TRƯỜNG ĐH CÔNG NGHỆ THÔNG TIN
KHOA MẠNG MÁY TÍNH & TRUYỀN THÔNG

Digital Forensics
Pháp chứng Kỹ thuật số

#3: Disk Forensics

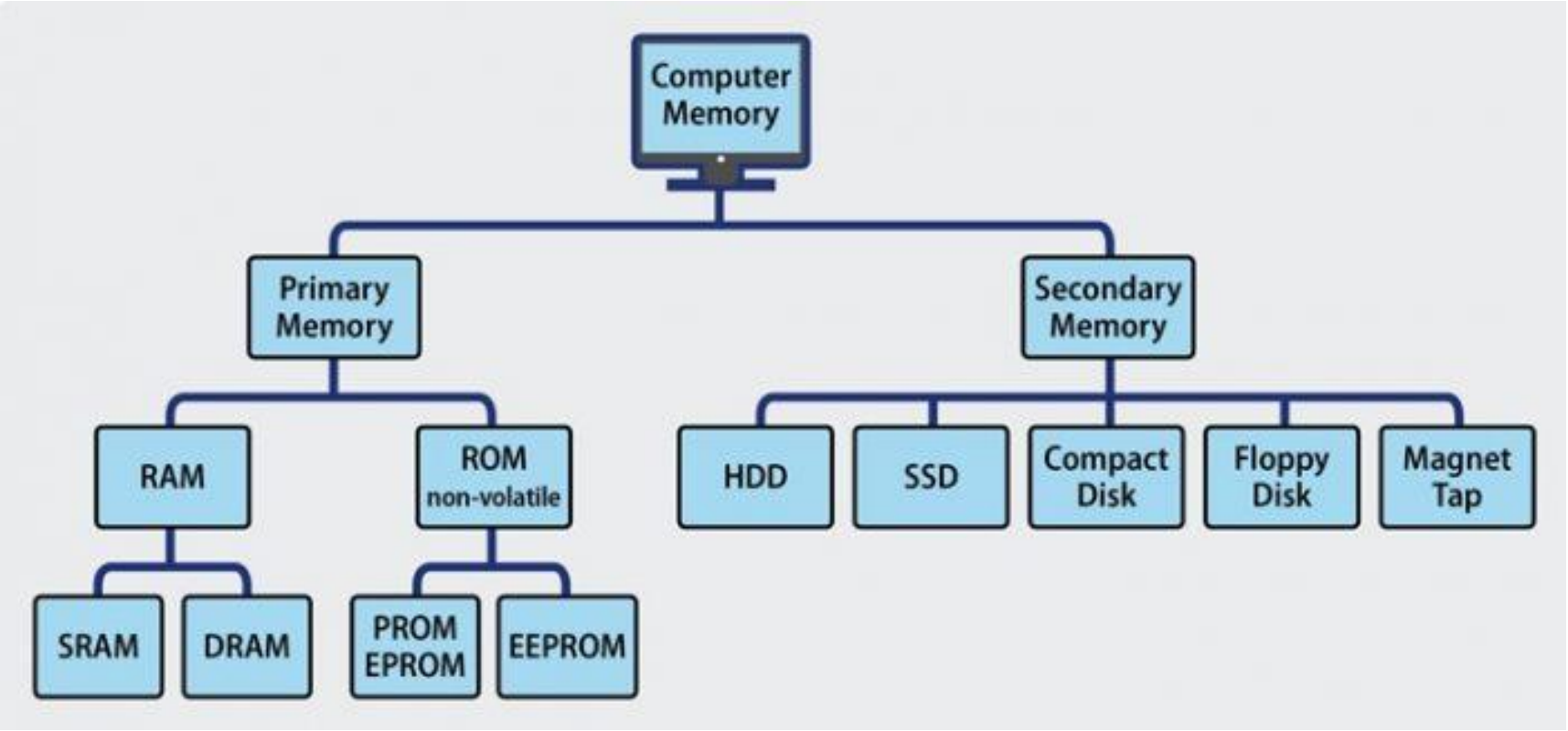
Spring 2022

ThS. Lê Đức Thịnh
thinhld@uit.edu.vn

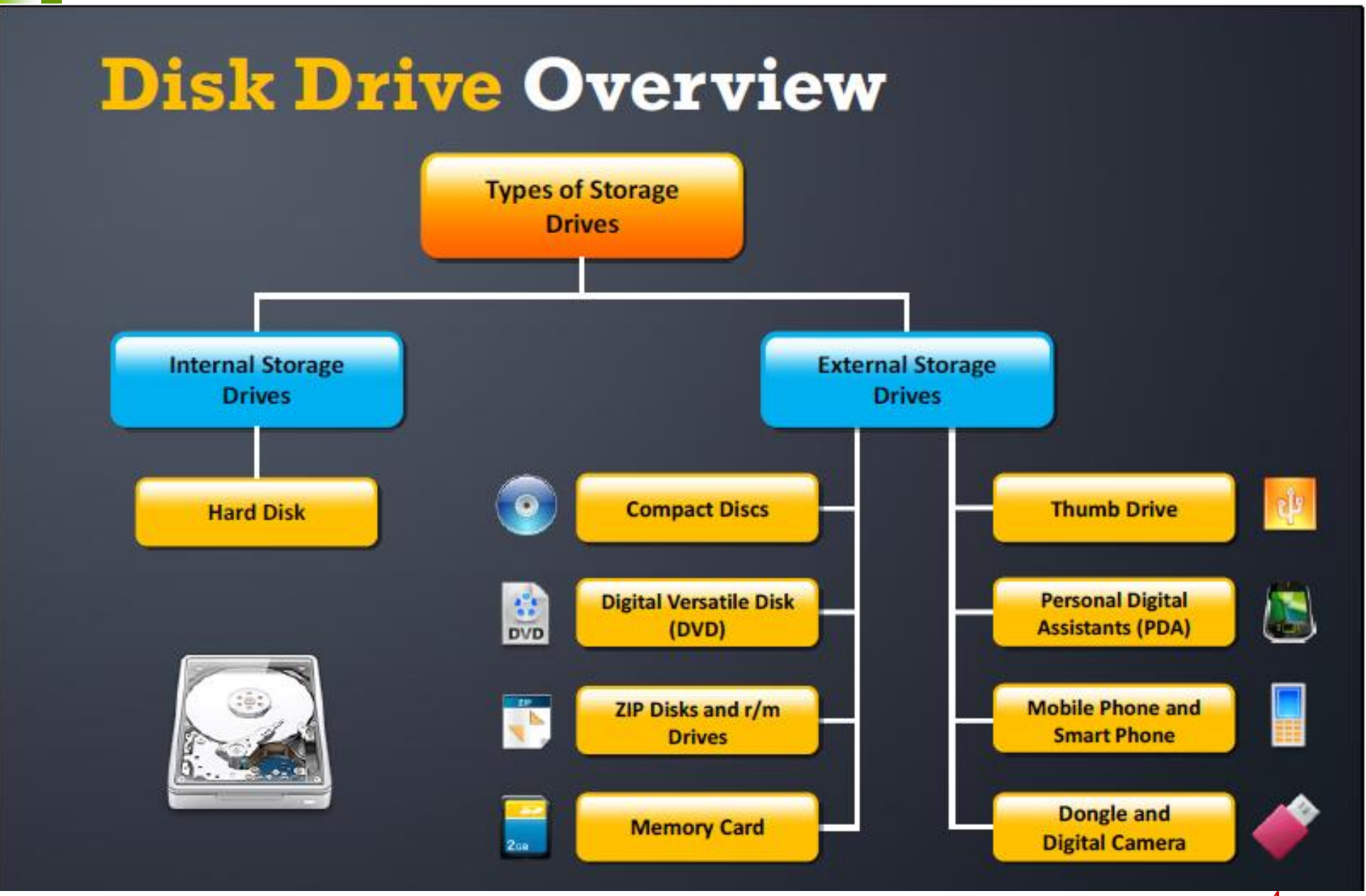
Nội dung trình bày

- Hiểu về đĩa cứng, phân vùng và tập tin hệ thống (file system)
- Tiến trình boot HĐH
- RAID
- Công cụ hỗ trợ điều tra đĩa cứng

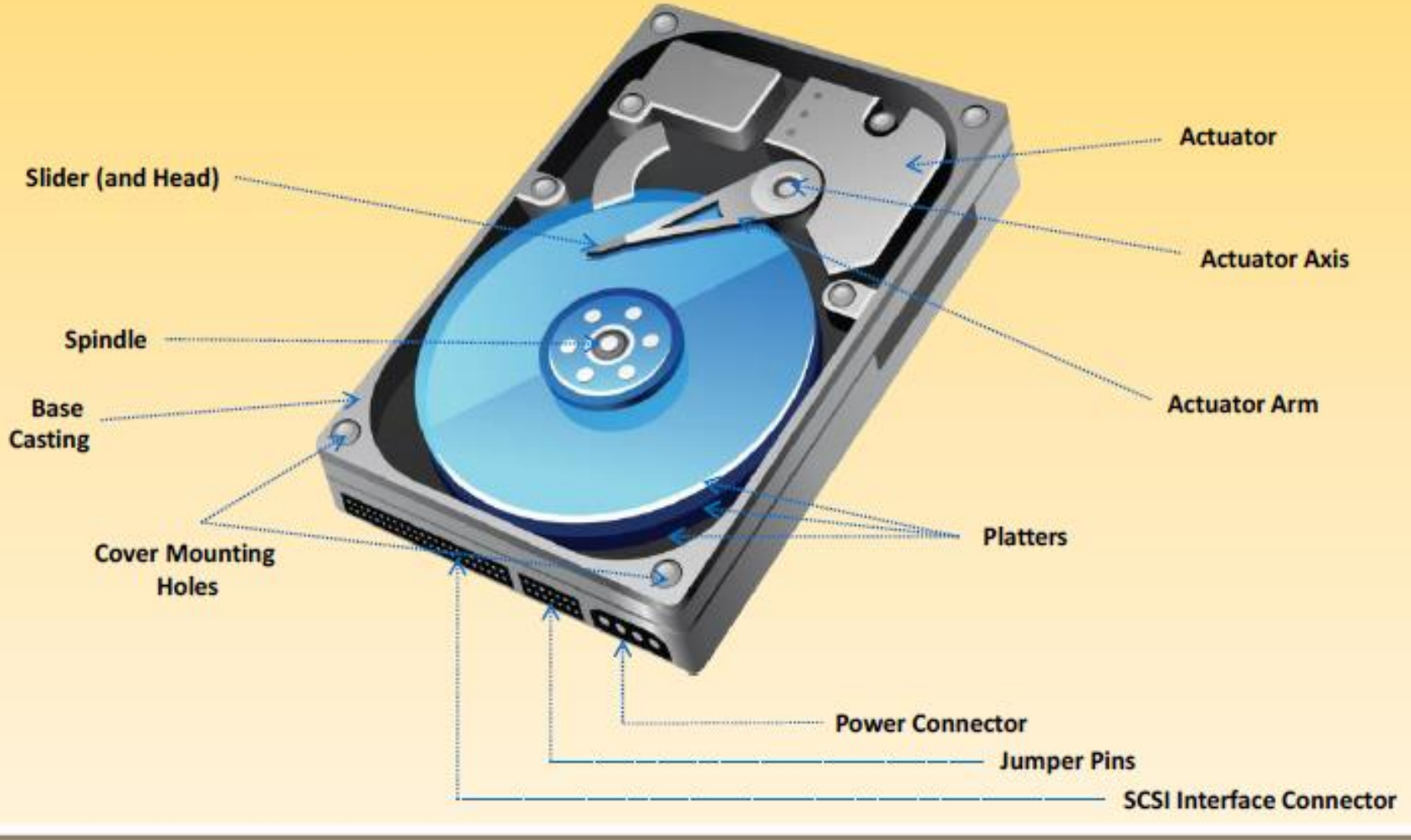
Phân loại bộ nhớ máy tính



Disk Drive Overview



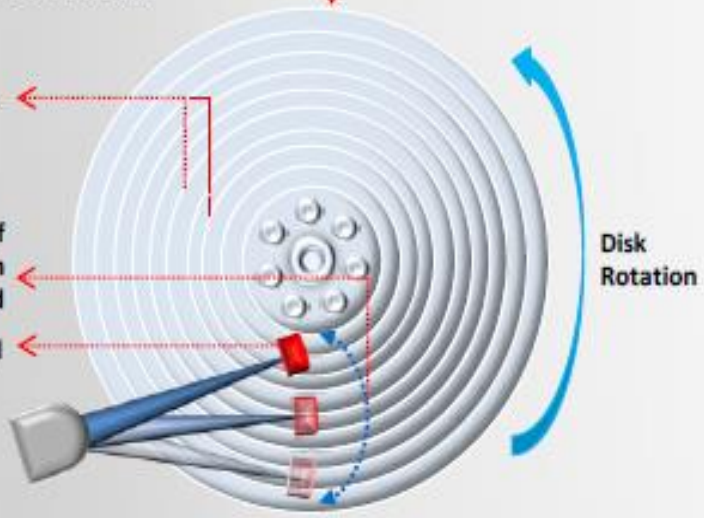
Physical Structure of a Hard Disk (Cont'd)

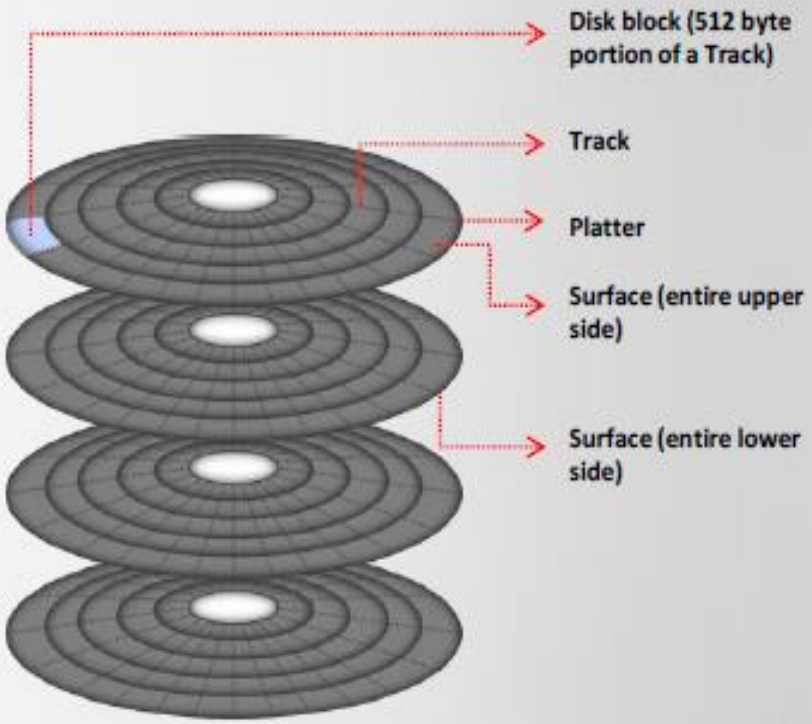


Physical Structure of a Hard Disk

Disk

Magnetized Data on Disk





Logical Structure of Hard Disk

The logical structure of a hard disk is nothing but the file system and software utilized to control access to the storage on the disk

Hard disk logical structure has significant influence on the performance, consistency, expandability, and compatibility of the storage subsystem of the hard disk

Different operating systems have different file systems and use different ways of arranging and controlling access to data on the hard disk

Slack Space (Cont'd)

1. Slack space is the free space on the cluster after writing data on that cluster

2. DOS and Windows utilize the fixed size clusters for the file's system

3. If the size of the stored data is less than the cluster's size, the unused area remains reserved for the file, resulting in slack space

4. DOS and FAT 16 (file allocation table) file system in the Windows utilizes large sized clusters

5. For example, if the partition size is 4 GB, each cluster will be 32 K. Even if a file needs only 10 K, the entire 32 K will be allocated, resulting in 22 K of slack space

Slack Space

Sector 1	Sector 2
File A	File A

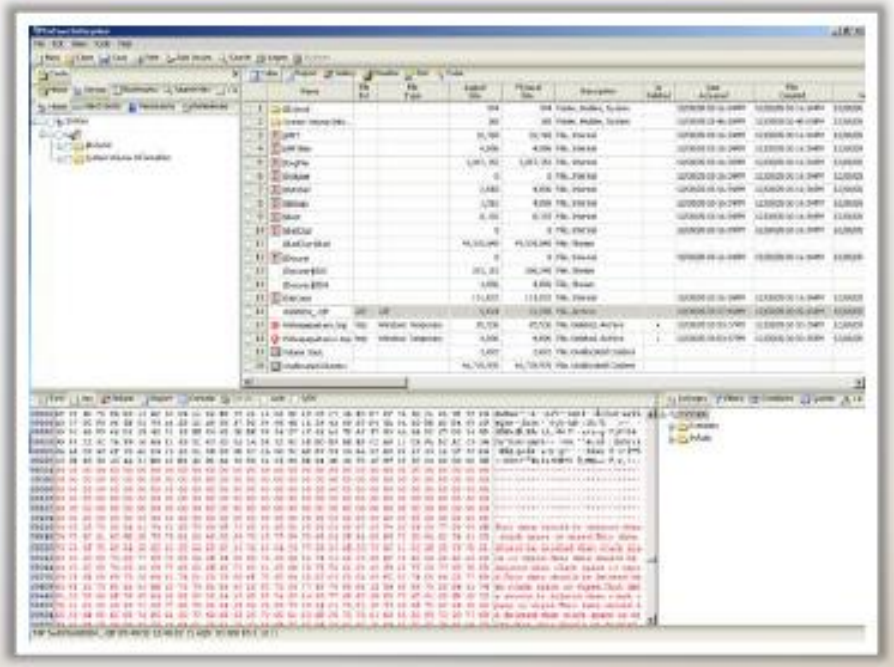
Sectors 1 and 2 are allocated to File A

Sector 1	Sector 2
File A	File A

File A was deleted – Sectors marked as unallocated

Sector 1	Sector 2	Sector 2 Slack
File B	File B	File A

File B written to unallocated space



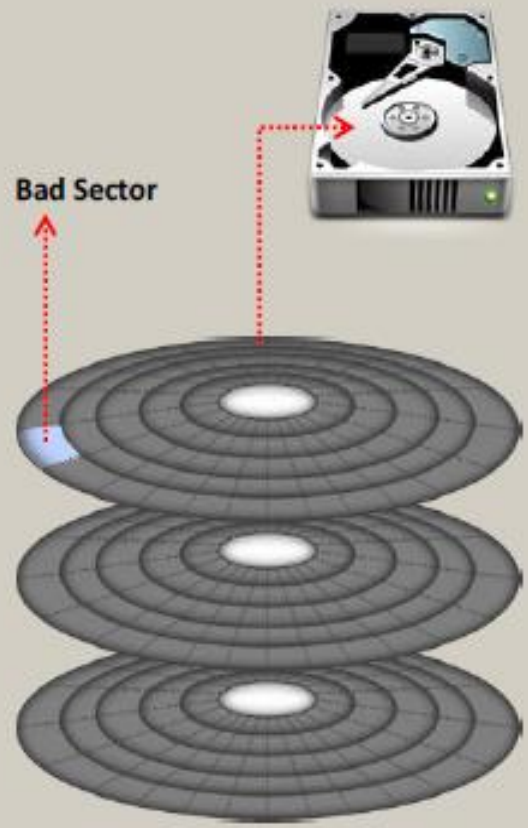
Bad Sector

Bad sector is a damaged portion of a disk on which no read/write operation can be performed

Formatting a disk enables the operating system to identify unusable sectors and mark them as bad

Bad sectors are due to configuration problems or any physical disturbances to the disk

Special software is used to recover the data on a bad sector



Types of Hard Disk Interfaces

SCSI

(Small Computer System Interface)

It enables connection of 15 peripheral devices to one PCI board

IDE/EIDE

(Integrated Drive Electronics/ Enhanced IDE)

It connects hard disks, optical disks, and tape drives to the PC

ATA

(Advanced Technology Attachment)

- Serial ATA
- Parallel ATA

USB

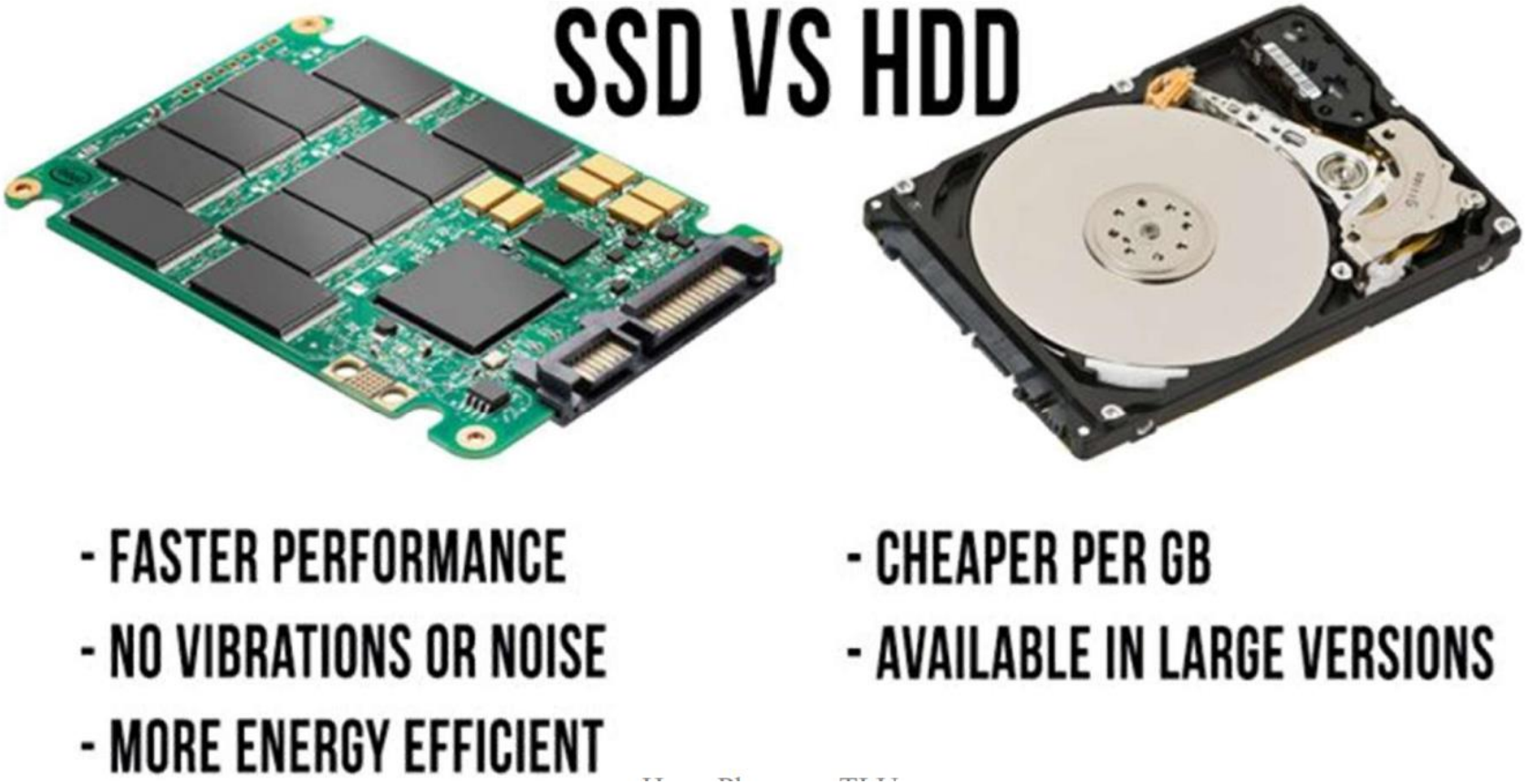
(Universal Serial Bus)

It permits electronic equipment to connect the computer

Fibre Channel

- Fibre Channel electrical interface
- Fibre Channel optical interface

Bộ nhớ Flash


SSD VS HDD
- FASTER PERFORMANCE
- NO VIBRATIONS OR NOISE
- MORE ENERGY EFFICIENT
- CHEAPER PER GB
- AVAILABLE IN LARGE VERSIONS

Bộ nhớ Flash

- Được sử dụng phổ biến trong các sản phẩm điện tử: smartphones, máy ảnh, máy nghe nhạc, USB, … Hiện nay, chi phí và hiệu năng phát triển có thể thay thế HDD.
- 2 loại bộ nhớ flash đặc biệt:
  - NOR
  - NAND

SSD (Solid State Drive)

- Sử dụng chip nhớ kiến trúc NAND
- Cấu trúc flash NAND được chia theo mô hình lưới, cơ bản là cell (ô nhớ), page (trang) và block (khối). Nhiều cell hợp thành một page, kích thước thường từ 2 - 16KB. Tương tự nhiều page sẽ tạo thành một block, gồm 128 đến 256 page với kích thước từ 256KB - 4MB
- Có ba công nghệ flash NAND hiện đang sử dụng phổ biến trong SSD là:
  - SLC (single-level cell)
  - MLC (multi-level cell)
  - TLC (triple-level cell)

SSD (Solid State Drive)

	SLC	MLC	TLC	HDD	RAM
P/E cycles	100k	10k	5k	*	*
Bits per cell	1	2	3	*	*
Seek latency (µs)	*	*	*	9000	*
Read latency (µs)	25	50	100	2000-7000	0.04-0.1
Write latency (µs)	250	900	1500	2000-7000	0.04-0.1
Erase latency (µs)	1500	3000	5000	*	*
Notes	* metric is not applicable for that type of memory				
Sources	P/E cycles [20]				

SSD (Solid State Drive)

- Hai vấn đề đối với SSD:
  - Hiệu năng SSD giảm dần khi sử dụng theo thời gian.
  - Có số lần ghi nhất định

File Systems

- Windows
- Linux
- MacOS

→ Sinh viên làm bài tập theo yêu cầu

What is the Booting Process?

Booting refers to the process of starting or resetting operating systems when the user turns on a computer system

It loads the operating system (stored in the hard disk) to the RAM (Working memory)

Types of Booting

- **Cold boot (Hard boot)**
  It is the process of starting a computer from a powered-down or off state
- **Warm boot (Soft boot)**
  It is the process of restarting a computer that is already turned on through the operating system

Essential Windows System Files

File Names	Description
Ntoskrnl.exe	Executive and kernel
Ntkrnlpa.exe	Executive and kernel with support for Physical Address Extension (PAE)
Hal.dll	Hardware abstraction layer
Win32k.sys	Kernel-mode part of the Win32 subsystem
Ntdll.dll	Internal support functions and system service dispatch stubs to executive functions
Kernel32.dll	Win32 subsystem DLL files
Advapi32.dll	
User32.dll	
Gdi32.dll	

Windows 7 Boot Process (Cont'd)

Below is process that occurs within the system when it is switched ON.

1. When the system is switched ON, CPU sends a Power Good signal to motherboard and checks for computer's BIOS firmware

2. BIOS starts a Power-On Self-Test (POST) which checks if all the hardware required for system boot are available and load all the firmware settings from nonvolatile memory on the motherboard

3. If post is successful, add-on adapters perform a self-test for integration with the system

4. The pre-boot process will complete with POST detecting a valid system boot disk

Windows 7 Boot Process (Cont'd)

5 After POST, the computer's firmware scans boot disk and loads the master boot record (MBR) which search for basic boot information in Boot Configuration Data (BCD)

6 MBR triggers Bootmgr.exe which locates Windows loader (Winload.exe) on the Windows boot partition and triggers Winload.exe

7 Windows loader loads the OS kernel ntoskrnl.exe

8 Once the Kernel starts running, the Windows loader loads HAL.DLL, boot-class device drivers marked as BOOT_START and the SYSTEM registry hive into the memory

9 Kernel passes the control of boot process to the Session Manager Process (SMSS.exe) which load all other registry hives and drivers required to configure Win32 subsystem run environment

10 Session Manager Process triggers Winlogon.exe which presents the user logon screen for user authorization

Windows 7 Boot Process

11

- **Session Manager Process** initiates **Service control manager** which starts all the services, rest of the non-essential device drivers, the security subsystem LSASS.EXE and Group policy scripts

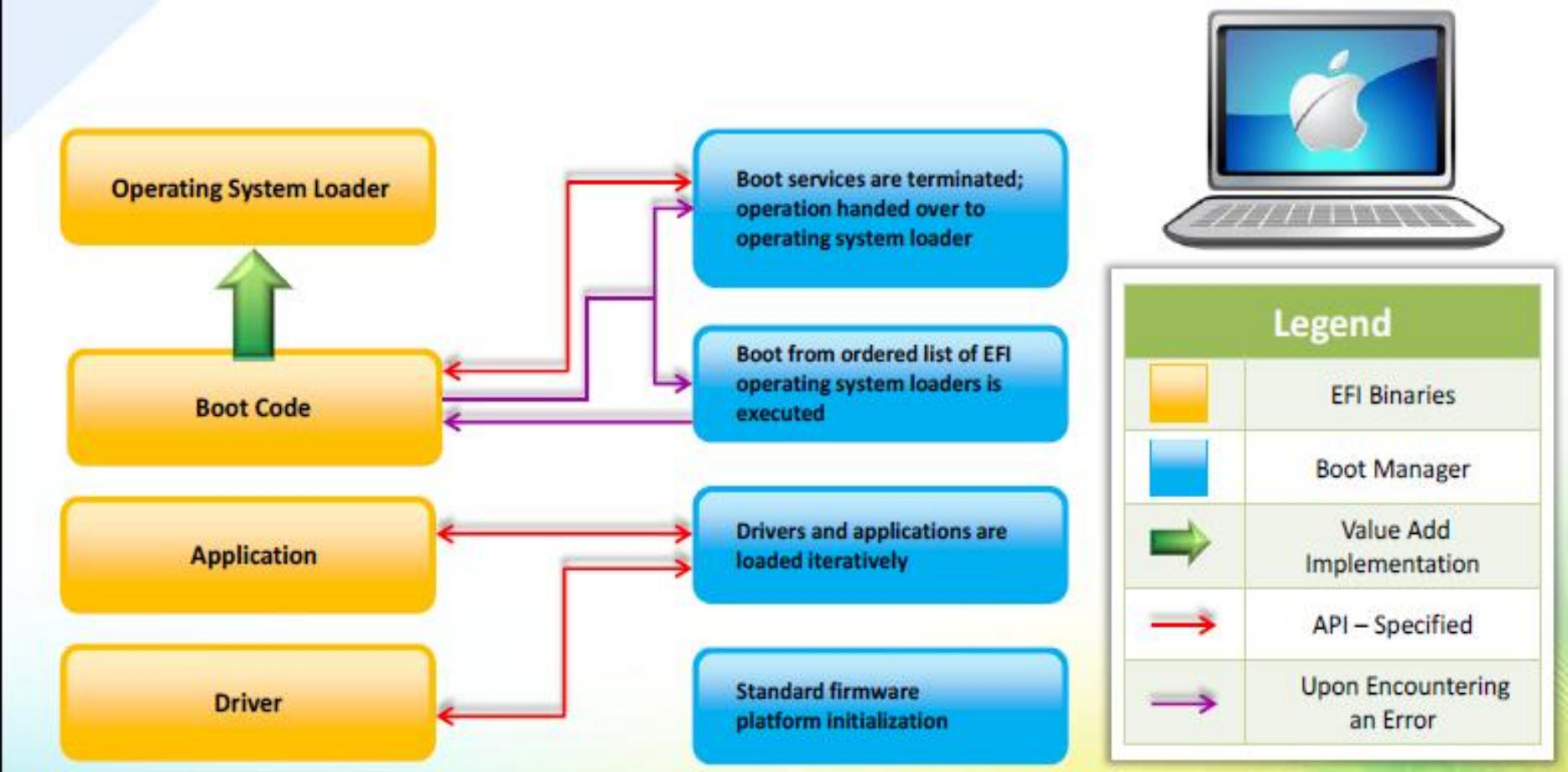
12

- Once user logs in, a session is created for the user by Windows

13

- Service control manager starts the **Explorer.exe** and initiates the Desktop Window Manager (DMW) process which set the desktop for the user

Macintosh Boot Process (Cont'd)



Macintosh Boot Process (Cont'd)

1. Macintosh boot process starts with the activation of **BootROM** which initializes system hardware and selects an operating system to run
2. Once you power on the Macintosh, BootROM performs **POST (Power-On Self Test)** to test some hardware interfaces required for startup
3. On PowerPC-based Macintosh computers, **OpenFirmware** initializes the rest of the hardware interfaces
4. On Intel-based Macintosh computers, **EFI** initializes the rest of the hardware interfaces
5. Once the hardware interfaces are initialized, it selects which **operating system** to use

Macintosh Boot Process (Cont'd)

6

If system contains multiple operating systems then it allows user to choose the particular operating system by holding down the Option key

7

Once BootROM operation is finished, the control passes to the BootX (PowerPC) or boot.efi (Intel) boot loader which is located at the `/System /Library/ CoreServices` directory

Boot loader loads a pre-linked version of the kernel which is located at `/System/Library/Caches /com.apple.kernelcaches`

8

If the pre-linked kernel is missing, the boot loader attempts to load mkext cache file, which contains a set of device drivers

9

Macintosh Boot Process

10 If mkext cache is also missing, the boot loader searches for drivers in the `/System/Library /Extensions` directory

11 Once the essential drivers have loaded, the boot loader starts the kernel initialization procedure, which initializes Mach and BSD data structure and then I/O kit

12 The I/O kit uses the device tree to link the loaded drivers into the kernel

13 Once the kernel finds the root device, it roots BSD off of it

14 The mach_init process has been replaced by "launchd" which runs start up items and prepares the system for user

RAID (Redundant Array of Independent Disks)

- Sử dụng nhiều đĩa vật lý để tạo thành đĩa logic (ảo hóa), nâng cao hoạt động
- Gồm 7 cấp (0-6), cấp không thể hiện mối quan hệ thứ bậc mà thể hiện kiến trúc thiết kế khác nhau. Gồm các điểm chung:
  - Tập hợp nhiều đĩa vật lý được OS xem như một/vài ổ đĩa logic
  - Dữ liệu được phân bố trên các đĩa vật lý của một manrng theo cơ chế phân dải – striping
  - Dung lượng đĩa dư thừa được sử dụng để lưu trữ parity, đảm bảo khôi phục dữ liệu trong trường hợp đĩa bị hỏng.

RAID Level 0: Disk Striping

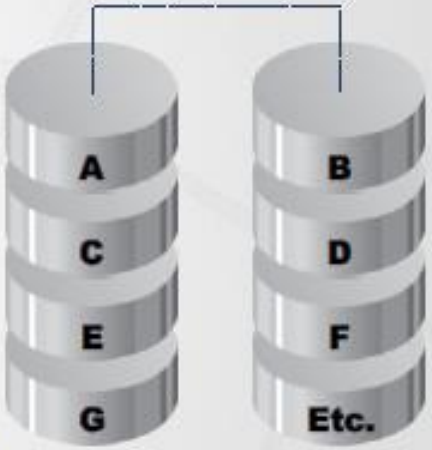
Data is split into blocks and written equally across **multiple hard drives**

It improves I/O performance by spreading the **I/O load** across many channels and disk drives

If any **drive fails**, data recovery is not possible

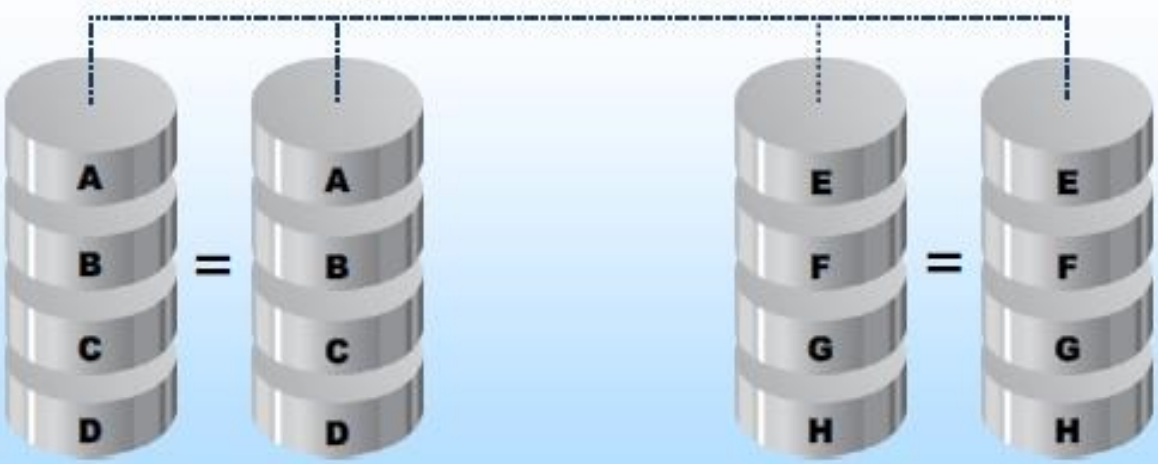
It does not provide **data redundancy**

It requires **minimum two drives** for set up

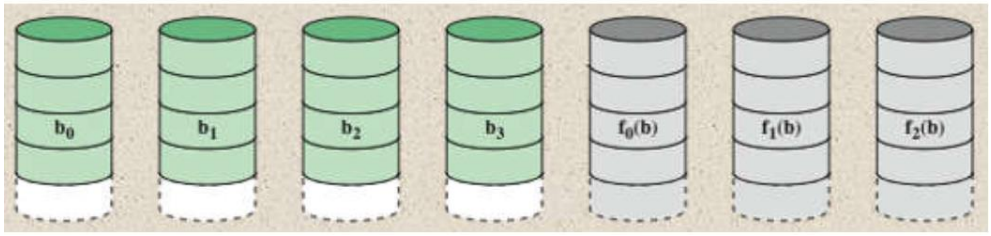


RAID Level 1: Disk Mirroring

- Multiple copies of data are written to multiple drives at the same time
- It provides data redundancy by completely duplicating the drive data to multiple drives
- If one drive fails, data recovery is possible
- It requires minimum two drives for set up



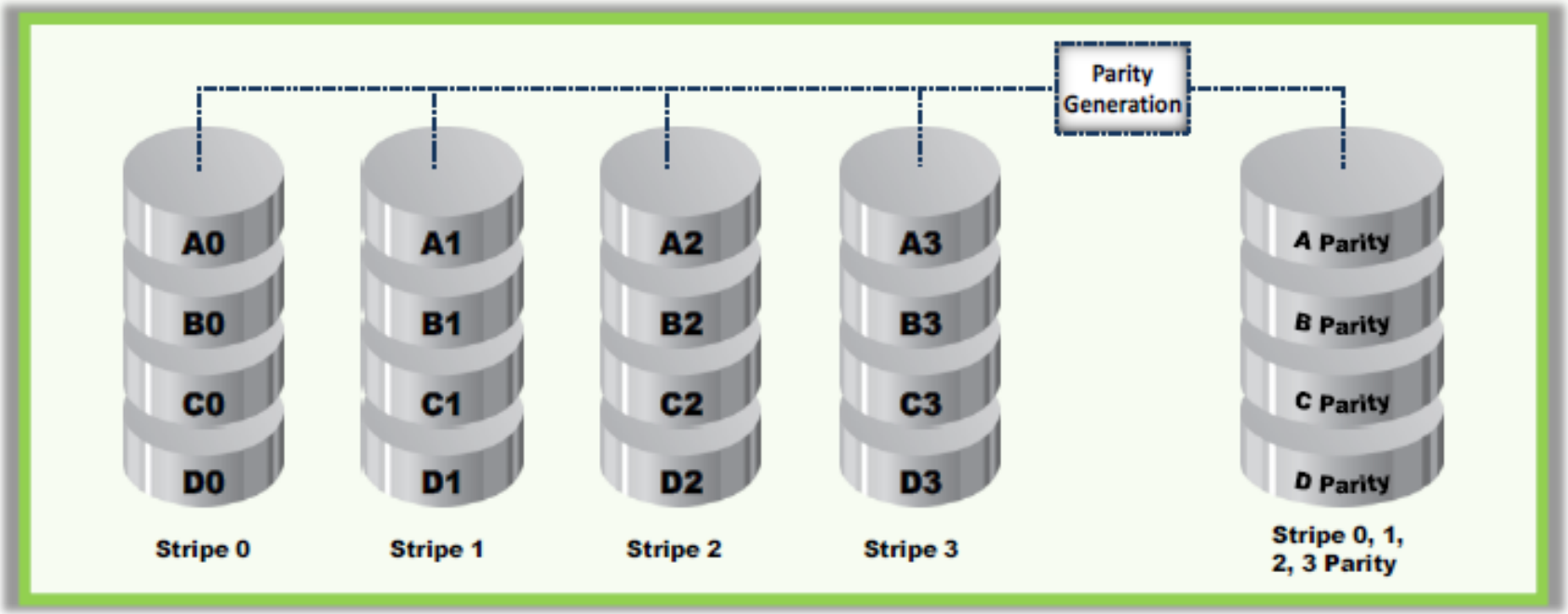
RAID Level 2



- Truy cập song song: tất cả các đĩa đều tham gia xử lý I/O
- Phân dải dữ liệu nhỏ, thường bằng byte
- Mã sửa lỗi được tính từ các bit tương ứng trên mỗi đĩa dữ liệu.
- Thường sử dụng mã Hamming

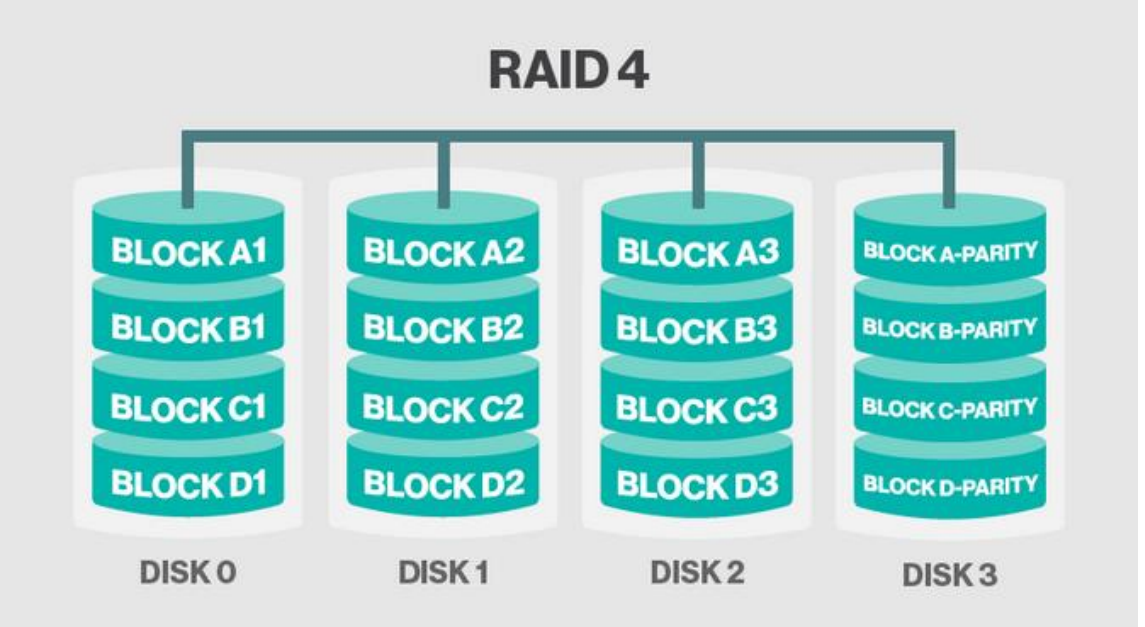
RAID Level 3: Disk Striping with Parity

- Data is striped at a byte level across multiple drives and one drive is set to store parity information
- If any drive fails, data recovery and error correction is possible through the parity drive
- Parity drive stores all the information about the data on multiple drives



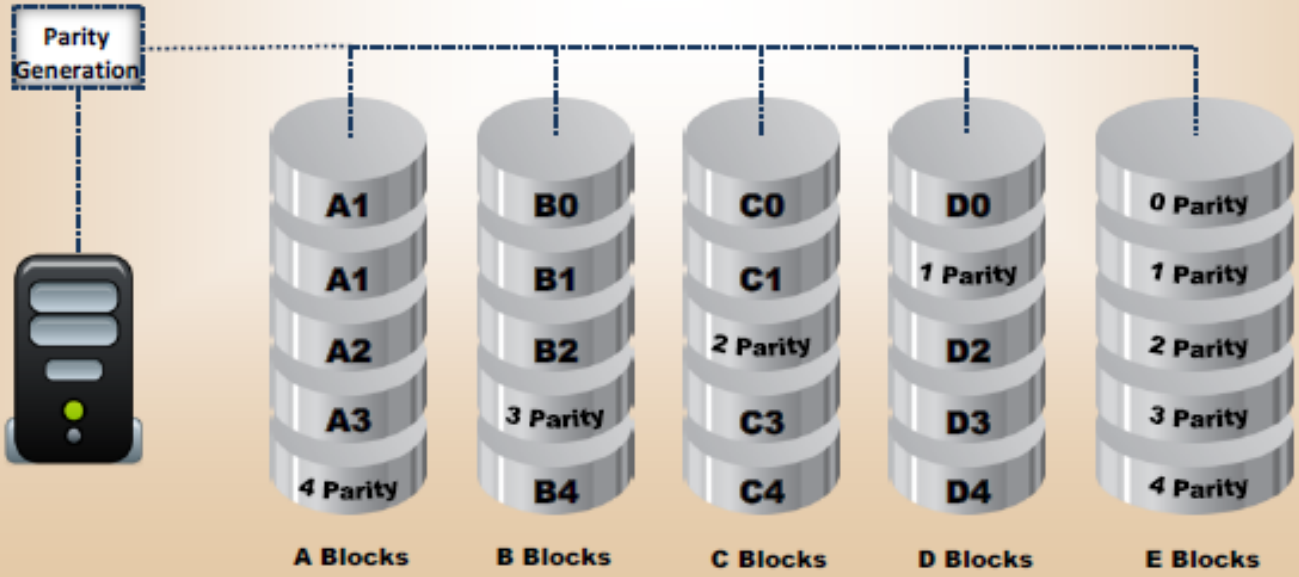
RAID Level 4

- Giống RAID 3 nhưng cải tiến hiệu suất bằng cách stripe dữ liệu qua nhiều đĩa theo khối (3 stripe theo byte)
- Giống RAID 5 nhưng sử dụng ổ đĩa parity riêng thay vì phân bố như 5

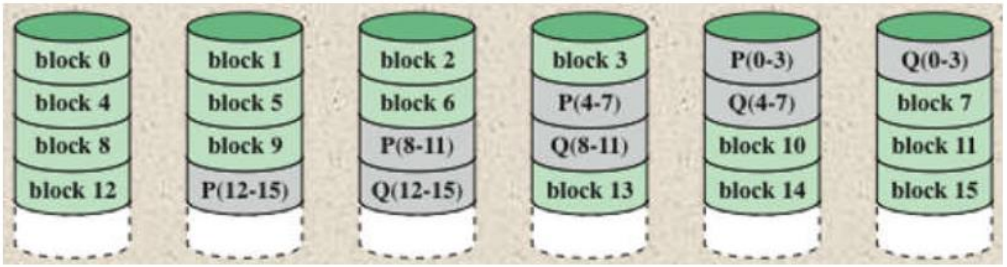


RAID Level 5: Block Interleaved Distributed Parity

- Data is striped at a byte level across multiple drives and parity information is distributed among all member drives
- Data writing process is slow
- It requires a minimum of three drives for setup



RAID Level 6



- Tương tự RAID 5 nhưng dùng đến 2 parity
- Có tính sẵn sàng dữ liệu rất cao
- Dữ liệu chỉnh mất khi cả 3 ổ đĩa bị hỏng.

Hybrid RAID

- RAID 10
- RAID 30
- RAID 50
- …

Các kiểu thu thập dữ liệu

- Static Acquisition
- Live acquisition

Data Acquisition

- Data Acquisition is the process of imaging or otherwise obtaining information from a digital device and its peripheral equipment and media

Static Acquisition

- It is defined as acquiring data that remains unaltered when the system is powered off or shutdown
- This type of data is termed as non-volatile and is usually recovered from hard drives, it can also exist in slack space, swap files and unallocated drive space
- Other sources of non-volatile data include CD-ROMs, USB thumb drives, smart phones and PDAs

Live Acquisition

- It is defined as acquiring data from a running computer (computer that is already "powered on" when encountered at a crime scene) that would be lost when the system is powered off
- This type of data is termed as volatile and resides in registries, cache, and RAM
- As RAM and other volatile data are dynamic, collection of this information should occur in real time

Recover Data from Unallocated Space Using File Carving Process

I. File carving is a process used to recover files from unallocated space of the hard disk

II. This technique is generally used by the investigator during the digital investigation to extract the files from unallocated space

III. Tools used for file carving process:

- PhotoRec
- EnCase

Các công cụ hỗ trợ

- EnCase
- ProDiscover
- Autopsy (thực hiện chi tiết trong bài thực hành)

References

- CHFIv8 Slides
- Kiến trúc máy tính, Hang-Phuong Nguyen
- Physical Memory Forensics, Mariusz Burdach
- Wikipedia
- Memory Forensics, Phulc

Q&A

TRƯỜNG ĐH CÔNG NGHỆ THÔNG TIN
KHOA MẠNG MÁY TÍNH & TRUYỀN THÔNG

Digital Forensics
Pháp chứng Kỹ thuật số

#3: Disk Forensics
Spring 2022

ThS. Lê Đức Thịnh
thinhld@uit.edu.vn