

Cơ sở dữ liệu nâng cao

Nguyễn Kim Anh
anhnk-fit@mail.hut.edu.vn

Bộ môn Hệ thống Thông tin, SoICT

1

Giới thiệu

- Tổng quan về các hệ CSDL
- CSDL phân tán
- CSDL suy diễn
- CSDL hướng đối tượng
- OLAP

2

Mở đầu

- Các hệ CSDL quan hệ cổ điển, tập trung đã thu được những thành tựu to lớn về cả phương diện lý thuyết và thực hành.
- Một số hạn chế của CSDL quan hệ: ngữ nghĩa của mô hình quan hệ chưa đủ phong phú, việc quản lý các dữ liệu động và phức tạp kém hiệu quả, vấn đề khai thác dữ liệu hạn chế và không phù hợp với các tổ chức xí nghiệp có quy mô lớn và rất lớn

➡ Một số hướng nghiên cứu mở rộng được đề xuất: CSDL phân tán, CSDL suy diễn, CSDL hướng đối tượng, OLAP,...

3

Tài liệu tham khảo

- J.D.Ullman, Principle of Knowledge-base and Database Systems, 1988.
- W.Kim, Introduction to Object-Oriented Database Systems
- S.Ceri&G.Pelagatti, Distributed Databases, Principles&Systems, 1984.
- M.T. Özsu & P. Valduriez, Principles of Distributed Database Systems. Prentice-Hall, 1999.

4

Chương 1. Tổng quan về các hệ CSDL

- Các hệ xử lý tệp truyền thống
- Các hệ CSDL
- Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL
- Các xu thế phát triển trong lĩnh vực CSDL

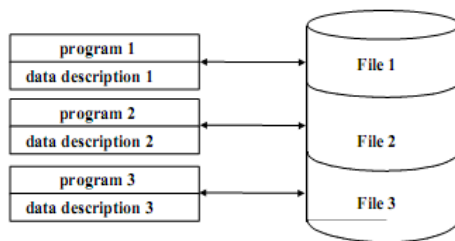
5

Các hệ xử lý tệp truyền thống

- **Cách tiến hành:**
 - Tin học hóa từng phần: tập chương trình ứng dụng và tệp dữ liệu ứng với các ứng dụng được tin học hóa
 - Không kế thừa và sử dụng lại được các chương trình ứng dụng và các tệp đã có
- **Nhược điểm:**
 - Dư thừa dữ liệu → không đảm bảo tính nhất quán dữ liệu
 - Các chương trình ứng dụng là phụ thuộc dữ liệu

6

File Systems



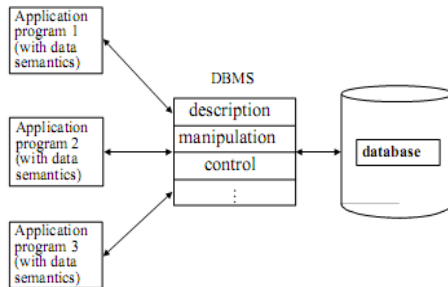
7

Các hệ CSDL

- Các thành phần của 1 hệ CSDL
- Phân loại các hệ CSDL
- Lợi ích của các hệ CSDL

8

Database Management



9

Các thành phần của 1 hệ CSDL

- CSDL hợp nhất:
 - không dư thừa và
 - được sử dụng chung
- Người sử dụng :
 - người lập trình ứng dụng viết các chương trình ứng dụng
 - người phân tích dữ liệu
 - người dùng cuối
- Phần mềm DBMS
- Phần cứng

10

Phân loại các hệ CSDL

- Theo mô hình dữ liệu
 - Hệ CSDL quan hệ
 - Hệ CSDL mạng
 - Hệ CSDL phân cấp
 - Hệ CSDL hướng đối tượng

11

Phân loại các hệ CSDL

- Theo qui mô xí nghiệp
 - Các hệ CSDL tập trung
 - Hệ CSDL cá nhân
 - Hệ CSDL trung tâm
 - Hệ CSDL Client/Server
 - Các hệ CSDL phân tán
 - Hệ CSDL phân tán thuần nhất
 - Hệ CSDL phân tán không thuần nhất

12

Phân loại các hệ CSDL

- Theo nhu cầu truy cập và khai thác CSDL
 - OLTP (On-Line Transaction Processing)
 - OLAP (On-Line Analytical Processing)
 - KDD (Knowledge Discovery from Databases)

13

Lợi ích của các hệ CSDL

- Quản lý xí nghiệp một cách hệ thống
- Áp dụng các tiêu chuẩn thống nhất đối với toàn xí nghiệp
- Giảm dư thừa dữ liệu, tránh sự không nhất quán dữ liệu
- Đảm bảo tính độc lập dữ liệu đối với các chương trình ứng dụng
- Cho phép nhiều người sử dụng truy cập và khai thác CSDL đồng thời

14

Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL

- Thiết kế CSDL cho xí nghiệp
 - **Xác định** tập các dữ liệu tác nghiệp của xí nghiệp cần lưu trữ trong CSDL → Xây dựng ERD-biểu diễn thể giới thực ở mức khái niệm
 - Lựa chọn **mô hình dữ liệu**
 - **Thiết kế và chuẩn hóa** CSDL

15

Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL

- Tạo lập CSDL cho xí nghiệp
 - **Thu thập** tập các dữ liệu tác nghiệp của xí nghiệp cần lưu trữ trong CSDL (CSDL có thể chứa các dữ liệu rõ chính xác, dữ liệu mờ không chính xác, không chắc chắn → CSDL chứa thông tin đầy đủ hay chứa thông tin không đầy đủ về xí nghiệp)
 - Lựa chọn **DBMS ứng với mô hình dữ liệu** đã chọn
 - Định nghĩa CSDL và tạo lập CSDL sử dụng DBMS đã chọn

16

Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL

- Tổ chức lưu trữ vật lý
 - Tổ chức tệp băm
 - Tổ chức tệp chỉ dẫn
 - Tổ chức cây cân bằng

17

Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL

- Xử lý và tối ưu hóa câu hỏi
 - Dựa trên Đại số quan hệ (ĐSQH) và Các phép biến đổi tương đương của ĐSQH
 - Dựa trên ngữ nghĩa

18

Các vấn đề quan tâm khi thiết kế, xây dựng hệ CSDL

- Quản trị giao dịch và điều khiển tương tranh
 - Kỹ thuật khóa
 - Kỹ thuật gán nhãn thời gian
- Giao diện truy vấn
 - Giao diện dựa trên ngôn ngữ lệnh
 - Giao diện dựa trên form
 - Giao diện đồ họa
 - Giao diện ngôn ngữ tự nhiên

19

Các xu thế phát triển trong lĩnh vực CSDL

- Sự phát triển về qui mô của tổ chức, xí nghiệp:
 - Địa bàn hoạt động rộng lớn tại nhiều vị trí khác nhau
 - Số người sử dụng tăng nhanh và sức ép từ người sử dụng về quyền tự trị của người sử dụng
 - CSDL rất lớn
 - Hệ CSDL phân tán

20

- Sự phong phú và đa dạng của các dữ liệu trong CSDL
 - Dữ liệu không có cấu trúc chặt chẽ
 - Mối quan hệ ràng buộc giữa các dữ liệu phức tạp và chồng chéo
- Hệ CSDL hướng đối tượng

21

Các xu thế phát triển trong lĩnh vực CSDL

- Sự phát triển về nhu cầu khai thác các CSDL
 - Hệ CSDL suy diễn: suy diễn ra các sự kiện mới
 - OLAP: phân tích tổng hợp dữ liệu
 - Phát hiện tri thức từ các CSDL rất lớn

22