Chapter 10 - Objectives

- Main components of an information system.
- Main stages of database system development lifecycle.
- Main phases of database design: conceptual, logical, and physical design.
- How to evaluate and select a DBMS.
- Benefits of CASE tools.
- Purpose and tasks associated with data administration and database administration.

Software Depression

- Last few decades have seen proliferation of software applications, many requiring constant maintenance involving:
  - correcting faults,
  - implementing new user requirements,
  - modifying software to run on new or upgraded platforms.
- Effort spent on maintenance began to absorb resources at an alarming rate.
Software Depression

- As a result, many major software projects were
  - late,
  - over budget,
  - unreliable,
  - difficult to maintain,
  - performed poorly.

- In late 1960s, led to 'software crisis', now refer to as the 'software depression'.

Software Depression

- Major reasons for failure of software projects includes:
  - lack of a complete requirements specification;
  - lack of appropriate development methodology;
  - poor decomposition of design into manageable components.

- Structured approach to development was proposed called Information Systems Lifecycle (ISLC).

Information System (IS)

- IS are resources that enable collection, management, control, and dissemination of information throughout an organization.

- Database is fundamental component of IS, and its development/usage should be viewed from perspective of the wider requirements of the organization.

Contents

- Main components of an information system.
  - Main stages of database system development lifecycle.
    - Planning and system definition
    - Analysis and design
    - Implementation
    - Data conversion and testing
    - Maintenance
  - CASE tools
  - Purpose and tasks associated with data administration and database administration.
Database System Development Lifecycle

- Database planning
- System definition
- Requirements collection and analysis
- Database design
- DBMS selection (optional)
- Application design
- Prototyping (optional)
- Implementation
- Data conversion and loading
- Testing
- Operational maintenance

Database Planning

- Management activities that clarifies purpose of the database project and provides clearer path towards the efficient and effective creation of required database system.
- Each objective should identify a particular task that the database must support.
- Database planning should also include development of standards that govern:
  - how data will be collected,
  - how the format should be specified,
  - what necessary documentation will be needed,
  - how design and implementation should proceed.

System Definition

- Describes scope and boundaries of database system and the major user views.
- Database application may have one or more user views.
  - User view defines what is required of a database system from perspective of:
    - a particular job role (such as Manager or Supervisor) or
    - enterprise application area (such as marketing, personnel, or stock control).
  - User views also allows requirements to be broken down into manageable pieces.
Requirements Collection and Analysis

- Process of collecting and analyzing information about the part of organization to be supported by the database system, and using this information to identify users' requirements of new system.
- Information is gathered for each major user view including:
  - a description of data used or generated;
  - details of how data is to be used/generated;
  - any additional requirements for new database system.
- Requirements for each user view are merged into a single set of requirements.
Database Design

- Process of creating a design for a database that will support the enterprise’s mission statement and mission objectives for the required database system.
- Main purposes of data modeling include:
  - To assist in understanding the meaning of the data;
  - To facilitate communication about the information requirements.

Three Phases of Database Design

1) Conceptual database design
   - Process of modelling of the data used, independent of all physical considerations.
   - Data model is built using the information in users’ requirements specification.
   - Conceptual data model is source of information for logical design phase.
Three Phases of Database Design

2) Logical database design
   - Process of constructing a model of the data based on a specific data model (e.g. relational), but independent of a particular DBMS.
   - Conceptual data model is refined and mapped on to a logical data model.

Criteria to Produce an Optimal Data Model

<table>
<thead>
<tr>
<th>Structure validity</th>
<th>Non-redundancy</th>
<th>Extensibility</th>
<th>Simplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency with the enterprise defines and organizes information.</td>
<td>Exclusion of extraneous information; the representation of any one piece of information exactly once.</td>
<td>Ability to evolve to support new requirements with minimal effects on existing users.</td>
<td>Easy to understand by IS professionals and non-technical users. Should be represented with a commonly used diagrammatic notation.</td>
</tr>
</tbody>
</table>

Three Phases of Database Design

3) Physical database design.
   - Process of producing a description of the database implementation on secondary storage, tailored to a specific DBMS system.
   - Describes base relations, file organizations, indexes used to achieve efficient access to data, and any associated integrity constraints and security measures.

Three-Level ANSI-SPARC Architecture and Phases of Database Design

- External schema
- Conceptual schema
- Internal schema
- Physical storage

Logical/conceptual database design

Physical database design
DBMS Selection

- Selection of an appropriate DBMS to support the database system.
  - Undertaken at any time prior to logical design provided sufficient information is available regarding system requirements.
- Main steps to selecting a DBMS:
  - define Terms of Reference (ToR) of study;
  - shortlist two or three products;
  - evaluate products;
  - recommend selection and produce report.

DBMS Evaluation Features

<table>
<thead>
<tr>
<th>Accessibility</th>
<th>Transaction handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query language SQL 2013/2019 compliant</td>
<td>Backup and recovery routines</td>
</tr>
<tr>
<td>Interfacing to 3GLs</td>
<td>Checkpointing facility</td>
</tr>
<tr>
<td>Multi-user</td>
<td>Logging facility</td>
</tr>
<tr>
<td>Security</td>
<td>Granularity of concurrency</td>
</tr>
<tr>
<td>- Office Access controls</td>
<td>Deadlock resolution strategy</td>
</tr>
<tr>
<td>- Authorization mechanism</td>
<td>Advanced transaction models</td>
</tr>
<tr>
<td>Authorization mechanism</td>
<td>Parallel query processing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>File structures available</td>
</tr>
<tr>
<td>File structure maintenance</td>
</tr>
<tr>
<td>Ease of reorganization</td>
</tr>
<tr>
<td>Indexing</td>
</tr>
<tr>
<td>Variable length fields/records</td>
</tr>
<tr>
<td>Data compression</td>
</tr>
<tr>
<td>Encryption routines</td>
</tr>
<tr>
<td>Memory requirements</td>
</tr>
<tr>
<td>Storage requirements</td>
</tr>
</tbody>
</table>

Utilities

- Performance measuring
- Tuning
- Load/unload facilities
- User usage monitoring
- Database administration support

Development

- 4GL /5GL tools
- CASE tools
- Windows capabilities
- Stored procedures, triggers, and rules
- Web development tools
### DBMS Evaluation Features

**Other features**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upgradability</td>
<td>Interoperability with other DBMSs and other systems</td>
</tr>
<tr>
<td>Vendor stability</td>
<td>Web integration</td>
</tr>
<tr>
<td>User base</td>
<td>Replication utilities</td>
</tr>
<tr>
<td>Training and user support</td>
<td>Distributed capabilities</td>
</tr>
<tr>
<td>Documentation</td>
<td>Portability</td>
</tr>
<tr>
<td>Operating system required</td>
<td>Hardware required</td>
</tr>
<tr>
<td>Cost</td>
<td>Network support</td>
</tr>
<tr>
<td>Online help</td>
<td>Object-oriented capabilities</td>
</tr>
<tr>
<td>Standards used</td>
<td>Architecture (2- or 3-tier client/server)</td>
</tr>
<tr>
<td>Version management</td>
<td>Performance</td>
</tr>
<tr>
<td>Extensible query optimization</td>
<td>Transaction throughput</td>
</tr>
<tr>
<td>Scalability</td>
<td>Maximum number of concurrent users</td>
</tr>
<tr>
<td>Support for analytical tools</td>
<td>XML support</td>
</tr>
</tbody>
</table>

### Example - Evaluation of DBMS Product

**DBMS:** Sample product  
**Vendor:** Sample vendor

#### Physical Definition Group

<table>
<thead>
<tr>
<th>Features</th>
<th>Comments</th>
<th>Rating</th>
<th>Weighting</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>File structures available</td>
<td>Choice of 4</td>
<td>8</td>
<td>0.15</td>
<td>1.2</td>
</tr>
<tr>
<td>File structure maintenance</td>
<td>NOT self-regulating</td>
<td>6</td>
<td>0.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Ease of reorganization</td>
<td>4</td>
<td>0.25</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Indexing</td>
<td>6</td>
<td>0.15</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Variable length fields/records</td>
<td>5</td>
<td>0.15</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Data compression</td>
<td>Specify with file structure</td>
<td>7</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td>Encryption routines</td>
<td>Choice of 2</td>
<td>4</td>
<td>0.05</td>
<td>0.2</td>
</tr>
<tr>
<td>Memory requirements</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Storage requirements</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>41</td>
<td>1.0</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>Physical definition group</td>
<td>5.75</td>
<td>0.25</td>
<td>1.44</td>
<td></td>
</tr>
</tbody>
</table>

### Assignment

- Compare 3 major DBMS products in market today.
  - With at least 4 criteria from each and every group, except only 2 each from the utilities and development groups.

### Application Design

- Design of user interface and application programs that use and process the database.
  - Database design and application design are parallel activities.

- Includes two important activities:
  - User interface design;
  - Transaction design.
Application Design Activities

- User interface design
  - Ease of use
  - Flow of control
  - Error checking and avoidance
  - Good looking
- Transaction design
  - An action, or series of actions, carried out by a single user or application program, which accesses or changes content of the database.
  - The designer should define and document the high-level characteristics of the transactions required.

Prototyping

- Building working model of a database system.

  - Purpose
    - to identify features of a system that work well, or are inadequate;
    - to suggest improvements or even new features;
    - to clarify the users’ requirements;
    - to evaluate feasibility of a particular system design.

Implementation

- Physical realization of the database and application designs.
  - Use DDL to create database schemas and empty database files.
  - Use DDL to create any specified user views.
  - Use 3GL or 4GL to create the application programs.
    - This will include the database transactions implemented using the DML, possibly embedded in a host programming language.

Contents

- Main components of an information system.
- Main stages of database system development lifecycle.
  - Planning and system definition
  - Analysis and design
    - Implementation
  - Data conversion and testing
  - Maintenance
- CASE tools
- Purpose and tasks associated with data administration and database administration.

@2012 by Chukiat Worasucheep
Example: Embedding Database into 3GL

```java
import java.sql.*; // Use classes in java.sql package

public class JdbcSelectTest { // JDK 7 and above
    public static void main(String[] args) { try {
        // Step 1: Allocate a database "Connection" object
        Connection conn = DriverManager.getConnection("jdbc:mysql://localhost:3306/bookshop", "username", "password"); // MySQL
        // connection conn = DriverManager.getConnection("jdbc:odbc:bookshopDBC"); // Access
        Statement stmt = conn.createStatement();
        }
        // Step 2: Allocate a "Statement" object in the Connection
        }
        // Step 3: Execute a SQL SELECT query, the query result
        // is returned in a "ResultSet" object.
        String strSelect = "select title, price, qty from books"
        System.out.println("The SQL query is: ", strSelect); // Echo for debugging
        System.out.println("Result size: ", stmt.executeQuery(strSelect));
    @2012 by Chukiat.Worasauchep.t = stmt.executeQuery(strSelect);
```

Data Conversion and Loading

- Transferring any existing data into new database and converting any existing applications to run on new database.
  - Only required when new database system is replacing an old system.
  - DBMS normally has utility that loads existing files into new database.
  - May be possible to convert and use application programs from old system for use by new system.

Testing

- Process of running the database system with intent of finding errors, according to a planned test strategy and realistic data.
- Testing cannot show absence of faults; it can show only that software faults are present.
- Demonstrates that database and application programs appear to be working according to requirements.
Operational Maintenance

• Process of monitoring and maintaining database system following installation.
• Monitoring performance of system.
  ◦ if performance falls, may require tuning or reorganization of the database.
• Maintaining and upgrading database application (when required).
• Incorporating new requirements into database application.

CASE Tools

• Support provided by CASE tools include:
  ◦ data dictionary to store information about database system’s data;
  ◦ design tools to support data analysis;
  ◦ tools to permit development of corporate data model, and conceptual and logical data models;
  ◦ tools to enable prototyping of applications.

Contents

• Main components of an information system.
• Main stages of database system development lifecycle.
  ◦ Planning and system definition
  ◦ Analysis and design
  ◦ Implementation
  ◦ Data conversion and testing
  ◦ Maintenance
• CASE tools
  ◦ Purpose and tasks associated with data administration and database administration.
**CASE Tools and Database System Development Lifecycle**

Some Database CASE tools

- [http://erwin.com/](http://erwin.com/)

**Contents**

- Main components of an information system.
- Main stages of database system development lifecycle.
  - Planning and system definition
  - Analysis and design
  - Implementation
  - Data conversion and testing
  - Maintenance
- CASE tools
  - Purpose and tasks associated with data administration and database administration.

**Data Administration and Database Administration**

- The Data Administrator (DA) and Database Administrator (DBA) are responsible for managing and controlling the corporate data and corporate database, respectively.
Data Administration

- DA is more concerned with early stages of database system development lifecycle and DBA is more concerned with later stages.
- Management of data resource including:
  - database planning,
  - development and maintenance of standards, policies and procedures, and conceptual and logical database design.

Database Administration

- Major responsibility of database administration (DBA) includes the management of physical realization of a database system including:
  - physical database design and implementation,
  - setting security and integrity controls,
  - monitoring system performance, and reorganizing the database.

What we’ve learnt

- Main components of an information system.
- Main stages of database system development lifecycle.
  - Planning and system definition
  - Analysis and design
  - Implementation
  - Data conversion and testing
  - Maintenance
- CASE tools
- Purpose and tasks associated with data administration and database administration.