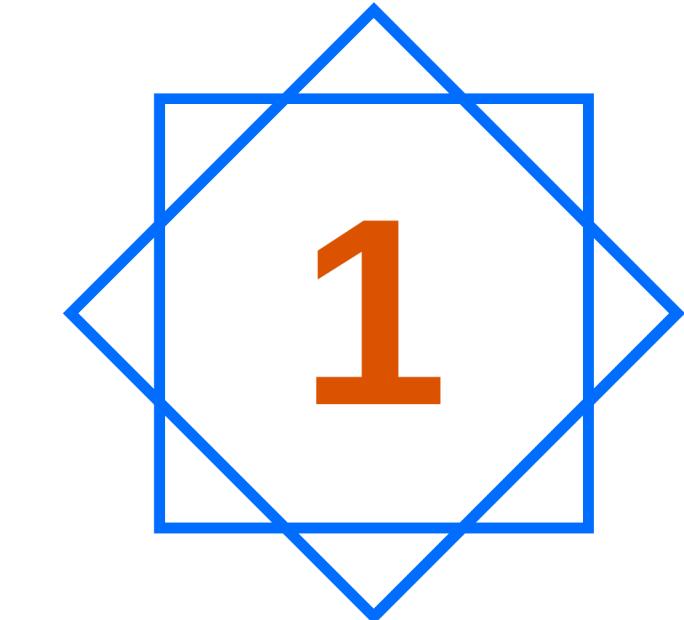


DATA WAREHOUSE

GROUP 8



MDX DEFINITION
MDX QUERY SYNTAX



DEMO MDX



META DATA



A CRITICAL NEED
IN DATA WAREHOUSE

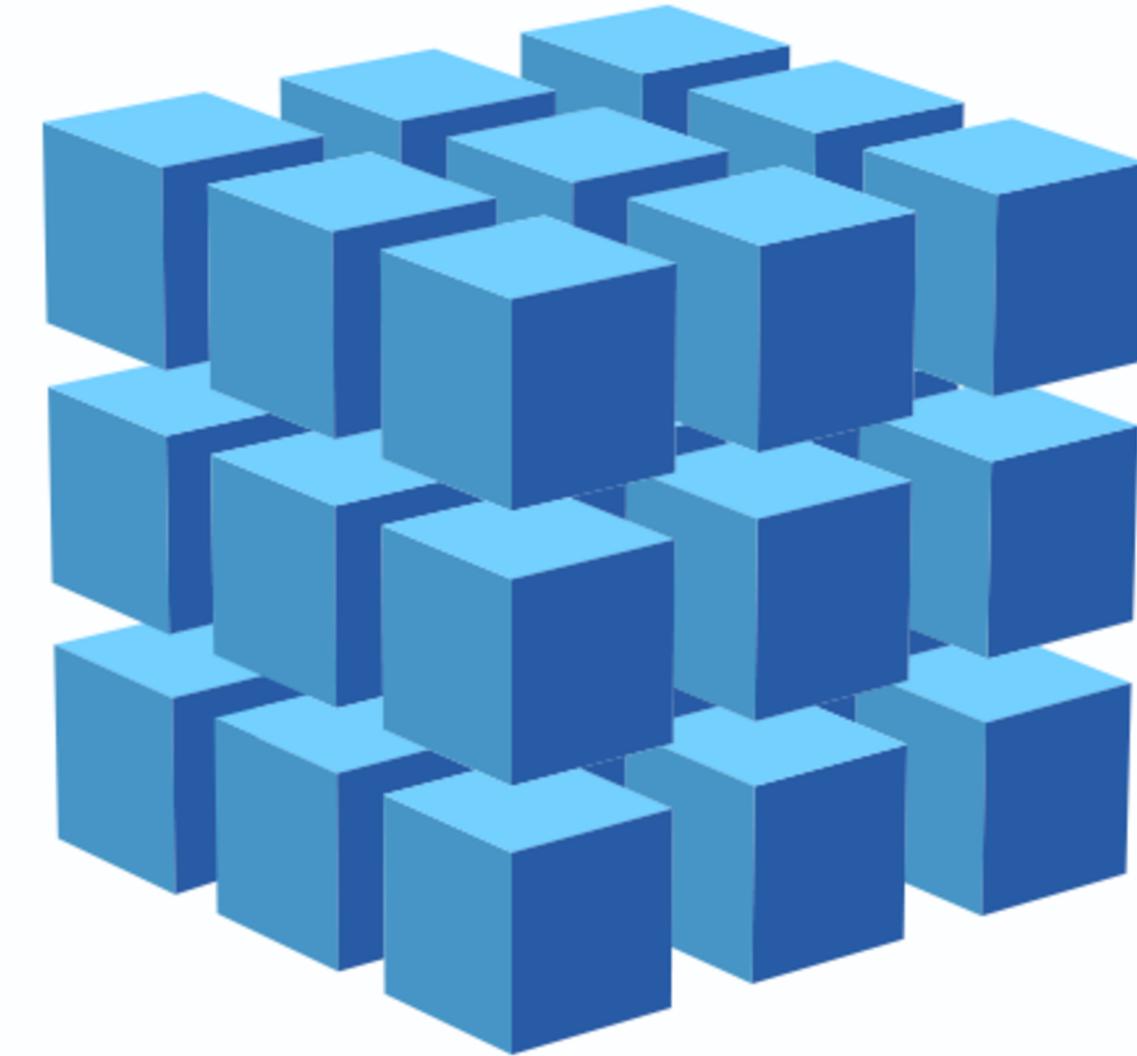


RAID TECHNOLOGY

TIẾN ĐẠT

1

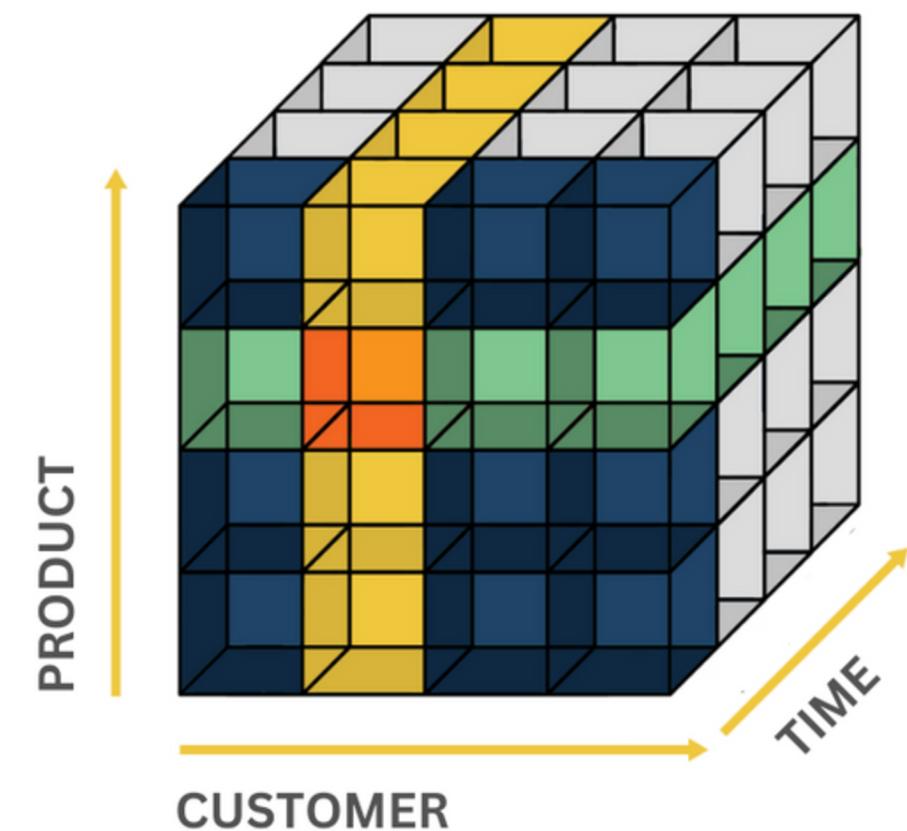
MDX DEFINITION MDX QUERY SYNTAX



TIẾN ĐẠT

MULTIDIMENSIONAL EXPRESSIONS (MDX):

- MultiDimensional Expressions (MDX) là ngôn ngữ truy vấn sử dụng trong các hệ thống quản lý và phân tích dữ liệu đa chiều.
- Được sử dụng để truy vấn dữ liệu trong các cube OLAP.
- Cung cấp các cú pháp và hàm tính toán để giúp truy vấn và phân tích dữ liệu đa chiều một cách nhanh chóng và hiệu quả.
- Sử dụng phổ biến trong các ứng dụng phân tích dữ liệu; là một phần quan trọng của hệ thống OLAP.



SYNTAX (BASIC)

VÍ DỤ:

```
SELECT [Measures].[Sales] ON  
COLUMNS,  
[Product].[Category].Members ON ROWS  
FROM [SalesCube]  
WHERE [Time].[Year].&[2022]
```



Số lượng sản phẩm bán ra được phân loại
theo category trong năm 2022.

```
SELECT <measures|dimensions>  
[on columns|rows]  
FROM <Cube>  
[WHERE <Slicing conditions>]
```

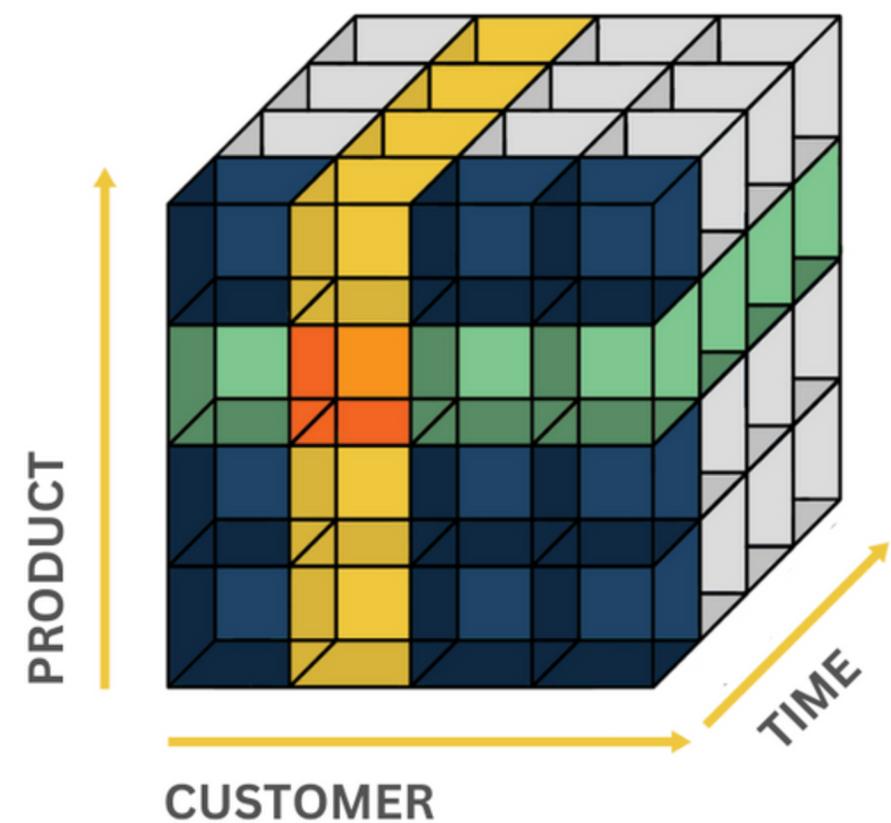
TRONG VÍ DỤ NÀY:

SELECT: Chọn measures là Sales để truy vấn.
ON COLUMNS: Sử dụng kết quả hiển thị trên trực ngang.
ON ROWS: Sử dụng dimensions Product.Category để hiển thị kết quả trên trực dọc.
FROM: Chỉ định cube được truy vấn là SalesCube.
WHERE: Chỉ yêu cầu các dữ liệu liên quan đến năm 2022.

TIẾN ĐẠT

HOW TO USE MDX?

- Truy cập database OLAP qua Microsoft Excel, Power BI, Tableau,...
- Sau khi truy cập, sử dụng MDX để truy vấn dữ liệu theo cú pháp đã đề cập ở các câu hỏi trước -> phân tích cấu trúc của cube đa chiều (kiểm tra dimensions, measures, levels) và thiết kế các truy vấn MDX để truy xuất dữ liệu.
- Khi sử dụng MDX, lựa chọn measures và dimensions để phân tích dữ liệu theo nhiều cách khác nhau, hiển thị trên trực ngang, trực dọc và lọc dữ liệu theo các điều kiện chỉ định.
- Sau khi thực hiện truy vấn và xuất kết quả, tiến hành phân tích kết quả bằng các biểu đồ, bản đồ, và các công cụ trực quan khác để có cái nhìn toàn diện hơn về dữ liệu đang được truy vấn.

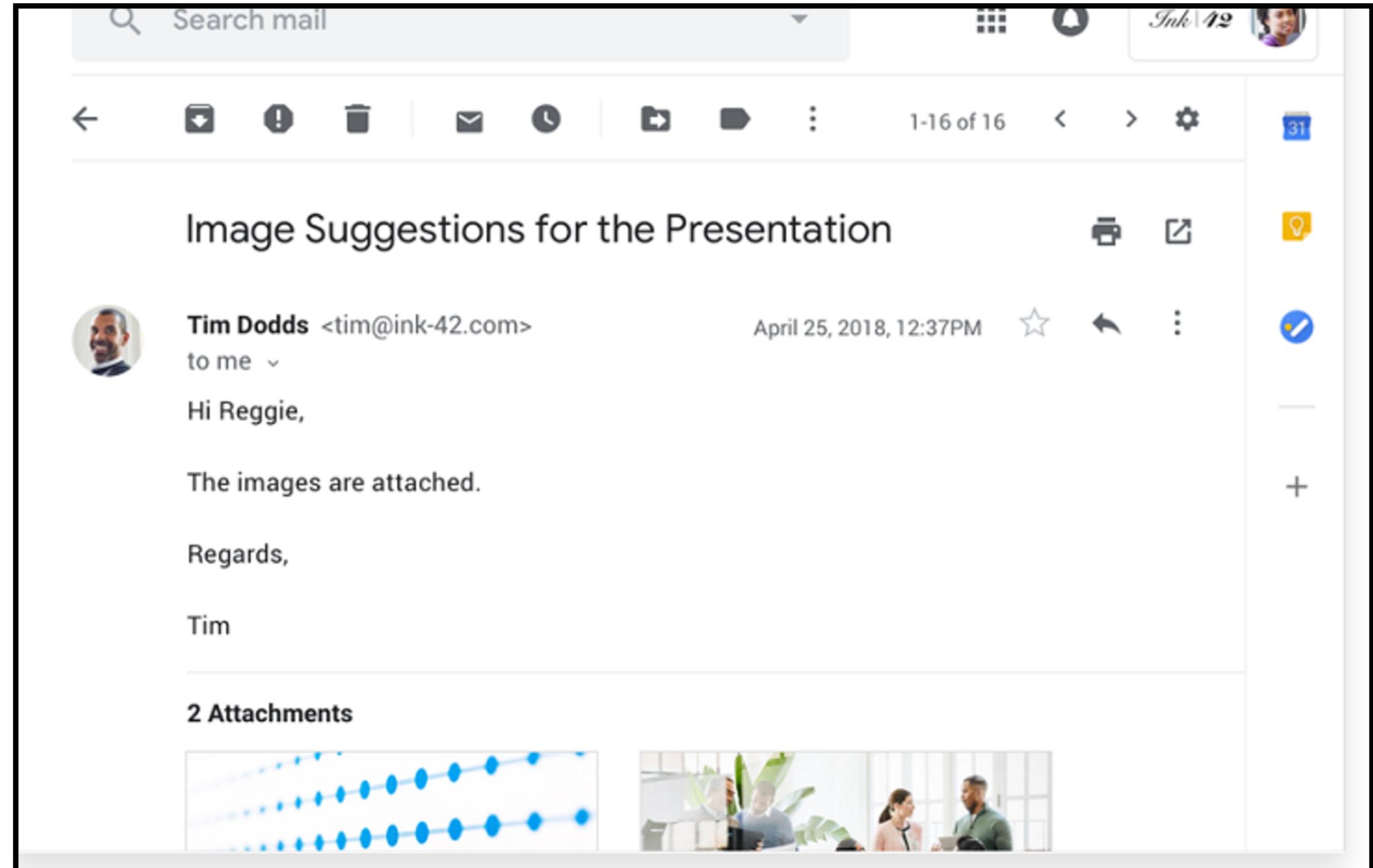


DEMO MDX

VŨ THỊ HÀ

3

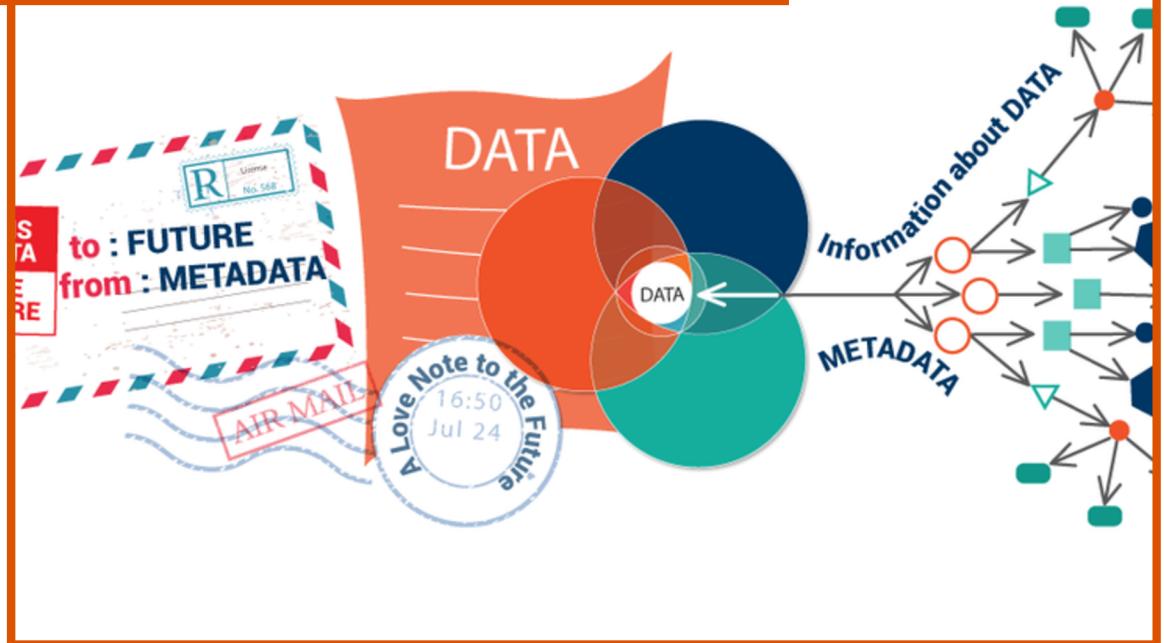
META DATA



KIỀU DOAN

META DATA

WHAT ?



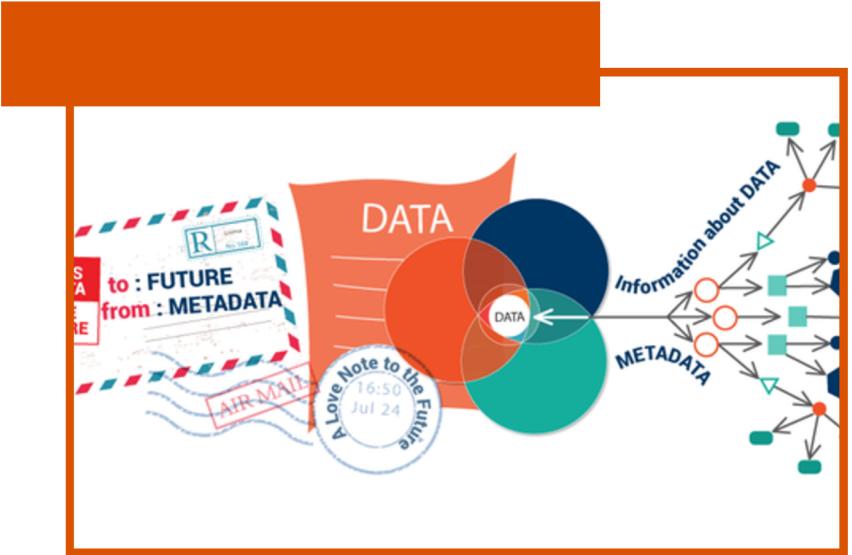
Metadata là dữ liệu về dữ liệu hoặc tài liệu về thông tin được người dùng yêu cầu. Còn được coi là các thông tin mô tả, định nghĩa, hoặc tổ chức về dữ liệu.

Bao gồm:

- Vị trí và mô tả của hệ thống kho và các thành phần.
- Tên, định nghĩa, cấu trúc và nội dung của kho dữ liệu và chế độ xem của người dùng.
- Nguồn dữ liệu đáng tin cậy.
- Quy tắc tích hợp và chuyển đổi được dùng để điền dữ liệu và cung cấp thông tin cho các công cụ phân tích của người dùng.
- Thông tin đăng ký để người đăng ký phân tích.
- Số liệu để phân tích việc sử dụng và hiệu suất kho.
- Ủy quyền bảo mật, danh sách kiểm soát truy cập,...



META DATA



Metadata trong kho dữ liệu cung cấp thông tin chi tiết về dữ liệu trong kho dữ liệu.

- Dữ liệu mô tả về dữ liệu
- Mục lục của dữ liệu
- Danh mục dữ liệu
- Bản đồ kho dữ liệu
- Lộ trình kho dữ liệu
- Thư mục kho dữ liệu
- Chất keo liên kết các nội dung trong kho dữ liệu với nhau

Metadata giúp làm rõ và nhất quán thông tin; tạo điều kiện thuận lợi cho việc phát hiện thông tin; tìm kiếm và truy xuất tài nguyên.



4

A CRITICAL NEED IN DATA WAREHOUSE



CÔNG HIẾU

A CRITICAL NEED IN DATA WAREHOUSE

For Using

- Lấy thông tin từ DW.
- Tự động tạo các truy vấn đặc biệt và chạy chúng dựa trên DW.
- Định dạng báo cáo riêng của mình.
- Trước khi có thể tạo và chạy truy vấn, cần biết về dữ liệu trong DW.

→ *Cần Metadata*

For Building

- Metadata cần thiết để xây dựng DW trong mọi hoạt động và nhiệm vụ.
- Xác định các nguồn hệ thống và dữ liệu cấu trúc của chúng.
 - Xác định cấu trúc và nội dung trong DW.
 - Xác định ánh xạ và chuyển đổi dữ liệu.

For Administering

Metadata đóng vai trò quan trọng trong việc quản trị kho dữ liệu hiện đại; đơn giản hóa các quy trình; nâng cao chất lượng dữ liệu (hướng dẫn tổ chức DL, cung cấp thông tin về nguồn DL).

A CRITICAL NEED IN DATA WAREHOUSE

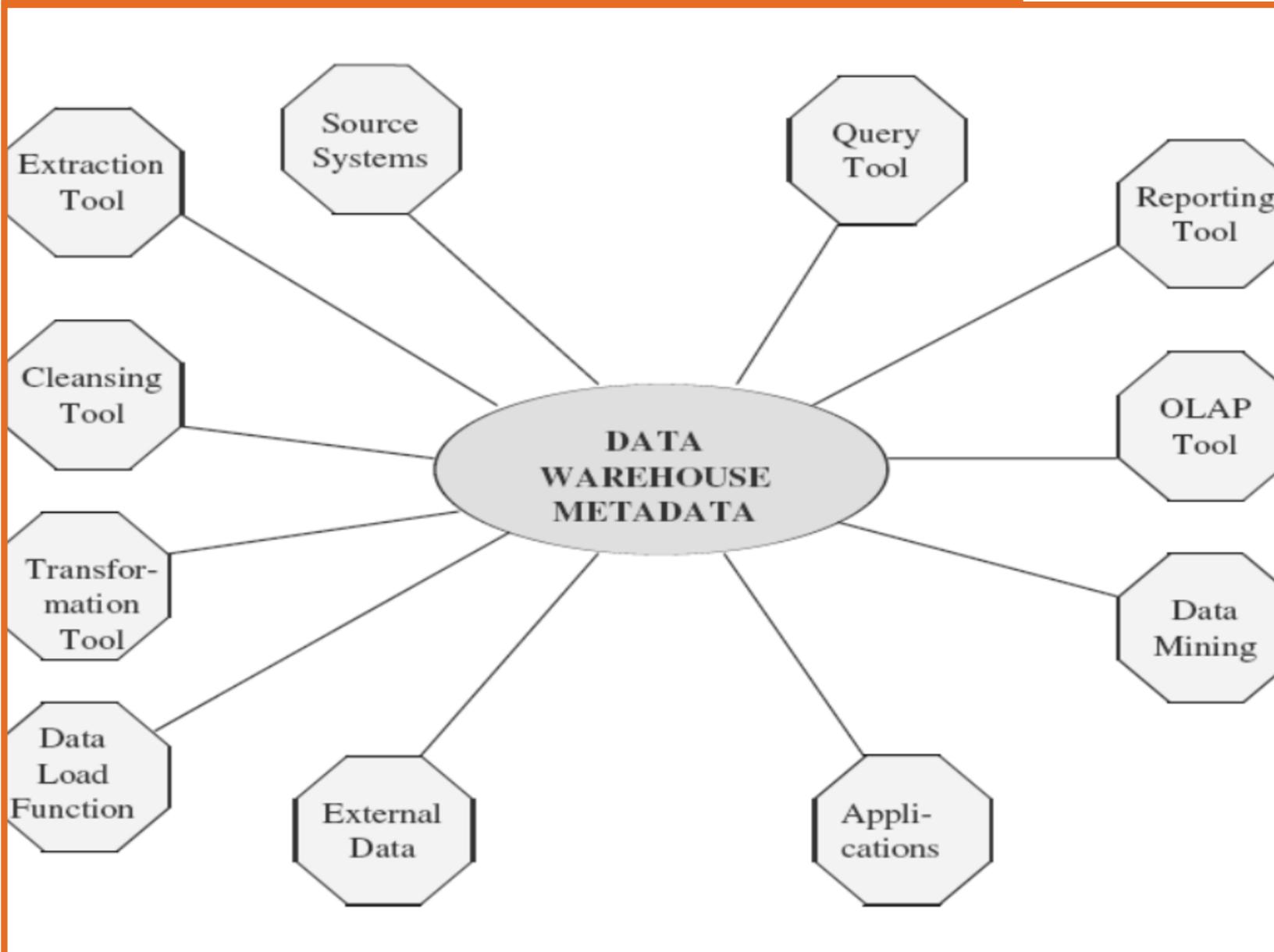
Who Needs Metadata?

	IT Professionals	Power Users	Casual Users
<i>Information Discovery</i>	Databases, Tables, Columns, Server Platforms	Databases, Tables, Columns	List of Predefined Queries and Reports, Business Views
<i>Meaning of Data</i>	Data Structures, Data Definitions, Data Mapping, Cleansing Functions, Transformation Rules	Business Terms, Data Definitions, Data Mapping, Cleansing Functions, Transformation Rules	Business Terms, Data Definitions, Filters, Data Sources, Conversion , Data Owners
<i>Information Access</i>	Program Code in SQL, 3GL,4GL, Front-end Applications, Security	Query Toolsets, Database Access for Complex Analysis	Authorization Requests, Information Retrieval into Desktop Applications such as Spreadsheets

CÔNG HIẾU

A CRITICAL NEED IN DATA WAREHOUSE

Metadata is Like a Nerve Center



Các quy trình khác nhau trong quá trình xây dựng và quản lý DW tạo ra các phần metadata của DW.

Trong DW, metadata đảm nhận một vị trí quan trọng và cho phép liên lạc giữa các phương pháp khác nhau.

Nó hoạt động như một trung tâm thần kinh trong DW.

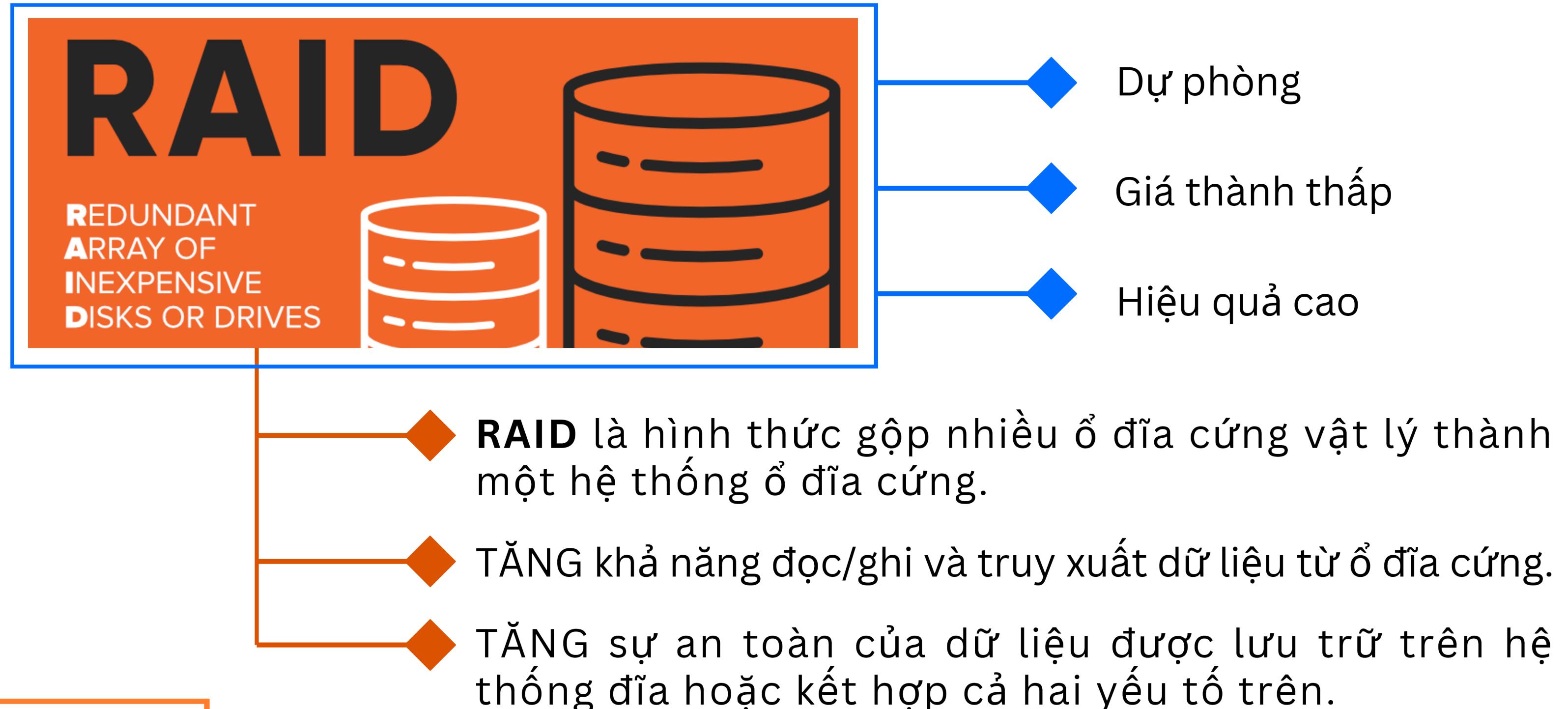
5

RAID TECHNOLOGY



PHƯƠNG UYÊN

RAID TECHNOLOGY



RAID TECHNOLOGY

Parity: ghi chẵn lẻ, tự động sao chép nội dung đã mất khi đĩa hỏng



Striping: chia sẻ dữ liệu ngẫu nhiên sang nhiều đĩa khác nhau

Mirroring: ghi dữ liệu sao chép

PHƯƠNG UYÊN

RAID TECHNOLOGY



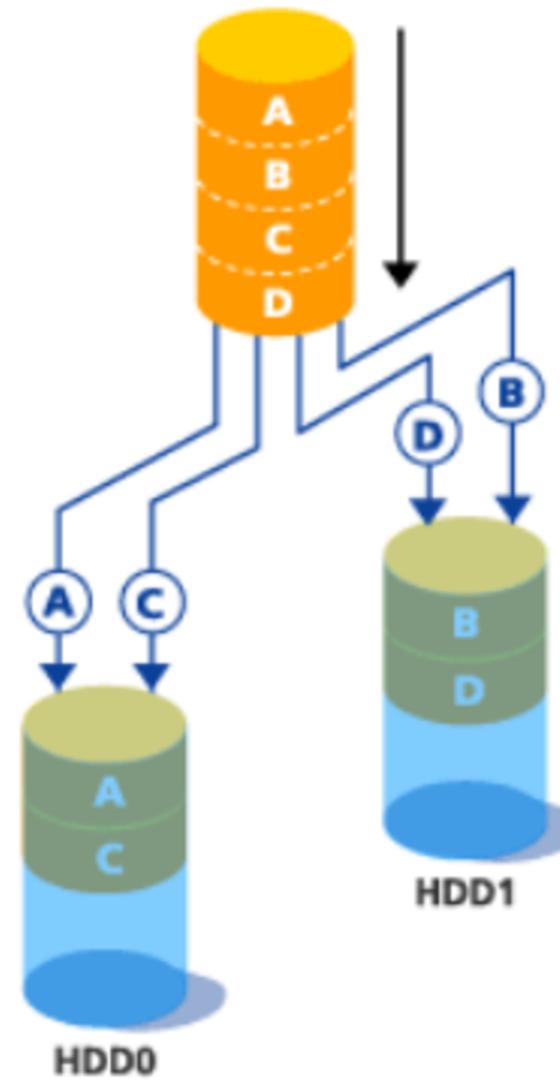
RAID 0

RAID 1

RAID 10

RAID 5

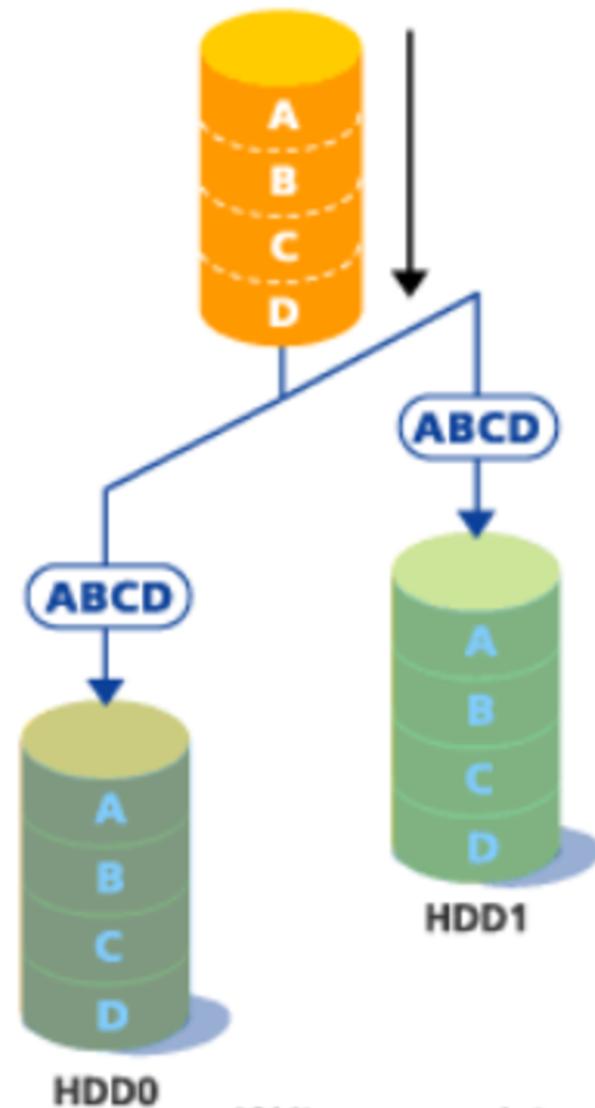
RAID TECHNOLOGY



RAID 0 - Striping

- ❖ Cần ít nhất 2 ổ đĩa ($n=2$) cùng loại.
Dữ liệu sẽ được chia ra thành
nhiều phần bằng nhau.
- ❖ **Ưu điểm:** Tăng tốc độ đọc/ghi ổ đĩa ($1/n$)
- ❖ **Nhược điểm:** Tính an toàn thấp
- ❖ Chỉnh sửa hình ảnh/video.

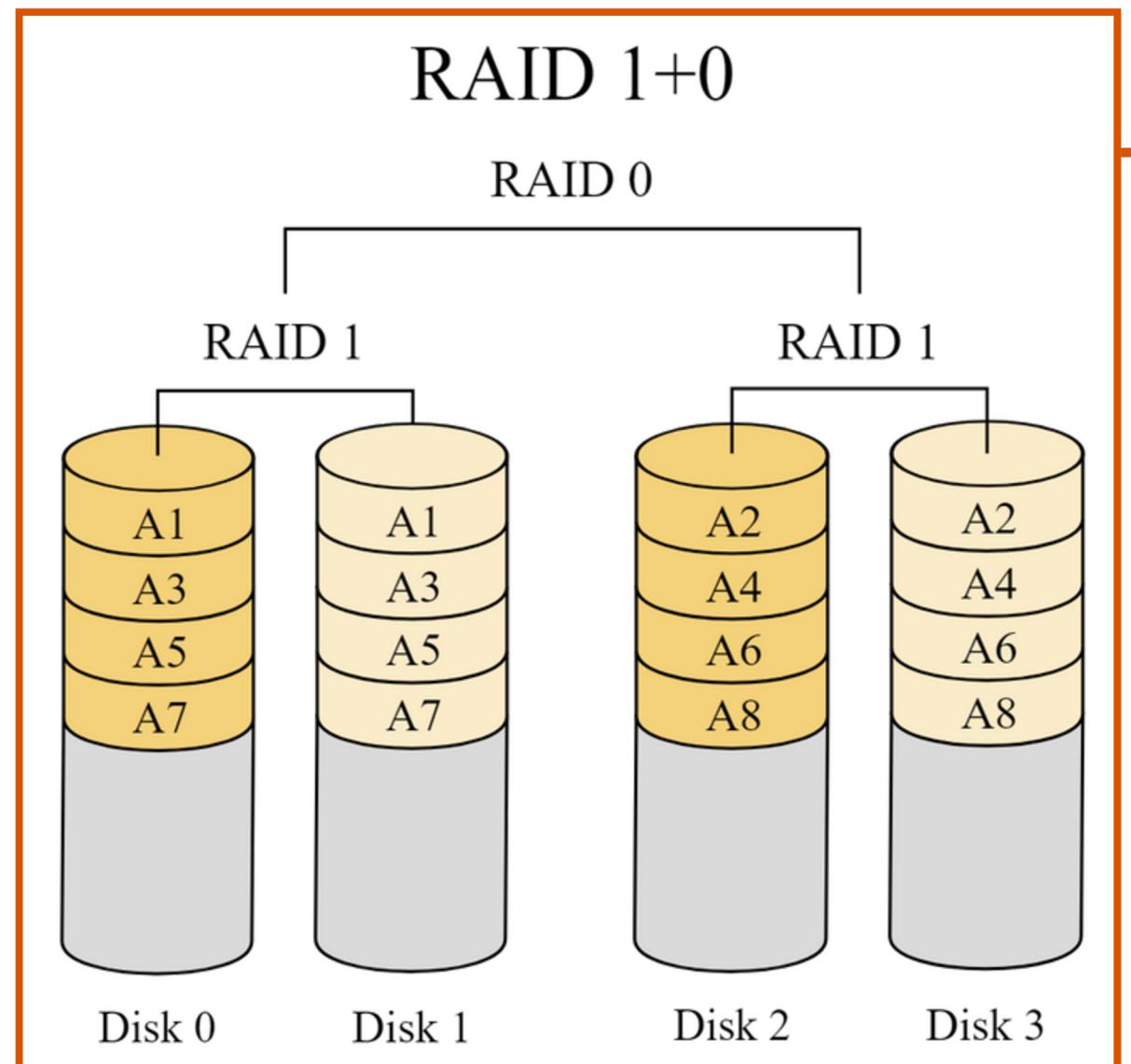
RAID TECHNOLOGY



RAID 1 - Mirroring

- ❖ Cần ít nhất 2 ổ đĩa, dữ liệu được ghi vào 2 ổ giống hệt nhau.
- ❖ **Ưu điểm:** An toàn về dữ liệu, nếu 1 trong 2 ổ đĩa bị hỏng thì dữ liệu trong ổ còn lại.
- ❖ **Nhược điểm:** Hiệu suất không cao, tốn kém

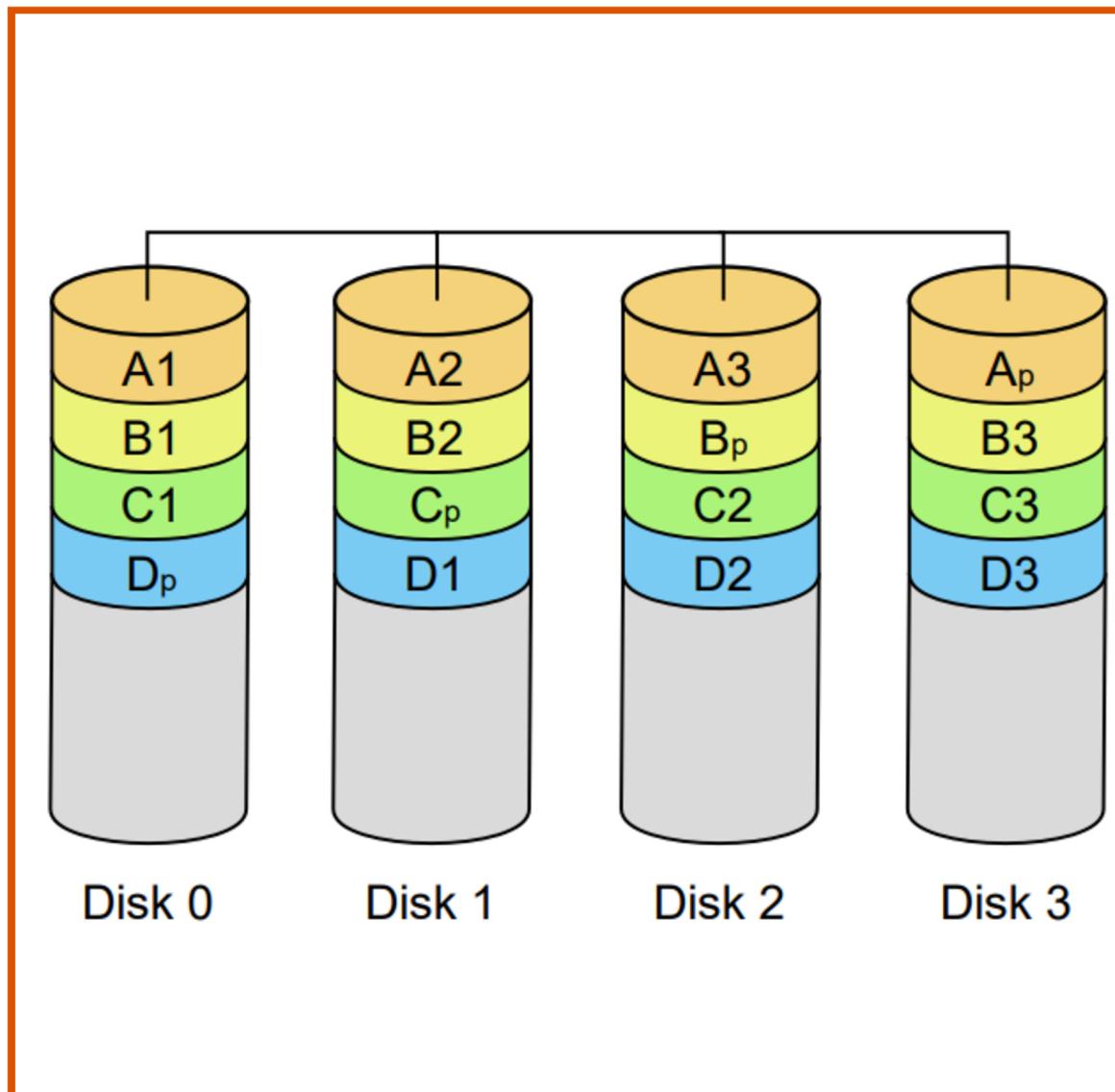
RAID TECHNOLOGY



RAID 10 - (Stripe + Mirror)

- ❖ Là sự kết hợp giữa Raid 1 (Mirror) và Raid 0 (Stripe), cần ít nhất 4 ổ đĩa.
- ❖ **Ưu điểm:** Tăng hiệu suất mà vẫn đảm bảo dữ liệu không bị thất thoát khi 1 trong số 4 ổ cứng bị hỏng.
- ❖ **Nhược điểm:** Chi phí cao
- ❖ Thích hợp với tất cả các đối tượng sử dụng

RAID TECHNOLOGY



RAID 5 - (Parity + Stripe)

- ❖ Lưu trữ các ổ cứng riêng biệt và vẫn có dự phòng.
- ❖ Cần ít nhất 3 ổ đĩa.
- ❖ **Ưu điểm:** Tăng hiệu suất, an toàn dữ liệu cao
- ❖ **Nhược điểm:** Tốn kém
- ❖ Những website, ứng dụng có số lượng truy cập và yêu cầu tài nguyên từ nhỏ đến vừa và lớn.



**THANKS FOR
ÜHATCHING**

toy có tình, khởi chỉnh (((=