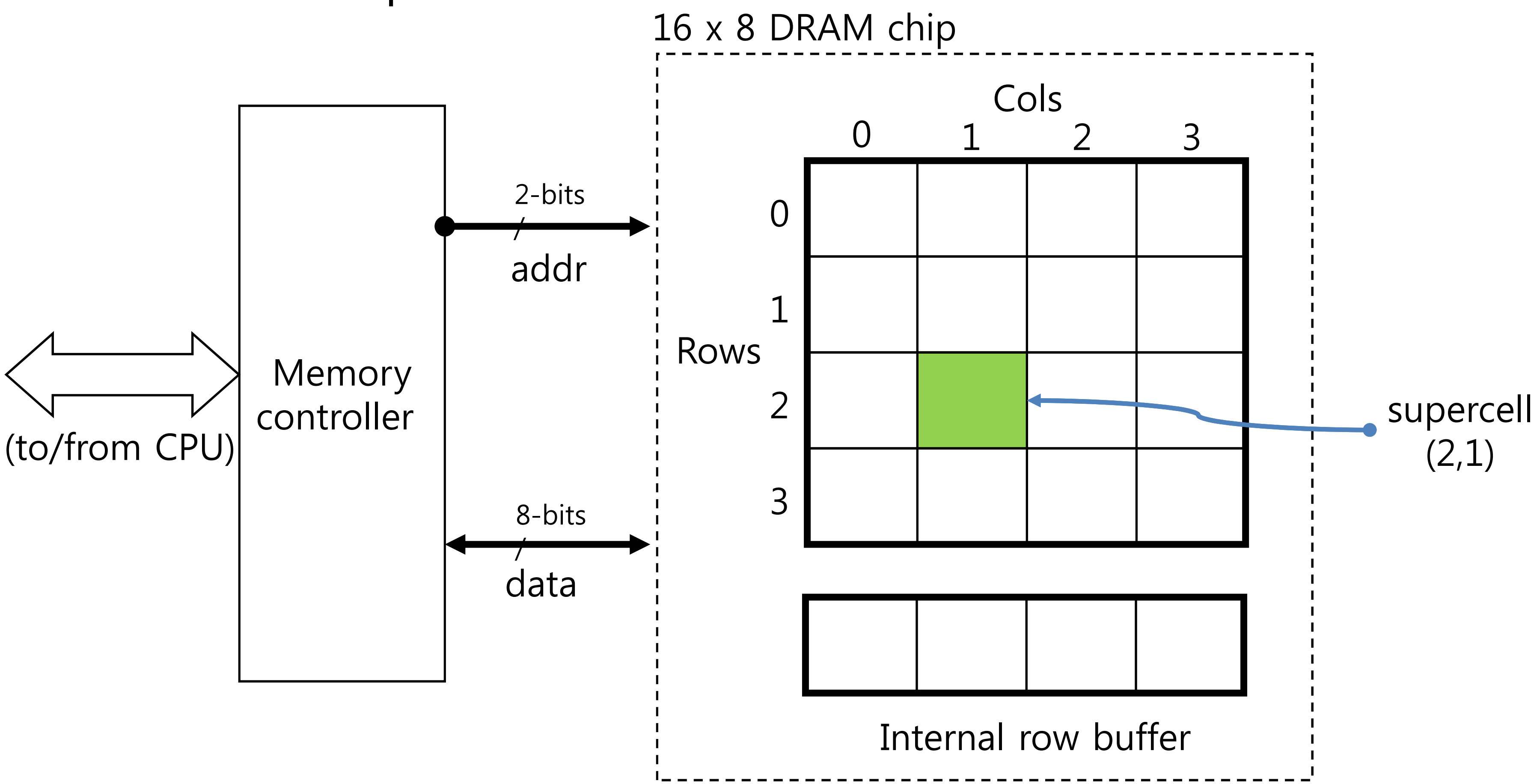
- DRAM chip
  - Partitioned into **d** supercells, each consisting of **w** DRAM cells
  - A **d**×**w** DRAM can store **d**·**w** bits
  - Supercells organized as a rectangular array with **r** rows and **c** columns (**d** = **r**·**c**)

- MC(Memory Controller)
  - Transfers **w** bits at a time to/from DRAM chip
  - For reading supercell (**i**, **j**),
    - ✓ MC sends row address **i** (RAS(Row Access Strobe) request), and then column address **j** (CAS(Column Address Strobe) request)
    - ✓ RAS and CAS shares the same address pins

Storage Technologies

■ Conventional DRAMs

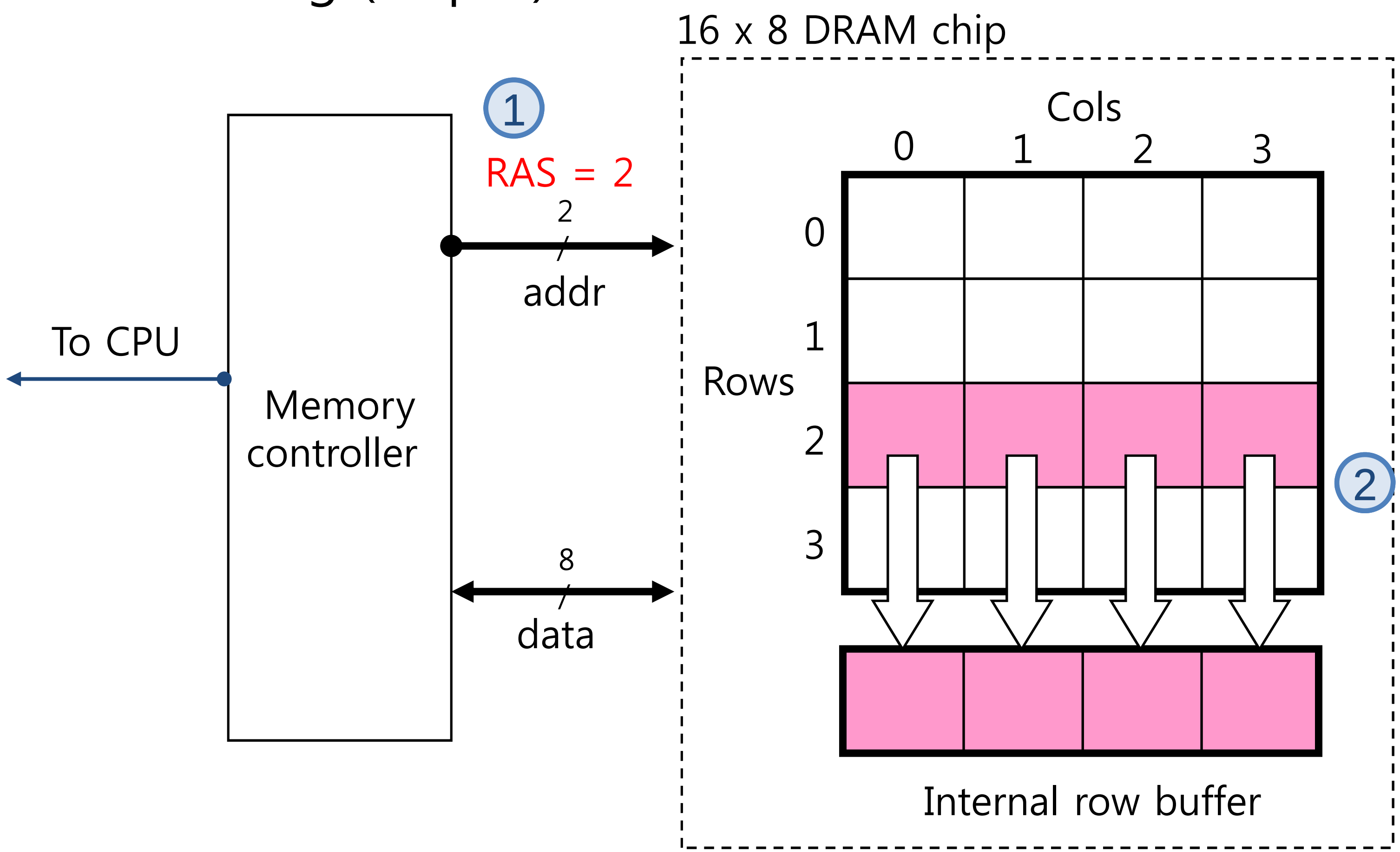
- DRAM chip



Storage Technologies

Conventional DRAMs

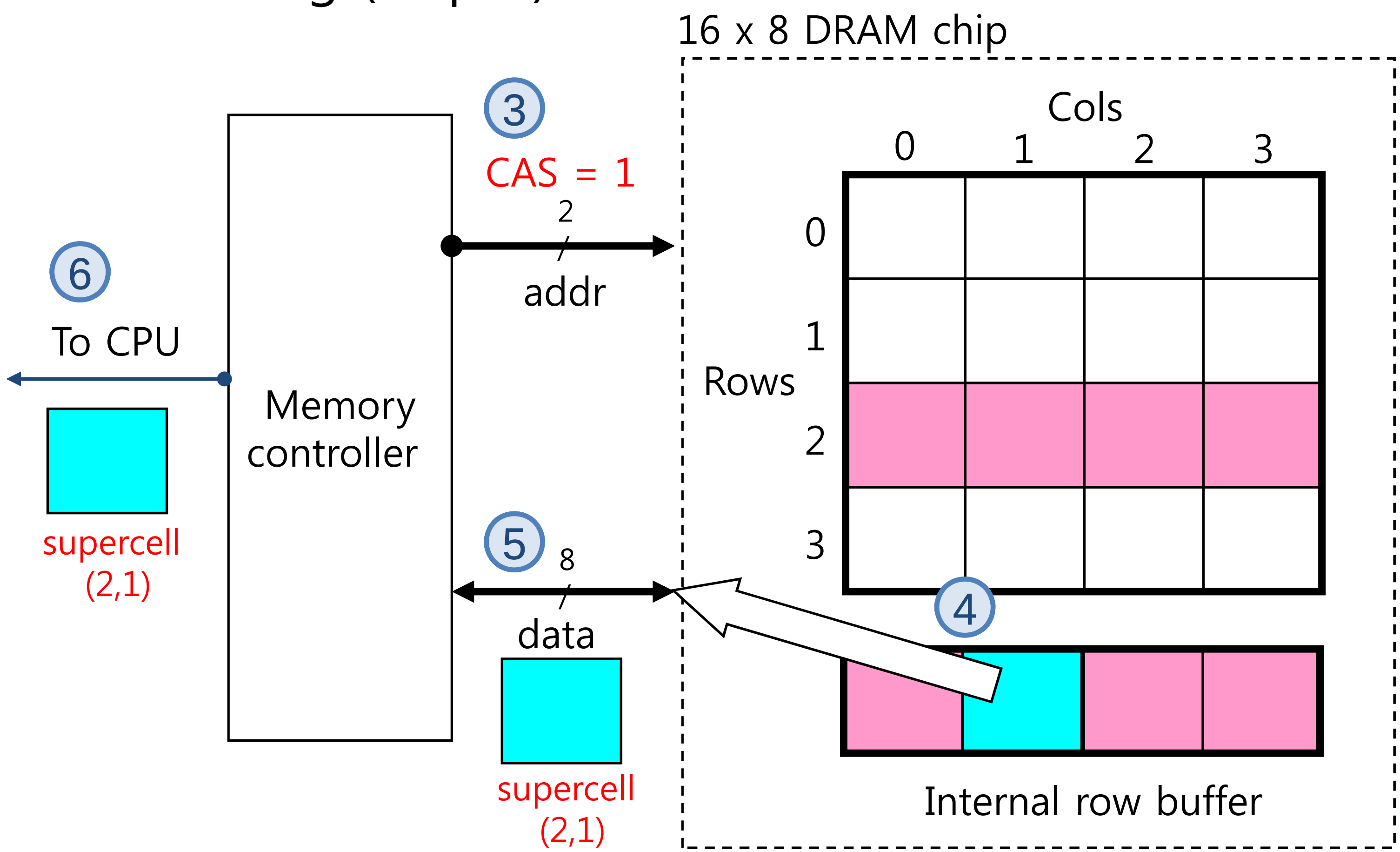
- Reading (step-1)



Storage Technologies

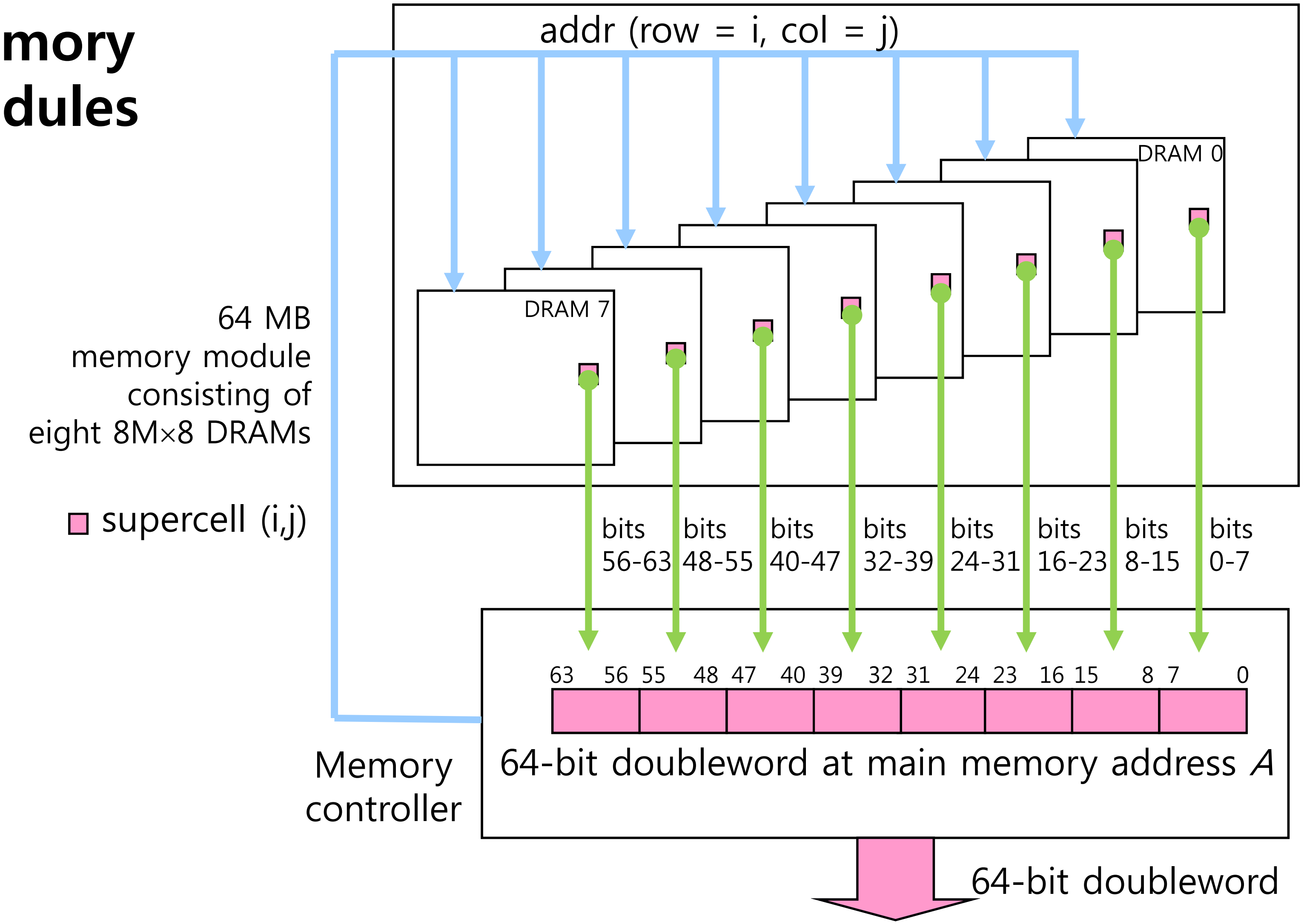
■ Conventional DRAMs

- Reading (step-2)



Storage Technologies

- **Memory modules**



Storage Technologies

- **Enhanced DRAMs**
  - FPM DRAM (Fast page mode)
    - Allows reuse of row addresses (quick access)
      - ✓ Eg) RAS → CAS → CAS → CAS
  - EDO DRAM (Extended data out)
  - SDRAM (Synchronous)
    - Uses a conventional clock signal instead of asynchronous control
  - DDR SDRAM (Double data-rate synchronous)
    - Double edge clocking sends two bits per cycle per pin
    - DDR2, DDR3 (used in Intel Core i7)
  - VRAM (Video)
    - Used in frame buffers of graphics system

Storage Technologies

- **Nonvolatile memories**
  - Retain their information even when they are powered off
  - ROMs
    - PROM, EPROM, EEPROM
  - Flash memory (SSD)
  - Hard disk (HDD)
  - NVRAM(Non-Volatile RAM)
    - PRAM, STT-MRAM

Storage Technologies

- **Nonvolatile memories: ROMs (Read-Only Memories)**
  - ROM
    - Programmed during production
  - PROM (Programmable ROM)
    - Can be programmed exactly once
  - EPROM (Erasable PROM)
    - Can be erased and reprogrammed on the order of 1,000 times
    - Erasing by shining ultraviolet light
    - Programming is done by using a special device
  - EEPROM (Electrically Erasable PROM)
    - Can be programmed in-place on printed circuit cards
      - ✓ No separate programming device
    - Can be reprogrammed on the order of 10,000 times

Storage Technologies

- **Nonvolatile memories: Flash memories**
  - Fast and durable nonvolatile storages
    for many electronic devices, including smart phones,
    tablets, laptops, desktops, and digital cameras
  - Asymmetric read-write speed
  - No in-place writes
  - Wares out after about 100,000 erasings

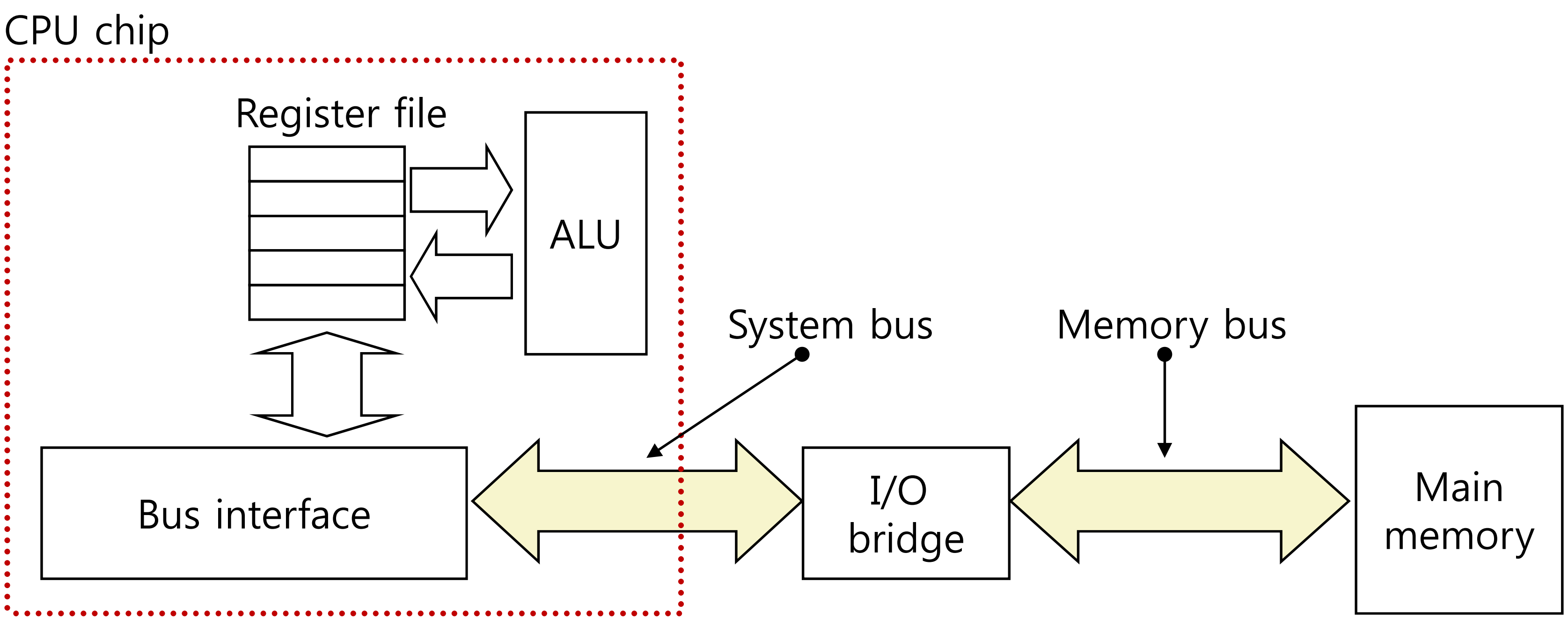
  - NAND flash
  - NOR flash

  - SSD (Solid State Disk)

Storage Technologies

Bus structure

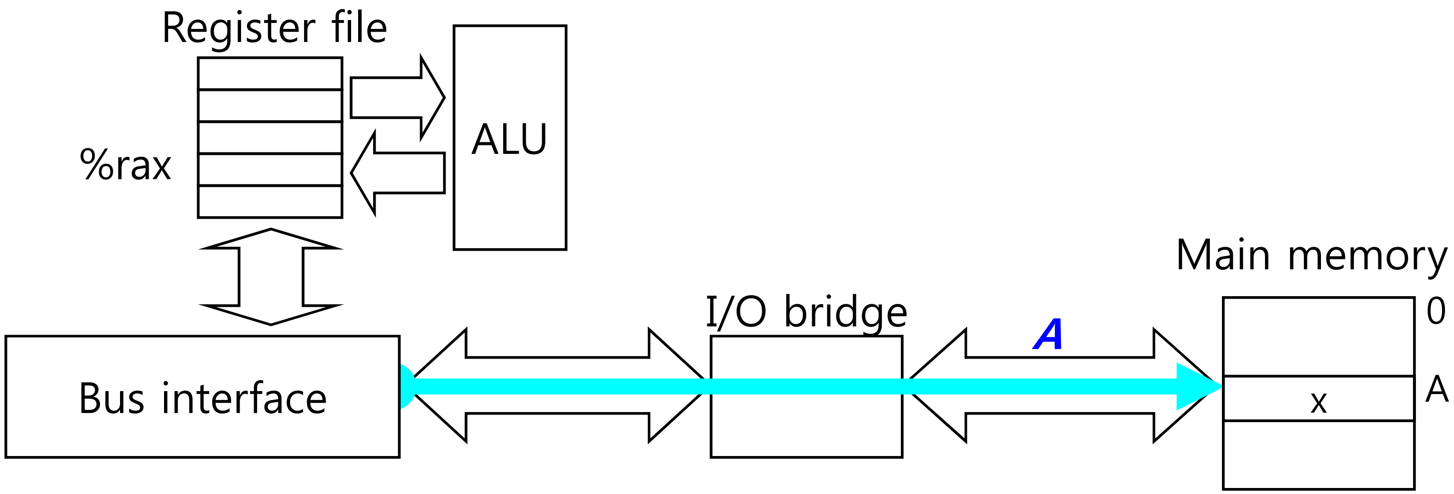
- Bus
  - Collection of parallel wires
    that carry address, data, and control signals
  - Typically shared by multiple devices



Storage Technologies

Bus structure

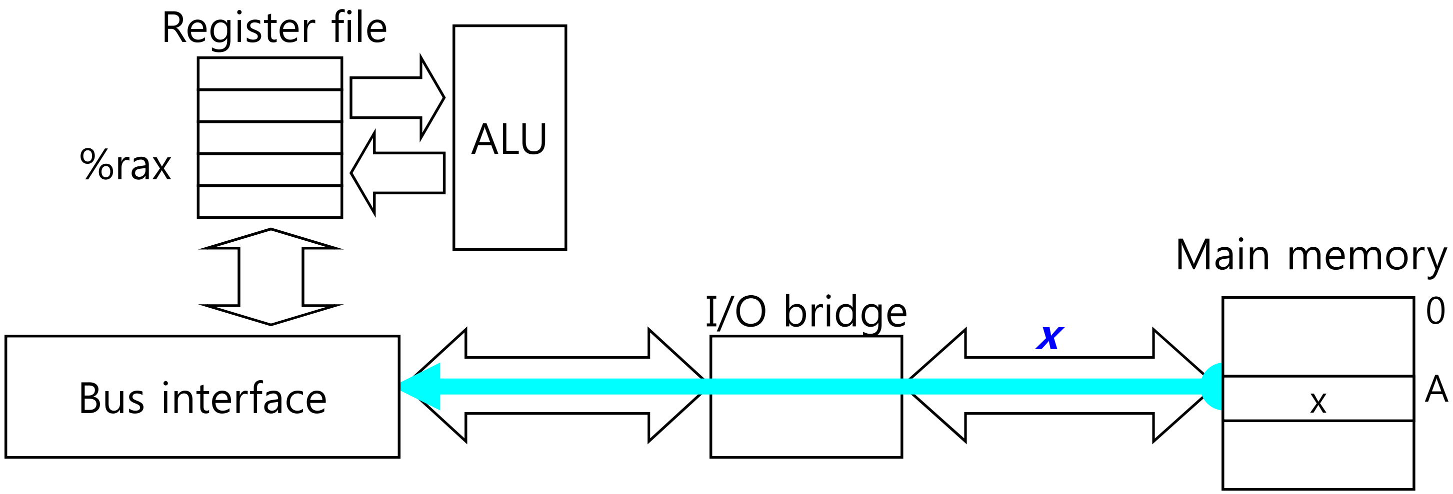
- Read transaction (**movq A,%rax**)
  - (Step-1) CPU places the address **A** on the system bus



Storage Technologies

Bus structure

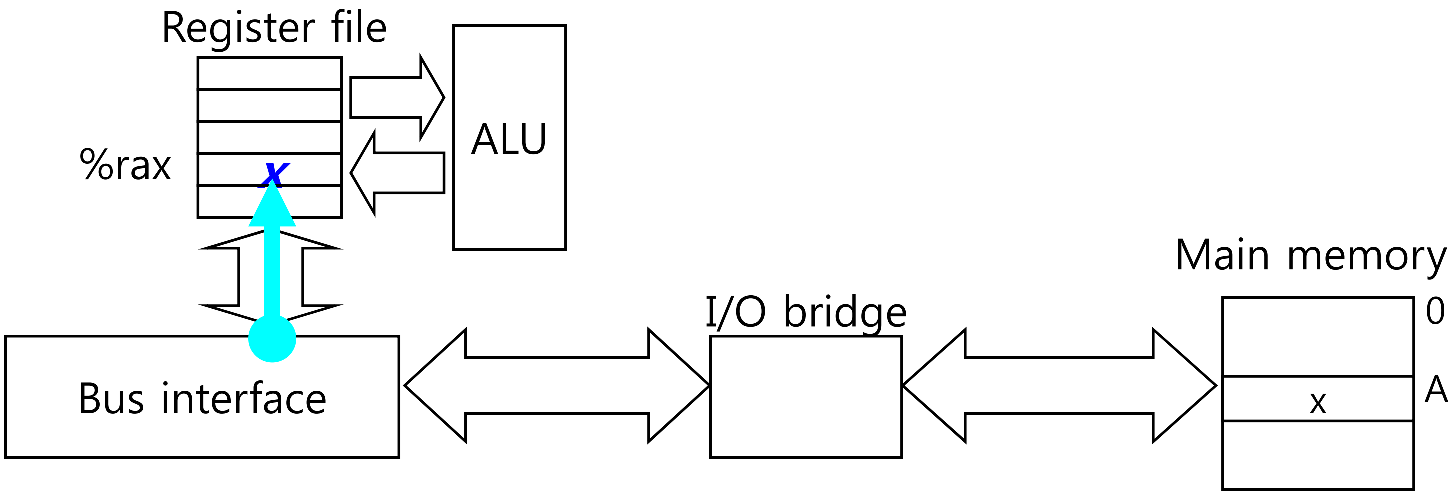
- Read transaction (**movq A,%rax**)
  - (Step-2) Main memory reads **A** from the memory bus, retrieves word **x**, and places it on the bus



Storage Technologies

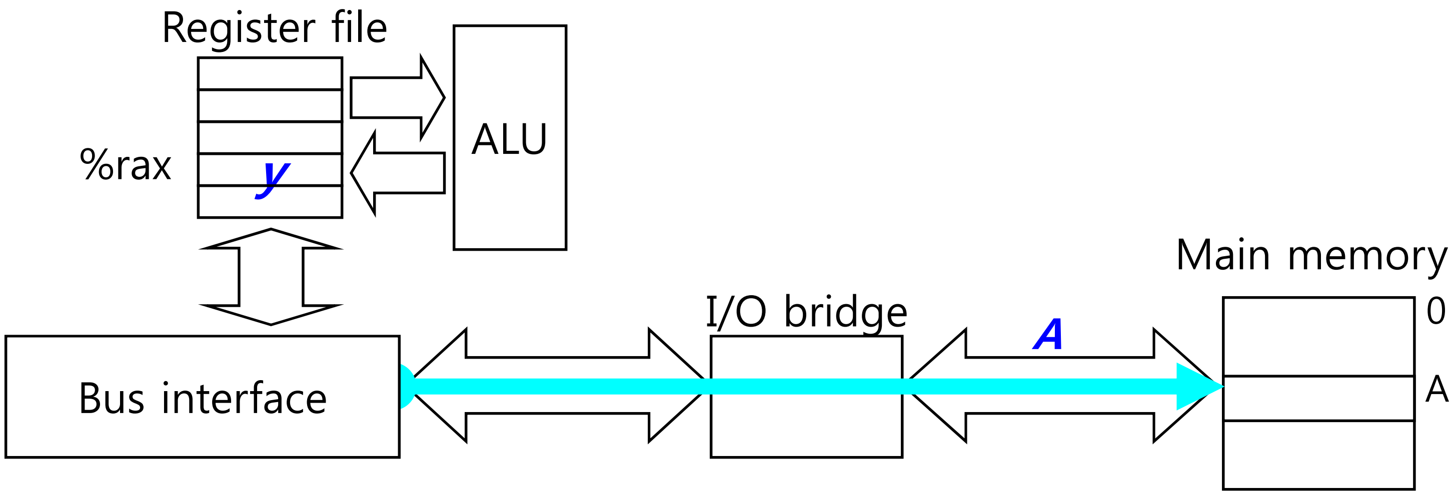
■ Bus structure

- Read transaction (**movq A, %rax**)
  - (Step-3) CPU reads word **x** from the bus and copies it into register **%rax**



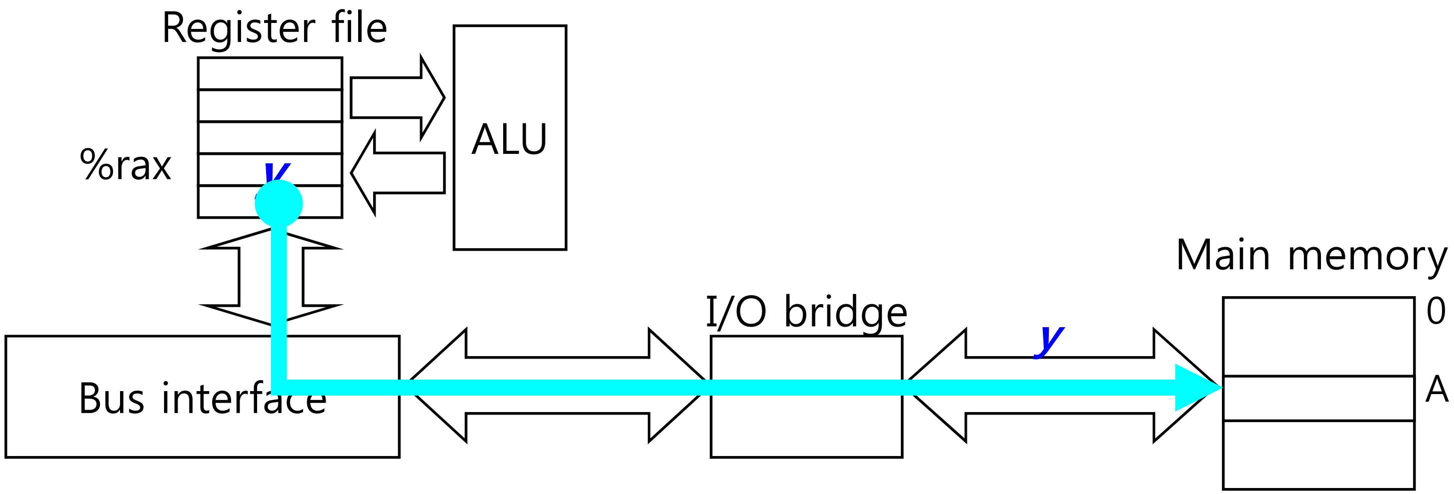
Storage Technologies

- **Bus structure**
  - Write transaction (**movq %rax,A**)
    - (Step-1) CPU places address **A** on bus; Main memory reads it and waits for the corresponding data word to arrive



Storage Technologies

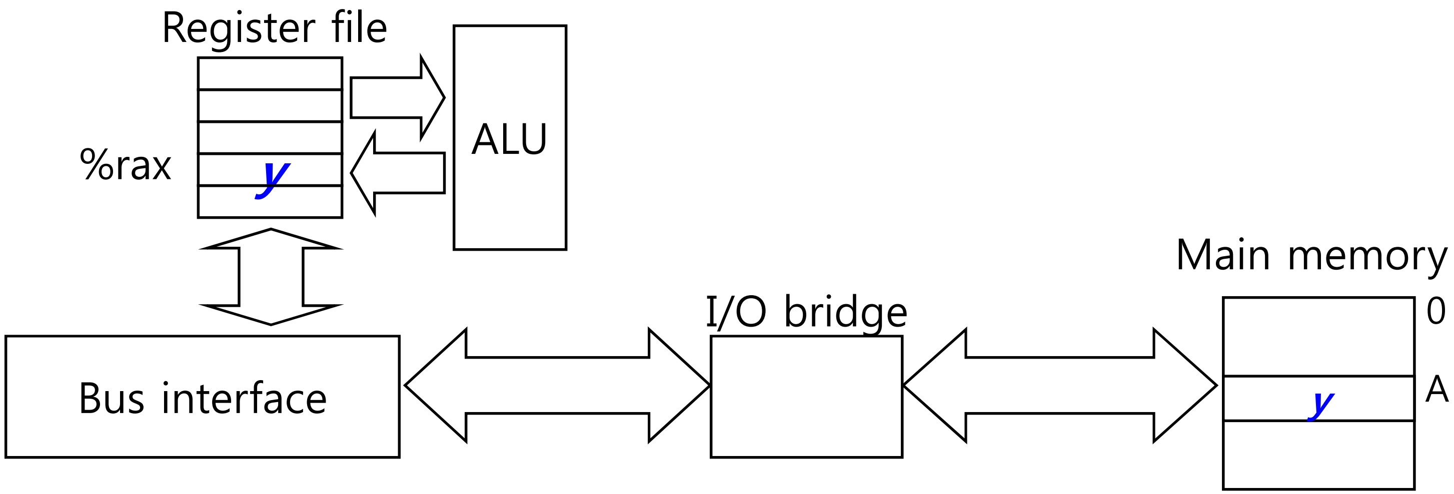
- **Bus structure**
  - Write transaction (**movq %rax,A**)
    - (Step-2) CPU places data word **y** on the bus



Storage Technologies

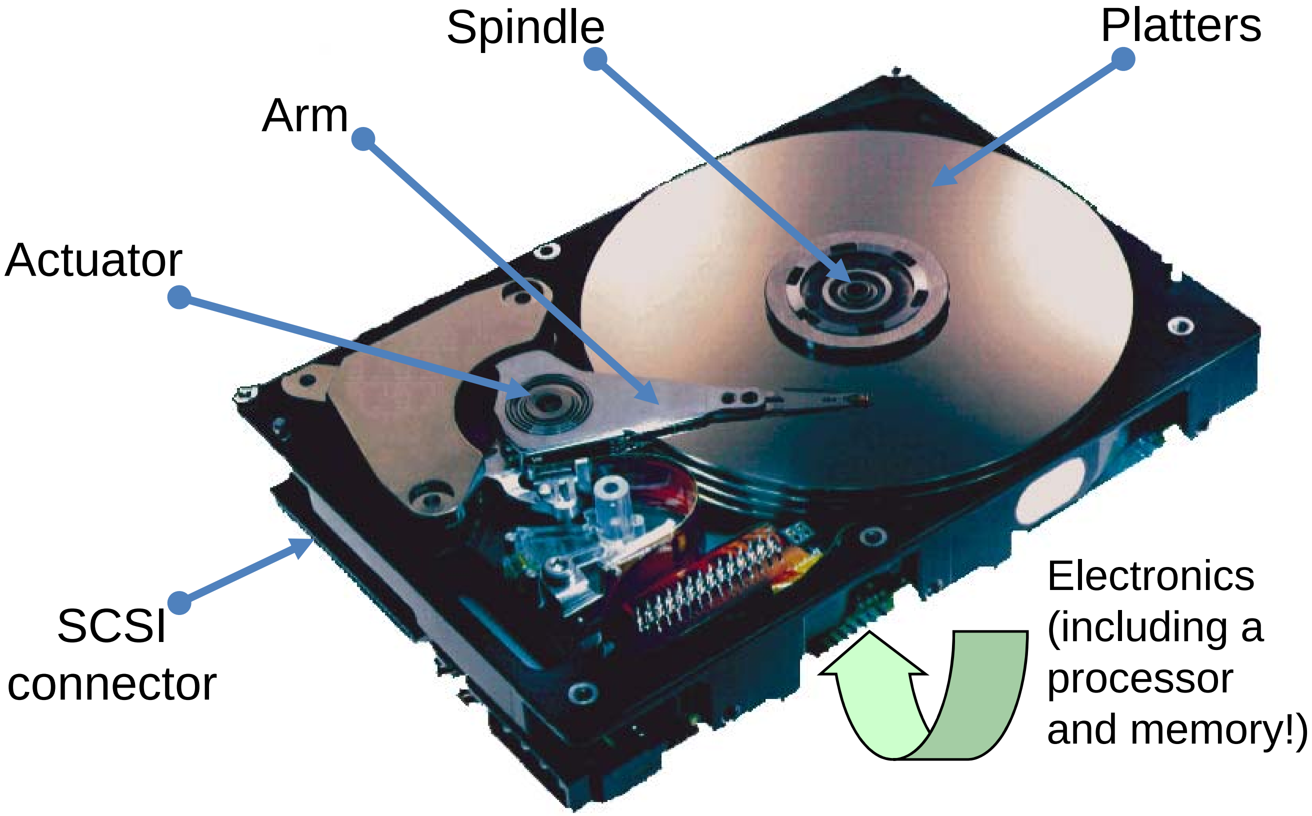
Bus structure

- Write transaction (**movq %rax,A**)
  - (Step-3) Main memory reads data word **y** from the bus and stores it at address A



Storage Technologies

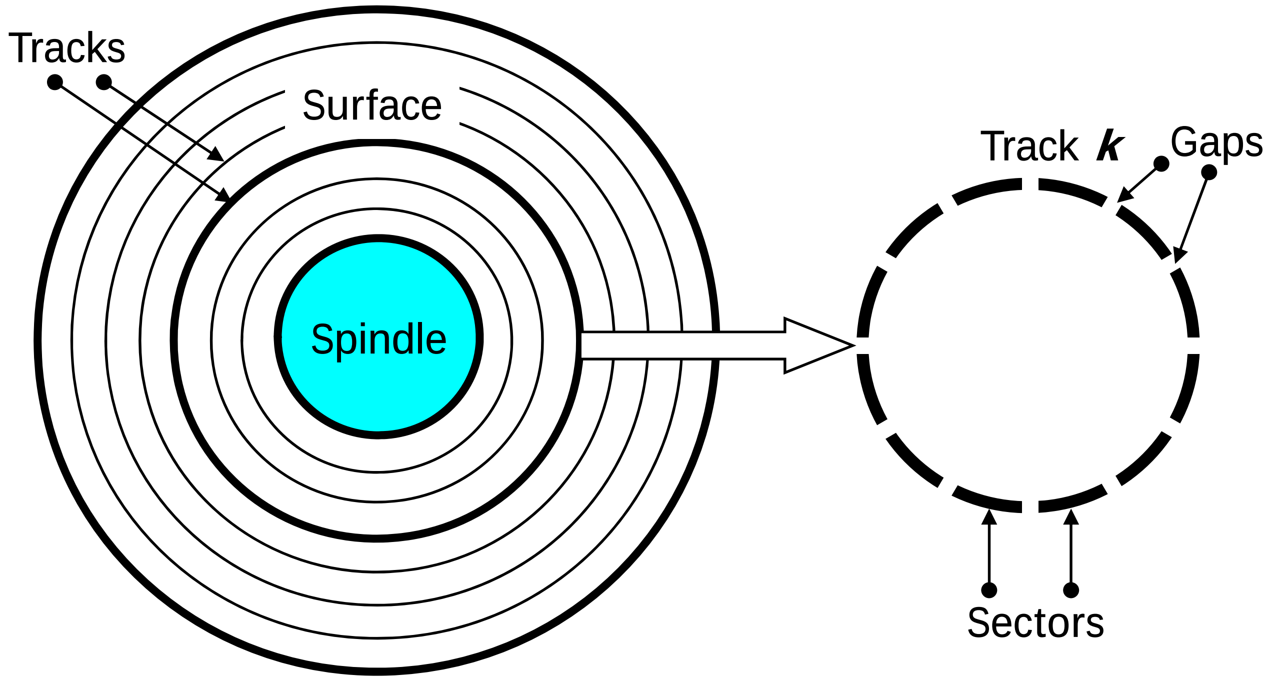
- **HDDs (Hard Disk Drives)**
  - **Inside the HDDs**



[Image courtesy of Seagate Technology]

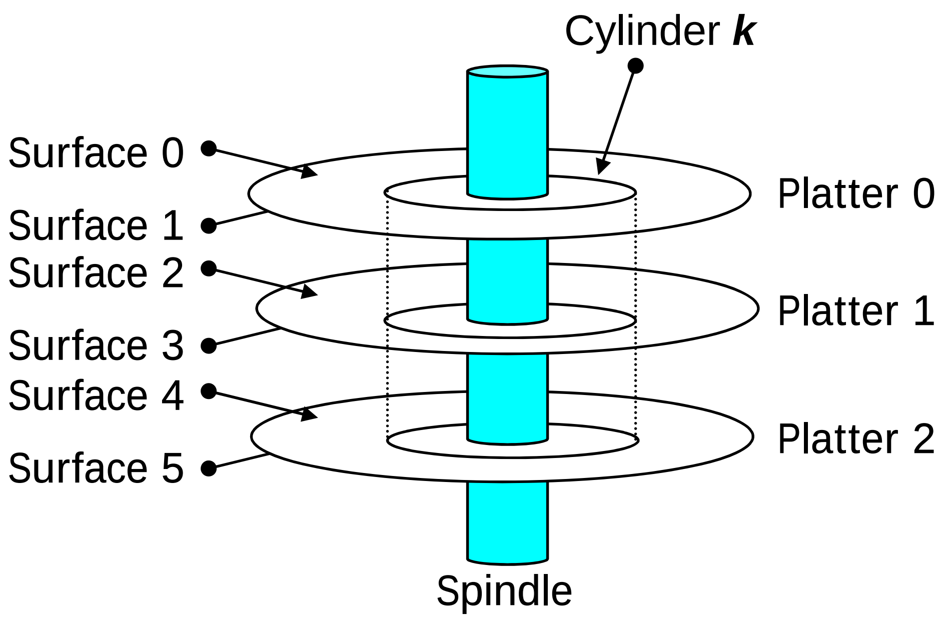
Storage Technologies

- **HDDs**
  - Disks consist of **platters**, each with two **surfaces**
  - Each surface consists of concentric rings called **tracks**
  - Each track consists of **sectors** separated by **gaps**



Storage Technologies

- **HDDs**
  - Aligned tracks form a **cylinder**



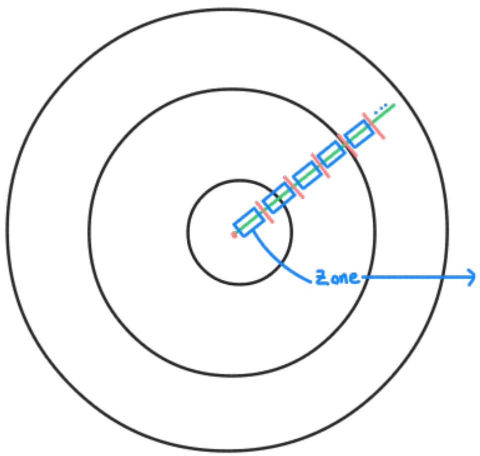
Storage Technologies

- **Disk capacity**
  - Maximum number of bits that can be stored
    - Vendors express capacity in units of gigabytes (GB), where 1GB $\cong 10^9$ Bytes
    - Capacity factors
      - ✓ Recording density (bits/in): number of bits that can be squeezed into a 1 inch segment of a track
      - ✓ Track density (tracks/in): number of tracks that can be squeezed into a 1 inch radial segment
      - ✓ Areal density (bits/in^2): product of recording and track density

Storage Technologies

■ Disk capacity

- Multiple zone recording
  - Modern disks partition tracks into disjoint subsets called **recording zones**
    - ✓ Each track in a zone has the same number of sectors, determined by the circumference of innermost track in the zone
    - ✓ Each zone has a different number of sectors per track



각 zone별 용량을 동일하게 만듦 : Multiple Zone Recording

Storage Technologies

Disk capacity

Capacity $= (\#\text{ bytes/sector}) \times (\text{avg. }\#\text{ sectors/track}) \times (\#\text{ tracks/surface}) \times (\#\text{ surfaces/platter}) \times (\#\text{ platters/disk})$

Example)

- 512 bytes/sector
- 300 sectors/track (on average)
- 20,000 tracks/surface
- 2 surfaces/platter
- 5 platters/disk

Capacity = 512 x 300 x 20,000 x 2 x 5 = 30,720,000,000 = 30.72 GB

Storage Technologies

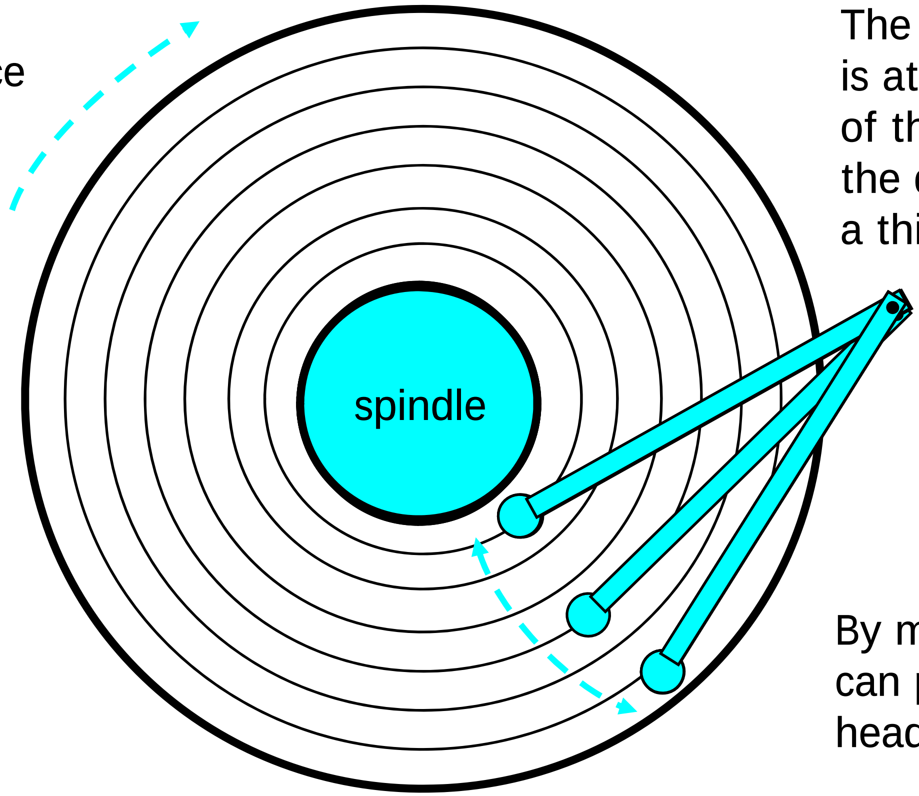
Disk operation

- Average time to access some target sector
  - $T_{access} = T_{avg\text{-}seek} + T_{avg\text{-}rotation} + T_{avg\text{-}transfer}$
  - Seek time ($T_{avg\text{-}seek}$)
    - ✓ Time to position heads over cylinder containing target sector
    - ✓ Typical $T_{avg\text{-}seek}$ is 3~9ms (Max 20ms)
  - Rotational latency ($T_{avg\text{-}rotation}$)
    - ✓ Time waiting for first bit of target sector to pass under head
    - ✓ $T_{avg\text{-}rotation} = 1/2 \times 1/\text{RPMs} \times 60\text{sec}/1\text{min}$
    - ✓ Typical $T_{avg\text{-}rotation}$ = 1~4ms
  - Transfer time ($T_{avg\text{-}transfer}$)
    - ✓ Time to read the bits in the target sector
    - ✓ $T_{avg\text{-}transfer} = 1/\text{RPM} \times 1/(\text{avg. \# sectors/track}) \times 60\text{secs}/1\text{min}$

Storage Technologies

■ Disk operation

The disk surface
spins at a fixed
rotational rate

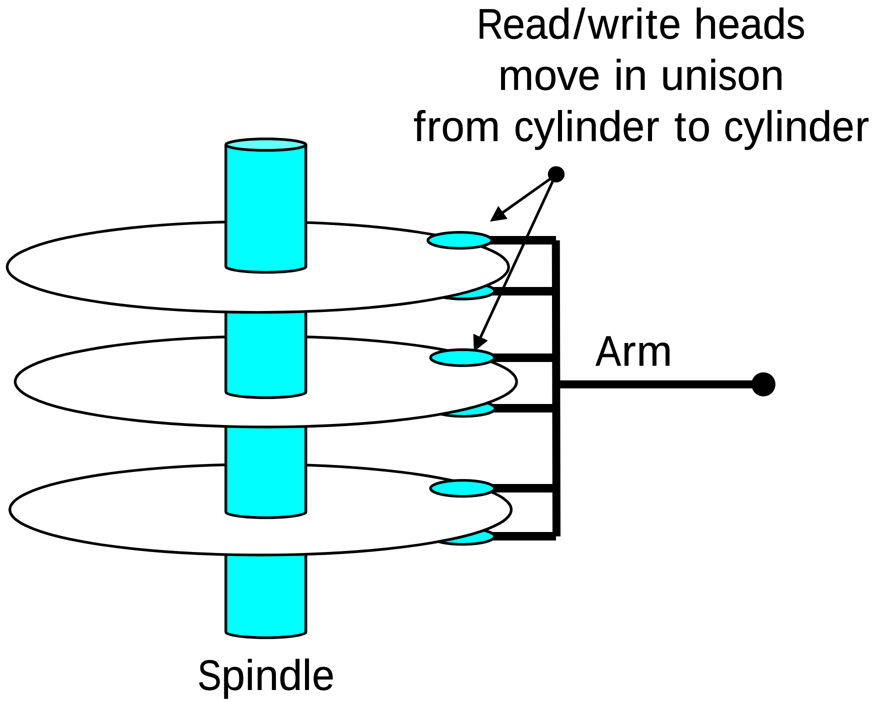


The read/write head
is attached to the end
of the arm and flies over
the disk surface on
a thin cushion of air

By moving radially, the arm
can position the read/write
head over any track

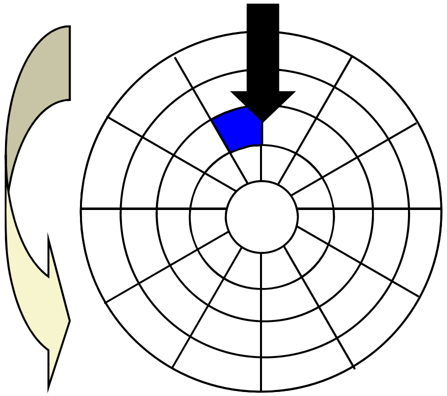
Storage Technologies

■ **Disk operation**



Storage Technologies

■ Disk operation: A scenario

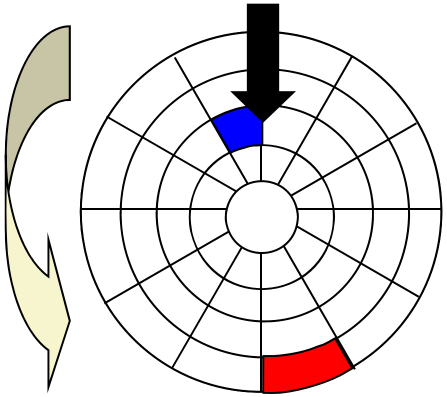


After BLUE read

After reading blue sector

Storage Technologies

■ Disk operation: A scenario

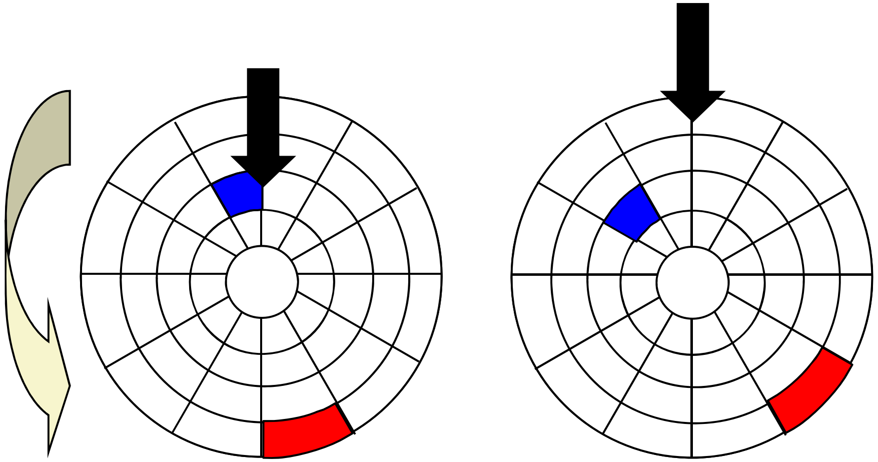


After BLUE read

Red request scheduled next

Storage Technologies

- **Disk operation: A scenario**

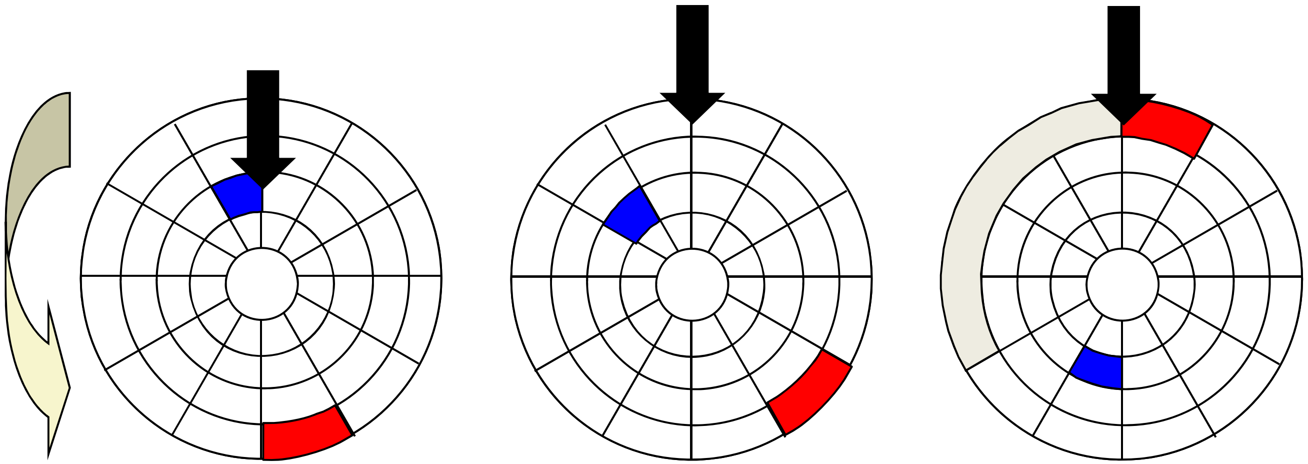


After BLUE read　　Seek for RED

Seek to red's track

Storage Technologies

- **Disk operation: A scenario**

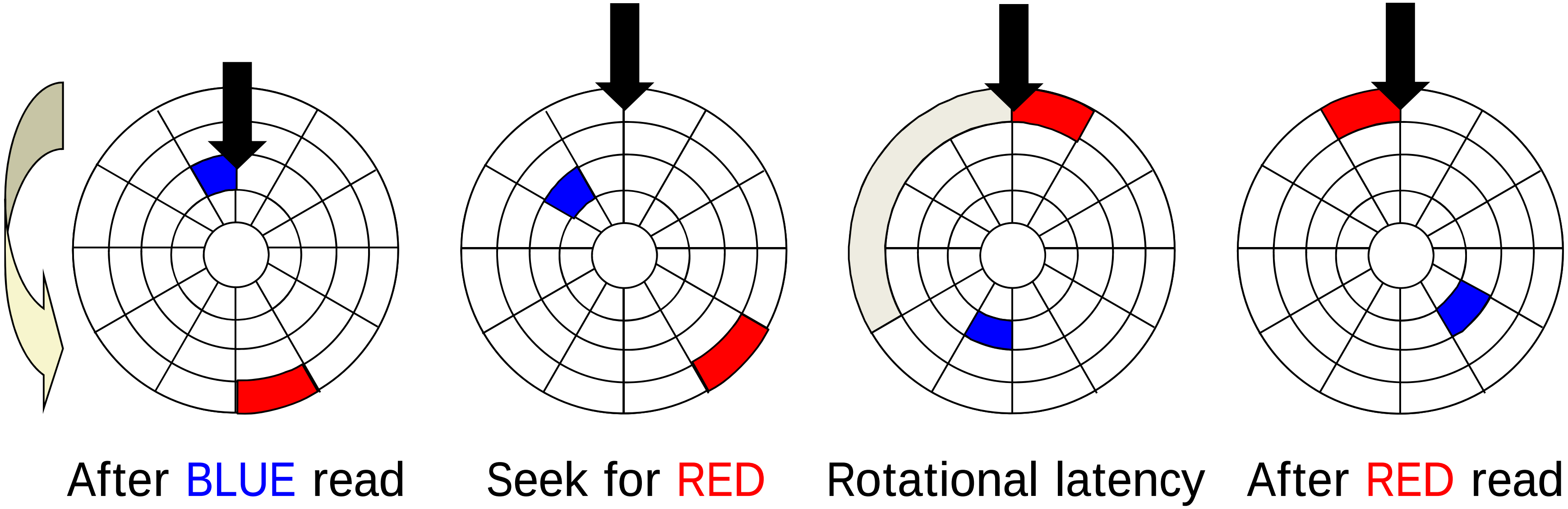


After BLUE read   Seek for RED   Rotational latency

Wait for red sector to rotate around

Storage Technologies

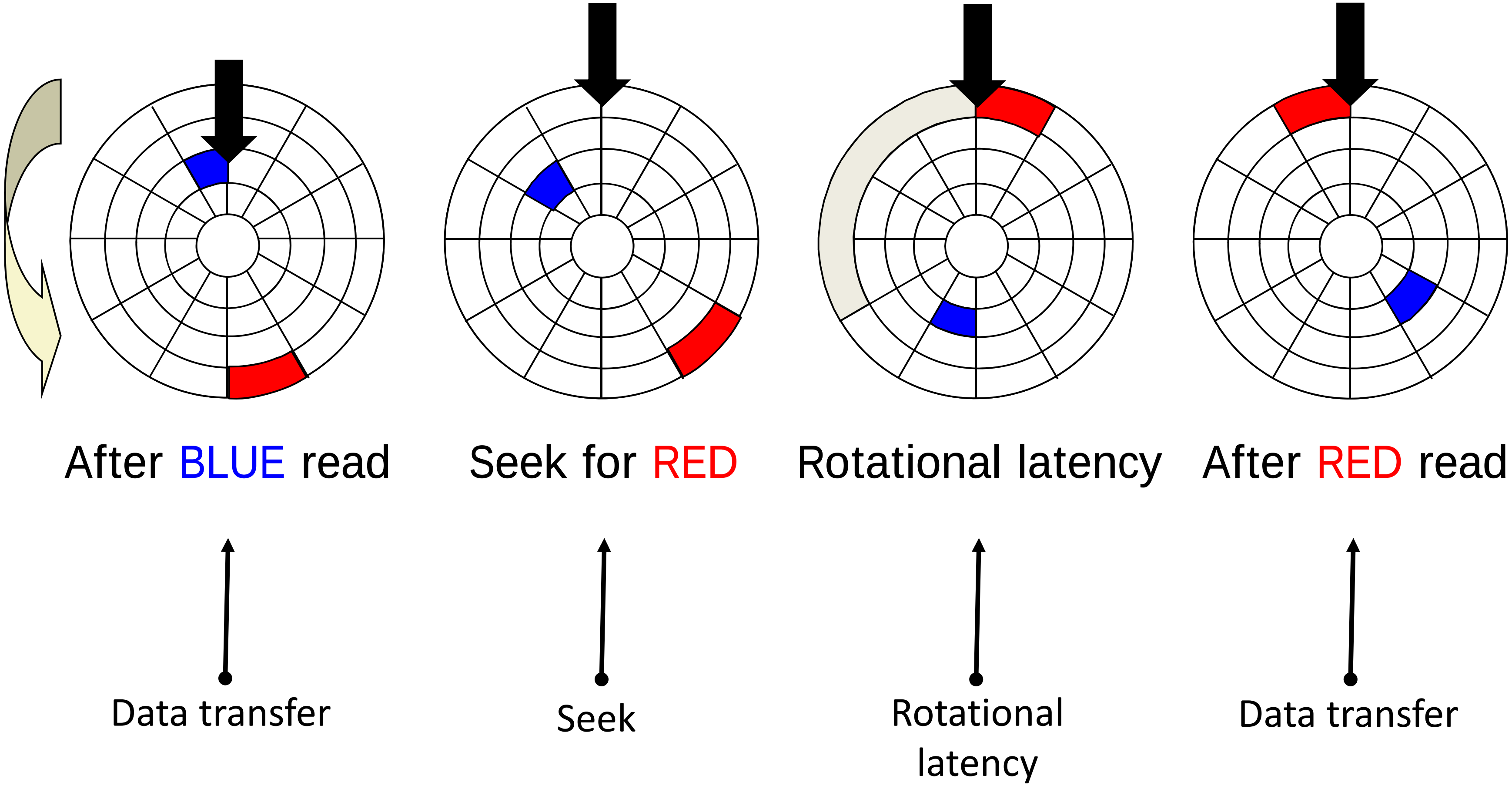
■ Disk operation: A scenario



Complete read of red

Storage Technologies

■ **Disk operation: A scenario**



Storage Technologies

Disk operation

- Example)
  - Given
    - ✓ Average seek time = 9ms
    - ✓ Rotational rate = 7,200 RPM
    - ✓ Avg. # sectors/track = 400
  - Derived
    - ✓ $T_{avg\text{-}rotation}$ = 1/2 × (60secs/7200RPM) × 1000ms/sec = 4.17ms
    - ✓ $T_{avg\text{-}transfer}$ = 1/(400sectors/track) × 60/7200RPM × 1000ms/sec = 0.02 ms
    - ✓ T_{access} = 9ms + 4.17ms + 0.02ms = 13.19ms

Storage Technologies

- **Disk operation**
  - Notes)
    - Access time dominated by seek time and rotational latency
    - SRAM access time is about 4ns/doubleword (256ns for 512B), DRAM about 60ns (about 4,000ns for 512B)
      - ✓ Disk, which has roughly 10ms access time, is about 40,000 times slower than SRAM, and 2,500 times slower then DRAM

Storage Technologies

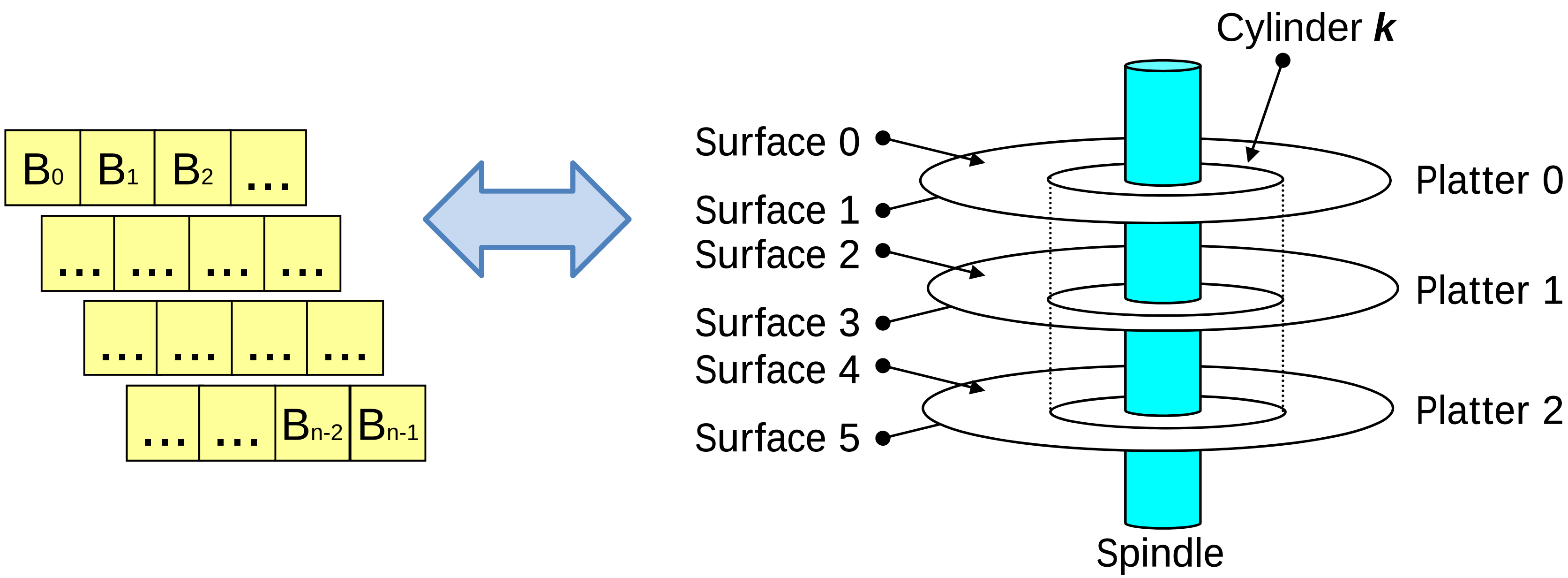
- **Logical disk blocks**
  - Modern disks present a simpler abstract view of the complex sector geometry
    - The set of available sectors is modeled as a sequence of **b**-sized logical blocks (0, 1, 2, …)

B_0	B_1	B_2	…	B_{n-2}	B_{n-1}

Storage Technologies

Logical disk blocks

- Mapping between logical blocks and physical sectors
  - Maintained by hardware/firmware device called disk controller
  - Converts requests for logical blocks into (track#, surface#, sector#) triples



Storage Technologies

- **Disk formatting**
  - $\boxed{\text{Filling in the gaps}}$ between sectors with information that identifies the sector
  - Identifying cylinders with surface defects and taking them out of action (bad sector를 out of action)
  - Setting aside a set of cylinders in each zone as spares
    - Called into action when one or more cylinders in the zone goes to bad
  - Accounts for the difference in formatted capacity and maximum capacity

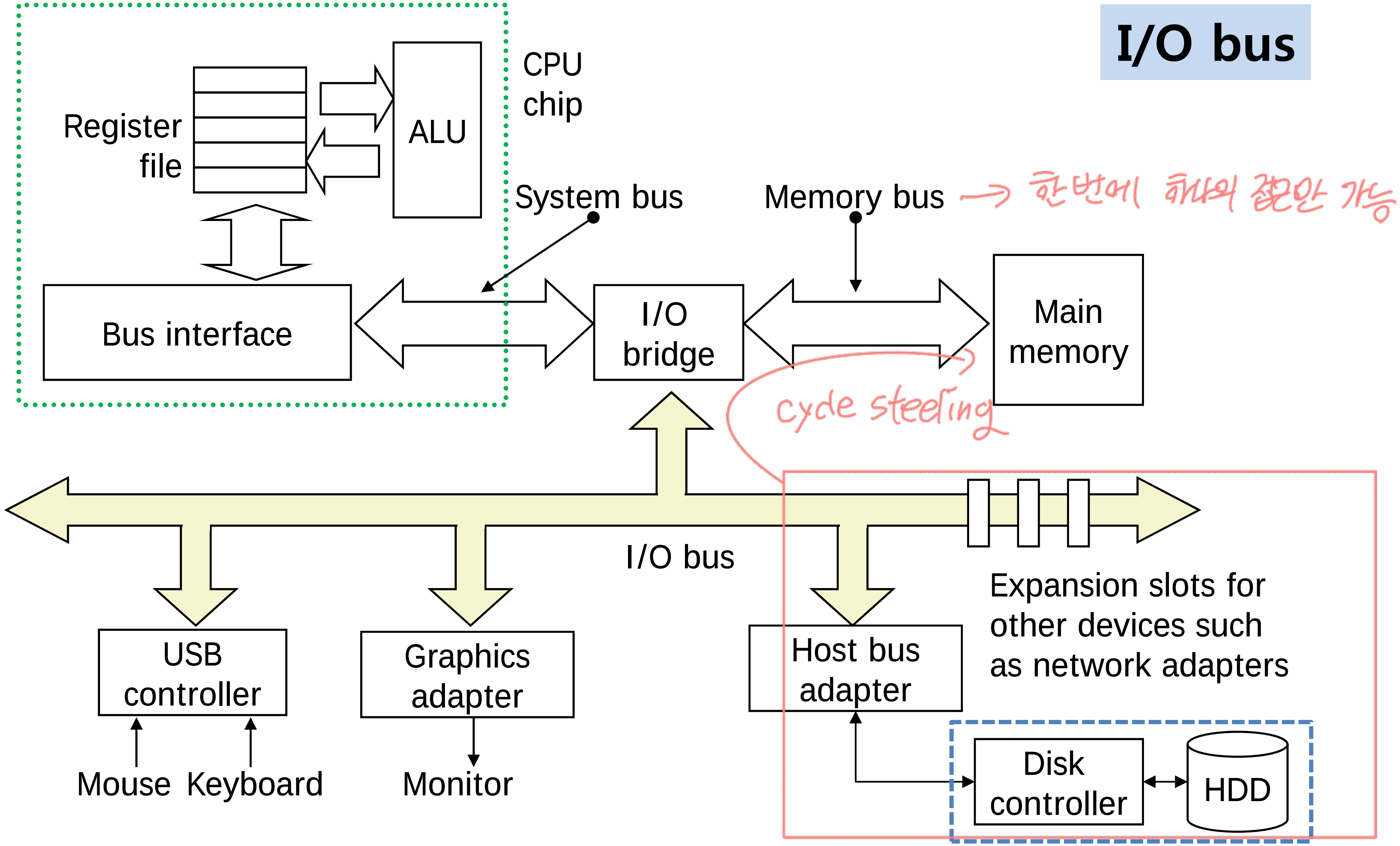
Storage Technologies

- **I/O bus**
  - Connects I/O devices to the CPU and memory
  - Designed to be independent of the underlying CPU
  - Can accommodate a wide variety of 3rd-party I/O devices
  - Eg) Intel's PCI (Peripheral Component Interconnect) bus

- **Host bus adapter**
  - Connects one or more disks to the I/O bus
  - Host bus interface
    (Communication protocol for the connection)
    - SATA (supports only one drive)
    - SCSI (supports multiple drives, faster and more expensive)

Storage Technologies

I/O bus

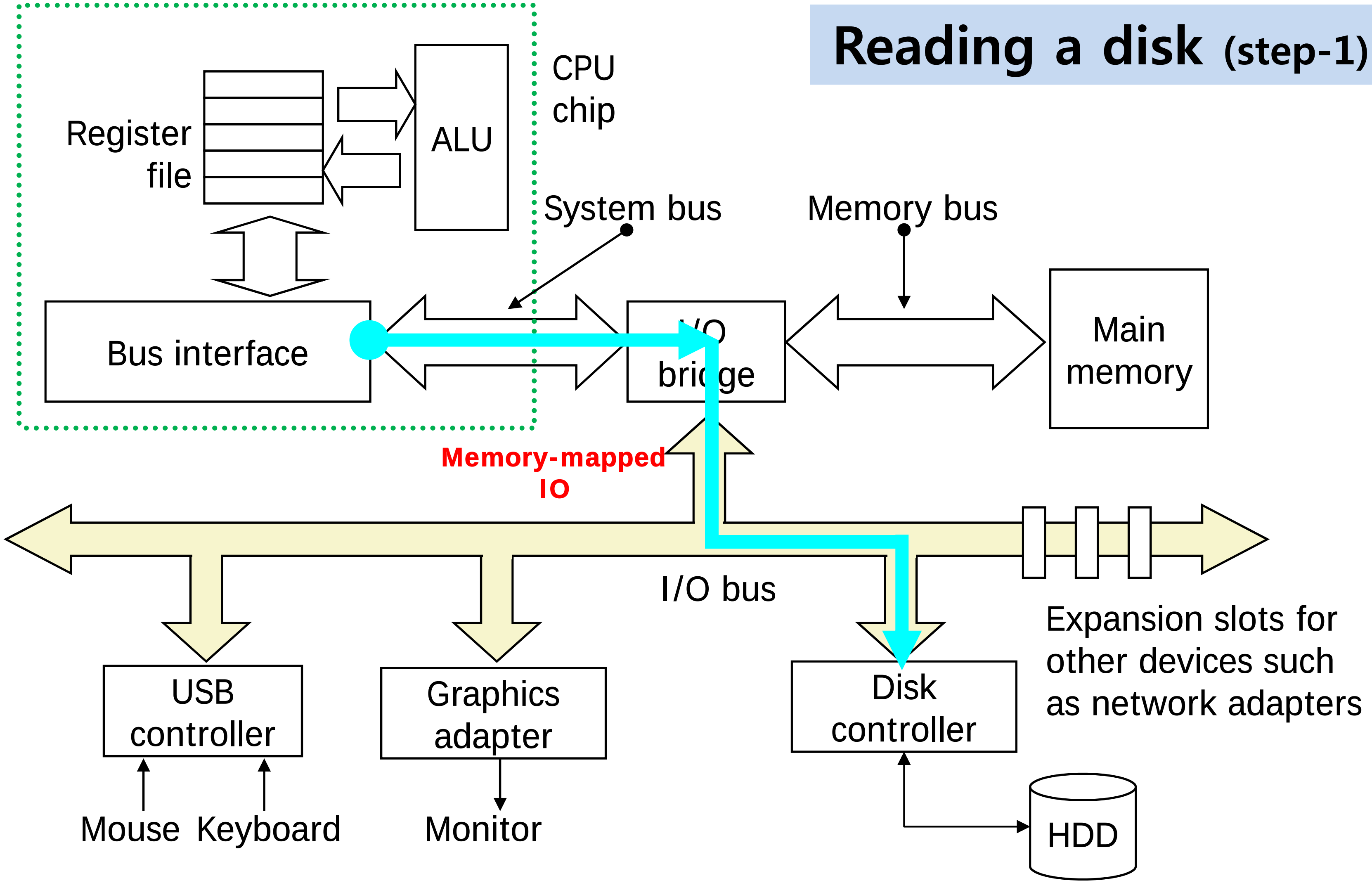


Storage Technologies

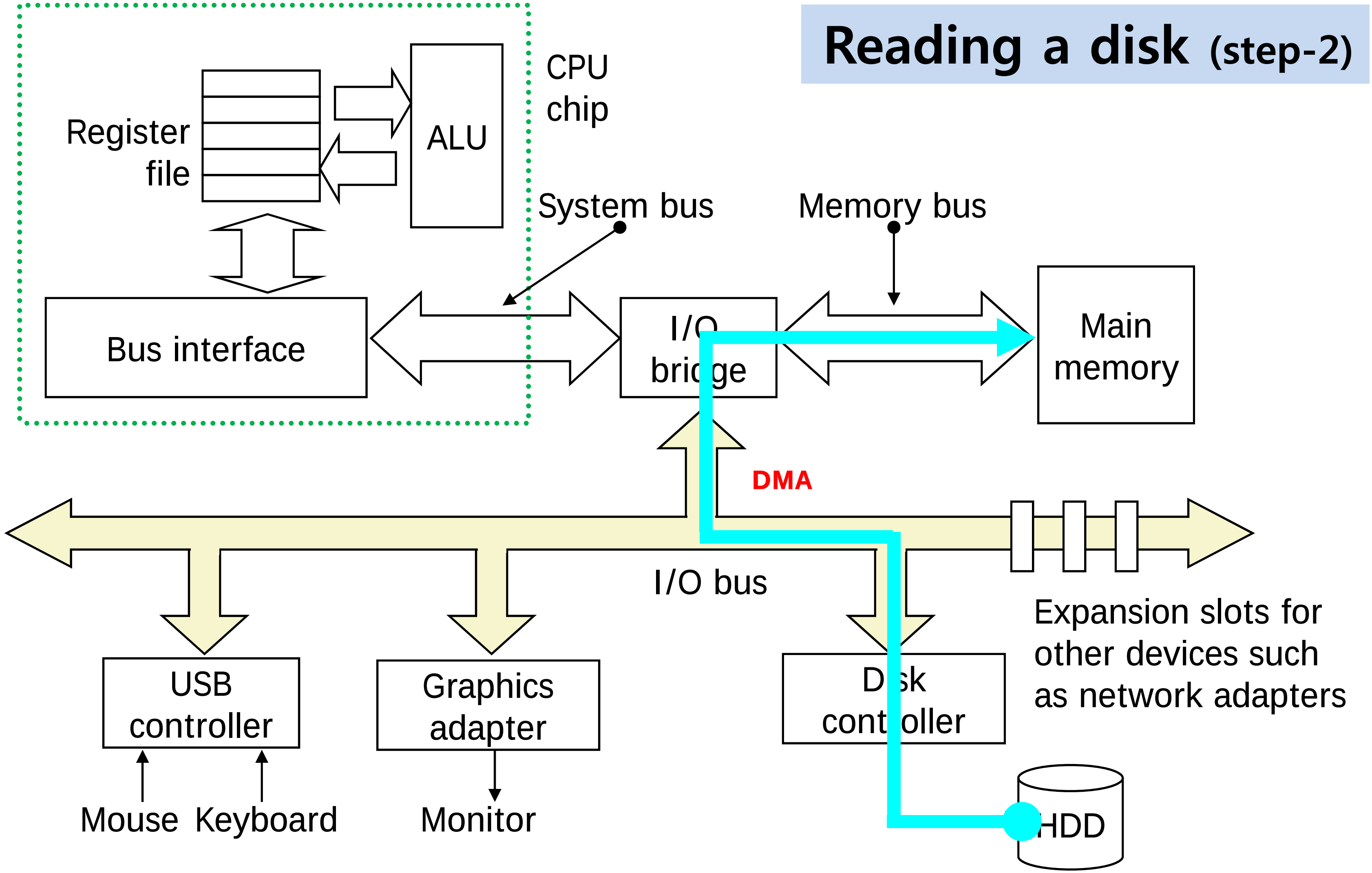
- **I/O bus**
  - Reading a disk sector
    - [Step-1] CPU initiates a disk read by writing a command, logical block number, and destination memory address to a **port** (address) associated with disk controller (**Memory-mapped I/O**)
    - [Step-2] Disk controller reads the sector and performs a **DMA (Direct Memory Access)** transfer into main memory
    - [Step-3] When the DMA transfer completes, the disk controller notifies the CPU with an **interrupt** (i.e., asserts a special "interrupt" pin on the CPU)

Storage Technologies

Reading a disk (step-1)

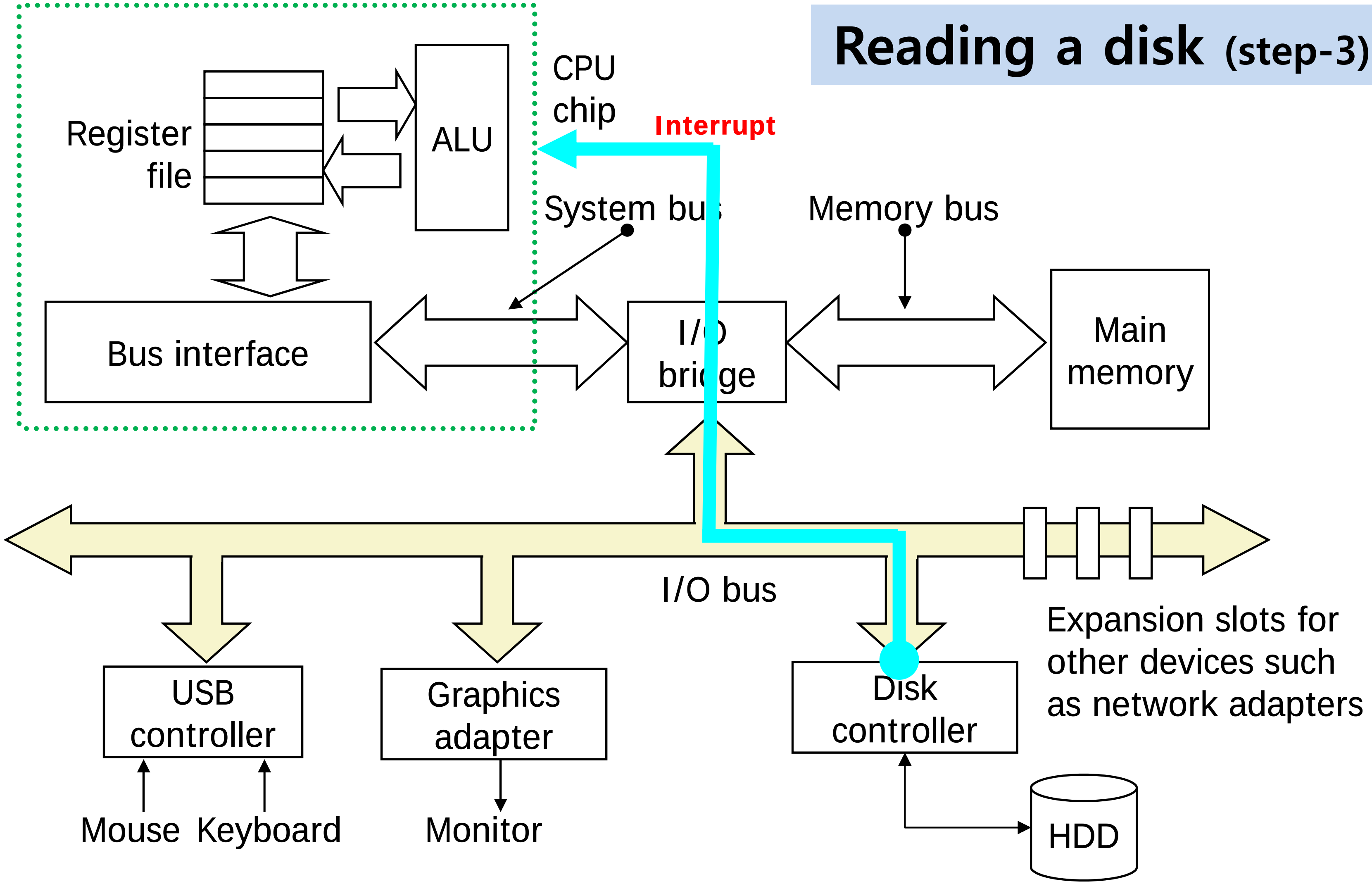


Storage Technologies



Storage Technologies

Reading a disk (step-3)



Storage Technologies

- **Commercial disks**
  - Seagate Cheetah 15K.4

Geometry attribute	Value
Platters	4
Surfaces (read/write heads)	8
Surface diameter	3.5 in.
Sector size	512 bytes
Zones	15
Cylinders	50,864
Recording density (max)	628,000 bits/in.
Track density	85,000 tracks/in.
Areal density (max)	53.4 Gbits/sq. in.
Formatted capacity	146.8 GB

Performance attribute	Value
Rotational rate	15,000 RPM
Avg. rotational latency	2 ms
Avg. seek time	4 ms
Sustained transfer rate	58–96 MB/s

Storage Technologies

Tool DIXtrac

- Developed at CMU
- Automatically discovers a wealth of low-level information about the geometry and performance of SCSI disks
- Example) **DIXtrac** on Seagate disk

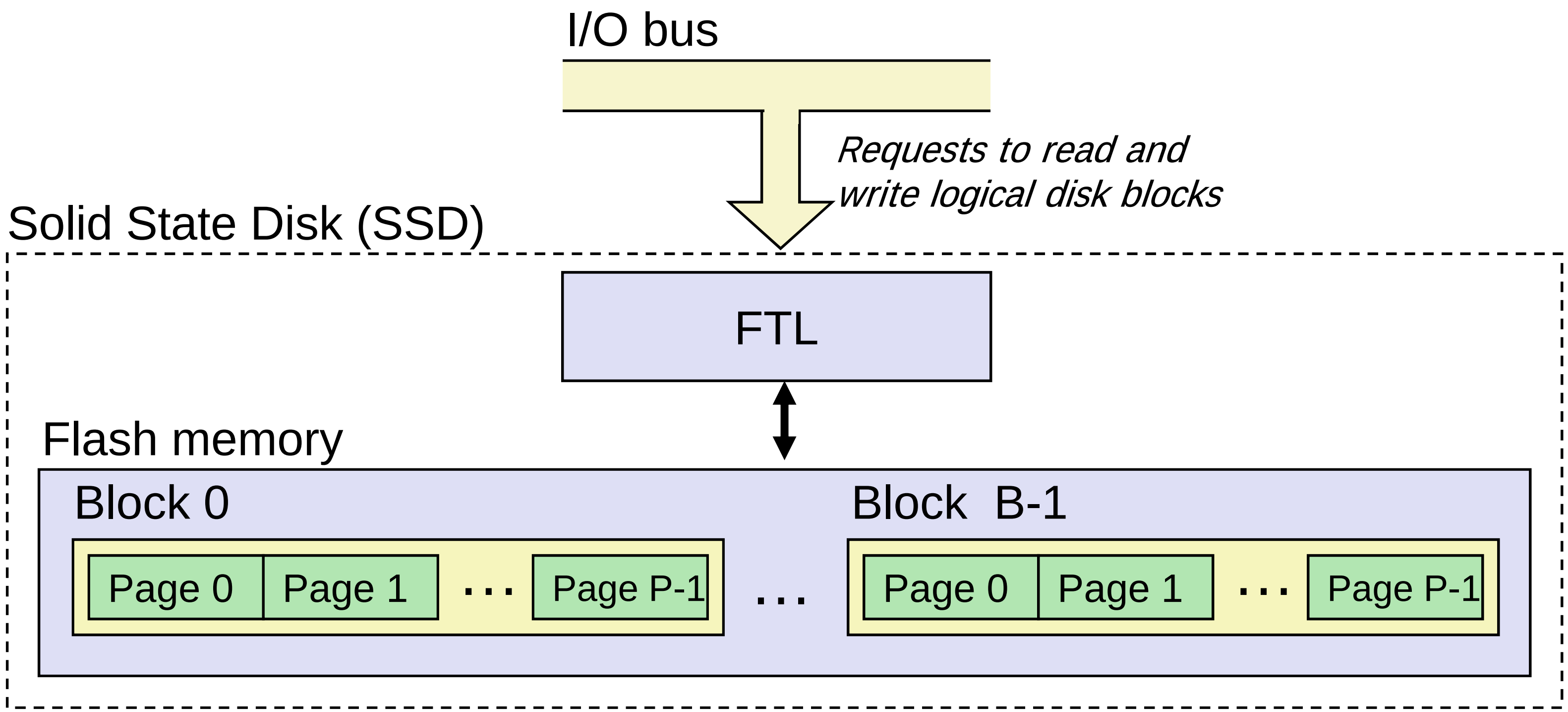
Zone number	Sectors per track	Cylinders per zone	Logical blocks per zone
(outer) 0	864	3201	22,076,928
1	844	3200	21,559,136
2	816	3400	22,149,504
3	806	3100	19,943,664
4	795	3100	19,671,480
5	768	3400	20,852,736
6	768	3450	21,159,936
7	725	3650	21,135,200
8	704	3700	20,804,608
9	672	3700	19,858,944
10	640	3700	18,913,280
11	603	3700	17,819,856
12	576	3707	17,054,208
13	528	3060	12,900,096
(inner) 14	—	—	—

Storage Technologies

- **SSDs (Solid State Disks)**
  - Based on flash memory
    - Consists of one or more flash memory chips and FTL
    - FTL(Flash Translation Layer) plays the role of disk controller
      - ✓ Translating requests for logical blocks into accesses of the underlying physical device
  - Plugs into the standard disk slot on the I/O bus (USB or SATA)
  - Behaves like any other disk

Storage Technologies

- **SSDs**



Storage Technologies

Characteristics of SSDs

- Page size
  - 512B ~ 4KB
- Block size
  - 32 ~ 128 pages (16KB ~ 512KB)
- Data is read and written in units of pages
  - A page can be written only after its block has been erased
- A block wears out after 100,000 repeated writes

Storage Technologies

■ Characteristics of SSDs

- Performance characteristics (Intel SSD 730)

Sequential read tput	550 MB/s
Random read tput (IOPS)	89,000
Random read tput (MB/s)	365 MB/s
Avg seq'l read access time	50 μs

Sequential write tput	470 MB/s
Random write tput (IOPS)	74,000
Random write tput (MB/s)	303 MB/s
Avg seq'l write access time	60 μs

- Why are random writes so slow?
  - Erasing a block is slow (around 1ms)
  - Writing a page triggers a copy of all useful pages in the block
    - ✓ Find a new block and erase it, if necessary
    - ✓ Write the page into the new block
    - ✓ Copy other pages from old block to the new block

Storage Technologies

Characteristics of SSDs

- Advantages
  - No moving parts → faster, less power, more rugged
- Disadvantages
  - Have the potential to wear out
    - ✓ Mitigated by "wear leveling logic" in FTL
  - About 30× more expensive per byte than in HDD
- Applications
  - Smart phones, laptops, digital cameras
  - Desktops and servers

Storage Technologies

■ Storage trends

SRAM

Metric	1985	1990	1995	2000	2005	2010	2015	*2015:1985*
$/MB	2,900	320	256	100	75	60	25	116 ↓
access (ns)	150	35	15	3	2	1.5	1.3	115 ↓

DRAM

Metric	1985	1990	1995	2000	2005	2010	2015	*2015:1985*
$/MB	880	100	30	1	0.1	0.06	0.02	44,000 ↓
access (ns)	200	100	70	60	50	40	20	10 ↓
typical size (MB)	0.256	4	16	64	2,000	8,000	16,000	62,500 ↑

HDD

Metric	1985	1990	1995	2000	2005	2010	2015	*2015:1985*
$/GB	100,000	8,000	300	10	5	0.3	0.03	3,333,333 ↓
access (ms)	75	28	10	8	5	3	3	25 ↓
typical size (GB)	0.01	0.16	1	20	160	1,500	3,000	300,000 ↑

Storage Technologies

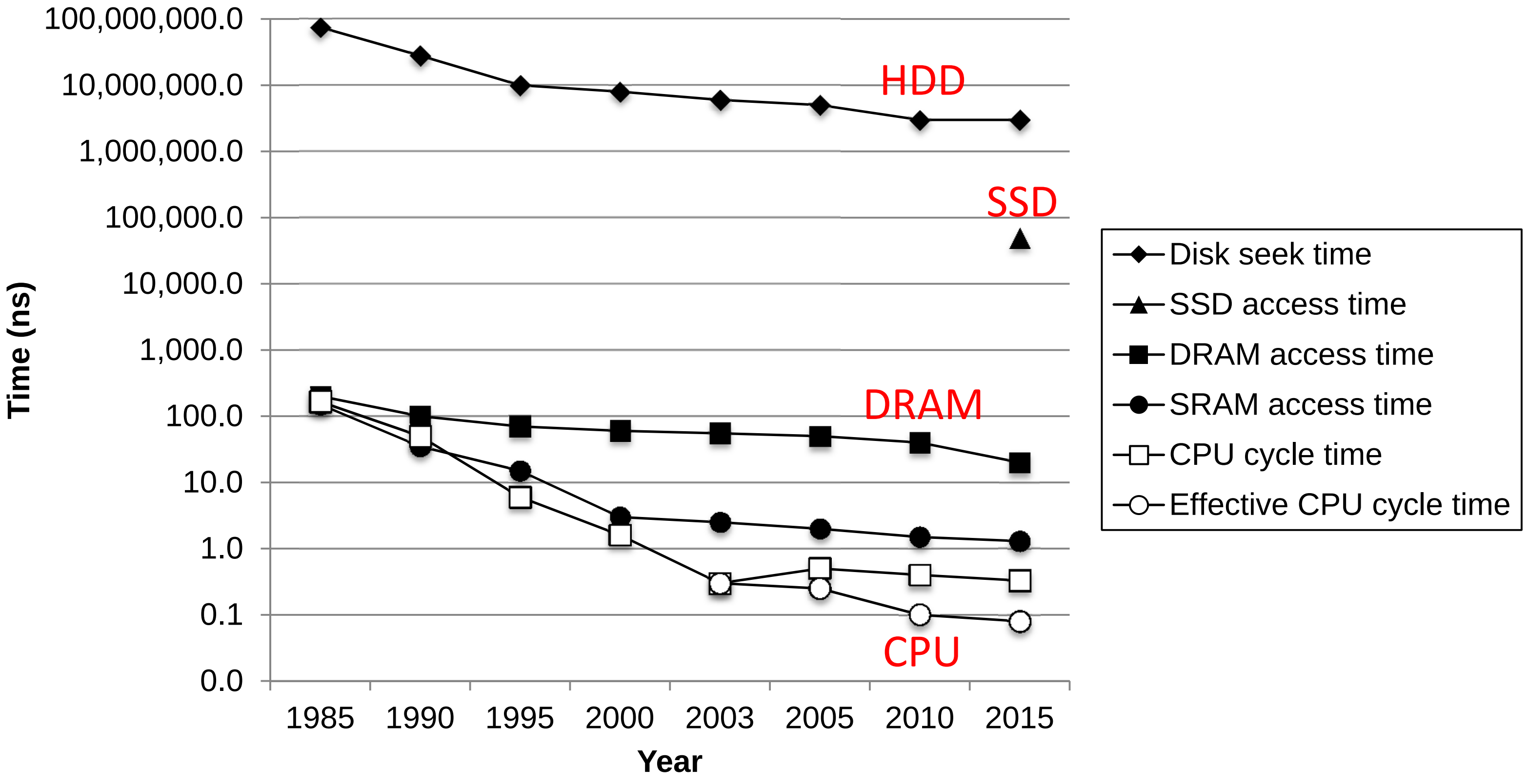
■ CPU clock rates

Inflection point in computer history when designers hit the "Power Wall"

	1985	1990	1995	2003	2005	2010	2015	2015:1985
CPU	80286	80386	Pentium	P-4	Core 2	Core i7(n)	Core i7(h)	---
Clock rate (MHz)	6	20	150	3,300	2,000	2,500	3,000	500 ↑
Cycle time (ns)	166	50	6	0.3	0.5	0.4	0.33	500 ↓
Cores	1	1	1	1	2	4	4	4 ↑
Effective cycle time (ns)	166	50	6	0.3	0.25	0.1	0.08	2,075 ↓

Storage Technologies

■ The CPU-memory gap



Storage Technologies

- **Notes)**
  - The gap between DRAM - disk performance and CPU performance is widening

  - Locality to the rescue!
    - The key to bridging this CPU-memory gap is a fundamental property of computer programs known as **locality**

Summary