Practical Database Design Methodology

© Department of Computer Science
Northern Illinois University
September 2004
Introduction to Information Systems

- Database systems are now an intricate part of information technology (IT) and information resource management (IRM) systems
- IT and IRM are essential to the successful management of business needs
Using Database Systems

• Data is regarded as a vital corporate resource
• More and more functions within an organization are computerized
• AND the functions need to be computerized to keep up with competitors
Using Database Systems

- Data is becoming more complex thus making modeling of data a necessity for understanding the data.
- Business costs are being reduced by letting end-users perform business functions (i.e., travel, retail orders, etc.)
Necessary Characteristics of Database Systems

• Data Independence
  – Allows application programs to still run even with changes to the logical or physical schemas

• External Schemas
  – Allow the same data to be used in multiple applications
New Capabilities of Database Systems

- Integration of data across multiple applications (one database)
- Simplicity of developing new applications using high-level languages
- Can support browsing and querying by managers
Information Repositories

• Mini DBMSs that manage the metadata
  – Data that describes the data
    • Structure
    • Constraints
    • Applications
    • Authorizations
Information Repositories

• Store information on
  – Schemas of the database system
  – Physical database design
    • Storage structure
    • Access paths
    • Record sized
  – Database users
    • Names
    • Responsibilities
    • Access rights
Information Repositories

• Store information on
  – Database transactions
  – Applications
  – Relationships
    • Between transactions and applications
    • Between transactions and data items
  – Usage statistics
    • Frequency of queries, transactions, etc.
Information System Life Cycle

- Feasibility analysis
- Requirements collection and analysis
- Design
- Implementation
- Validation and acceptance testing
- Deployment, operation and maintenance
Database Application
System Life Cycle

- System definition
- Database design
- Database implementation
- Loading or data conversion
- Application conversion
- Testing and validation
- Operation
- Monitoring and maintenance
System Definition

• Defining the
  – Scope of the database system
  – The users
  – The applications

• Identifying
  – User interfaces
  – Response time constraints
  – Storage and processing needs
Database Design

• Creating a complete
  – Conceptual design
  – Logical design
  – Physical design

Requires choosing the DBMS first.
Database Implementation

• Specifying
  – The logical schema
  – The physical schema
• Creating empty database files
• Implementing software applications
  – Design, coding, implementation
Loading or Data Conversion

• Populate the database
  – Loading data
  – Conversion of existing files
Application Conversion

- Old software programs are converted to the new system at this time
Testing and Validation

• The new system is thoroughly tested and validated
• Sometimes parallel testing is performed
Operation

- Put the new database system into operation
- May run parallel systems for a while
Monitoring and Maintenance

• Constantly monitor the system for
  – Defects
  – Improvements that are needed

• Perform maintenance
  – Expanding system
  – Fixing defects
  – Reorganization of data and/or database
Database Design Process

• Problem
  – Design the logical and physical structure of one or more databases to accommodate the information needs of the users in an organization for a defined set of applications.
Database Design Process

• Goals
  – Satisfy the information requirements of the users and applications
  – Provide easy-to-understand structuring of the information
  – Support
    • Processing requirements
    • Performance objectives
      – Response time
      – Processing time
      – Storage space
Database Design Process

• Phases
  – Requirements collection and analysis
  – Conceptual database design
  – Choice of DBMS
  – Data model mapping
  • Conceptual schema to logical schema
  – Physical database design
  – Database system implementation and tuning
Phase 1: Requirements Collection and Analysis

- Identify
  - Major users
  - Major applications
  - Major inputs and outputs
  - Key individuals and/or committees
Phase 1: Requirements Collection and Analysis

- Study
  - Existing documentation
  - Existing applications
  - Policy manuals
  - Forms
  - Reports
  - Organization charts
Phase 1: Requirements Collection and Analysis

• Also study
  – Current operating environment
  – Planned use of information
    • Types of transactions with frequencies
    • Flow of information within the system
  – Geographic characteristics of
    • Users
    • Transactions
    • Reports (destination)
Phase 2: Conceptual Database Design

• Phase 2a: Conceptual schema design goal
  – A complete understanding of
    • The database structure
    • Meaning
    • Interrelationships
    • Constraints

• Independent of a DBMS
Phase 2: Conceptual Database Design

- Conceptual schema is a “stable description” of the database contents
- A good understanding of the conceptual schema is crucial for database users and application designers
- Use of a diagrammatic description of the conceptual schema is an excellent communication tool among
  - Database users
  - Designers
  - Analysts
Phase 2: Conceptual Database Design

• Phase 2b: Transaction Design
  – Design the characteristics of known database transactions independent of a DBMS
  – Using a transaction look at
    • Input
    • Functional behavior
    • Output
Phase 3: Choice of DBMS

- Consider a number of factors
  - Technical
  - Economic
  - Politics
Phase 3: Choice of DBMS

- Consider the following costs
  - Software acquisition cost
  - Maintenance cost
  - Hardware acquisition cost
  - Database creation and conversion cost
  - Personnel cost
  - Training cost
  - Operating cost
Phase 4: Data Model Mapping

• System independent logical database design
  – Translate an ER schema into a relational schema

• DBMS specific logical database design
  – Use specific modeling features and constraints of the chosen DBMS
Phase 5: Physical Database Design

• Choose with goal of good performance
  – Storage structures
  – Access structures
  – Access paths

• Criteria for choosing
  – Response time needed
  – Space utilization requirements
  – Transaction throughput stated by user
Phase 6: Database System Implementation and Tuning

- Loading a new database (i.e., populating the database)
- Set up monitoring for daily operations
- Tune the database as needed