VIETNAM NATIONAL UNIVERSITY HANOI (VNU)

INFORMATION TECHNOLOGY INSTITUTE

Computer Architecture

Lecture 6: External memories

Duy-Hieu Bui

VNU Information Technology Institute

Laboratory for Smart Integrated System (SISLAB)

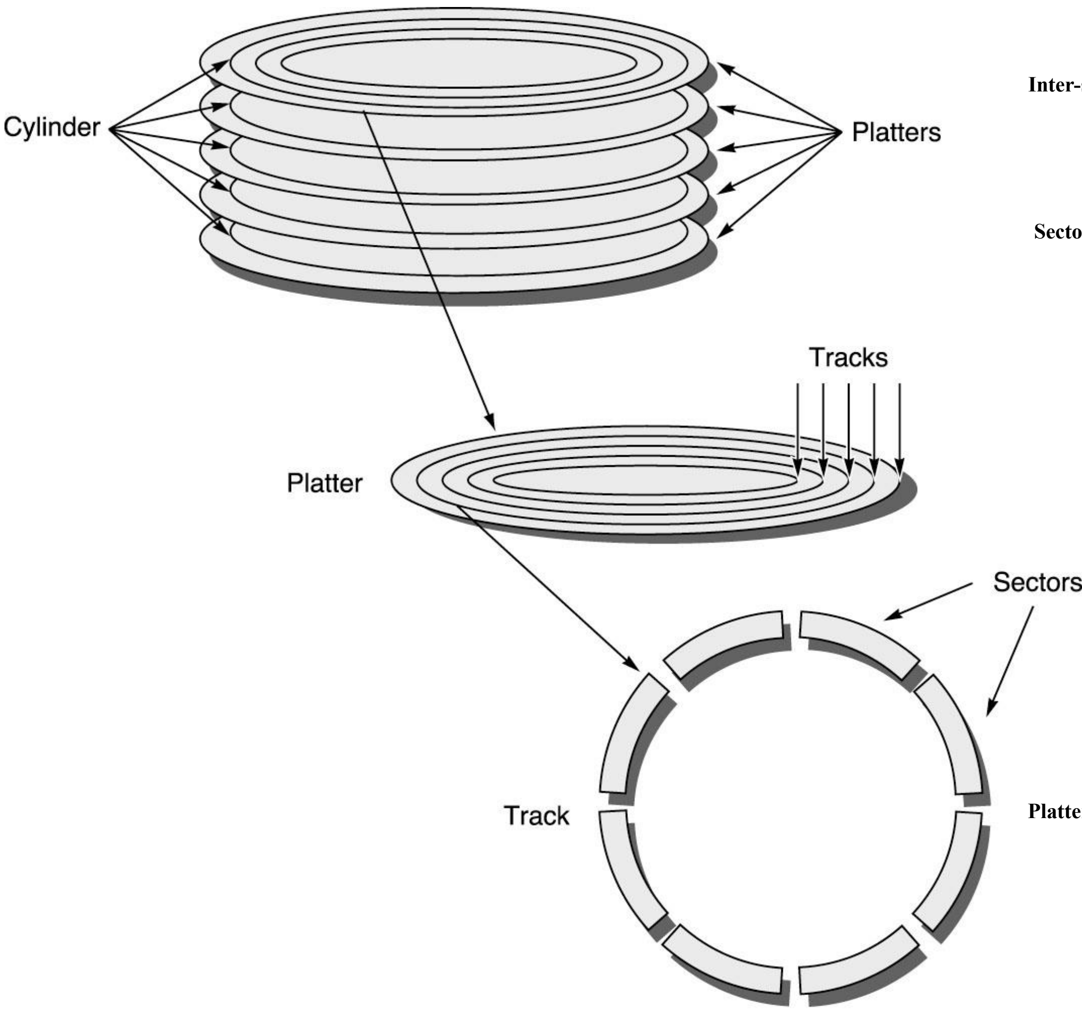
Email: hieubd@vnu.edu.vn

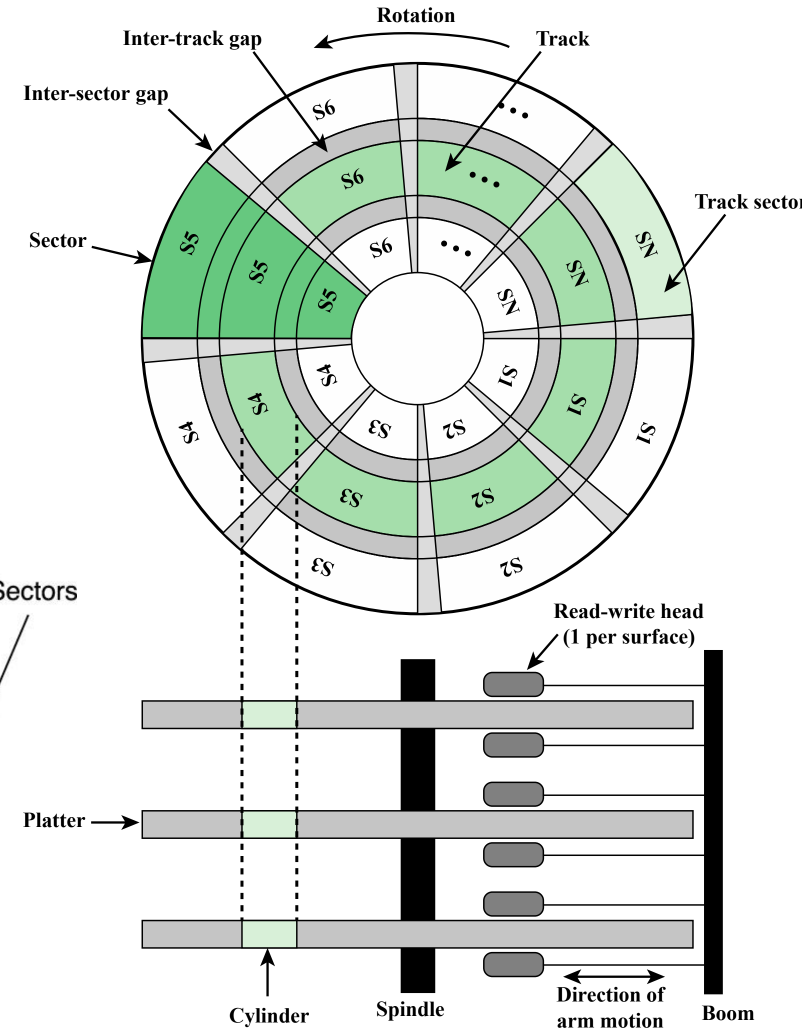
https://duyhieubui.github.io

4. Bộ nhớ ngoài

- RAM for storage
  - Flash
- Đĩa từ - Magnetic Disk
  - HDD
  - RAID
  - Removable
- Đĩa quang - Optical
  - CD-ROM
  - CD-Recordable (CD-R)
  - CD-R/W
  - DVD
- Băng từ - Magnetic Tape

Magnetic Disks





HDD sector layout

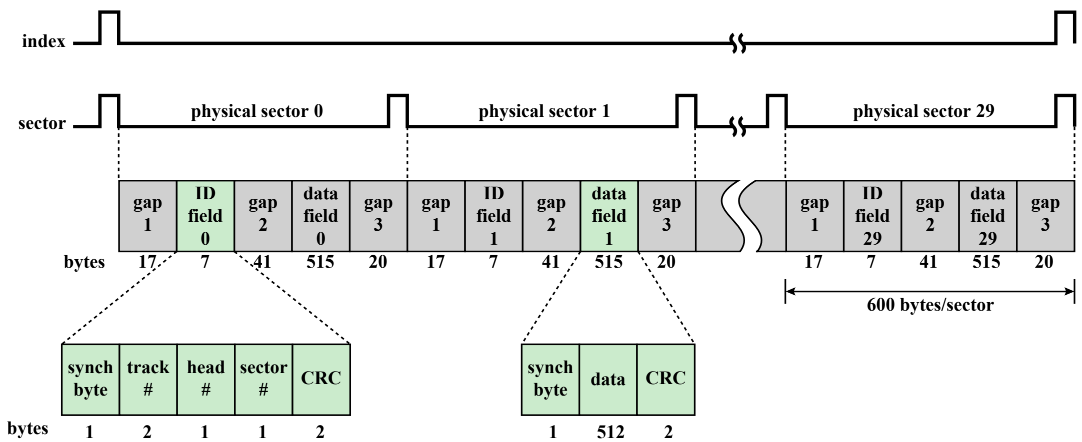


Figure 6.4 Winchester Disk Format (Seagate ST506)

HDD characteristics

Head Motion
- Fixed head (one per track)
- Movable head (one per surface)

Disk Portability
- Nonremovable disk
- Removable disk

Sides
- Single sided
- Double sided

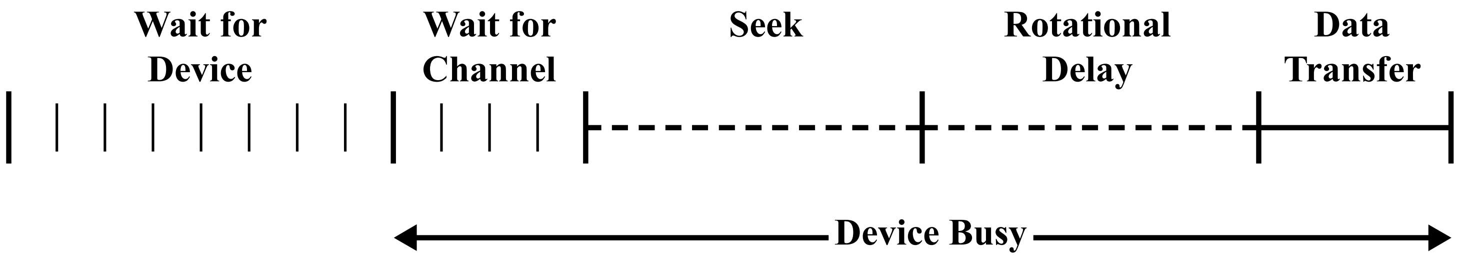
Platters
- Single platter
- Multiple platter

Head Mechanism
- Contact (floppy)
- Fixed gap
- Aerodynamic gap (Winchester)

Typical HDD parameters

Characteristics	Seagate Enterprise	Seagate Barracuda XT	Seagate Cheetah NS	Seagate Laptop HDD
Application	Enterprise	Desktop	Network attached storage, application servers	Laptop
Capacity	6 TB	3 TB	600 GB	2 TB
Average seek time	4.16 ms	N/A	3.9 ms read 4.2 ms write	13 ms
Spindle speed	7200 rpm	7200 rpm	10, 075 rpm	5400 rpm
Average latency	4.16 ms	4.16 ms	2.98	5.6 ms
Maximum sustained transfer rate	216 MB/s	149 MB/s	97 MB/s	300 MB/s
Bytes per sector	512/4096	512	512	4096
Tracks per cylinder (number of platter surfaces)	8	10	8	4
Cache	128 MB	64 MB	16 MB	8 MB

Timing of a disk I/O transfer



Magnetic Disks

- Ưu điểm: rẻ($/MB), độ tin cậy chấp nhận được
  - Primary storage, memory swapping
- Nhược: chỉ có thể đọc/ghi toàn bộ một sector
  - ➔ không thể truy cập trực tiếp như bộ nhớ chính
- Thời gian truy cập đĩa
  - Queuing delay
    - Wait until disk gets to do this operation
  - Seek time
    - Head moves to correct track
  - Rotational latency
    - Correct sector must get under the head
  - Data transfer time and controller time

Khuynh hướng của đĩa từ

- Dung lượng: ~gấp đôi sau mỗi năm

- Thời gian truy cập trung bình
  - 5-12ms (việc cải thiện tốc độ tương đối chậm)

- Latency quay trung bình (1/2 full rotation)
  - 5,000 RPM to 10,000 RPM to 15,000 RPM
  - Cải thiện chậm, phức tạp (reliability, noise)

- Tốc độ truyền dữ liệu
  - Phục thuộc vào bề mặt đĩa, dữ liệu trên một track

Đĩa quang

- Giới hạn bởi các chuẩn
  - CD and DVD capacity fixed over years
  - Technology actually improves, but it takes time for it to make it into new standards

- Kích thước bé, dễ thay thế
  - Good for backups and carrying around

Băng từ

- Thời gian truy cập rất chậm
  - Must rewind tape to correct place for read/write
- Chi phí thấp ($/MB)
  - It’s just miles of tape!
  - But disks have caught up anyway…
- Được sử dụng để sao lưu dữ liệu (secondary storage)
  - Large capacity & Replaceable

RAM for Storage

- Disks are about 100 times cheaper ($/MB)

- DRAM is about 100,000 faster (latency)

- Solid-State Disks
  - Actually, a DRAM and a battery
    - Much faster than disk, more reliable
    - Expensive (not very good for archives and such)

- Flash memory
  - Much faster than disks, but slower than DRAM
  - Very low power consumption
  - Can be sold in small sizes (few GB, but tiny)

SSD Compared to HDD

- SSDs have the following advantages over HDDs:

■ High-performance input/output operations per second (IOPS)

■ Durability

■ Longer lifespan

■ Lower power consumption

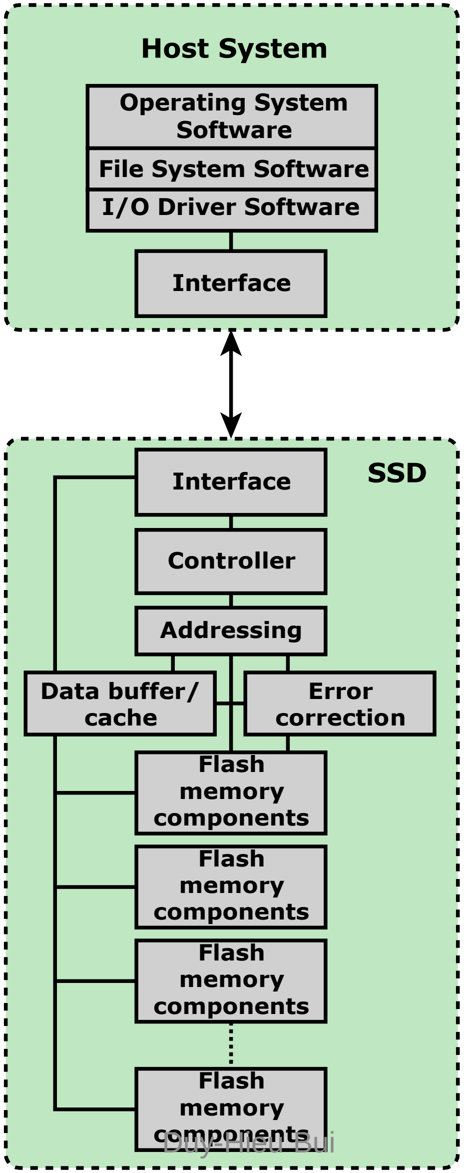
■ Quieter and cooler running capabilities

■ Lower access times and latency rates

Comparison of SSD and HDD

	NAND Flash Drives	Seagate Laptop Internal HDD
File copy/write speed	200—550 Mbps	50—120 Mbps
Power draw/battery life	Less power draw, averages 2–3 watts, resulting in 30+ minute battery boost	More power draw, averages 6–7 watts and therefore uses more battery
Storage capacity	Typically not larger than 512 GB for notebook size drives; 1 TB max for desktops	Typically around 500 GB and 2 TB maximum for notebook size drives; 4 TB max for desktops
Cost	Approx. $0.50 per GB for a 1-TB drive	Approx $0.15 per GB for a 4-TB drive

Solid State Drive Architecture



Practical Issues

There are two practical issues peculiar to SSDs that are not faced by HDDs:

- SDD performance has a tendency to slow down as the device is used
  - 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

- Flash memory becomes unusable after a certain number of 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

Một số vấn đề lỗi

- Phân loại theo nguyên nhân
  - Hardware Faults
    - Hardware devices fail to perform as designed
  - Design Faults
    - Faults in software and some faults in HW
    - E.g. the Pentium FDIV bug was a design fault
  - Operation Faults
    - Operator and user mistakes
  - Environmental Faults
    - Fire, power failure, sabotage, etc.

Khắc phục: cải thiện độ tin cậy

- Fault Avoidance
  - Prevent occurrence of faults by construction
- Fault Tolerance
  - Prevent faults from becoming failures
  - Typically done through redundancy
- Error Removal
  - Removing latent errors by verification
- Error Forecasting
  - Estimate presence, creation, and consequences of errors

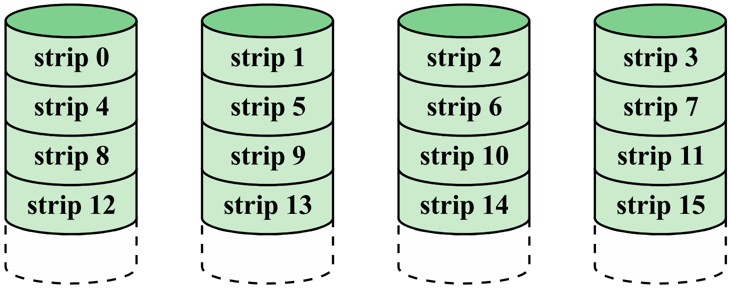
Disk Fault Tolerance with RAID

- Redundant Array of Independent/Inexpensive Disks
  - Several smaller disks play a role of one big disk
- Can improve performance
  - Data spread among multiple disks
  - Accesses to different disks go in parallel
- Can improve reliability
  - Data can be kept with some redundancy

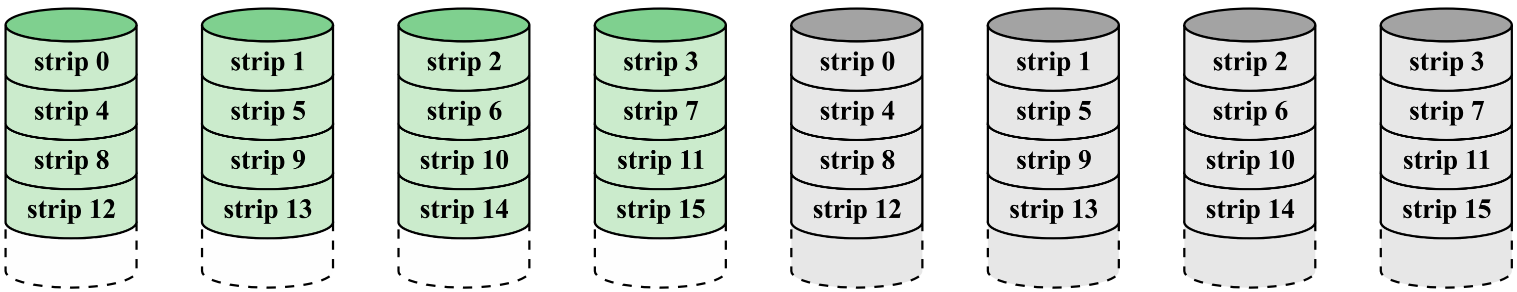
Raid Levels

Category	Level	Description	Disks Required	Data Availability	Large I/O Data Transfer Capacity	Small I/O Request Rate
Striping	0	Nonredundant	N	Lower than single disk	Very high	Very high for both read and write
Mirroring	1	Mirrored	$2N$	Higher than RAID 2, 3, 4, or 5; lower than RAID 6	Higher than single disk for read; similar to single disk for write	Up to twice that of a single disk for read; similar to single disk for write
Parallel access	2	Redundant via Hamming code	$N + m$	Much higher than single disk; comparable to RAID 3, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
	3	Bit-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
Independent access	4	Block-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 5	Similar to RAID 0 for read; significantly lower than single disk for write	Similar to RAID 0 for read; significantly lower than single disk for write
	5	Block-interleaved distributed parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 4	Similar to RAID 0 for read; lower than single disk for write	Similar to RAID 0 for read; generally lower than single disk for write
	6	Block-interleaved dual distributed parity	$N + 2$	Highest of all listed alternatives	Similar to RAID 0 for read; lower than RAID 5 for write	Similar to RAID 0 for read; significantly lower than RAID 5 for write

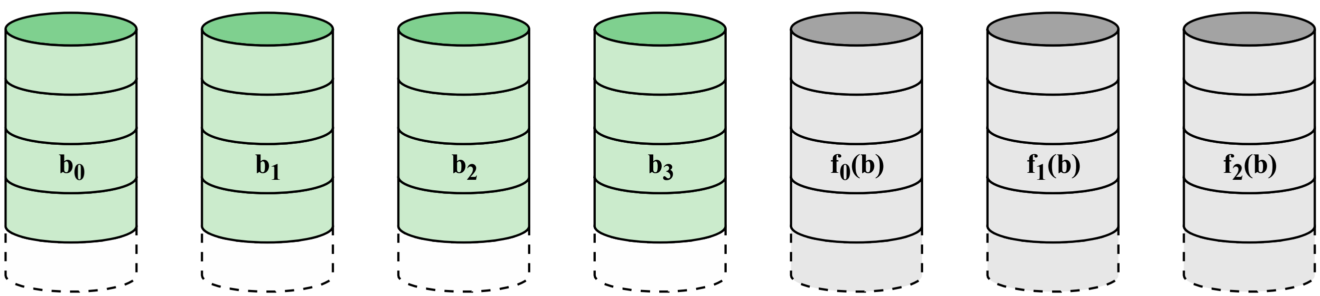
RAID 0, 1, 2



(a) RAID 0 (non-redundant)

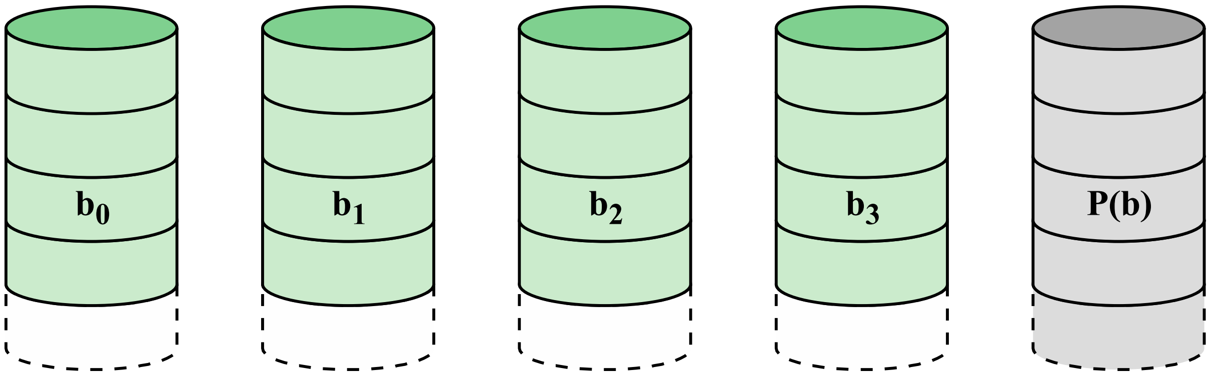


(b) RAID 1 (mirrored)

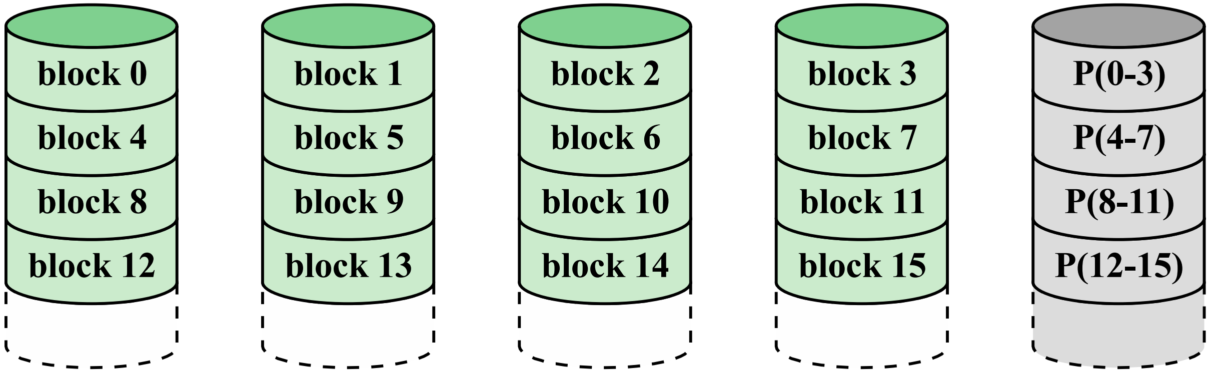


(c) RAID 2 (redundancy through Hamming code)

RAID 3 & 4

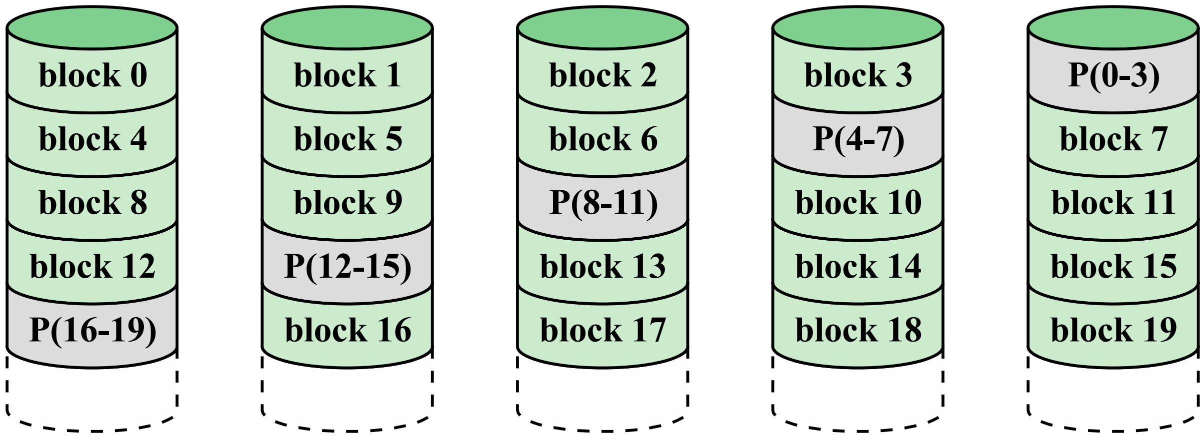


(d) RAID 3 (bit-interleaved parity)

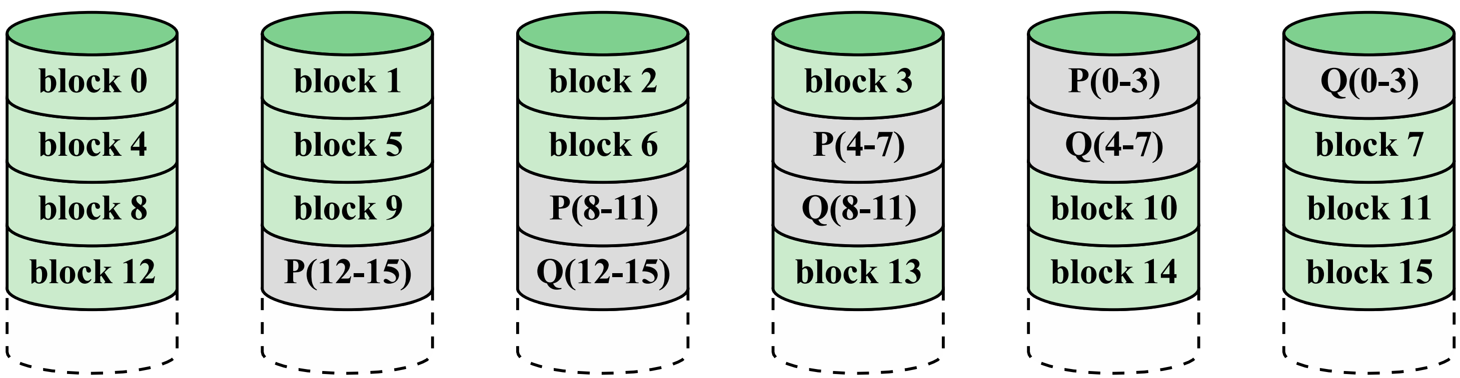


(e) RAID 4 (block-level parity)

RAID 5 & 6

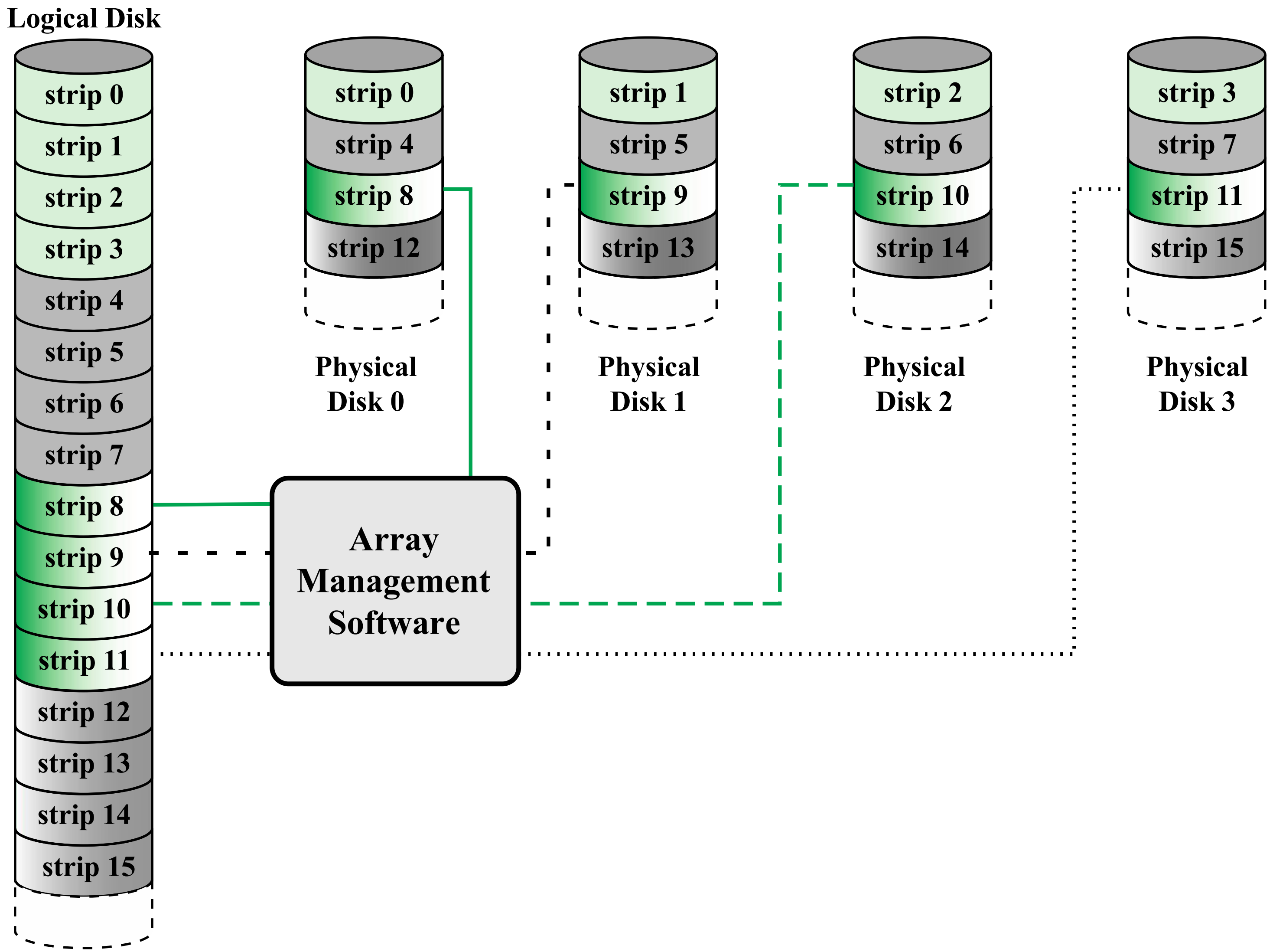


(f) RAID 5 (block-level distributed parity)



(g) RAID 6 (dual redundancy)

Data mapping for raid level 0



Raid summary

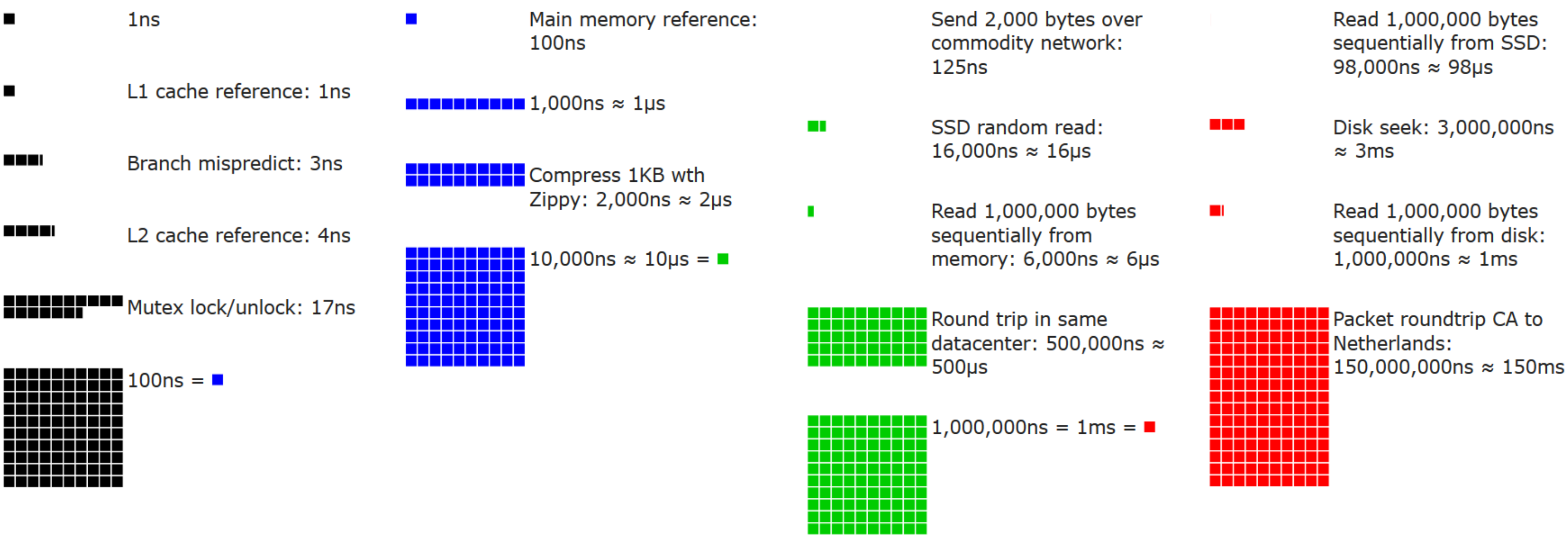
Level	Advantages	Disadvantages	Applications
0	I/O performance is greatly improved by spreading the I/O load across many channels and drives No parity calculation overhead is involved Very simple design Easy to implement	The failure of just one drive will result in all data in an array being lost	Video 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 summary (cont.)

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 Prepress applications 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 viable
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

Latency Numbers Every Programmer Should Know

	Time (ns)	Comment
L1 cache reference	1	> 2 ALU instruction latency
Branch mispredict	3	
L2 cache reference	4	4x L1 cache
Mutex lock/unlock	17	
Main memory reference	100	25x L2 cache, 100x L1 cache



http://people.eecs.berkeley.edu/~rcs/research/interactive_latency.html

Tổng kết

- Khái niệm về bộ nhớ trong máy tính, các đặc điểm chính, …
- Bộ nhớ chính
  - Nguyên tắc
  - Phân loại
  - Tổ chức bộ nhớ lớn
- Bộ nhớ cache
  - Nguyên tắc chung
  - Các phương pháp ánh xạ
  - Các giải thuật thay thế, chính sách ghi, …
- Bộ nhớ ngoài: ổ cứng, quang, RAID, …