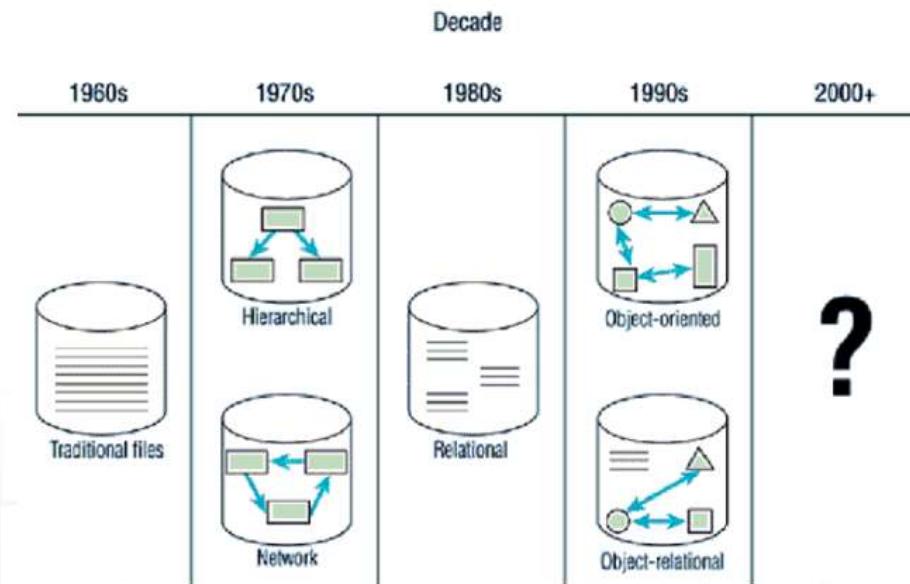


ĐẶT VĂN ĐỀ

- Ứng dụng có sử dụng CSDL rất phổ biến hiện nay, bao phủ hầu hết trong các hoạt động kinh tế, xã hội, giáo dục, y tế → Tầm quan trọng của một công cụ trị CSDL
- Lịch sử phát triển của mô hình CSDL cũng qua nhiều giai đoạn:



- Tương ứng với sự phát triển của mô hình lưu trữ dữ liệu là sự phát triển của phần mềm cài đặt mô hình dữ liệu đó (HOTCSDL)

ĐẶT VĂN ĐỀ



ORACLE®

Môn học: HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU



Các phần mềm
này hoạt động
như thế nào ? Tại
sao ? Có những
thành phần nào ?



MỤC TIÊU MÔN HỌC

LÝ
THUYẾT

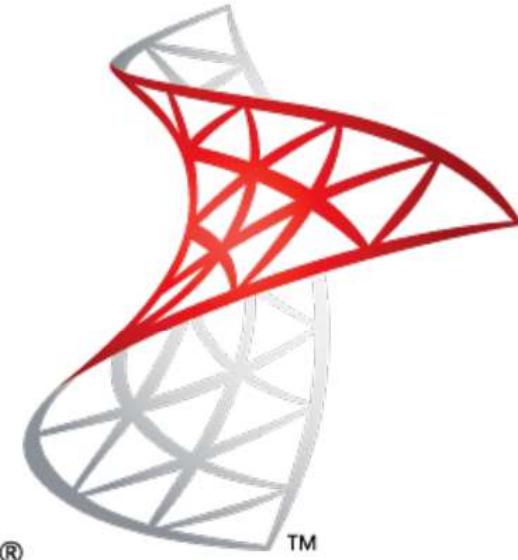
Cung cấp cho sinh viên kiến thức nền tảng về các Hệ quản trị Cơ sở dữ liệu (HQTCSQL): Các thành phần của một HQTCSQL và các chức năng của chúng, các cơ chế **quản lý truy xuất đồng thời, an toàn** và **an ninh dữ liệu, tối ưu hóa câu hỏi**, các cấu trúc **tổ chức lưu trữ** bên trong.

THỰC
HÀNH

Tìm hiểu và vận dụng các kỹ thuật **quản lý truy xuất đồng thời** của một HQTCSQL cụ thể: **MS SQL Server**



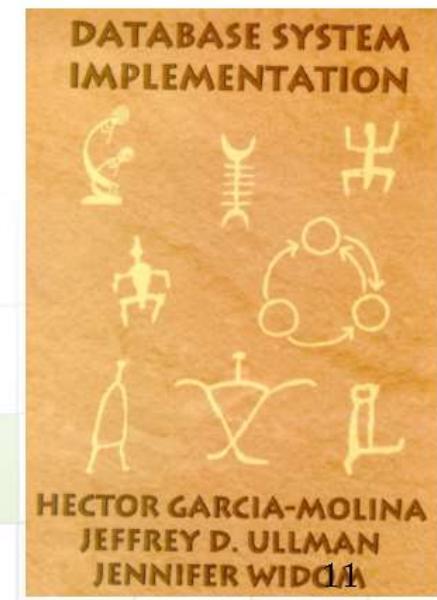
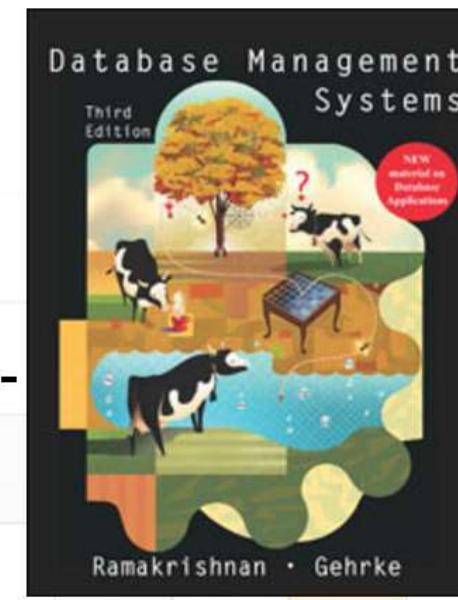
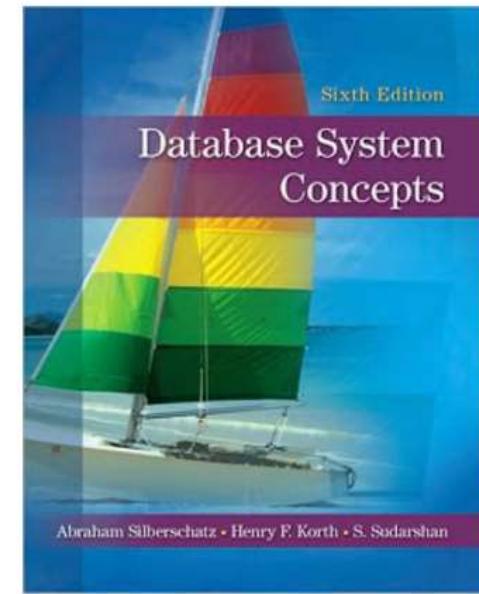
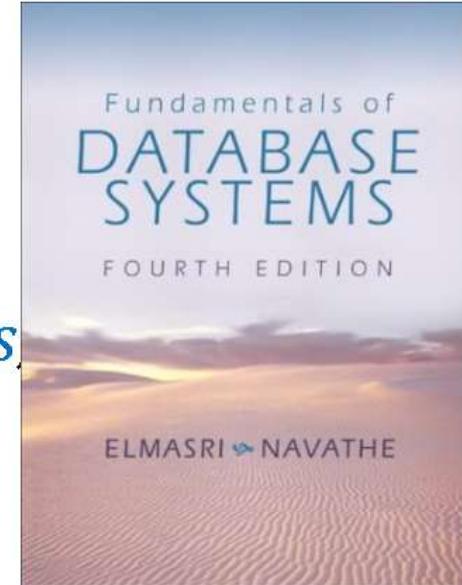
PHẦN MỀM



Microsoft®
SQL Server®

TÀI LIỆU THAM KHẢO

- *Fundamentals of Database Systems*, 4th Edition, Elmasri
Navathe
- *Database Management Systems*,
3rd Edition, Raghu
Ramakrishnan and Johannes
Gehrke
- *Database System Concepts*, 4th
Edition, Silberschatz–Korth
–Sudarshan
- *Database Systems Implementation*, Hector Garcia-Molina, D. Ullman, Jennifer D. Widom



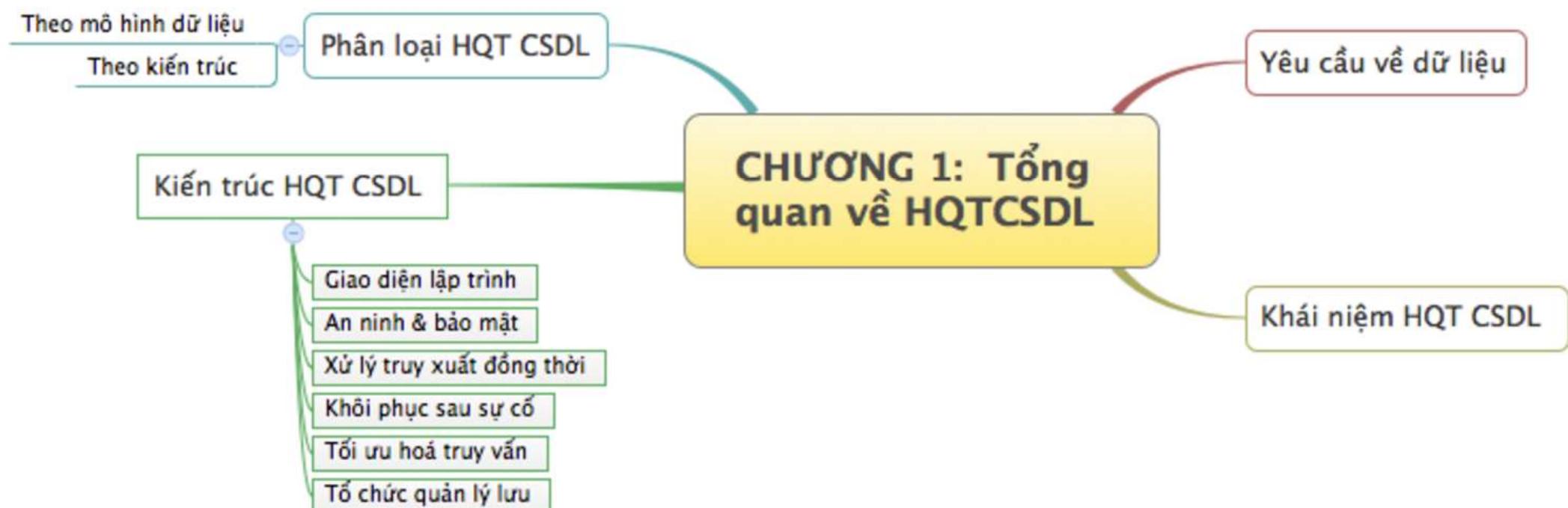


HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

Chương 1:

TỔNG QUAN VỀ HQT CSDL

Nội dung





NỘI DUNG

- ❖ Yêu cầu về dữ liệu trong CSDL
- ❖ Khái niệm HQT CSDL
- ❖ Kiến trúc của một HQT CSDL
- ❖ Phân loại HQT CSDL



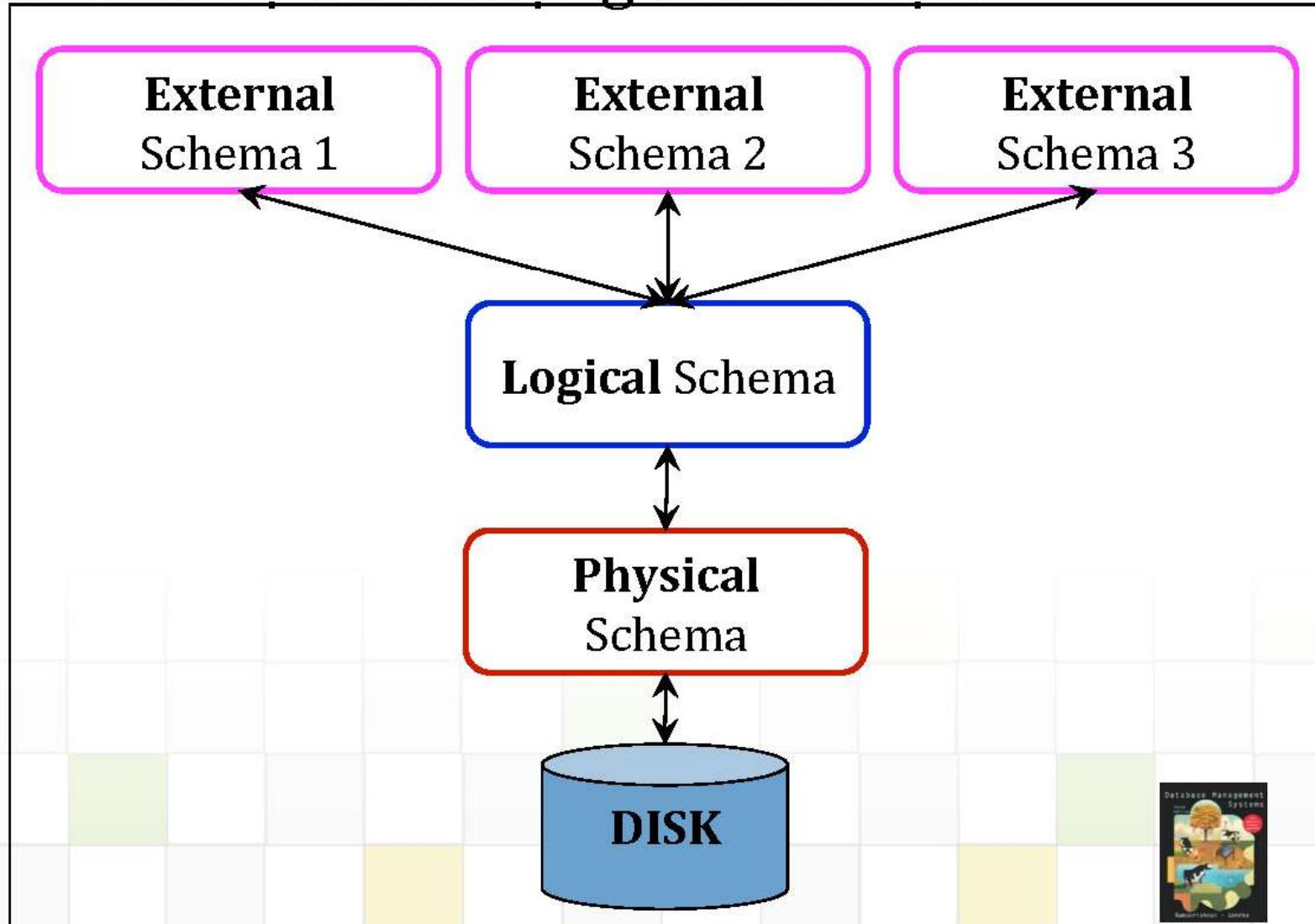
Yêu cầu về dữ liệu trong CSDL

- Dữ liệu trong CSDL phải được thể hiện ở các mức độ trừu tượng khác nhau (3 mức độ):
 - **Mức ngoài** (External level)
 - Mô tả một phần của CSDL mà một đối tượng / một nhóm người dùng được quyền tiếp cận
 - **Mức luận lý** (Logic level)
 - Mô tả những thông tin gì được lưu trữ trong CSDL và những mối quan hệ giữa những thông tin đó
 - **Mức vật lý** (Physical level)
 - Dữ liệu được lưu trữ như thế nào trên thiết bị lưu trữ.

→ Làm tăng tính độc lập (*data independence*) của cách thức lưu trữ dữ liệu, thiết kế dữ liệu và chương trình sử dụng dữ liệu.

Yêu cầu về dữ liệu trong CSDL

- Các mức độ trừu tượng của dữ liệu:

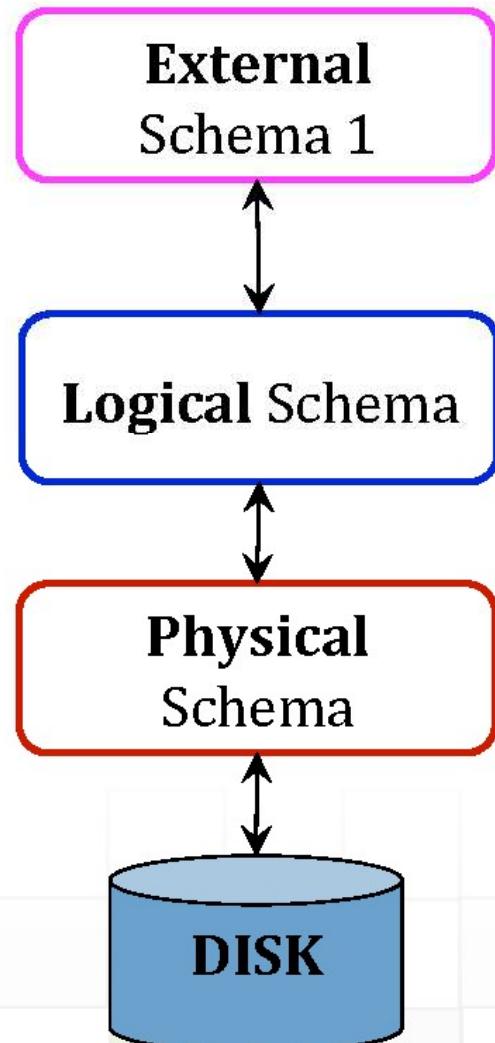




Yêu cầu về dữ liệu trong CSDL

- Dữ liệu trong CSDL cần có các đặc trưng:
 - Ít hoặc không trùng lặp dữ liệu
 - Chia sẻ cho nhiều người dùng mà không gây ra xung đột
 - An ninh, bảo mật
 - Khôi phục khi có sự cố
 - Độc lập dữ liệu
 - Độc lập luận lý: Khả năng thay đổi lược đồ mức luận lý mà không làm ảnh hưởng đến lược đồ ngoài cũng như chương trình ứng dụng.
 - Độc lập vật lý: Khả năng thay đổi tổ chức vật lý của CSDL mà không làm ảnh hưởng đến lược đồ luận lý.
- Vì vậy cần có một hệ thống quản lý hiệu quả dữ liệu trong CSDL.

Lợi ích của tính độc lập dữ liệu



- Độc lập luận lý:
 - Cho phép thêm bớt thuộc tính, bảng, các mối quan hệ mà không cần phải viết lại chương trình, ...

- Độc lập vật lý:
 - Cho phép thay đổi thiết bị lưu trữ, cách thức lưu trữ, các cấu trúc dữ liệu, các tổ chức tập tin khác nhau, các kiểu tổ chức chỉ mục khác nhau, ...



NỘI DUNG

- ❖ Yêu cầu về dữ liệu trong CSDL
- ❖ Khái niệm HQT CSDL
- ❖ Kiến trúc của một HQT CSDL
- ❖ Phân loại HQT CSDL



Khái niệm HQT CSDL

- Là một **hệ thống phần mềm** cung cấp các công cụ để xây dựng, khai thác và quản lý cơ sở dữ liệu.
 - **Xây dựng** (Sử dụng ngôn ngữ DDL): Định nghĩa cấu trúc CSDL, lưu trữ dữ liệu
 - **Khai thác** (Sử dụng ngôn ngữ DML): Truy vấn dữ liệu, Cập nhật dữ liệu
 - **Quản lý**:
 - Quản lý an toàn và bảo mật
 - Điều khiển truy xuất đồng thời.
 - Khôi phục khi có sự cố.
 - ...
- Một số HQTCSQL: MS SQL Server, Oracle, DB2, ...



Lịch sử phát triển của các HQT CSDL

Decade of RDBMS

1960s	1970s	1980s – 1990s	2000s
Mô hình mạng <i>CODASYL</i> Mô hình phân cấp <i>IMS</i>	Mô hình quan hệ QUEL SEQUEL	Mô hình đối tượng SQL	No SQL Database
<i>SABRE system</i>	<i>Ingres</i> <u>PostgreSQL</u> <u>dBASE</u> Ingres Corp Sybase MS SQL Server <i>System R</i> Non-Stop SQL SQL/DS DB2 Allbase Oracle	Prototypes for ODBMS	MongoDB, Oracle NoSQL Database, Apache Cassandra , ...



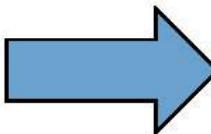
NỘI DUNG

- ❖ Yêu cầu về dữ liệu trong CSDL
- ❖ Khái niệm HQT CSDL
- ❖ Kiến trúc của một HQT CSDL
- ❖ Phân loại HQT CSDL



Kiến trúc của một HQT CSDL

- Các thành phần chính:





Thành phần Giao diện lập trình

- HQTCSQL cung cấp giao diện lập trình dễ sử dụng với một ngôn ngữ lập trình CSDL:
 - Giao diện: tương tác dòng lệnh (command line), đồ họa (GUI)
 - Ngôn ngữ: SQL, T-SQL
 - VD: MS SQL Server cung cấp ngôn ngữ Transacion SQL (T-SQL)
- Các loại ngôn ngữ sử dụng trong HQTCSQL:
 - *Ngôn ngữ định nghĩa dữ liệu (DDL – Data Definition Language)*: Giúp người dùng ra lệnh cho HQTCSQL tạo ra các cấu trúc dữ liệu của CSDL (Cách tổ chức dữ liệu và mối liên hệ giữa các đối tượng dữ liệu).
 - *Ngôn ngữ thao tác CSDL (DML – Data Manupulation Language)* : Giúp người dùng tích luỹ, hiệu chỉnh và khai thác dữ liệu



Thành phần An ninh và bảo mật

- **Bảo mật dữ liệu:** HQTCSDL hỗ trợ các tính năng về chứng thực, phân quyền giúp kiểm soát tốt những người dùng hợp pháp của hệ thống..
- **An ninh dữ liệu:** HQTCSDL hỗ trợ các phương pháp mã hóa dữ liệu để ngăn chặn các tấn công của những đối tượng tin tặc (đánh cắp thông tin trên đường truyền, đánh cắp nội dung CSDL).



Thành phần Khôi phục sau sự cố

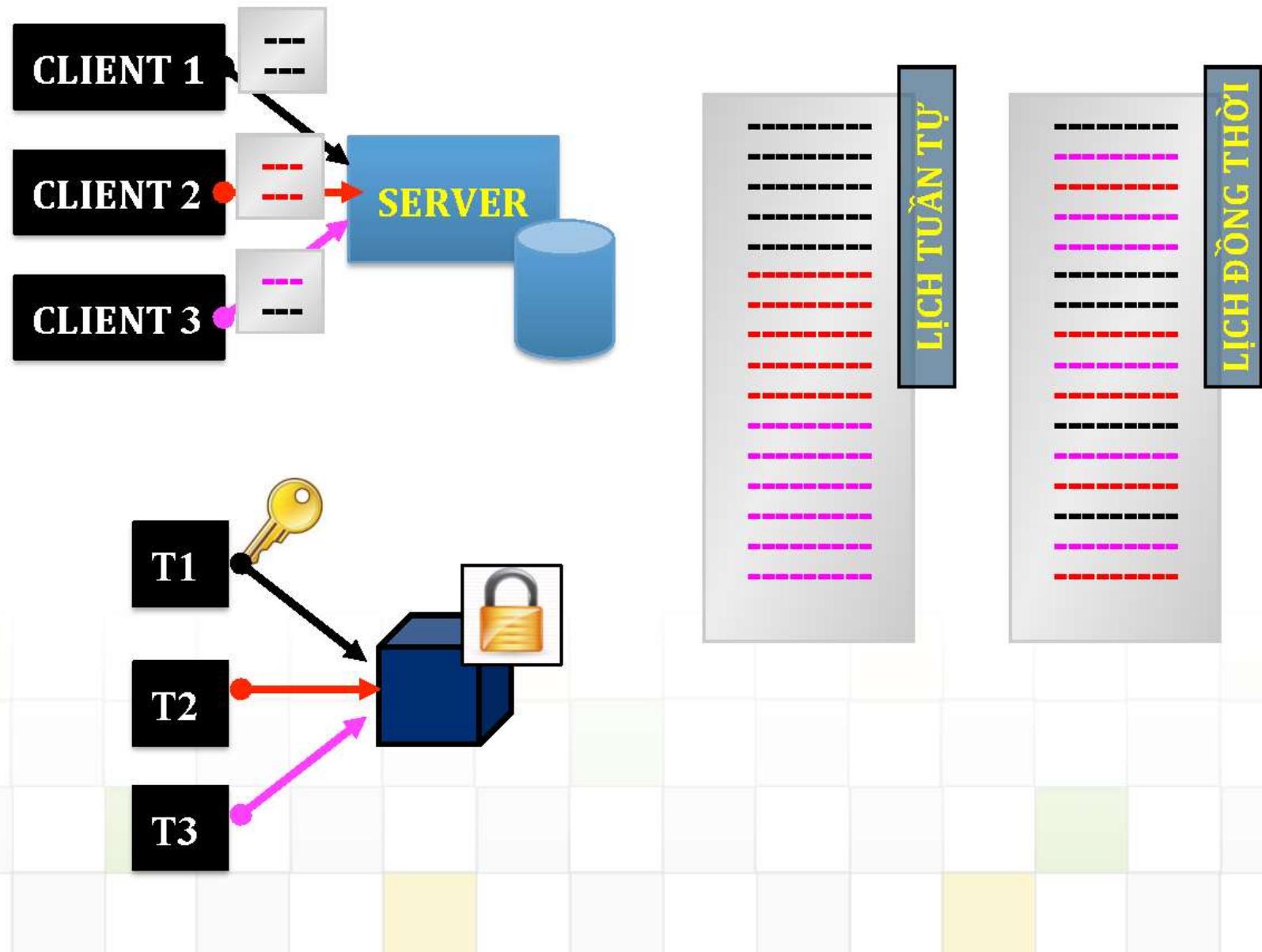
- Mục tiêu: Đảm bảo sự tổn thất, sai sót về mặt dữ liệu là ít nhất có thể.
- Cách tiếp cận: Để đảm bảo tính bền vững của CSDL, mọi thay đổi lên CSDL phải được ghi nhận lại trong nhật ký (Log)
- Các thành phần hỗ trợ quá trình khôi phục sau sự cố:
 - *Bộ phận quản lý nhật ký (Log manager)*: đảm bảo ghi nhận đầy đủ và chính xác mọi thay đổi trên CSDL vào nhật ký.
 - *Bộ phận quản lý khôi phục sự cố (Recovery Manager)*: dựa vào nhật ký để phục hồi lại CSDL về trạng thái nhất quán trước đó (Trạng thái thoả tất cả RBTV của CSDL)



Xử lý truy xuất đồng thời

- Mục tiêu:
 - Đảm bảo các xử lý có thể được thực hiện đồng thời mà làm không làm cho dữ liệu bị mất tính nhất quán (vi phạm các ràng buộc toàn vẹn)
- Các thành phần con: Bộ phận quản lý giao tác (Transaction Manager & Locking Manager)
- Phương pháp:
 - Sử dụng khái niệm giao tác (**transaction**) để biểu diễn một đơn vị xử lý, một giao tác bao gồm các hành động mà được thực hiện tòn bộ hoặc không có hành động nào được thực hiện.
 - Bộ lập lịch (**scheduler**) có nhiệm vụ lập 1 lịch thực hiện từ n giao tác không tách biệt về thời gian sao cho kết quả không vi phạm tính nhất quán của CSDL.
 - Sử dụng cơ chế khóa (**lock**) để khóa các đơn vị dữ liệu nào đó khi cần → ngăn 2 giao tác cùng thao tác lên 1 đơn vị dữ liệu ấy tại cùng 1 điểm → Hỗ trợ để lập lịch.

Điều khiển đồng thời (tt)





Điều khiển đồng thời (tt)

- Vấn đề deadlock
 - Do sử dụng cơ chế khóa nên các giao tác sẽ phải chờ khi cần truy xuất 1 đơn vị dữ liệu đang bị khóa.
 - Tình huống chờ vĩnh viễn mà vẫn không được truy xuất đơn vị dữ liệu bị khóa gọi là Deadlock (khoá chết)
 - Các giao tác chờ đợi lẫn nhau để được cấp phát tài nguyên và không giao tác nào có thể hoàn tất.
 - Thành phần quản lý giao tác sẽ phải can thiệp vào:
 - Hoặc hủy bỏ một trong các giao tác gây deadlock
 - Hoặc ngăn chặn từ trước để không bao giờ xảy ra deadlock



Xử lý truy vấn

- Biểu diễn câu truy vấn ở dạng ngôn ngữ cấp cao (SQL) và thực hiện câu truy vấn có hiệu quả.
- Query compiler – biên dịch truy vấn

Query parser

- Xây dựng cấu trúc phân tích câu truy vấn dưới dạng cây

Query
preprocessor

- Kiểm tra ngữ nghĩa của câu truy vấn
- Chuyển đổi cấu trúc cây sang ngôn ngữ đại số quan hệ

Query
optimizer

- Sắp xếp các phép toán nhằm mục đích tối ưu hóa câu truy vấn



Quản lý lưu trữ

- Thành phần có nhiệm vụ điều khiển việc đọc/ghi dữ liệu qua lại giữa bộ nhớ và thiết bị lưu trữ
- Làm việc với các khái niệm:
 - Tập tin dữ liệu
 - Từ điển dữ liệu
 - Lưu trữ các metadata (Siêu dữ liệu) về cấu trúc của CSDL, đặc biệt là lược đồ của CSDL
 - Chỉ mục
 - Giúp cho việc tìm kiếm Dữ liệu được nhanh chóng



NỘI DUNG

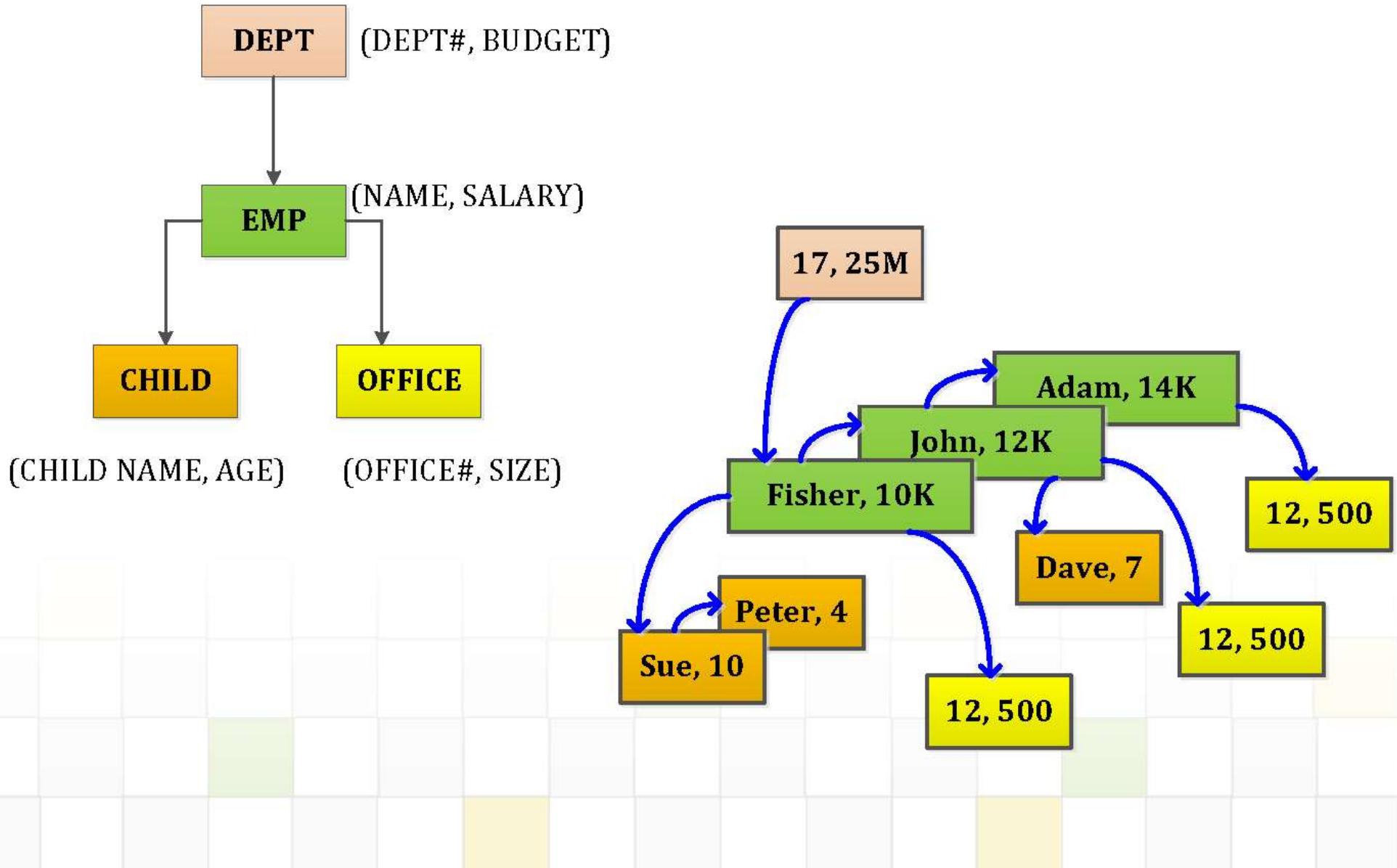
- ❖ Yêu cầu về dữ liệu trong CSDL
- ❖ Khái niệm HQT CSDL
- ❖ Kiến trúc của một HQT CSDL
- ❖ Phân loại HQT CSDL



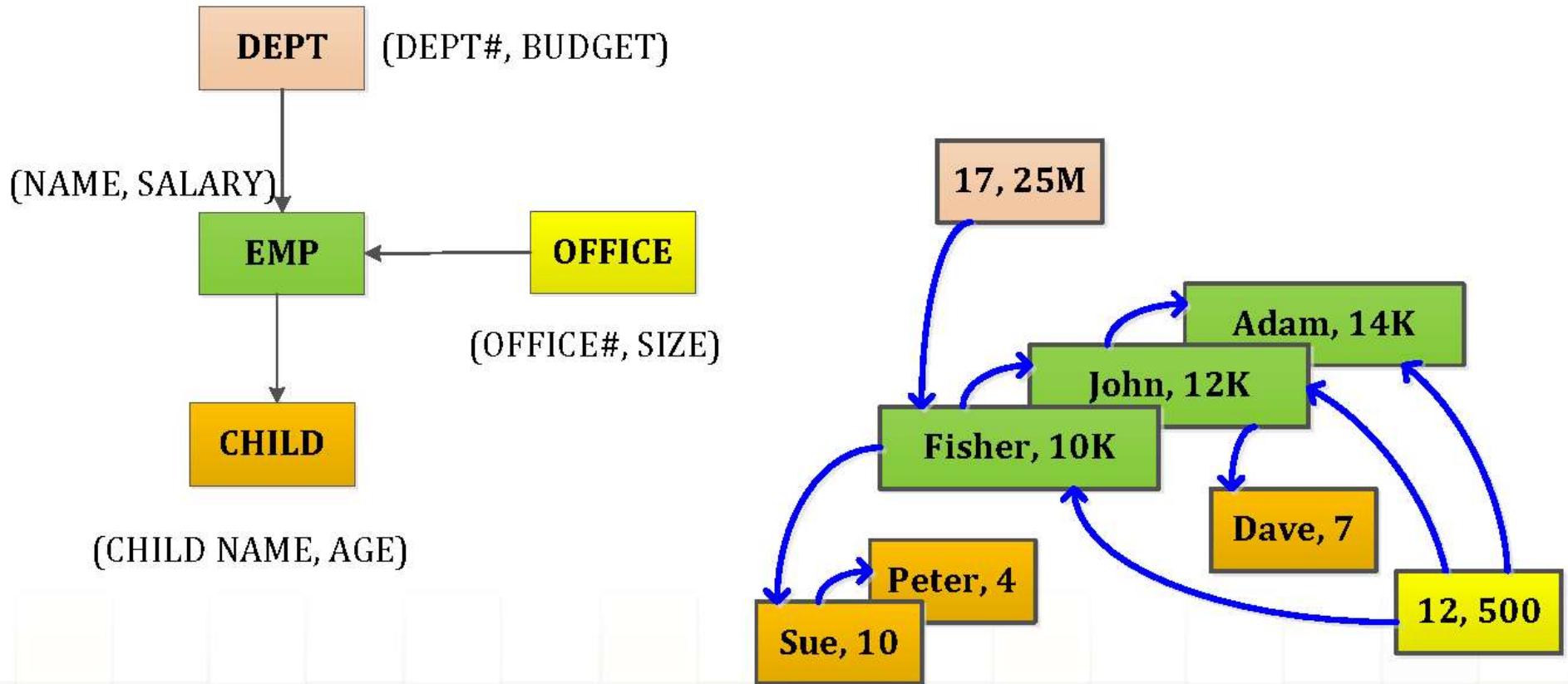
Phân loại HQT CSDL

- Theo mô hình dữ liệu:
 - Phân cấp
 - Mạng
 - Quan hệ
 - Đối tượng
- Theo kiến trúc tính toán:
 - Tập trung (Centralized database system)
 - Khách / chủ (Client server database system)
 - Phân tán (Distributed database system)
- Theo đặc tính:
 - HQTCSDL thời gian thực (real-time database system)
 - HQTCSDL chịu lỗi cao (high fault tolerance database system)
 - HQTCSDL đa phương tiện (multi-media database system)

Mô hình phân cấp



Mô hình mạng





Mô hình quan hệ

DEPT (DEPT #, BUDGET)

OFFICE (OFFICE #, SIZE)

EMP (NAME, SALARY)

CHILD (CHILD NAME, AGE)

WORKS (DEPT #, NAME)

OFFSPRING (NAME, CHILD NAME)

OCCUPIED (NAME, OFFICE #)

OCCUPIED		
Fisher	12	
John	12	
Adam	12	

DEPT		
17		25M
12		500

CHILD		
Sue	10	
Peter	4	
Dave	7	

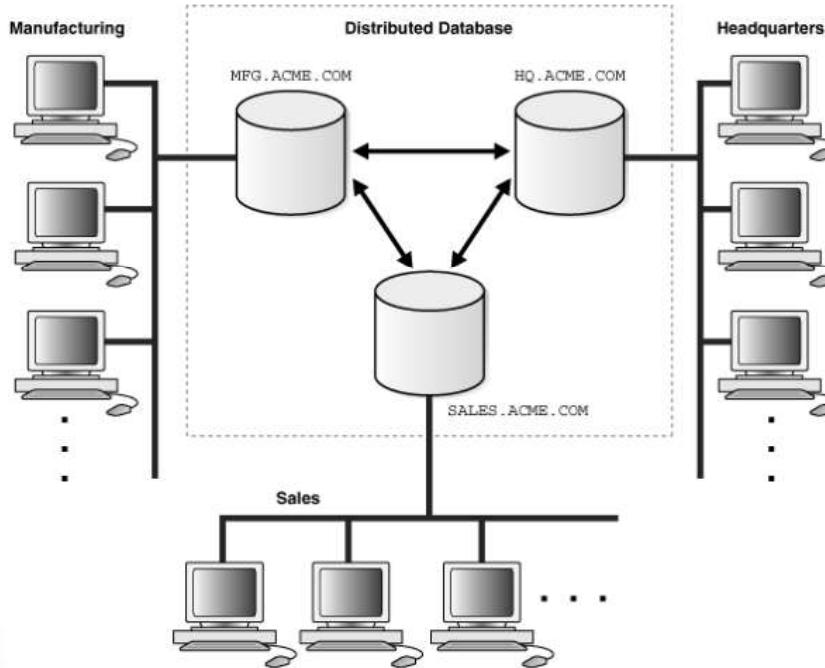
EMP		
Fisher	10K	
John	12K	
Adam	14K	

OFFSPRING		
Fisher	Sue	
Fisher	Peter	
Jone	Dave	

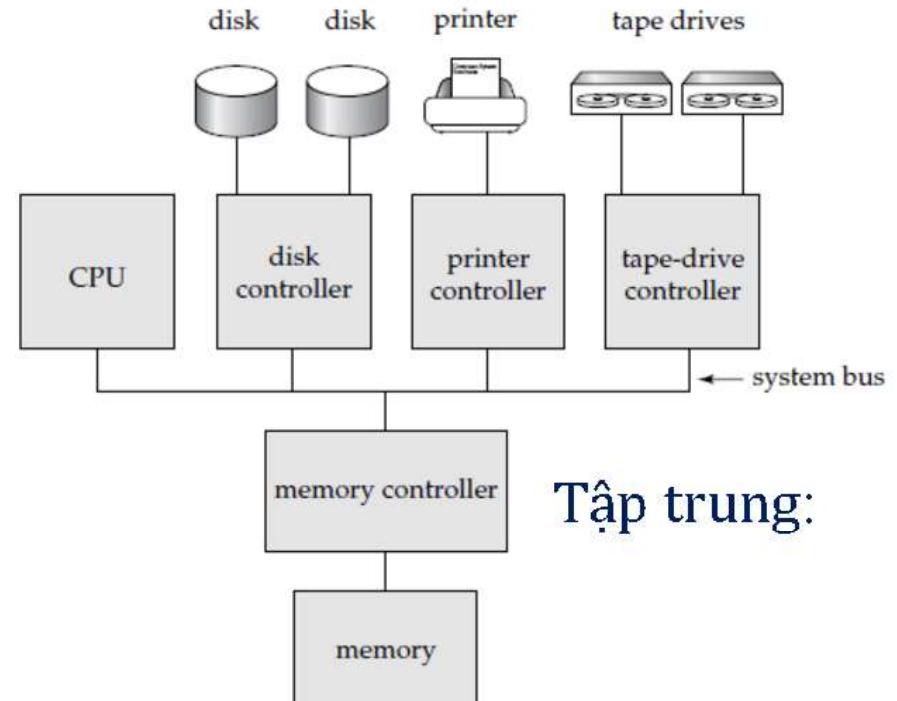
WORKS		
17		Fisher
17		John
17		Adam

Phân loại HQTCSDL

- Theo kiến trúc tính toán:



Phân tán



Tập trung:

Figure 18.1 A centralized computer system.

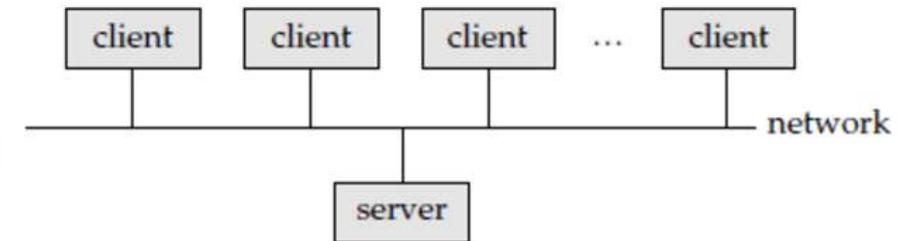


Figure 18.2 General structure of a client-server system.

Khách / chủ