



Hệ thống cơ sở dữ liệu

Bởi:
unknown

Định nghĩa CSDL

Cơ sở dữ liệu là một tập hợp các bảng dữ liệu có quan hệ với nhau sao cho cấu trúc của chúng cũng như các mối quan hệ bên trong giữa chúng là tách biệt với chương trình ứng dụng bên ngoài, đồng thời nhiều người dùng khác nhau cũng như nhiều ứng dụng khác nhau có thể cùng khai thác và cùng chia sẻ một cách có chọn lọc lúc cần.

Vi dụ:

Hệ thống bán vé máy bay của một hãng hàng không:

Mã chuyến bay	Loại máy bay	Sân bay đi	Sân bay đến	Ngày bay	Giờ bay	Giờ đến
VN272	ART72	SG	HN	30/06/02	7:20	08:05P
VN372	ART72	HN	HUE	30/06/02	8:45	09:40P
VNA32	A320A	HUE	SG	30/06/02	12:05	13:35P
VN472	ART72	HN	SG	30/06/02	12:25	13:25P
VNB77	BOE77	SG	HUE	30/06/02	14:05	15:45P
VNB67	BOE67	SG	DN	30/06/02	14:05	15:45P
VNT06	TU106	DN	HN	30/06/02	16:25	17:35P
...

Để lựa chọn chuyến bay, khách hàng có thể tham khảo lịch bay của hãng, thông tin về chuyến bay được lập theo bảng trên.

Mỗi chuyến bay là một mối quan hệ giữa các thuộc tính: mã chuyến bay, loại máy bay, sân bay đi, sân bay đến, ngày bay, giờ bay, giờ đến.

Bảng trên được gọi là bảng dữ liệu, tùy từng đối tượng mà khai thác dữ liệu trong bảng trên theo mục đích của mình chẳng hạn:

- Mỗi khách hàng muốn bay từ Hà Nội vào Sài Gòn, lúc đó anh ta chỉ quan tâm tới các dòng chứa thông tin về các chuyến bay từ Hà Nội tới Sài Gòn.
- Trong khi đó đối với một nhà quản lý, như tổng giám đốc có thể ông ta chỉ cần biết tới số lượng chuyến bay thực hiện trong ngày bằng cách đếm số dòng trên bảng.

Điều này có nghĩa là các dữ liệu của bảng trên độc lập với các xử lý tác động lên chúng, và được tổ chức thành một bảng gồm các cột và các hàng, mỗi cột được gọi là một thuộc tính (attribute) hay một trường, mỗi hàng (dòng) được gọi là một thể hiện (instance) một bộ (tuple) hay một bản ghi (record) của bảng dữ liệu.

Các thao tác thường áp dụng lên bảng dữ liệu là:

- Thêm một bản ghi
- Xoá, sửa một bản ghi.

Mục đích của CSDL

Tích hợp (intergration)

- Tập trung và quản lý các dữ liệu rời rạc.
- Dữ liệu không tồn tại riêng lẻ mà có sự phụ thuộc lẫn nhau (vd: dữ liệu nhân viên có quan hệ chặt chẽ với dữ liệu của phòng ban và gia đình)
- CSDL không chỉ quản lý dữ liệu mà còn quản lý các quan hệ của dữ liệu

Chia sẻ(sharing)

- Nhiều user cùng sử dụng đồng thời
- Dữ liệu có thể được sử dụng cho nhiều mục đích, nhiều ứng dụng

Các thành phần của hệ thống

CSLD hợp nhất

CSDL phải thoả mãn hai yêu cầu sau:

- Không dư thừa dữ liệu(trên thực tế là dư thừa ít nhất)
- Sử dụng dùng chung

Người sử dụng

Hệ thống cơ sở dữ liệu

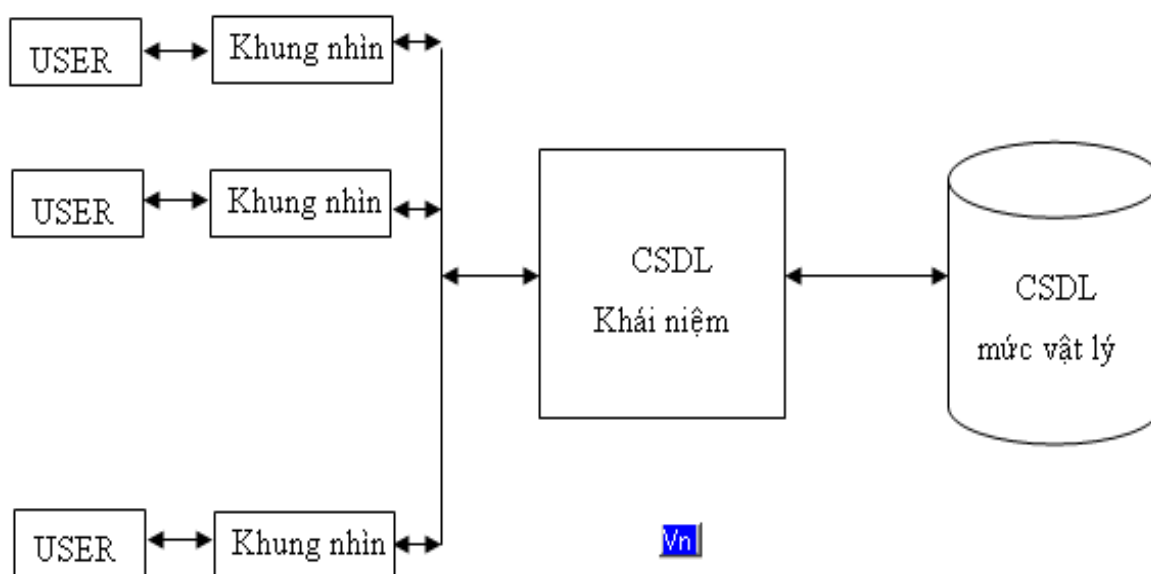
- Là người có nhu cầu truy nhập CSDL để thực hiện một thao tác nào đó
- Người sử dụng cuối (End-User): Là những người truy nhập vào CSDL từ một terminal, muốn tìm kiếm tra cứu thông tin.
- Người viết chương trình ứng dụng: Những người này ngoài những thao tác trên còn cần đến một ngôn ngữ lập trình (NNLT).
- Người quản trị CSDL: Là người có nhiệm vụ điều khiển toàn bộ hệ CSDL, là người có quyền cao nhất
- Hệ quản trị cơ sở dữ liệu: Đây chính là phần mềm của hệ CSDL

Phần cứng:

Là các thiết bị vật lý được sử dụng để lưu trữ dữ liệu

Các khái niệm cơ bản của hệ thống CSDL

Kiến trúc một hệ cơ sở dữ liệu



Các mức trừu tượng

Mức vật lý(Physical level)

- Tập hợp các tập tin, các chỉ mục, những cấu trúc lưu trữ khác dùng để truy xuất
- Tồn tại trong các thiết bị lưu trữ như đĩa từ
- Được quản trị bởi phần mềm quản trị CSDL

Mức khái niệm(Conceptual level)

- Trừu tượng hóa thế giới thực
- CSDL khái niệm được thiết kế như một thể thống nhất, bao gồm tất cả các dữ liệu được dùng bởi một tổ chức
- DBMS cho phép gộp các tập tin lại và đọc chúng theo phương cách riêng - phương cách được mô tả bằng CSDL khái niệm

Mức khung nhìn(View level)

- Là cách nhìn, là quan niệm của từng người sử dụng đối với CSDL mức khái niệm. Sự khác nhau giữa khung nhìn và mức khái niệm thực chất là không lớn

Tính độc lập dữ liệu, chia sẻ dữ liệu

Hai loại độc lập dữ liệu trong CSDL:

- Độc lập vật lý: Thay đổi tổ chức của cơ sở dữ liệu vật lý có thể thay đổi hiệu quả tính toán của chương trình, nhưng không đòi hỏi phải thay đổi lại chương trình. Nói một cách khác, mức quan niệm phải đảm bảo được các truy nhập đến CSDL từ phía người sử dụng
- Độc lập logic: sự thay đổi, thêm bớt thông tin về các thực thể ở mức quan niệm không đòi hỏi thay đổi các khung nhìn của NSD dẫn tới không cần thay đổi chương trình ứng dụng

Tính chia sẻ dữ liệu: Vì độc lập với chương trình ứng dụng nên nhiều chương trình ứng dụng cùng sử dụng một cơ sở dữ liệu.

Với hai tính chất trên, ta có thể trừu tượng hoá dữ liệu ở mức cao. Sự trừu tượng hoá không những tăng cường hiệu quả quản lý mà còn giúp tư duy tốt về CSDL.