

HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

Tổng quan về CSDL và Hệ quản trị CSDL

Trình bày: TS. Phạm Thị Thu Thúy
Khoa CNTT, Trường Đại học Nha Trang
thuthuy@ntu.edu.vn

Đặt vấn đề

Nhược điểm của cách tổ chức dữ liệu theo hướng tiếp cận cũ

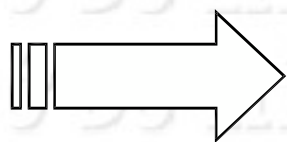
Dư thừa và không nhất quán

Dữ liệu có Format khác nhau

Chia sẻ thông tin giữa phòng ban ?

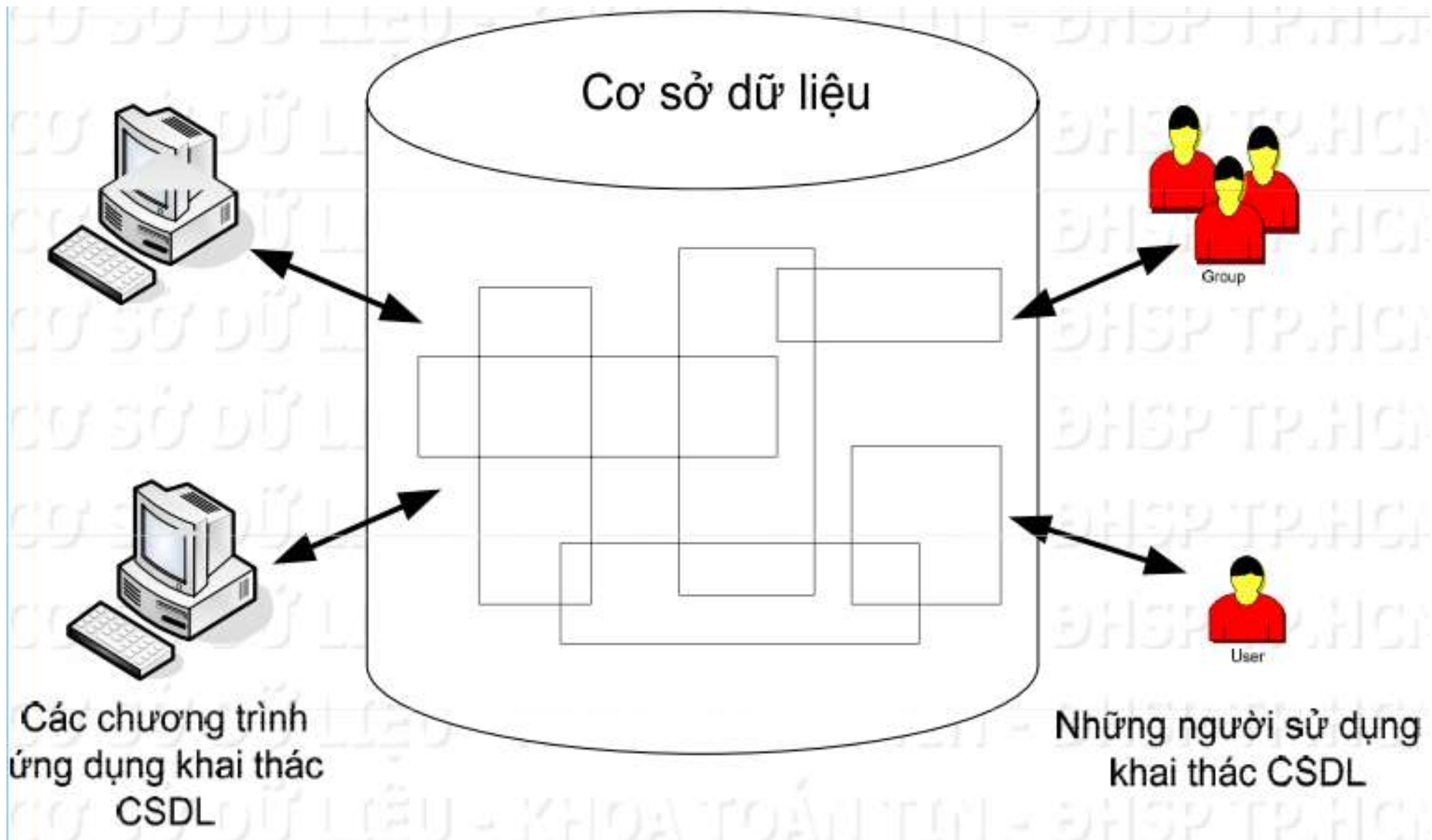
Nhu cầu truy xuất đồng thời ?

Vấn đề phân quyền ?



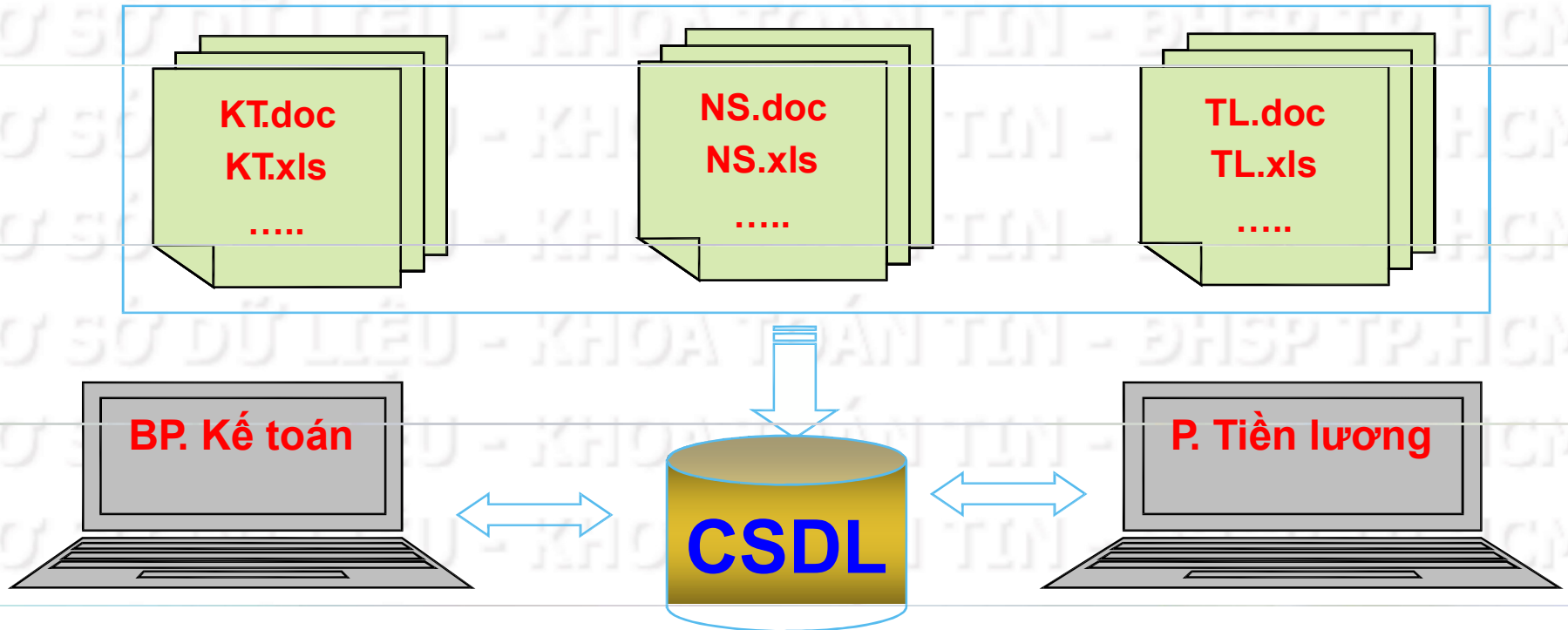
**CẦN TỔ CHỨC DỮ LIỆU
NHƯ THẾ NÀO ?**

Cơ sở dữ liệu - Database



Khái niệm về Cơ sở dữ liệu

Cơ sở dữ liệu là một tập hợp dữ liệu được tổ chức và lưu trữ theo một cấu trúc chặt chẽ nhằm phục vụ nhiều đối tượng sử dụng với nhiều mục đích khác nhau



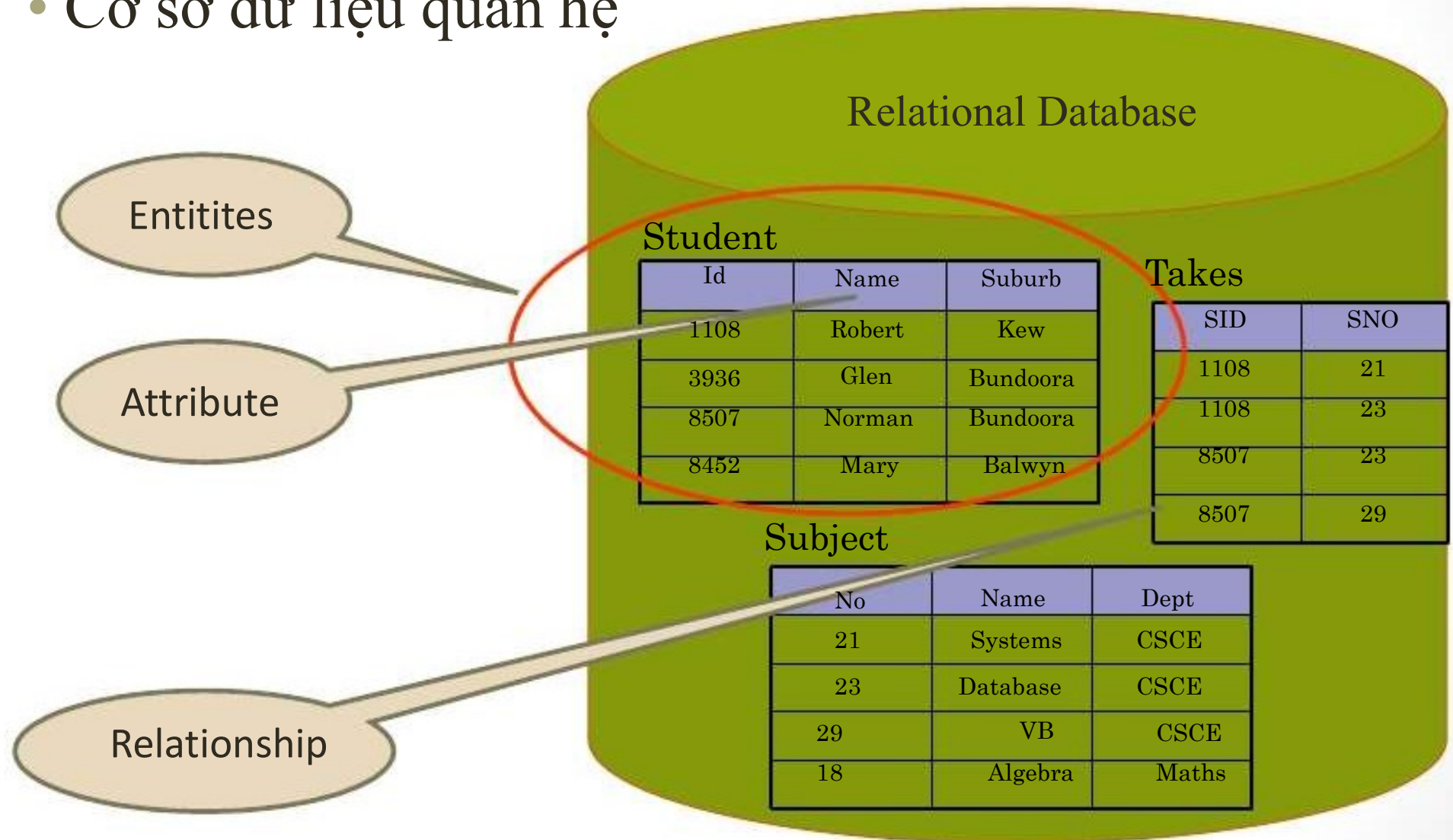


CƠ SỞ DỮ LIỆU SQL

MÔ HÌNH DỮ LIỆU QUAN HỆ

Mô hình dữ liệu quan hệ

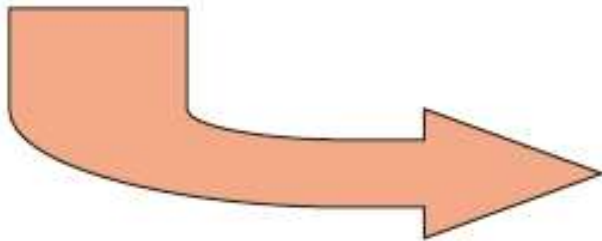
- Cơ sở dữ liệu quan hệ



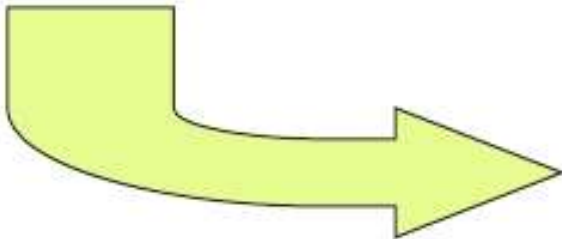
Mô hình dữ liệu quan hệ

- Các khái niệm:

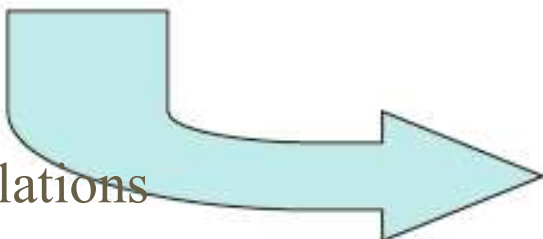
Entities



Attributes



Relations



Relationship

Mô hình dữ liệu quan hệ - Các khái niệm

- Thuộc tính:
 - đặc trưng của đối tượng.
 - được phân biệt qua tên gọi
 - thuộc một kdl nhất định (số, chuỗi, ngày tháng, lôgic, hình ảnh...).
 - mỗi thuộc tính chỉ chọn lấy giá trị trong một tập con của kiểu dữ liệu và tập hợp đó gọi là miền giá trị của thuộc tính đó.
 - Ví dụ điểm thi của sinh viên chỉ là các số nguyên từ 0 đến 10.

Mô hình dữ liệu quan hệ - Các khái niệm (tt)

- Lược đồ quan hệ
 - tập tất cả các thuộc tính cần quản lý của một đối tượng cùng với những mối liên hệ giữa chúng.
 - Ví dụ: ta có LĐQH sinh viên (đặt tên là SV) với các thuộc tính như sau:
SV(maSoSV, hoTenSV, ngaySinh, diemTB, mucHbg)

Mô hình dữ liệu quan hệ - Các khái niệm (tt)

- Nhiều lược đồ quan hệ cùng nằm trong một hệ thống quản lý được gọi là một lược đồ CSDL.
 - Ví dụ lược đồ CSDL để quản lý điểm của sinh viên có thể gồm những lược đồ quan hệ sau:

SV(maSV, hotenSV, ngaySinh, maLop, tinh, hocBong)

Lop(maLop, tenLop, siSo, maKhoa)

Khoa(maKhoa, tenKhoa, soCB)

MonHoc(maMH, tenMH, soTiet)

KetQua(maSV, maMH, diemThi)

Mô hình dữ liệu quan hệ - Các khái niệm (tt)

- Định nghĩa quan hệ (relation):
 - Sự thể hiện của lược đồ quan hệ ở một thời điểm nào đó được gọi là quan hệ.

Mô hình dữ liệu quan hệ - Các khái niệm (tt)

- *Primary keys (các khóa chính)*
 - là yếu tố chủ yếu cho các CSDL QH
 - làm cho các dòng là duy nhất

MAMONHOC	TENMONHOC	SODVHT
HO-001	Hoá đại cương	3
TI-001	Tin học đại cương	4
TI-002	Ngôn ngữ C	5
TI-003	Lý thuyết hệ điều hành	4
TI-004	Cấu trúc dữ liệu và giải thuật	4
TO-001	Đại số tuyến tính	4
TO-002	Giải tích 1	4
TO-003	Bài tập Đại số	2
TO-004	Bài tập Giải tích 1	2
VL-001	Vật lý đại cương	3

Mô hình dữ liệu quan hệ - Các khái niệm (tt)

- *Foreign keys (các khóa ngoại)*
 - là một hoặc nhiều cột tham chiếu đến các khóa chính của bảng khác.

MAKHOA	TENKHOA	DIENTHOAI
DHT01	Khoa Toán cơ - Tin học	054822407
DHT02	Khoa Công nghệ thông tin	054826767
DHT03	Khoa Vật lý	054823462
...

MALOP	TENLOP	KHOA	HEDAOTAO	NAMNHAPHOC	SISO	MAKHOA
C24101	Toán K24	24	Chính quy	2000	5	DHT01
C25101	Toán K25	25	Chính quy	2001	5	DHT01
C25102	Tin K25	25	Chính quy	2001	6	DHT02
C24102	Tin K24	24	Chính quy	2000	8	DHT03
...

Thiết kế các CSDL quan hệ

- Các bước:
 - Phân tích tình huống để tập hợp thông tin về csdl dự kiến
 - Quyết định các cột, kiểu dữ liệu và chiều dài dữ liệu
 - Chuẩn hóa dữ liệu bên trong các bảng
 - Tạo csdl và các bảng

Thiết kế các CSDL quan hệ (tt)

- Xét ví dụ sau:

CungCapSP(maNCC, tenNCC, diaChi, sanPham, gia)

maNCC	tenNCC	diaChi	sanPham	gia
S01	Sanyo	Tokyo, Nhật Bản	Tủ lạnh	250
S01	Sanyo	Tokyo, Nhật Bản	Máy giặt	270
S01	Sanyo	Tokyo, Nhật Bản	Máy lạnh	200
S02	Sharp	Hiroshima, Nhật Bản	Tivi	140
S03	Samsung	Hồng Kông, Đài Loan	LCD	200
S03	Samsung	Hồng Kông, Đài Loan	Tivi	120
M01	Maytag	New York, Mỹ	Tủ lạnh	260

- Sự trùng lặp thông tin sẽ làm tăng chi phí lưu trữ thông tin cũng như chi phí kiểm tra RBTV.
- Cách giải quyết: tách thành hai lược đồ như sau:
 - CungCap(maNCC, sanPham, gia) và NhaCC(maNCC, tenNCC, diaChi)

Hỏi: Khóa chính là gì?

Đáp: {maNV, tenKH}

maNV	tenNV	donVi	luong	tenKH	ngayKTKH
100	Nguyễn Văn An	P. Tiếp thị	480	Anh văn	31/12/2009
100	Nguyễn Văn An	P. Tiếp thị	480	NC thị trường	10/03/2010
140	Trần Thị Mỹ	P. Kế toán	550	Luật thuế	15/04/2010
110	Lê Văn Minh	P. Hệ thống thông tin	500	Anh văn	31/12/2009
110	Lê Văn Minh	P. Hệ thống thông tin	500	C++	31/01/2010
190	Trần Văn Ban	P. Tài chính	520		
150	Hồ Minh Nhật	P. Tiếp thị	450	Anh văn	31/12/2009
150	Hồ Minh Nhật	P. Tiếp thị	450	Java	10/01/2010

Sự bất thường ở bảng này?

- Thêm vào – không thể thêm vào 1 nhân viên mới mà không tham gia lớp nào
- Xoá đi – nếu xoá bỏ nhân viên 140, ta sẽ mất thông tin về lớp *Luật thuế*
- Hiệu chỉnh – để tăng lương cho nhân viên 100, đòi hỏi phải cập nhật 2 hàng

Thiết kế các CSDL quan hệ - Chuẩn hóa

- Có các dạng chuẩn sau:
 - Dạng chuẩn 1 (First Normal Form) – 1NF
 - Dạng chuẩn 2 (Second Normal Form) – 2NF
 - Dạng chuẩn 3 (Third Normal Form) – 3NF
 - Dạng chuẩn Boyce codd (Boyce codd Normal Form) – BCNF

Thiết kế các CSDL quan hệ - Chuẩn hóa (tt)

- Dạng chuẩn 1 (First Normal Form) – 1NF
- Quy định rằng một cột không thể chứa nhiều giá trị
- Dạng chuẩn 2 (First Normal Form) – 2NF)
- Đạt 1NF
- Mọi cột không tham gia vào khóa phải phụ thuộc vào toàn bộ khóa
- Dạng chuẩn 3 (Third Normal Form) – 3NF)
- Đạt 2NF
- Mọi cột không tham gia vào khóa không được phụ thuộc vào cột khác (không phải khóa)

Hệ quản trị CSDL – khái niệm

- Hệ quản trị cơ sở dữ liệu (HQTCSDL, database management system):
 - Là một tập hợp các ctr cho phép người dùng định nghĩa, tạo lập, bảo trì các CSDL và cung cấp các truy cập có điều khiển đến các CSDL này.
 - CSDL là một thành phần trong HQTCSDL.
 - Các HQTCSDL thường gặp như: Oracle, SQL Server, DB2, Paradox, MS Access, Sybase, Foxpro....

Hệ quản trị CSDL – lịch sử phát triển

- Các hệ quản trị CSDL đầu tiên ra đời vào những năm 1960 dựa trên mô hình dữ liệu phân cấp.
 - IMS (Information Management System) của hãng IBM.
- Giữa những năm 60 của thế kỷ 20, IDS (Integrated Data Store) ra đời đánh dấu sự xuất hiện đầu tiên của một hệ quản trị CSDL dựa trên mô hình dữ liệu mạng.
- Năm 1976, hệ quản trị CSDL đầu tiên dựa vào mô hình quan hệ của hãng IBM ra đời mang tên System-R.

Hệ quản trị CSDL – lịch sử phát triển (tt)

- Từ 1980 đến nay:
 - hệ quản trị CSDL kiểu QH ra đời cho cả môi trường máy tính lớn và máy tính cá nhân.
- Từ 1990, bắt đầu xây dựng các hệ quản trị CSDL kiểu hướng đối tượng và hướng đối tượng – quan hệ. Những hệ quản trị CSDL như vậy được coi là thuộc vào thế hệ thứ ba.

Hệ quản trị CSDL – chức năng

- Các chức năng của hệ quản trị csdl:
 - Cung cấp mô hình dữ liệu giúp user mô tả dữ liệu, cho phép khai báo cấu trúc dữ liệu, mô tả các mối liên hệ của dữ liệu cũng như các quy tắc quản lý áp đặt lên dữ liệu đó.
 - Cung cấp một giao diện giữa user và csdl và csdl với các hệ thống khác.
 - Cung cấp một ngôn ngữ bậc cao (phi thủ tục) giúp user truy xuất và thao tác csdl

Hệ quản trị CSDL – chức năng (tt)

- Quản lý giao tác, phân quyền và an toàn dữ liệu
- Điều khiển được sự tương hợp, tính toàn vẹn khi chuyển hóa dữ liệu và có sự cố của hệ thống
- Kiểm tra độ tin cậy của dữ liệu

Hệ quản trị CSDL

Để giải quyết tốt các tình huống mà CSDL đặt ra như bảo mật, cạnh tranh trong truy xuất,...cần phải có phần mềm chuyên dụng là hệ quản trị CSDL

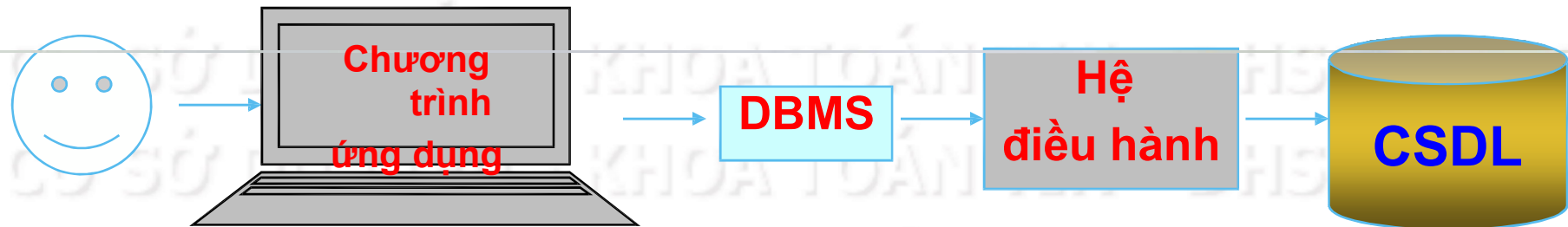
Hệ quản trị CSDL là một chương trình cho phép xây dựng và quản lý dữ liệu. Ví dụ: SQL Server, Oracle, Sybase, DB2,..

Hệ quản trị CSDL

Hệ quản trị CSDL phải có phương tiện cho phép thực hiện: định nghĩa CSDL, lưu trữ dữ liệu, thao tác trên dữ liệu

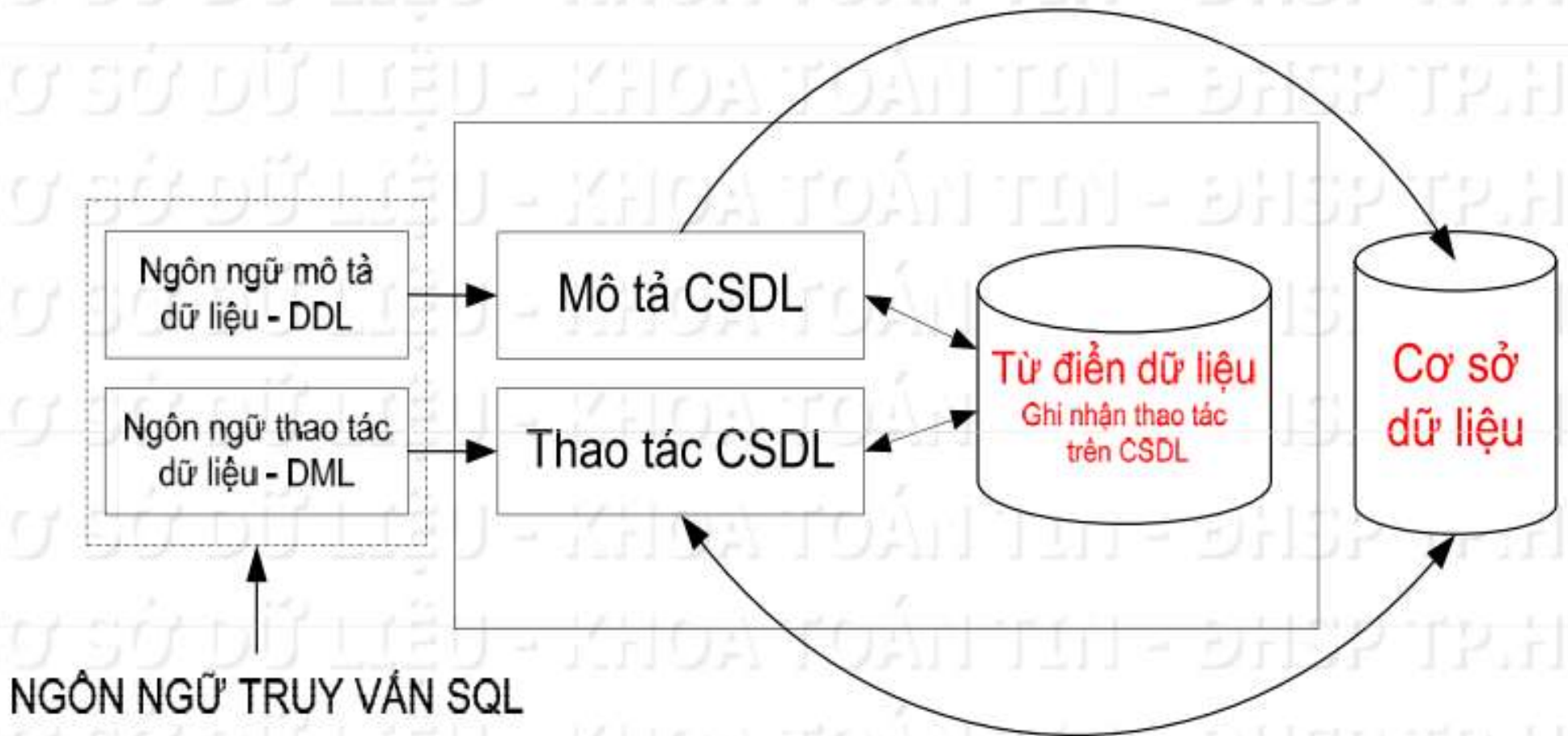
Định nghĩa CSDL: cho phép định nghĩa cấu trúc CSDL, xác định kiểu dữ liệu, các ràng buộc toàn vẹn

Thao tác trên dữ liệu: cho phép truy vấn dữ liệu, thêm, xóa, sửa dữ liệu

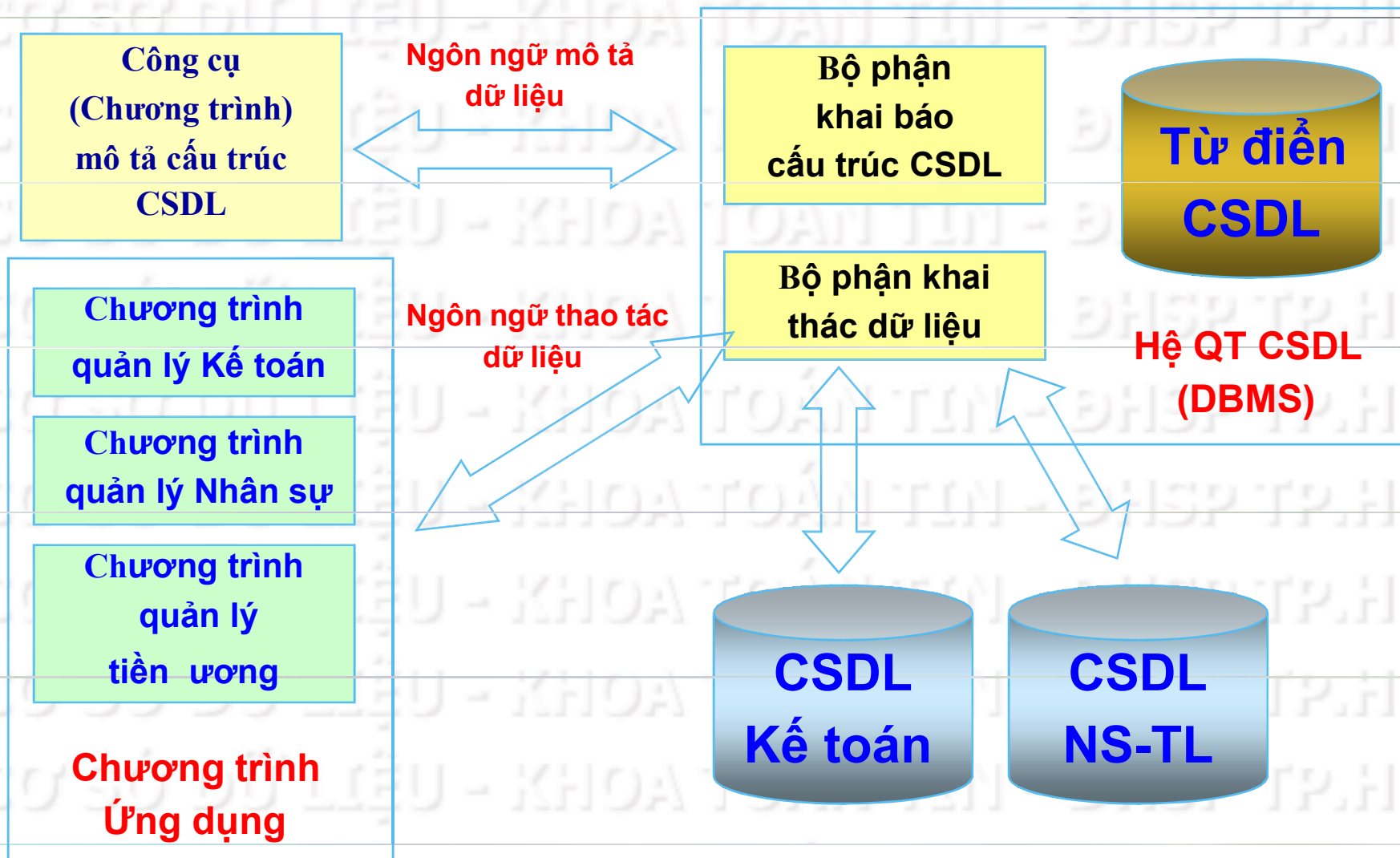


Kiến trúc của một DBMS

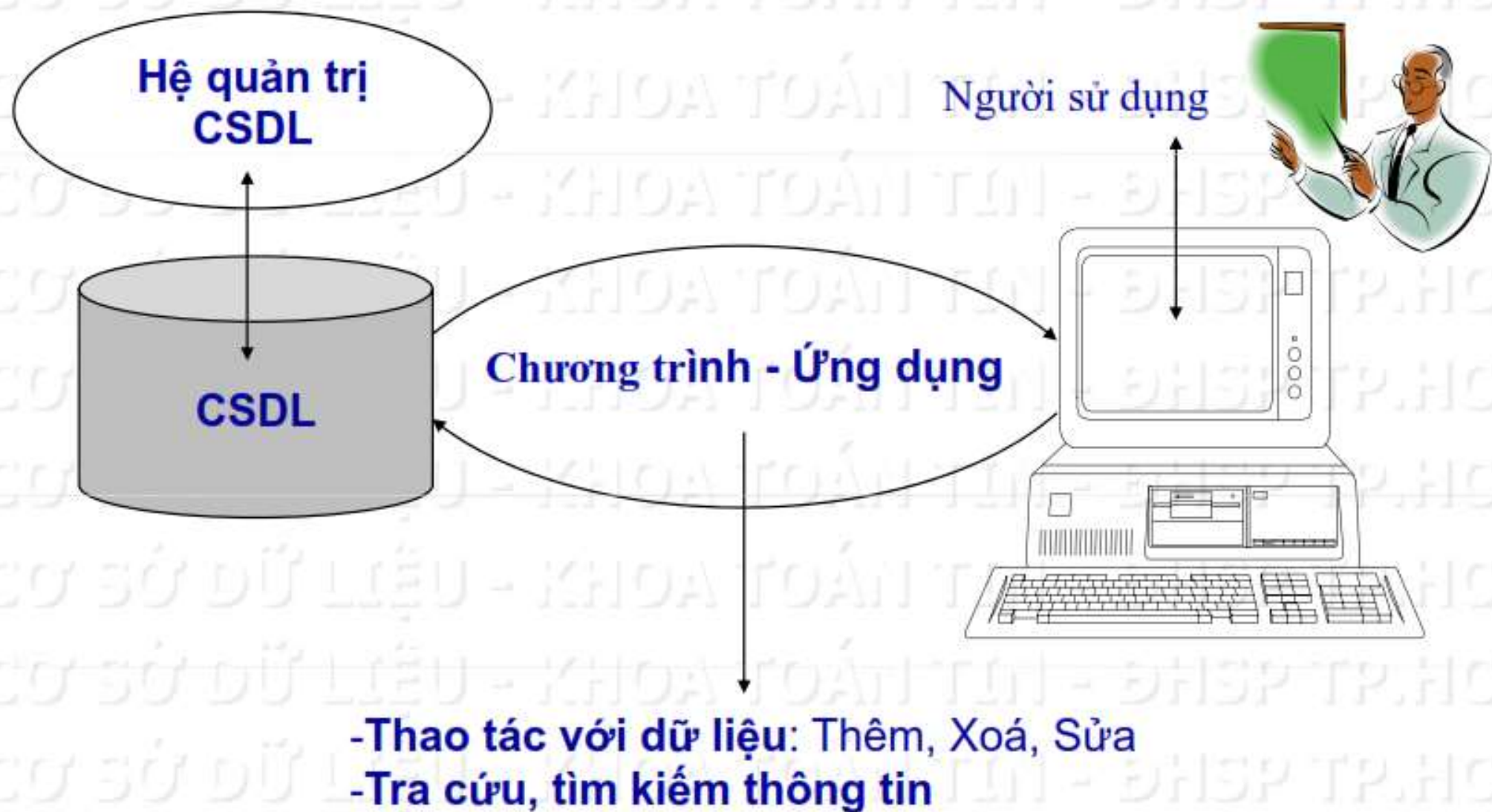
Hệ quản trị CSDL



Sơ đồ tổng quát của một DBMS



Quan hệ giữa CSDL với DBMS và với ứng dụng



CƠ SỞ DỮ LIỆU SQL SERVER

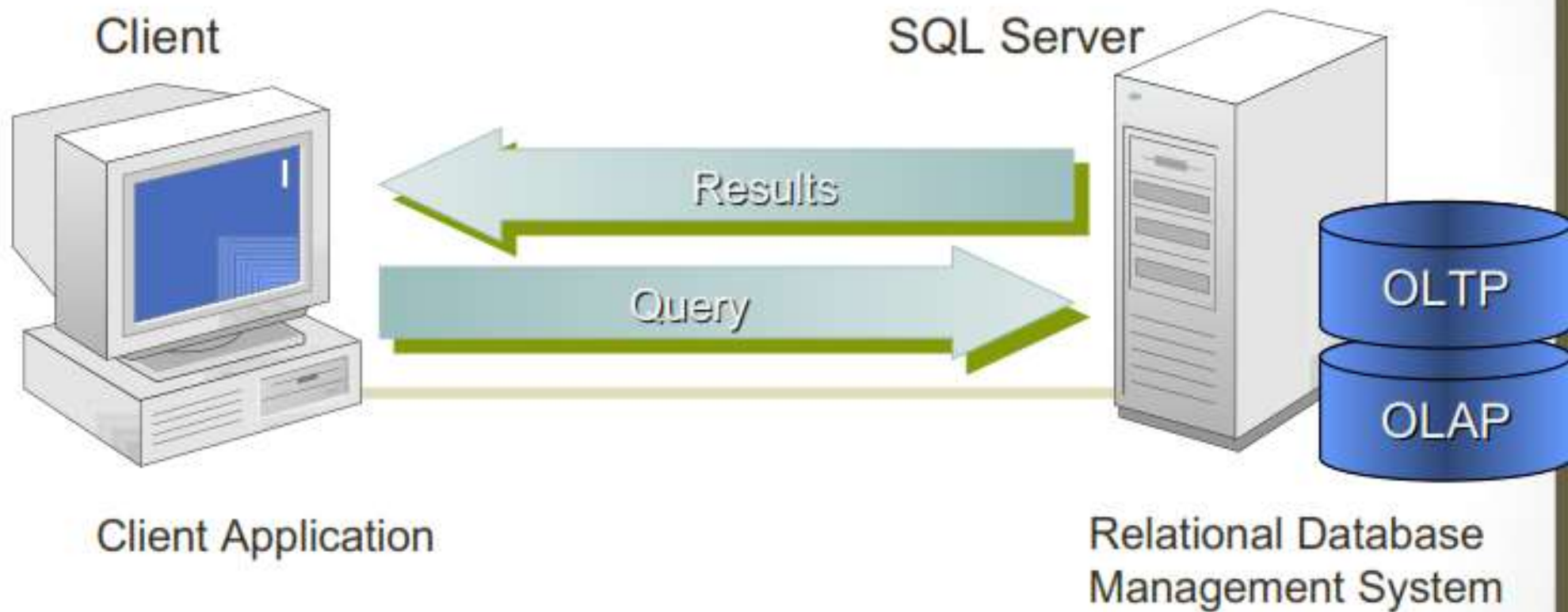


TỔNG QUAN VỀ SQL SERVER

Tổng quan về SQL Server

- Giới thiệu SQL Server
- Lịch sử phát triển SQL Server
- Các ấn bản SQL Server
- Một số khái niệm của SQL Server
- Các mô hình truy xuất dữ liệu
- Transact-SQL

Giới thiệu SQL Server



Giới thiệu SQL Server

- SQL Server:
 - là một hệ thống quản lý cơ sở dữ liệu (Relational Database Management System (RDBMS)) sử dụng Transact-SQL để trao đổi dữ liệu giữa Client computer và SQL Server computer.
 - được tối ưu để có thể chạy trên môi trường cơ sở dữ liệu rất lớn (Very Large Database Environment) lên đến Tera-Byte và có thể phục vụ cùng lúc cho hàng ngàn user.
 - có thể kết hợp "ăn ý" với các server khác như Microsoft Internet Information Server (IIS), E-Commerce Server....

Lịch sử phát triển SQL

- Phiên bản gốc của SQL, ban đầu được gọi là SEQUEL (Structured English Query Languages),
 - được thiết kế và cài đặt vào năm 1970 tại phòng nghiên cứu San Jose của hãng IBM.
- Qua thời gian, ngôn ngữ được bổ sung để có thể xây dựng CSDL và quản lý an toàn cho cơ cấu CSDL.
- Các phương tiện CSDL ngày nay sử dụng các phiên bản khác nhau của SQL. Microsoft SQL Server sử dụng phiên bản gọi là Transact-SQL

Lịch sử phát triển SQL (tt)

- SQL Server đầu tiên là của Microsoft và tập đoàn Sybase trên nền IBM OS/2
- SQL Server 7.0 là cái mốc để đánh dấu sự chấm dứt sử dụng mã lệnh Sybase cho nhóm cơ sở dữ liệu của Microsoft.

Lịch sử phát triển SQL (tt)

- SQL Server 2000:
 - làm nổi bật dòng mã lệnh mới, bổ sung các tính năng mới có ý nghĩa.
 - tăng cường khả năng đánh giá, độ tin cậy và giá trị của sản phẩm.

Lịch sử phát triển SQL (tt)

- SQL server 2005
 - Nâng cao bảo mật
 - Mở rộng T-SQL
 - Tăng cường hỗ trợ người phát triển
 - Tăng cường khả năng quản lý
 - Nâng cao độ sẵn sàng của CSDL
 - Nâng cao khả năng mở rộng CSDL
 - Tăng cường khả năng khai thác thông tin

Lịch sử phát triển SQL (tt)

- SQL server 2008
 - Nền tảng cho các nhiệm vụ then chốt
 - Sự phát triển động
 - Dữ liệu quan hệ mở rộng
 - Thông tin trong toàn bộ doanh nghiệp

Lịch sử phát triển SQL (tt)

- SQL Server 2017

Tính năng mới trong Database Engine

identity_cache

Cải thiện xử lý truy vấn thích ứng

Automatic Tuning

Tính năng mới trong Graph DB

Graph DB là gì?

Always Available (truy cập dữ liệu chéo)

Cải tiến về DTA

Hàm chuỗi (string) mới

TRANSLATE

CONCATE_WS

TRIM

STRING_AGG

Lịch sử phát triển SQL (tt)

- SQL Server 2019
 - [Big data clusters](#)
 - [Database engine](#)
 - Accelerated database recovery
 - Reduced recompilations for workloads using temporary tables across multiple scopes
 - Query Store plan forcing support for fast forward and static cursors...
 - [SQL Server Analysis Services \(SSAS\)](#): Calculation groups in tabular models reduce the number of measures by reusing calculation logic.

Các ấn bản của SQL Server (Editions)

- Enterprise (Developer) Edition
- Standard Edition
- SQL Server Express Edition

Một số khái niệm của SQL Server

- Data Type (Kiểu dữ liệu):
 - là các kiểu lưu trữ cơ bản về dữ liệu của bạn. Chúng ta có thể chọn ra các KDL khác nhau, chẳng hạn như ký tự, số hoặc ngày.
- Stored procedure (thủ tục thường trú – thủ tục lưu trữ):
 - mã lệnh Transact – SQL có thể được viết và lưu trữ mang một tên.

Một số khái niệm của SQL Server (tt)

- User – defined functions (các hàm do người dùng định nghĩa):
 - là mã lệnh Transact – SQL rất giống các mã lệnh của thủ tục lưu trữ.
 - có thể được gọi trong các truy vấn CSDL hoặc để chỉnh sửa một cột dữ liệu hoặc đóng vai trò như các bảng
 - Vd có thể viết các hàm date cho riêng mình để chỉnh sửa các cột dữ liệu có kiểu datetime.

Một số khái niệm của SQL Server (tt)

- Constraints (các ràng buộc)
 - là các cơ chế được hệ thống cài đặt dựa trên máy chủ nhằm bảo vệ tính toàn vẹn của dữ liệu.

Một số khái niệm của SQL Server (tt)

- Triggers (các bẫy lỗi – còn dịch là bộ kích khởi)
 - là các thủ tục lưu trữ kích hoạt trước hoặc sau khi bổ sung, sửa chữa hoặc xóa dữ liệu ra khỏi CSDL.
 - đảm bảo tuân theo các quy tắc kinh doanh hoặc các quy tắc toàn vẹn dữ liệu khác trong CSDL.
 - Ví dụ bộ kích khởi có thể đảm bảo rằng mỗi quyển sách trong hiệu sách có một nhà xuất bản hợp lệ gán cho nó.

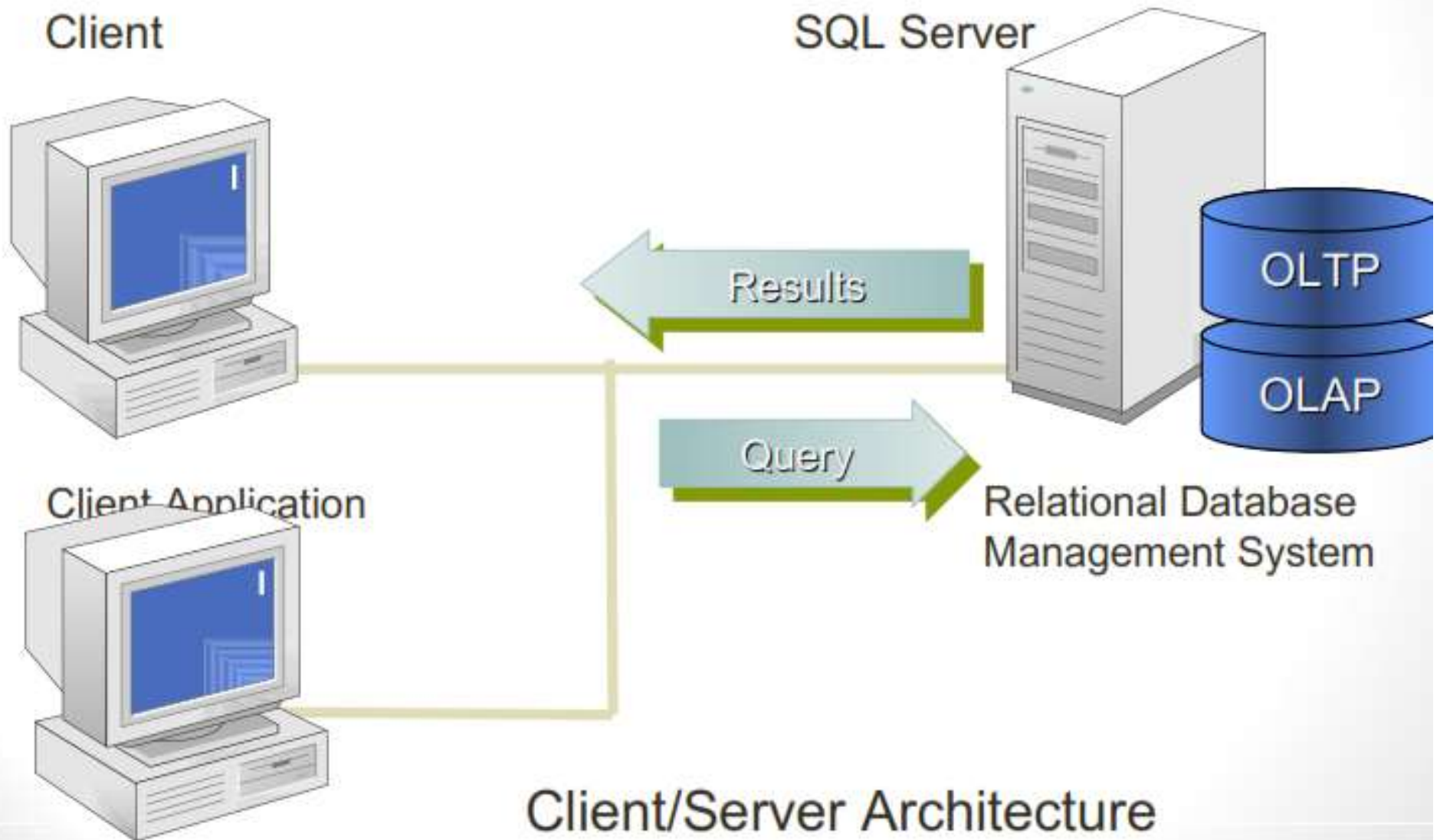
Một số khái niệm của SQL Server (tt)

- Views (các khung nhìn – hay bảng ảo)
 - là các truy vấn lưu trữ trong CSDL để có thể tham chiếu tới một hoặc nhiều bảng.
 - thường thể hiện hoặc không thể hiện các cột nào đó trong bảng hoặc là liên kết hai hoặc nhiều bảng với nhau
- Indexes (các chỉ mục)
 - giúp tổ chức lại dữ liệu, nên các truy vấn chạy nhanh hơn.

Các mô hình truy xuất dữ liệu

- Kiến trúc Client/ Server (Client / Server architecture)
 - Phía máy chủ: thường được cài đặt các hệ cơ sở dữ liệu bao gồm các bảng biểu, các thủ tục lưu trữ... và điều khiển các tiến trình sau:
 - Quản lý dữ liệu.
 - Bảo mật dữ liệu.
 - Thực hiện truy vấn, ràng buộc và các thủ tục lưu trữ.
 - Điều khiển lỗi.
 - Phía máy khách: Tạo giao diện người sử dụng (User Interface)

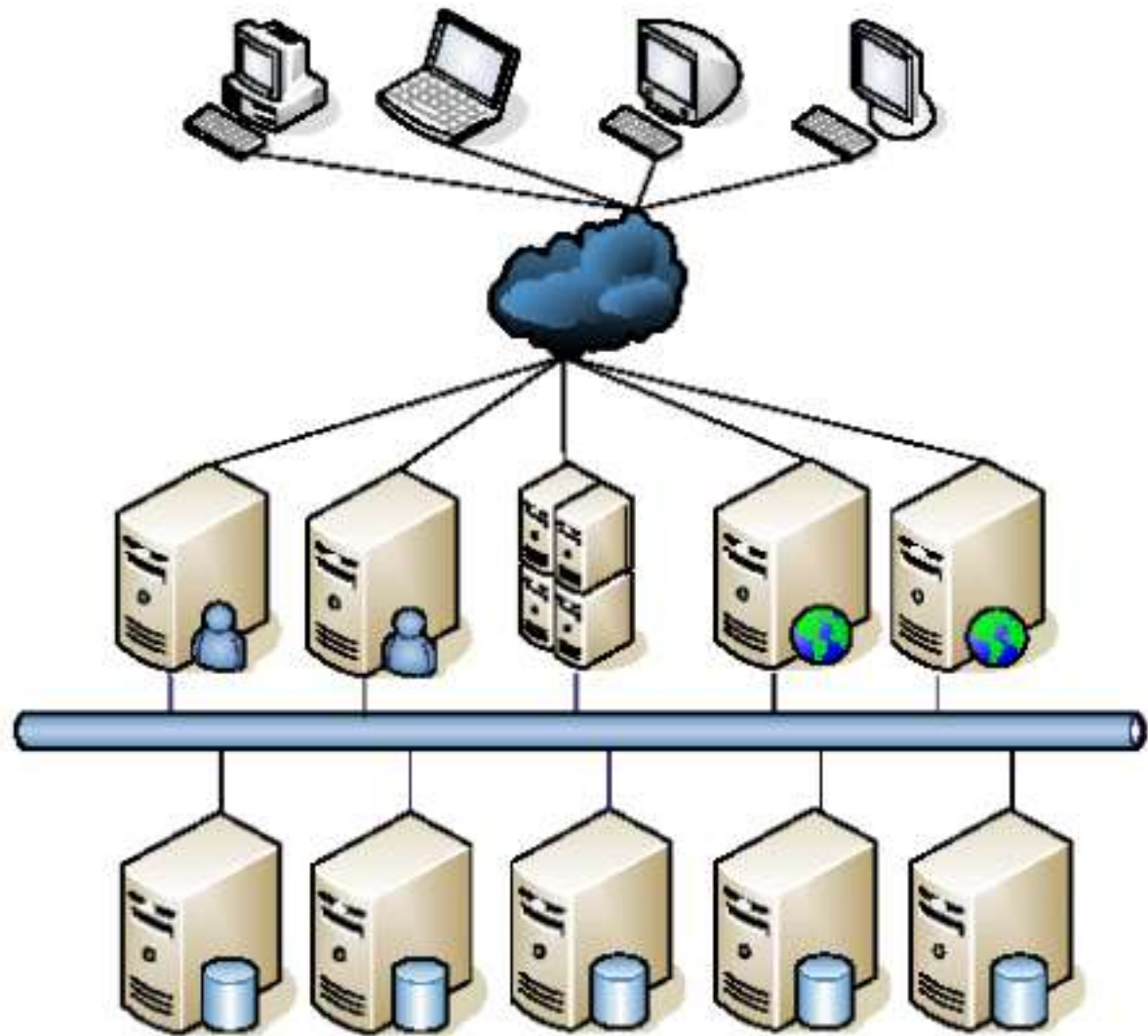
Các mô hình truy xuất dữ liệu (tt)



Các mô hình truy xuất dữ liệu (tt)

- Kiến trúc nhiều tầng (N-tier Architecture)
 - Quá trình xử lý được phân tán trên các lớp khác nhau với các chức năng riêng biệt → phù hợp cho việc tổ chức hệ thống thông tin trên mạng Internet và mạng Intranet hay trong các mạng cục bộ.
 - Các tầng (lớp): nguồn dữ liệu, truy cập dữ liệu, lớp giao dịch, lớp giao diện

Các mô hình truy xuất dữ liệu (tt)



N-tier Architecture

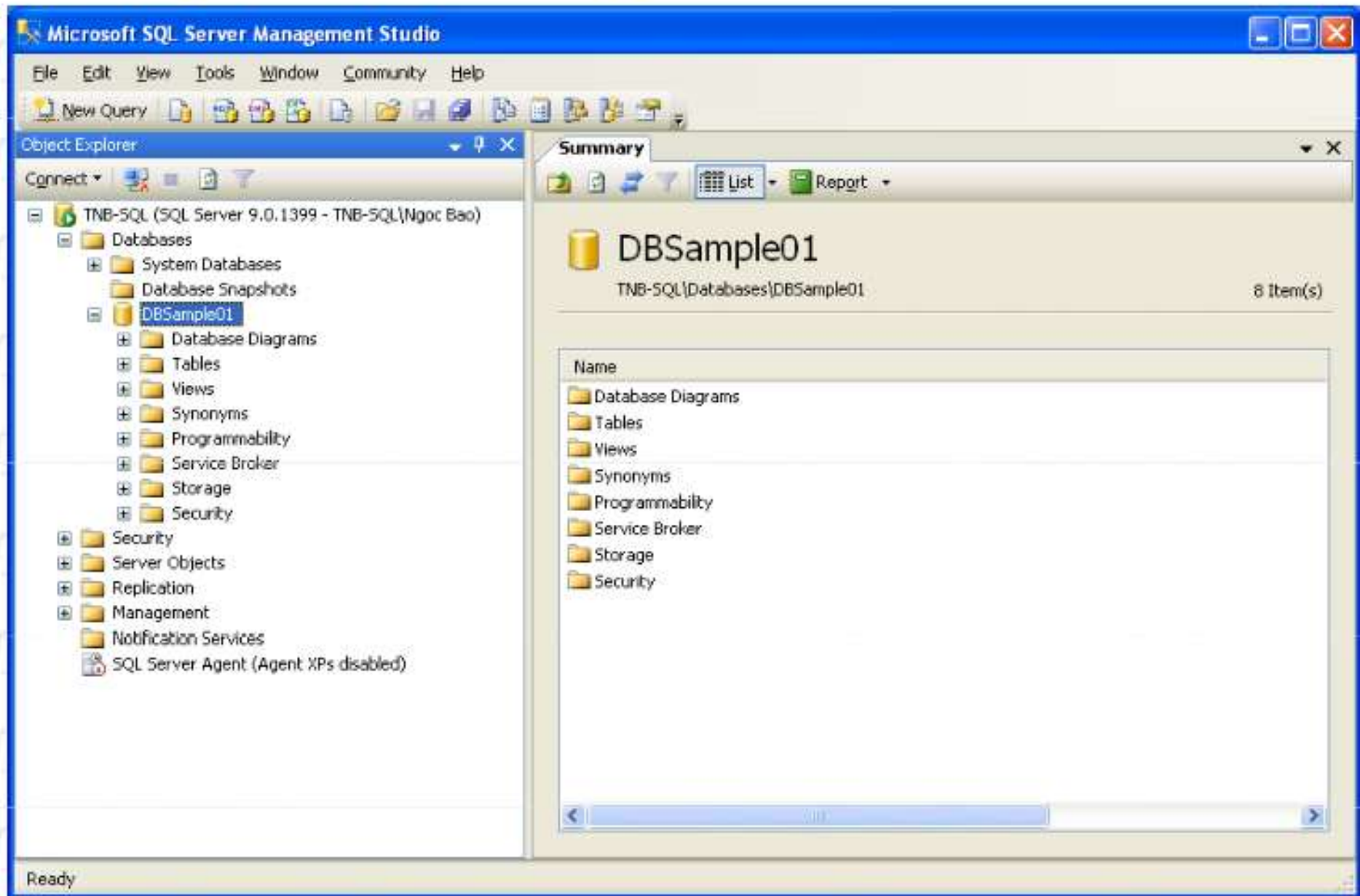
Transact - SQL (T-SQL)

- **Transact - SQL (T-SQL):**

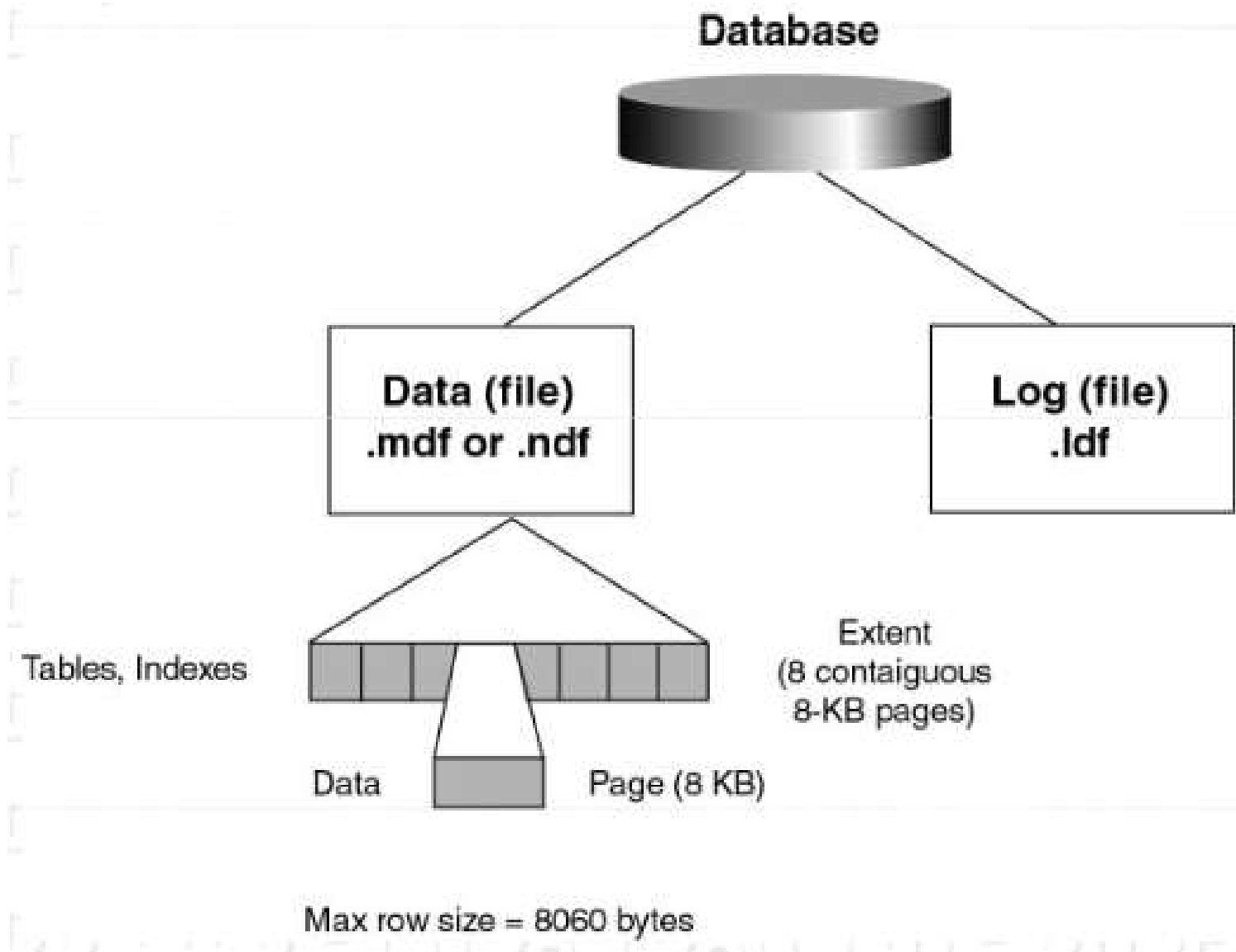
- Là ngôn ngữ SQL mở rộng dựa trên chuẩn ANSI
- T-SQL gồm các nhóm:
 - Data Definition Language (DDL): ngôn ngữ định nghĩa dữ liệu
 - CREATE, ALTER, DROP
 - Data Manipulation Language (DML): ngôn ngữ xử lý dữ liệu
 - SELECT, INSERT, UPDATE
 - Data Control Language Statements (DCL): ngôn ngữ điều khiển dữ liệu
 - DENY, GRANT, REVOKE

Công cụ SQL Server

- SQL Server Management Studio



Các thành phần cơ bản trong một CSDL trên SQL Server



Thông thường, các tập tin này được lưu tại địa chỉ: **C:\Program Files\Microsoft SQL Server\MSSQL12.MSSQLSERVER\MSSQL\DATA**

Xem chương tiếp theo:
Tạo và quản trị CSDL

Bài tập

- SV tự cài đặt Hệ quản trị CSDL SQL Server trên máy tính cá nhân, phiên bản phù hợp với cấu hình của máy tính.
- *Lưu ý:* Khi cài đặt xong phải kiểm tra có đủ công cụ **SQL Server Management Studio**.