

CHƯƠNG 1

TỔNG QUAN VỀ MỘT HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

Nhắc lại nội dung môn học

Chương I: Tổng quan về một HQT CSDL

Chương II: Quản lý giao tác và truy xuất đồng thời

Chương III: Phục hồi dữ liệu sau sự cố – An toàn dữ liệu

Chương IV: Lưu trữ dữ liệu và các phương thức truy xuất

Chương V: Xử lý và tối ưu hóa câu truy vấn

Mục tiêu & Nội dung chương 1

Mục tiêu:

- Giới thiệu kiến thức tổng quan liên quan đến một HQT CSDL.

Nội dung:

1. Giới thiệu về HQT CSDL.
2. Lịch sử phát triển của các HQT CSDL.
3. Kiến trúc một HQT CSDL.
4. Phân loại các HQT CSDL.

Nhắc lại

- Một số đặc tính của CSDL
 - Một CSDL là một tập hợp dữ liệu có quan hệ logic với nhau.
 - CSDL thể hiện một thế giới thực thu nhỏ.
 - CSDL được thiết kế, cài đặt và khai thác cho mục đích cụ thể.
 - Có kích thước và độ phức tạp khác nhau.
 - Được duy trì thủ công hoặc bằng máy tính.

Nhắc lại

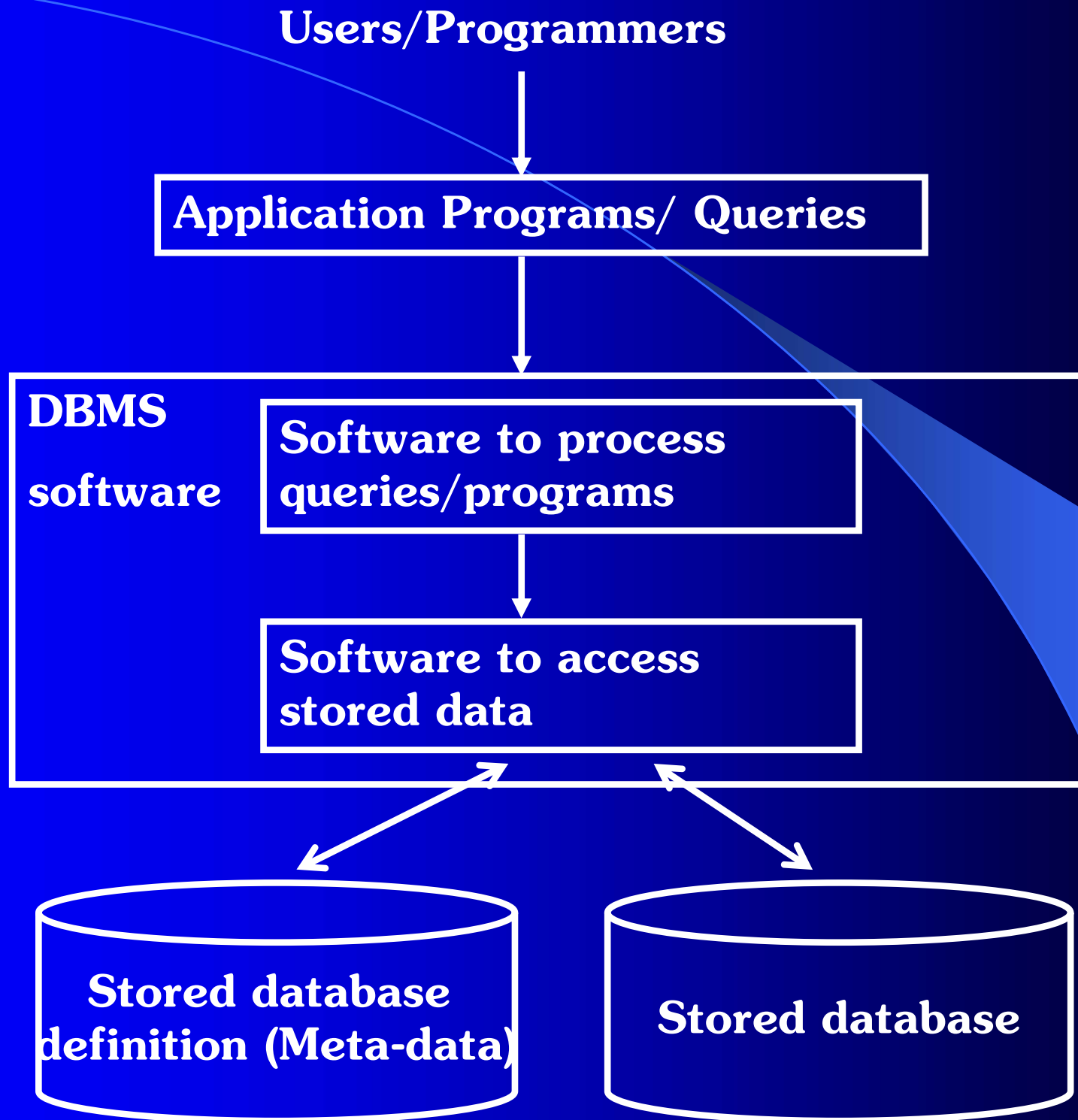
- **Đặc tính của dữ liệu trong một CSDL**
 - Không trùng lặp.
 - Có thể chia sẻ dữ liệu cho nhiều người dùng (tạo nhiều khung nhìn khác nhau).
 - Có thể đảm bảo an toàn dữ liệu.
 - Cho phép truy xuất đồng thời.
 - Khôi phục dữ liệu khi có sự cố.
 - Tính tự mô tả của CSDL.
 - Chia sẻ dữ liệu và xử lý giao tác liên quan đến nhiều người dùng.

Người dùng

- Các người dùng liên quan đến 1 CSDL
 - Người quản trị CSDL (dba).
 - Người thiết kế CSDL (Database designer).
 - Người dùng.
 - Phân tích viên hệ thống và Lập trình viên ứng dụng.

Định nghĩa

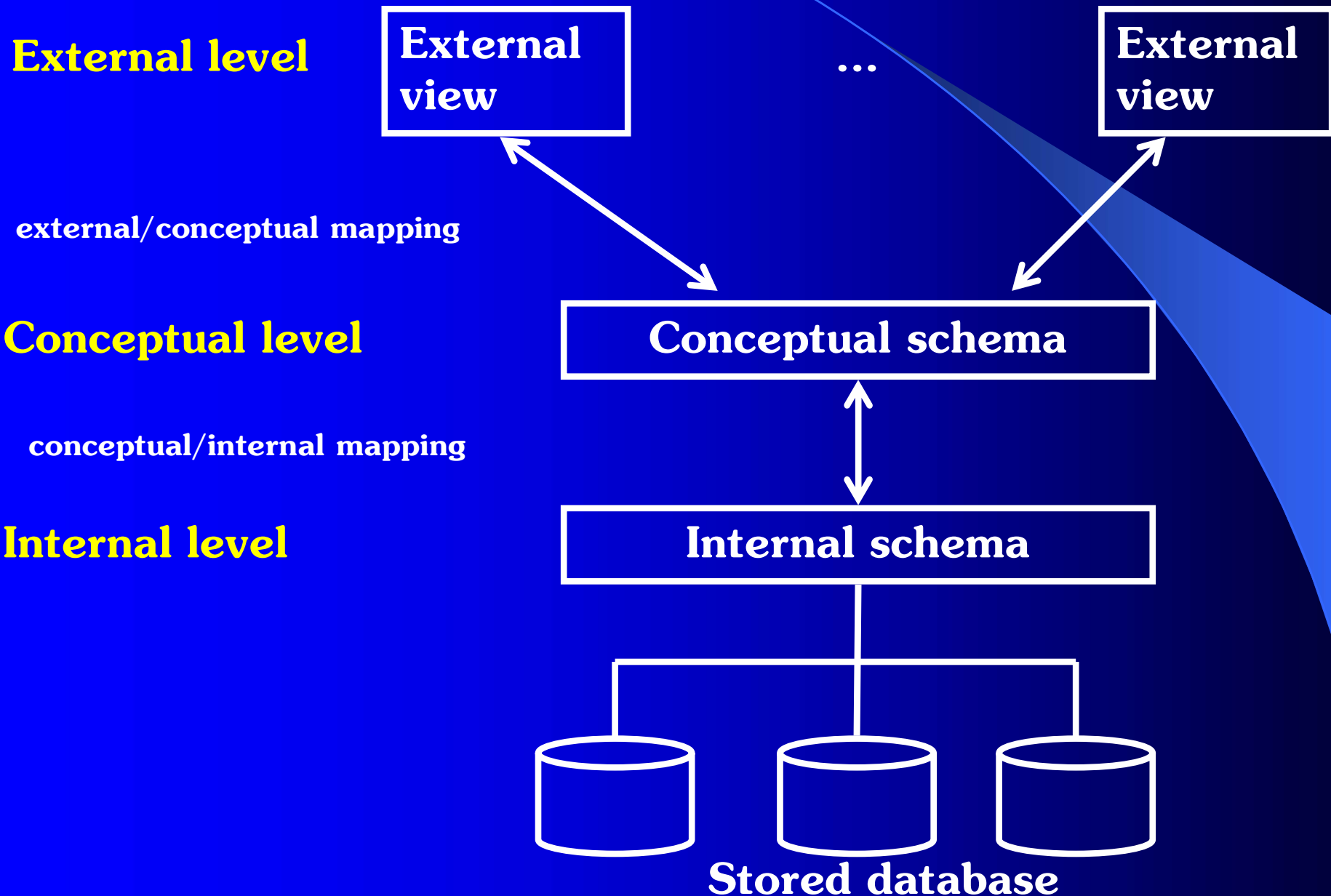
- **Database Management System : DBMS**
 - **Định nghĩa:** Là một chương trình cho phép người dùng định nghĩa, xây dựng, chia sẻ và bảo vệ dữ liệu.



Hệ cơ sở dữ liệu

- Database + DBMS = database system

Hệ CSDL: Kiến trúc 3 lược đồ của hệ CSDL



Hệ cơ sở dữ liệu

- **Độc lập dữ liệu (data independence):** thay đổi lược đồ mức này mà không phải thay đổi lược đồ ở mức cao hơn.
 - **Logical data independence:** có thể thay đổi lược đồ mức quan niệm (thêm/bớt kiểu mẫu tin, thêm/bớt mục dữ liệu, thay đổi ràng buộc) mà không cần thay đổi lược đồ ngoài/chương trình ứng dụng, chỉ cần thay đổi định nghĩa view và thay đổi ánh xạ.
 - **Physical data independence:** có thể thay đổi lược đồ trong (tạo lại/ thêm chỉ mục) mà không phải thay đổi lược đồ mức quan niệm, và vì vậy không cần phải thay đổi lược đồ ngoài.

Giới thiệu

- Ưu điểm của việc dùng HQT CSDL
 - Quản lý dư thừa.
 - Quản lý quyền truy cập.
 - Lưu trữ bền vững.
 - Hỗ trợ cấu trúc lưu trữ phục vụ truy vấn hiệu quả.
 - Sao chép và phục hồi dữ liệu.
 - Hỗ trợ giao diện đa người dùng.
 - Thể hiện mối quan hệ dữ liệu phức tạp.
 - Ép thỏa ràng buộc toàn vẹn.
 - Dễ quản lý, giảm thiểu thời gian phát triển ứng dụng, uyển chuyển, khả năng cập nhật thông tin.
 - Kinh tế.

Giới thiệu

- Có khi không nên dùng HQT CSDL do
 - Chi phí cao cho phần cứng, phần mềm, huấn luyện.
 - Hệ thống chậm do tính tổng quát của HQT khi định nghĩa, xử lý dữ liệu, quản lý bảo mật, điều khiển đồng thời, quản lý khôi phục dữ liệu,
- Dùng file để lưu thông tin khi:
 - CSDL đơn giản, không thay đổi thường
 - Không cần nhiều người dùng truy cập vào dữ liệu.

Lịch sử phát triển của HQT CSDL

- Tùy vào sự phát triển của ứng dụng sử dụng HQT CSDL
 - 1960s – 1980s: Mô hình phân cấp, mô hình mạng, dùng file.
 - Cuối 1970s – 1980s: RDBMS.
 - 1980s: ứng dụng hướng đối tượng.
 - 1990s: các ứng dụng Web và nhu cầu trao đổi dữ liệu tăng cao → dữ liệu XML. CSDL đa phương tiện, khai khoáng dữ liệu, dữ liệu thời gian thực, ... → cấu trúc dữ liệu phức tạp, kiểu dữ liệu mới, phương thức thao tác trên kiểu dữ liệu thay đổi, cách thức lưu trữ và cấu trúc chỉ mục thay đổi, ...

Các thành phần của HQT CSDL

Chương trình ứng dụng

Giao tiếp chương trình

Quản lý an toàn

Quản lý phục hồi dữ liệu

Quản lý giao tác

Quản lý truy xuất đồng thời

Quản lý lưu trữ

Giao tiếp chương trình

- Ngôn ngữ định nghĩa dữ liệu (DDL- Data Definition Language) dùng để tạo cấu trúc CSDL (*tạo bảng, chỉ mục*)
- Ngôn ngữ thao tác dữ liệu (DML- Data Manipulation Language): thêm, xóa, sửa, tìm kiếm dữ liệu, được nhúng trong các chương trình ứng dụng
- Giao diện HQT CSDL
 - Menu-based, cho web client, giao diện đồ họa, giao diện form, giao diện dành riêng cho DBA.

Quản lý an toàn

- Phân chia quyền truy xuất dữ liệu, bảo vệ dữ liệu, ngăn ngừa các truy xuất không hợp lệ.
- Ví dụ: CSDL quản lý dữ liệu về đăng ký học phần và điểm số của sinh viên:
 - Sinh viên có thể nhập mới học phần dự định học.
 - Giáo viên có thể thêm mới 1 học phần dự định mở.
 - Sinh viên không thể sửa điểm thi.

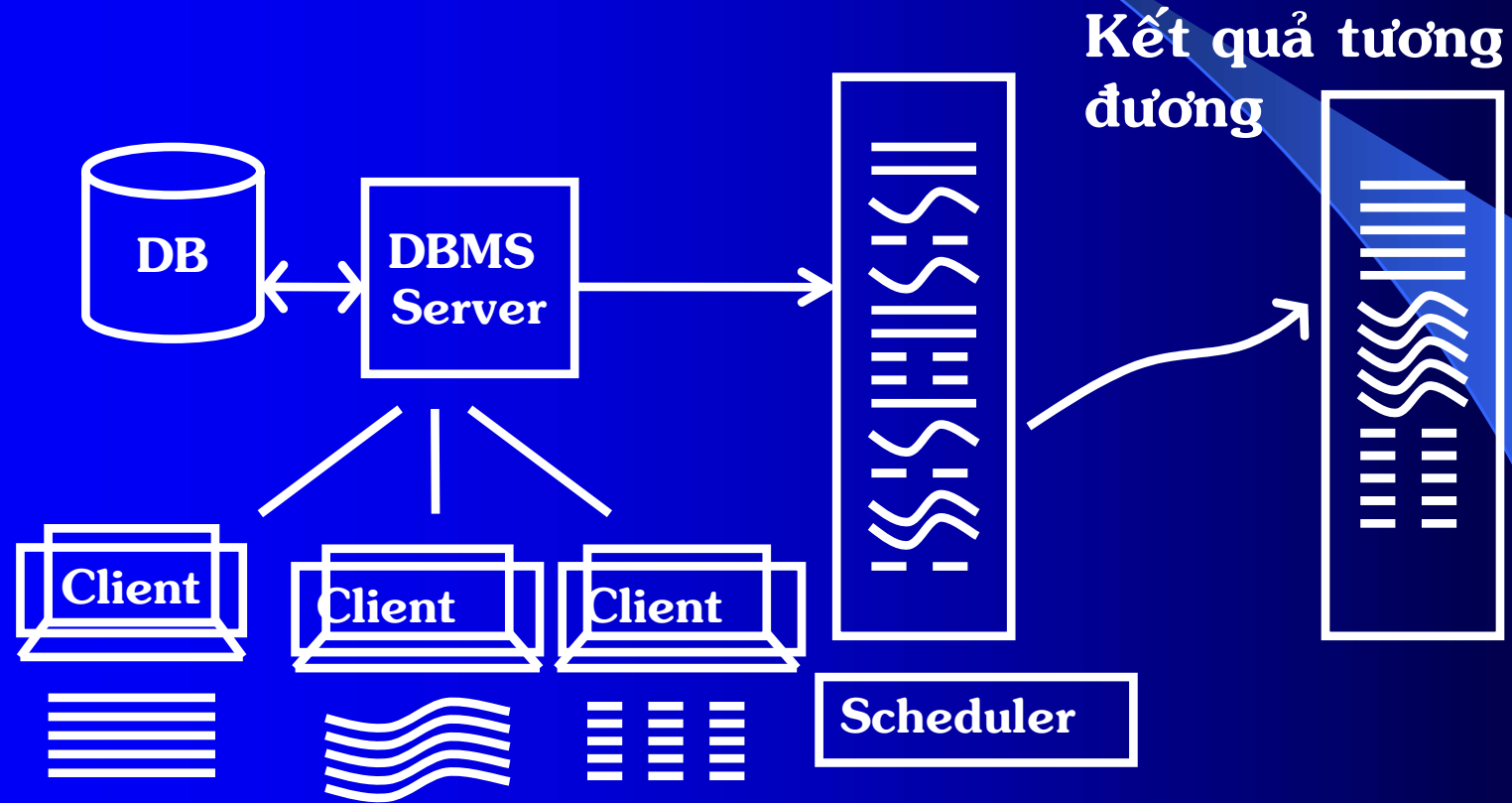
Phục hồi dữ liệu

- Phục hồi dữ liệu sau khi có sự cố xảy ra. Các thao tác của các giao tác chưa được hoàn tất sẽ được undo (*dựa vào log file*).
- Ví dụ: Hệ thống quản lý dữ liệu của 1 ngân hàng phải có cơ chế xử lý phù hợp khi mất điện: giao dịch đã hoàn tất thì phải được ghi nhận, giao dịch đang diễn ra và chưa hoàn tất thì không được ghi nhận.

Quản lý giao tác

- Giao tác (transaction) là 1 tập hợp có thứ tự các thao tác truy xuất dữ liệu trên CSDL thành 1 đơn vị công việc logic (*xem là 1 thao tác nguyên tố*), chuyển CSDL từ **trạng thái nhất quán này** sang **trạng thái nhất quán khác**.
- Quản lý giao tác điều hành và phối hợp việc thực hiện các giao tác trên CSDL.

Quản lý giao tác (tt)



Quản lý xử lý đồng thời

- Đơn vị cạnh tranh truy xuất là giao tác.
- Điều khiển cạnh tranh truy xuất nhằm phát hiện và ngăn ngừa các tình trạng làm cho một giao tác không thể tiếp tục thực hiện một cách bình thường.
Ví dụ: Hai khách hàng cùng mua 1 vé máy bay.

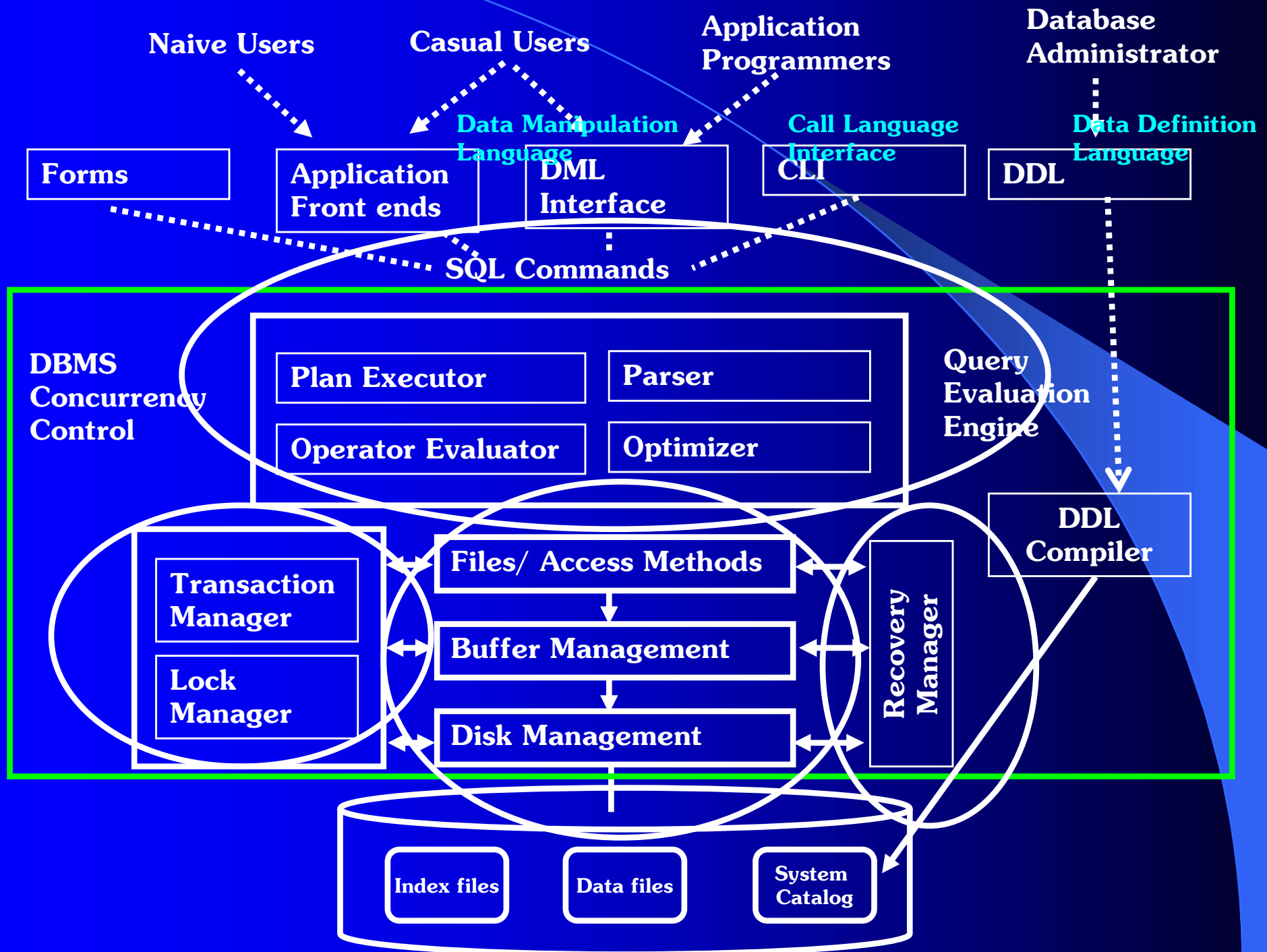
Quản lý lưu trữ

- Giao tiếp với hệ điều hành để truy xuất và lưu trữ dữ liệu.
- HQT CSDL sẽ quyết định sử dụng loại phương thức truy xuất trong khi hệ điều hành sẽ thực hiện việc truy xuất dữ liệu.

Meta data (Data Dictionary)

- Meta data còn được gọi là siêu dữ liệu, là dữ liệu để quản lý dữ liệu.
- Lưu trữ cấu trúc của các CSDL, các người dùng, mật khẩu, quyền truy xuất, chỉ mục, ...

Kiến trúc của DBMS



Phân loại HQT CSDL

- Theo mô hình dữ liệu
 - Mạng
 - Phân cấp
 - Quan hệ
 - Hướng đối tượng
 - XML
- Theo kiến trúc DB
 - Đơn người dùng
 - Khách / Chủ (Client/ Server)
 - Phân tán
- Theo loại dữ liệu
 - Dữ liệu Multimedia
 - Dữ liệu thời gian thực
- ...

Hết chương 1.