Overview of Database Systems

Yanlei Diao
University of Massachusetts Amherst

Outline

• An outside look: DB Application
• An inside look: Anatomy of DBMS

Database Management System (DBMS): a software package designed to store and manage a large amount of data.

An Outside Look: DB Application

High-level, declarative interface

Data-Intensive Application

• Persistent storage
• Performance
• Concurrency
• Automatic recovery
• Security
Case Study: The Internet Shop*

- DBDudes Inc.: a well-known database consulting firm
- Barns and Noble (B&N): a large bookstore specializing in books on horse racing
- B&N decides to go online, asks DBDudes to help with the database design and implementation
- Step 0: DBDudes makes B&N agree to
  - pay steep fees and
  - schedule a lunch meeting for requirements analysis

* The example and all related material was taken from "Database Management Systems" Edition 3.

---

Step 1: Requirements Analysis

- "I’d like my customers to be able to browse my catalog of books and place orders online."
  - Books:
    - For each book, B&N’s catalog contains its ISBN number, title, author, price, year of publication, ...
  - Customers:
    - Most customers are regulars with names and addresses registered with B&N.
    - New customers must first call and establish an account.
  - On the new website:
    - Customers identify themselves before browsing and ordering.
  - Shipping:
    - For each order, B&N ships all copies of a book together once they become available.

---

Step 2: Conceptual Design

- A high level description of the data in terms of the Entity-Relationship (ER) model.
- Design review:
  - What if a customer places two orders of the same book in one day?
  - Modification: add "ordernum" to Orders.

---

* Yanlei Diao, University of Massachusetts Amherst 4/10/07
Step 3: Logical Design

- Mapping the ER diagram to the relational model

CREATE TABLE Books
(isbn CHAR(10),
title CHAR(80),
author CHAR(80),
city_in_stock INTEGER,
price REAL,
year INTEGER,
PRIMARY KEY(isbn))

CREATE TABLE Customers
(cid INTEGER,
cname CHAR(80),
address CHAR(200),
PRIMARY KEY(cid))

CREATE TABLE Orders
(ordernum INTEGER,
isbn CHAR(10),
cid INTEGER,
cardnum CHAR(16),
qty INTEGER,
order_date DATE,
ship_date DATE,
FOREIGN KEY (isbn) REFERENCES Books,
FOREIGN KEY (cid) REFERENCES Customers)

CREATE VIEW OrderInfo
(isbn, cid, qty, order_date, ship_date)
AS SELECT O.isbn,
        O.cid,
        O.qty,
        O.order_date,
        O.ship_date
FROM Orders O

Step 4: Schema Refinement

Redundant Storage!

Step 5: Internet Application Development

Presentation tier
- Interface to the user
- Adapt to display devices

Application logic tier
- Business logic (actions, state between steps)
- Access multiple sources

Data management tier
- One/multiple DBMS(es)

Client Program
(Web Browser)

Application Server
(Apache Tomcat, ...)

Database System
(DB2, MySQL, ...)

HTML
JavaScript
Cookies

HTTP

JSP
Servlets
XSLT

JDBC

XML
stored procedures

B&N Client:
- User input
- Session state

B&N Business logic:
- Home page
- Login page
- Search page
- Cart page
- Confirm page

B&N Data:
- Books
- Customers
- Users
- Orders
- Orderlists
An Example Internet Store

Example SQL Queries

Step 6: Physical Design
Outline

• An outside look: DB Application
• An inside look: Anatomy of DBMS

An Inside Look: Anatomy of DBMS

Query Processor

- Syntax checking
- Internal representation
- Handling views
  - Logical/semantics rewriting
  - Flattening subqueries
- Building a query execution plan
  - Efficient, if not optimal
  - Define plan space
  - Cost estimation for each
  - Search algorithms
- Pull-based execution of a plan
  - Each operator is an iterator

Query Rewriter

- Query Optimizer

Query Executor

- IndexScan Customers
  - cname = "John"
- IndexScan Orders
  - order_data
  - C.cid = O.cid

IndexScan (Indexed Join)

(On-the-fly)
Transaction Manager

Access Methods
- Customers
- Orders
- Index Scan

Concurrency:
- 2PL

Replacement policy:
- Support for Concurrency & Recovery

Disk Manager

Buffer Manager
- Allocate/Deallocate a page
- Read/Write a page

Database
- Heap file
- Page format
- Record format

Indices

Catalog

DBMS: Theory + Systems

Query Parser
- Theory!

Query Optimizer

Query Executor

Lock Manager
- Access Methods
- Log Manager

DB

Systems!