

Trường Đại học Khoa học Tự nhiên

Khoa Công nghệ Thông tin

Bộ môn Hệ thống Thông tin

Quản trị Cơ sở Dữ liệu

Chương I: **TỔNG QUAN**

GV : TUẤN NGUYỄN HOÀI ĐỨC

E-mail : tnhduc@fit.hcmus.edu.vn

Nội dung trình bày

- ◆ Lược sử CSDL quan hệ
- ◆ Yêu cầu về dữ liệu trong CSDL
- ◆ Hệ quản trị CSDL
- ◆ Giới thiệu SQL Server

Lược sử CSDL quan hệ

Các loại CSDL trước CSDL quan hệ :

- ❖ Paper-based

- *Hồ sơ giấy được lưu trữ có hệ thống*

- ❖ Legacy mainframe

- *Dùng trong hệ thống IBM, chạy bằng các máy VSAM (Virtual Storage Access Method)*

- ❖ dBase

- *Dữ liệu lưu trữ thành tập tin, bắt đầu có tổ chức thành các table nhưng không quản lý được lượng dữ liệu kích thước lớn và không hỗ trợ đa người dùng.*

Lược sử CSDL quan hệ (tt)

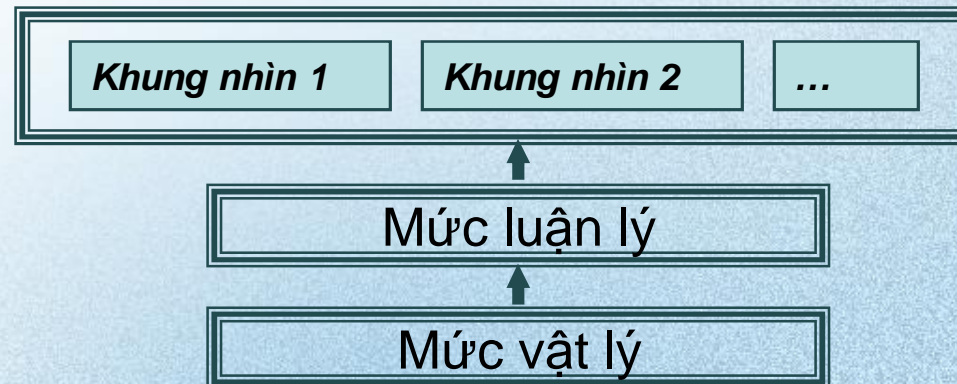
- ◆ Từ trước 1960, quy tắc cấu trúc dữ liệu thành các quan hệ và ngôn ngữ SQL đã được IBM giới thiệu
- ◆ Khắc phục được nhiều vấn đề của các mô hình CSDL trước đó
 - ***Giảm chi phí lưu trữ (tránh trùng lặp)***
 - ***Hỗ trợ đa người dùng***
 - ***Cải thiện hiệu quả khai thác (chỉ mục, tối ưu truy vấn...)***
- ◆ Tuy nhiên chưa đưa vào thực tế
- ◆ Đến thập niên 70, Oracle và Sybase tiên phong xây dựng hệ thống CSDL quan hệ thực sự đầu tiên
- ◆ Đến thập niên 80, viện tiêu chuẩn hợp chủng quốc Hoa Kỳ đưa ra những quy định chuẩn hoá chung cho SQL. Nhờ đó các công ty có thể phát triển các CSDL theo một chuẩn thống nhất và có khả năng giao tiếp với nhau.

Nội dung trình bày

- ◆ Lược sử CSDL quan hệ
- ◆ Yêu cầu về dữ liệu trong CSDL
- ◆ Hệ quản trị CSDL
- ◆ Giới thiệu SQL Server

Yêu cầu về Dữ liệu trong CSDL

- ◆ Dữ liệu trong CSDL cần thể hiện qua 3 cấp độ sau
 - **Mức quan niệm (Conceptual Level)**
 - Mô tả 1 phần của CSDL mà 1 đối tượng người dùng được quyền tiếp cận
 - **Mức luận lý (Logical level)**
 - Những thông tin gì được lưu trữ trong CSDL
 - Có những mối quan hệ nào giữa những thông tin đó
 - **Mức vật lý (Physical Level)**
 - Dữ liệu được lưu trữ như thế nào trên thiết bị lưu trữ



Yêu cầu về Dữ liệu trong CSDL (tt)

- ◆ Dữ liệu trong CSDL cần có các đặc trưng sau
 - *Ít hoặc không trùng lặp dữ liệu*
 - *Chia sẻ cho nhiều người sử dụng → Truy xuất đồng thời*
 - *An ninh, bảo mật*
 - *Khôi phục khi có sự cố*
 - *Độc lập*
 - Vật lý (physical)
 - Khi thay đổi tổ chức vật lý của CSDL thì không làm ảnh hưởng đến chương trình đang khai thác CSDL đó
 - Luận lý (logical)
 - Khi thay đổi mức quan niệm không cần thay đổi tổ chức vật lý đang lưu trữ dữ liệu của mức quan niệm ấy
- ◆ Vì vậy cần có một hệ thống quản lý hiệu quả Dữ liệu trong CSDL

Nội dung trình bày

- ◆ Lược sử CSDL quan hệ
- ◆ Yêu cầu về dữ liệu trong CSDL
- ◆ Hệ quản trị CSDL
- ◆ Giới thiệu SQL Server

Khái niệm DBMS

- ◆ Là hệ thống phần mềm cung cấp các công cụ để xây dựng, khai thác và quản lý CSDL
 - **Định nghĩa cấu trúc dữ liệu (DDL – Data Definition Language)**
 - **Cung cấp khả năng thao tác trên CSDL (DML – Data Manipulation Language)**
 - Cập nhật Dữ liệu
 - Truy vấn Dữ liệu
 - **Hỗ trợ lưu trữ dữ liệu**
 - **Điều khiển truy xuất dữ liệu đồng thời**
 - **Bảo mật và an ninh Dữ liệu**
- ◆ DBMS là đối tượng nghiên cứu của môn học

Các thành phần chính của DBMS



Giao diện lập trình

- ◆ DBMS cung cấp giao diện lập trình để sử dụng với một ngôn ngữ lập trình CSDL
 - **Vd : SQL Server cung cấp ngôn ngữ Transaction-SQL (T-SQL)**
- ◆ Ngôn ngữ bao gồm
 - ***Ngôn ngữ định nghĩa dữ liệu : Giúp người dùng ra lệnh cho DBMS 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 dữ liệu : giúp người dùng tích lũy, hiệu chỉnh và khai thác Dữ liệu***

An ninh và Bảo mật

❖ Bộ phận bảo đảm An toàn và Bảo mật thực hiện hai nhiệm vụ

- *Kiểm soát, quản lý tốt những người dùng hợp pháp của hệ thống → Bảo mật Dữ liệu*
- *Ngăn chặn các tấn công của những đối tượng tin tặc → An ninh Dữ liệu*

Khôi phục sau sự cố

- ◆ Đảm bảo tồn thất, sai sót về dữ liệu là ít nhất có thể
- ◆ Nhật ký
 - **Để CSDL được bền vững (durable), mọi thay đổi lên CSDL phải được ghi nhận lại (trong Nhật ký)**
 - **Log manager – ghi chép nhật ký**
 - Đảm bảo ghi nhận đầy đủ và chính xác mọi thay đổi trên CSDL
 - **Recovery manager – khôi phục**
 - Dựa vào nhật ký để phục hồi lại CSDL về trạng thái nhất quán trước đó (Trạng thái thỏa tất cả các RBTV của CSDL đó)

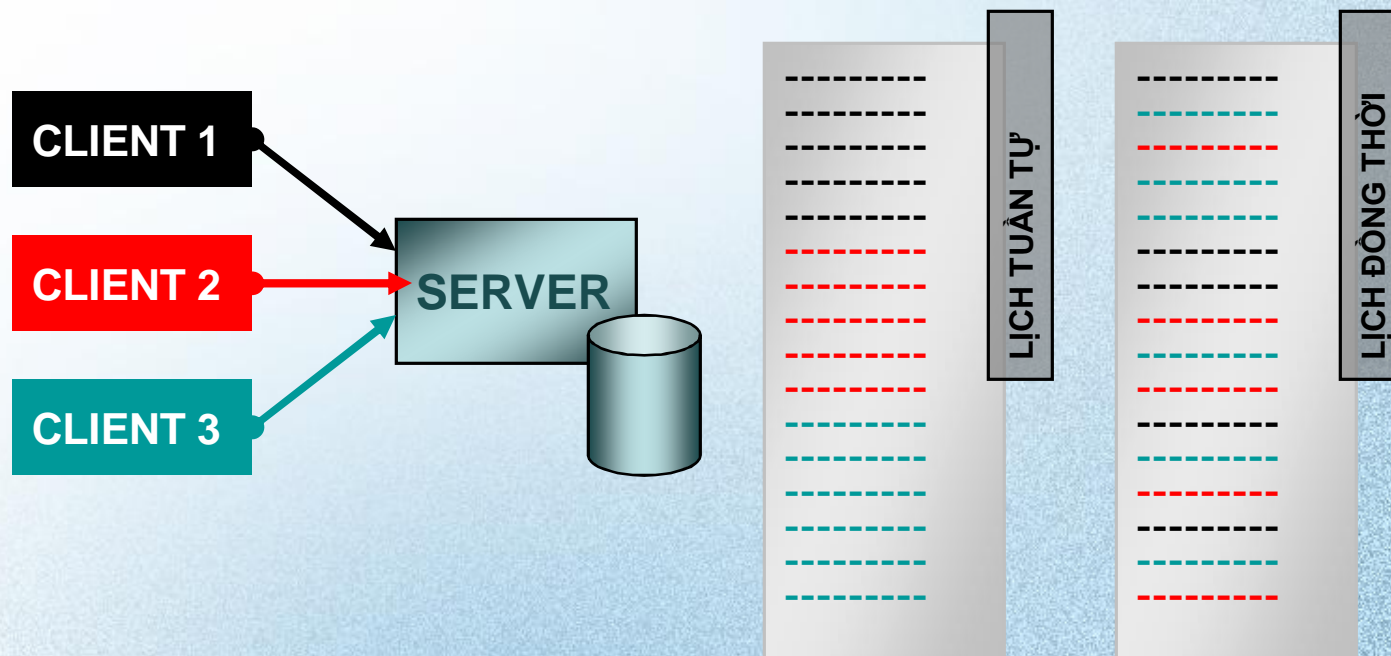
Xử lý truy xuất đồng thời

- ❖ Đối tượng của thành phần quản lý giao tác là các Giao tác
 - ***Giao tác là một nhóm các hành động mà hoặc được thực hiện hết toàn bộ, hoặc không thực hiện bất kỳ hành động nào trong nhóm ấy.***

Xử lý truy xuất đồng thời (tt)

◆ Điều khiển đồng thời

- **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.
- **Cơ chế khóa (lock)** trên 1 đơn vị dữ liệu nào đó → ngăn 2 giao tác cùng thao tác lên 1 đơn vị dữ liệu ấy tại cùng 1 điểm



Xử lý truy xuất đồng thời (tt)

◆ Giải quyết deadlock

- *Vì 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 (khóa 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 trình bày

- ◆ Lược sử CSDL quan hệ
- ◆ Yêu cầu về dữ liệu trong CSDL
- ◆ Hệ quản trị CSDL
- ◆ Giới thiệu SQL Server

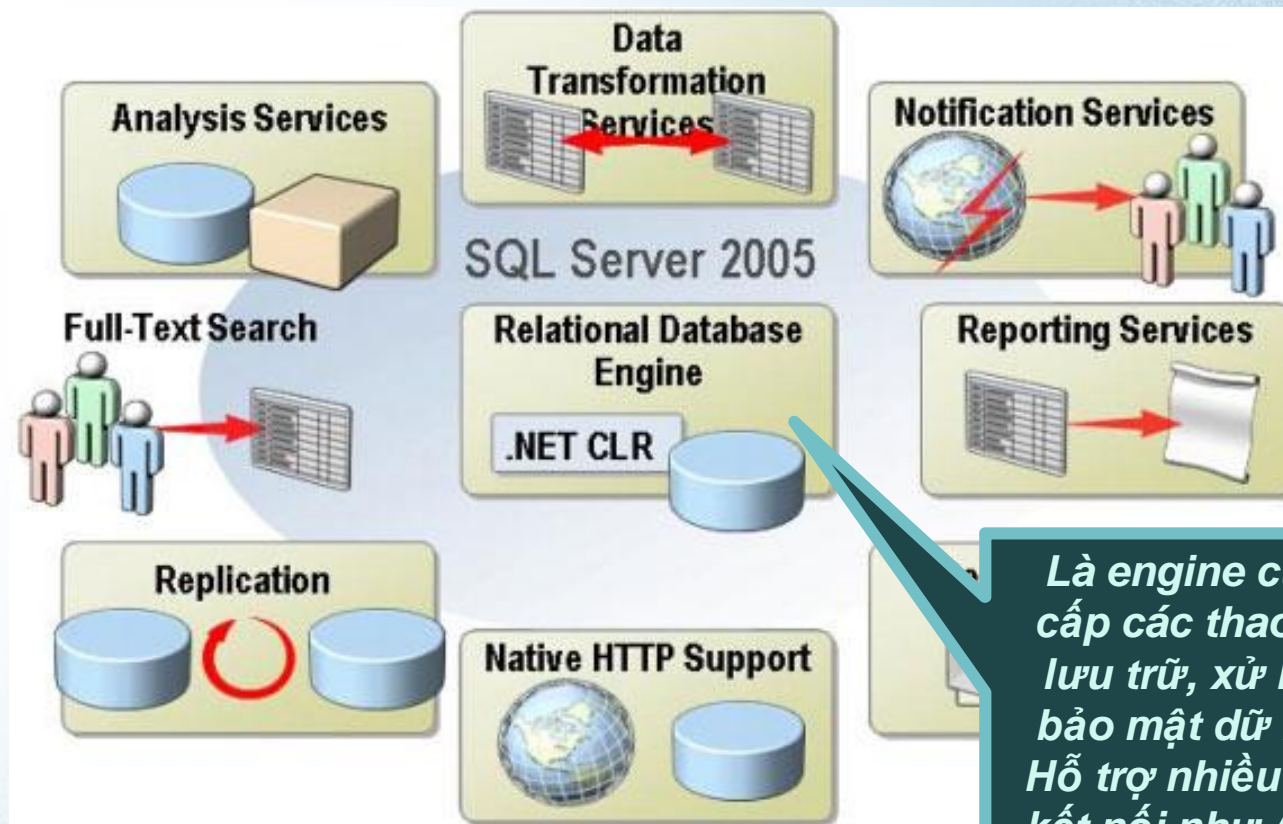
Giới thiệu SQL Server

- ❖ Tiền thân là Sybase SQL Server
- ❖ Năm 1989. Microsoft bắt đầu hợp tác với Sybase để phát triển phiên bản 4.2 (1993)
- ❖ Sau đó các phiên bản 6.5 rồi 7.0 lần lượt ra đời với những đóng góp nổi bật của công nghệ từ phía Microsoft.
- ❖ Khi công nghệ NT của Microsoft ra đời tạo bước nhảy vọt cho HĐH Windows thì MS SQL Server 2000 cũng dần hình thành và không còn sự tham gia của Sybase nữa.
- ❖ MS SQL Server được tung ra cùng với sự xuất hiện của HĐH Windows 2000. Phiên bản MS SQL Server 2000 cuối cùng là MS08-040 ra đời tháng 5/2008.

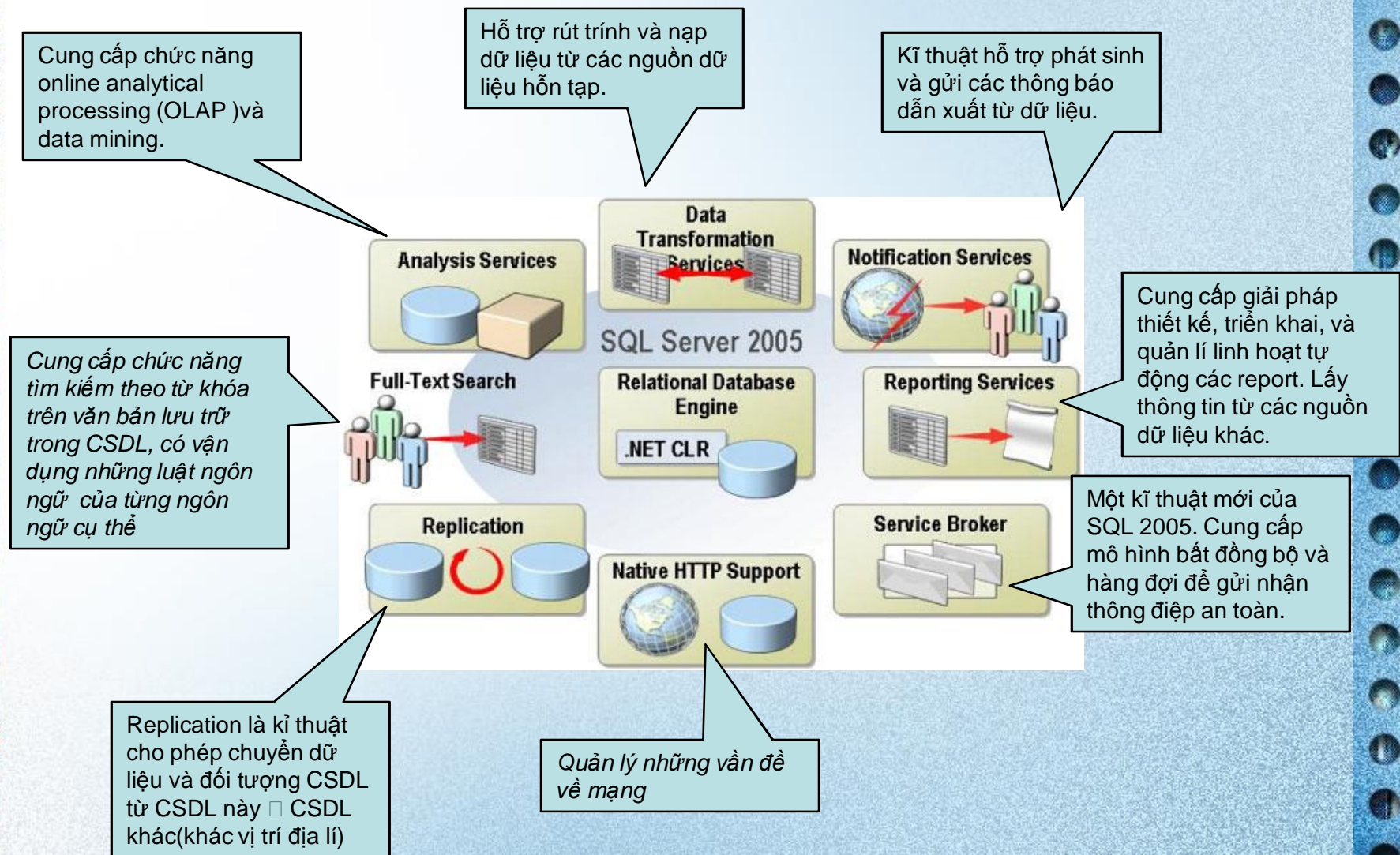
Giới thiệu SQL Server

- ❖ Các phiên bản 2005 (codenamed 'Yukon') lần lượt được tung ra từ tháng 1/2006 (RTM) đến tháng 3/2011 (SP4 CU-3).
 - **CU = Cumulative Update package.**
- ❖ Các phiên bản 2008 (codenamed 'Katmai') lần lượt được tung ra từ tháng 8/2008 (RTM) đến tháng 5/2012 (SP3 CU-5).
- ❖ Các phiên bản 2008 R2 (codenamed 'Kilimanjaro') lần lượt được tung ra từ tháng 4/2010 (RTM) đến tháng 6/2011 (SP1 CU-7).
- ❖ Các phiên bản 2012 (codenamed 'Denali') lần lượt được tung ra từ tháng 8/2011 (CTP 3.0) đến tháng 6/2012 (CU-2).

Các thành phần của SQL Server



Các thành phần của SQL Server



Giới thiệu SQL Server (tt)

- ◆ Các thành phần cấu thành một hệ thống CSDL
 - *Tables, kèm theo constraint*
 - *Diagrams*
 - *Views*
 - *Stored procedure*
 - *User defined function*
 - *Users*
 - *Roles*
 - *Rules*
 - *Defaults*
 - *User defined data types*

Giới thiệu SQL Server (tt)

- ◆ Cơ sở dữ liệu hệ thống (thận trọng khi làm việc với các CSDL này)
 - **Master : Chứa tất cả các bảng dữ liệu của hệ thống, dùng kiểm soát toàn bộ hoạt động của hệ thống → Tối quan trọng**
 - Ví dụ : Các bảng, thủ tục thường trú được lưu trong bảng SysObjects của DB Master.
 - **Model : Chứa các template dùng làm mẫu cho việc tạo CSDL mới → Tối quan trọng**
 - **Msdb : Lưu các lịch trình làm việc định sẵn (lịch backup dữ liệu, lịch thực thi định kỳ các SP...)**

Giới thiệu SQL Server (tt)

- ❖ Cơ sở dữ liệu hệ thống (thận trọng khi làm việc với các CSDL này)
 - *Tempdb : CSDL tạm dùng để tạo thử các DB tham khảo hay thực tập*
 - *Pubs : Chứa tất cả các hướng dẫn, trợ giúp, sách tham khảo của SQL server (bao gồm cả BOL)*
 - *Northwind : CSDL làm ví dụ mẫu để tham khảo, có đầy đủ dữ liệu ví dụ phục vụ các truy vấn và xử lý thử nghiệm trên SQL Server.*

Giới thiệu SQL Server (tt)

◆ Các công cụ chính :

- ***Books online (BOL): Tài liệu tham khảo (help)***
- ***Enterprise Manager (EM): Quản lý SQL Server bằng giao diện đồ họa.***
- ***Query Analyzer (QA): Trình soạn thảo và thực thi các lệnh SQL, các SP, UDF,...***
- ***Service Manager : Dùng để khởi động, ngưng hoặc refresh các dịch vụ***
 - SQL Server Executive : Thực thi lưu trữ và truy vấn dữ liệu
 - SQL Server Agent : Thực thi các công việc đã định sẵn trong các lịch trình định kỳ.

Giới thiệu SQL Server (tt)

- ◆ Các file lưu trữ cho 1 DB
 - *Một và chỉ một file lưu trữ dữ liệu sơ cấp – main data file (*.mdf)*
 - *Một hoặc nhiều file lưu trữ dữ liệu thứ cấp – secondary data file (*.ndf)*
 - *Một hoặc nhiều file nhật ký thao tác – transaction log file (*.ldf)*
- ◆ Một server chấp nhận tối đa 32767 DB.

Hết chương I

