Database Management: Introduction

Overview

- Database Concepts and Terminology
- Database Models

What is a Database?

Files and Databases

- File: A collection of records or documents dealing with one organization, person, area or subject. (Rowley)
  - Manual (paper) files
  - Computer files
- Database: A collection of similar records with relationships between the records. (Rowley)
  - bibliographic, statistical, business data, images, etc.

Database

- A Database is a collection of stored operational data used by the application systems of some particular enterprise. (C.J. Date)
    - Paper "Databases"
        - Still contain a large portion of the world's knowledge
    - File-Based Data Processing Systems
        - Early batch processing of (primarily) business data
    - Database Management Systems (DBMS)

Terms and Concepts

- Database Management System -- DBMS
  - Software system used to define, create, maintain and provide controlled access to the database and repository

Terms and Concepts

- Repository
  - AKA Data Dictionary
  - The place where all metadata for a particular database is stored
  - may also include information on relationships between files or tables in a particular database

# Terms and Concepts

- Metadata
  - Data about data
    - In DBMS means all of the characteristics describing the attributes of an entity, E.G.:
      - name of attribute
      - data type of attribute
      - size of the attribute
      - format or special characteristics
  - Characteristics of files or relations
    - name, content, notes, etc.

Why DBMS?

- History
  - 50's and 60's all applications were custom built for particular needs
  - File based
  - Many similar/duplicative applications dealing with collections of business data
  - Early DBMS were extensions of programming languages
  - 1970 - E.F. Codd and the Relational Model
  - 1979 - Ashton-Tate & first Microcomputer DBMS

File Based Systems
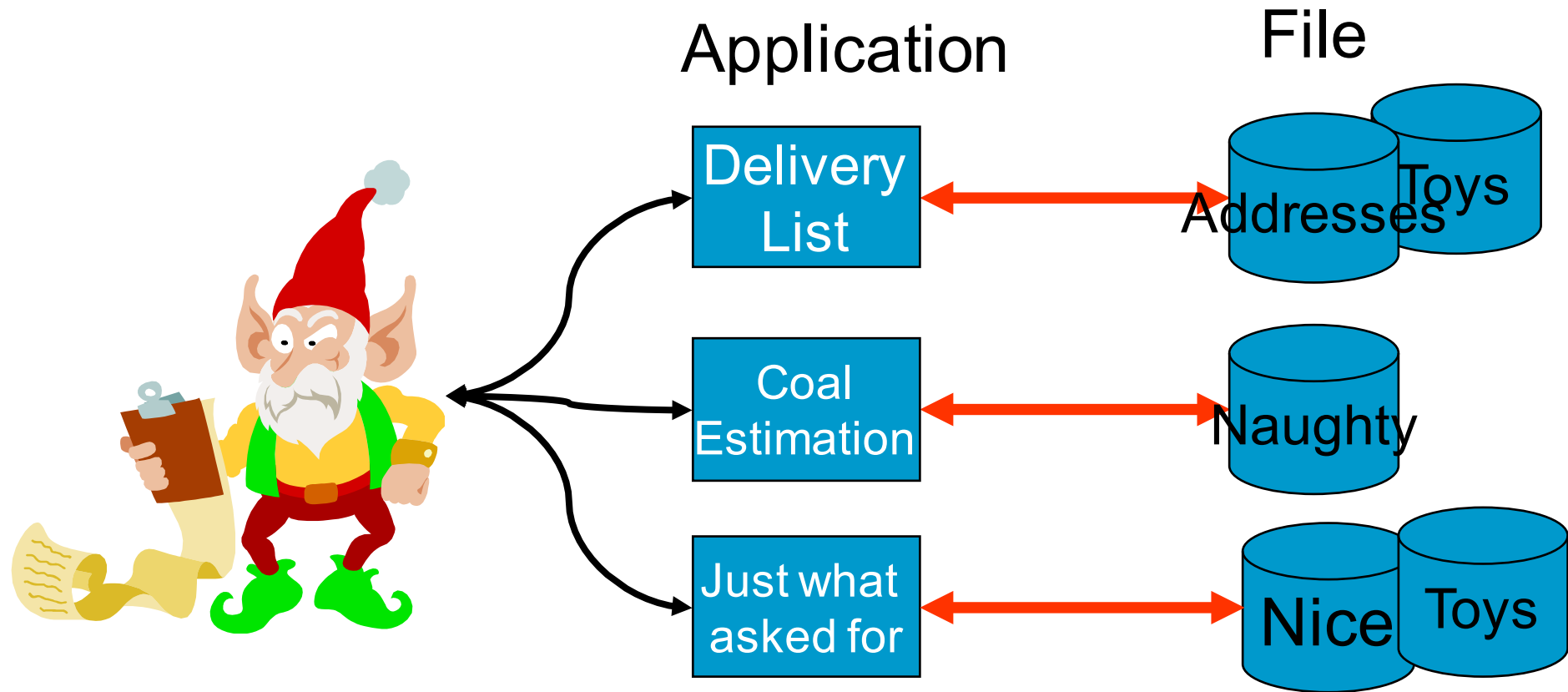
From File Systems to DBMS

- Problems with File Processing systems
  - Inconsistent Data
  - Inflexibility
  - Limited Data Sharing
  - Poor enforcement of standards
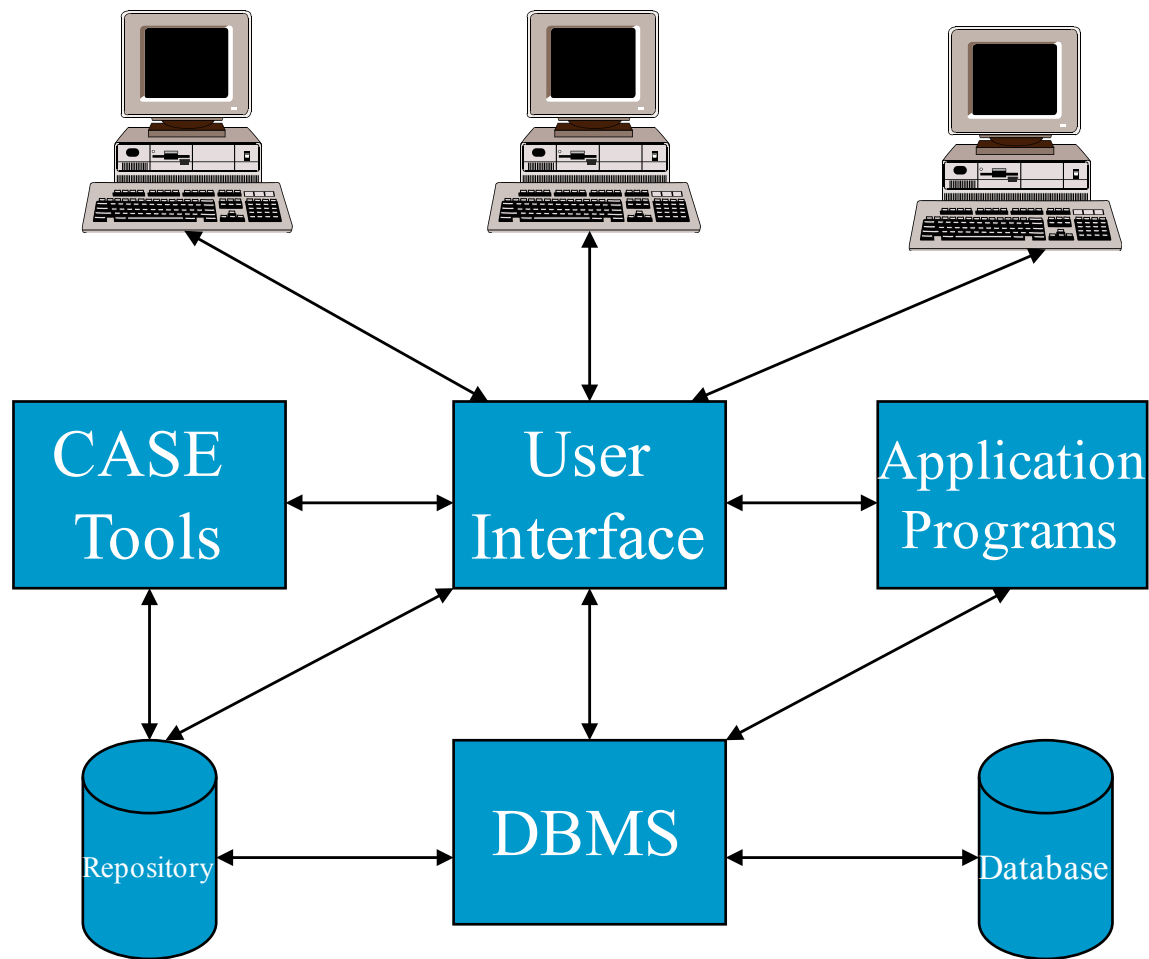 - Excessive program maintenance

DBMS Benefits

- Minimal Data Redundancy
- Consistency of Data
- Integration of Data
- Sharing of Data
- Ease of Application Development
- Uniform Security, Privacy, and Integrity Controls
- Data Accessibility and Responsiveness
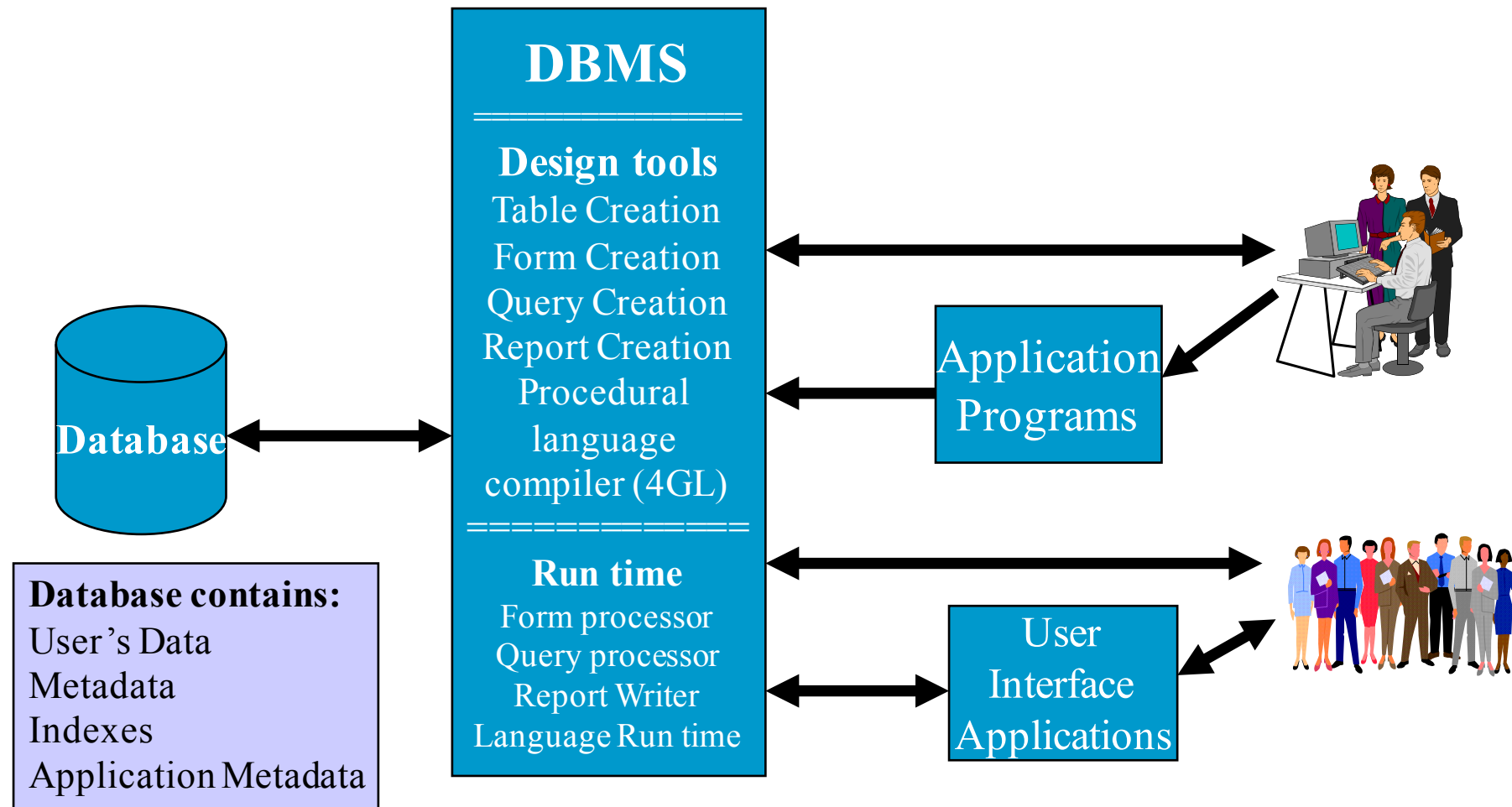- Data Independence
- Reduced Program Maintenance

Terms and Concepts

- ***Data Independence***
  - Physical representation and location of data and the use of that data are separated
    - The application doesn't need to know how or where the database has stored the data, but just how to ask for it.
    - Moving a database from one DBMS to another should not have a material effect on application program
    - Recoding, adding fields, etc. in the database should not affect applications

Database Environment

Database Components

DBMS
========

Design tools
Table Creation
Form Creation
Query Creation
Report Creation
Procedural
language
compiler (4GL)
============

Run time
Form processor
Query processor
Report Writer
Language Run time

Database

Database contains:
User's Data
Metadata
Indexes
Application Metadata

Application
Programs

User
Interface
Applications

Types of Database Systems

- PC Databases
- Centralized Database
- Client/Server Databases
- Distributed Databases
- Database Models

PC Databases
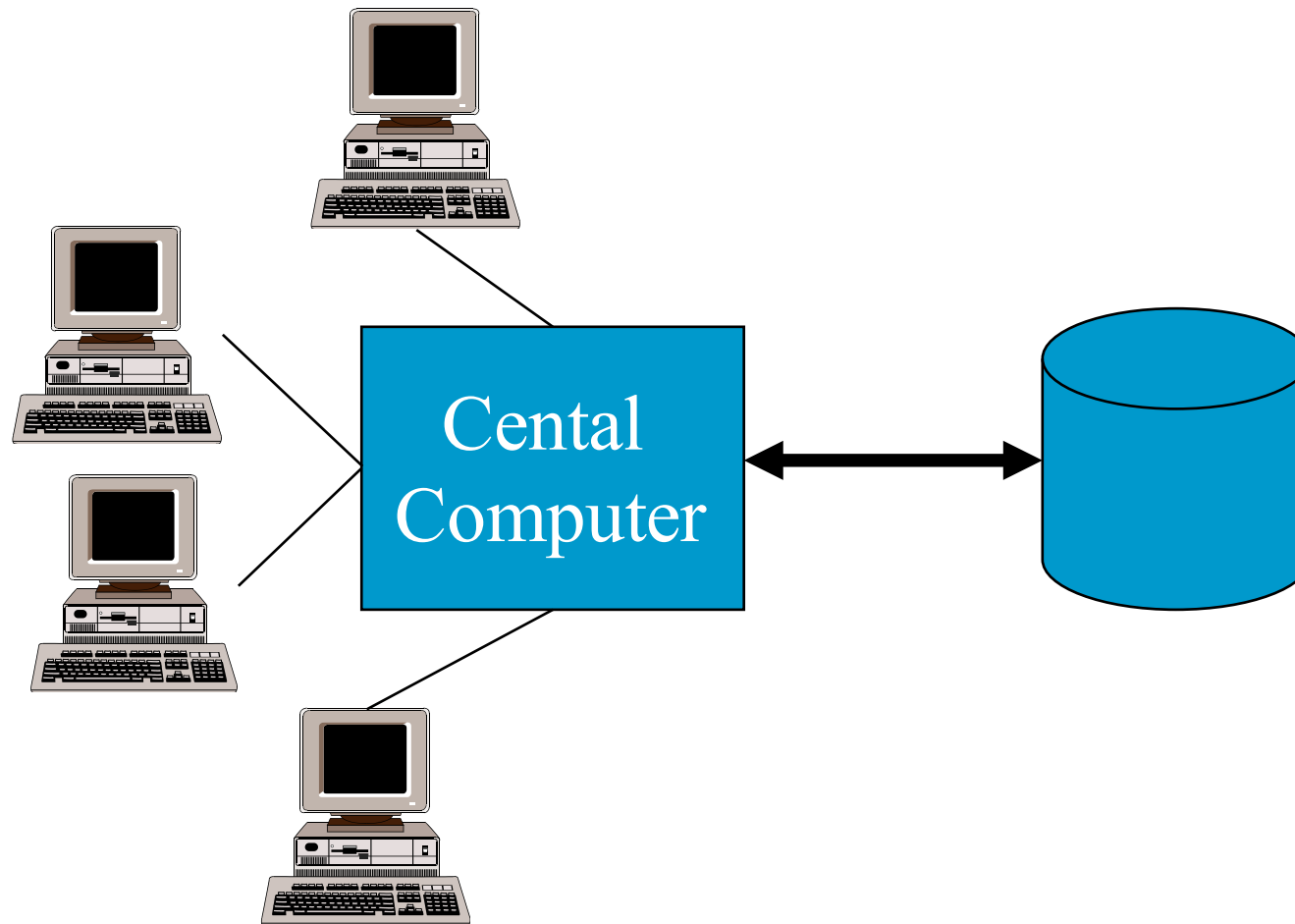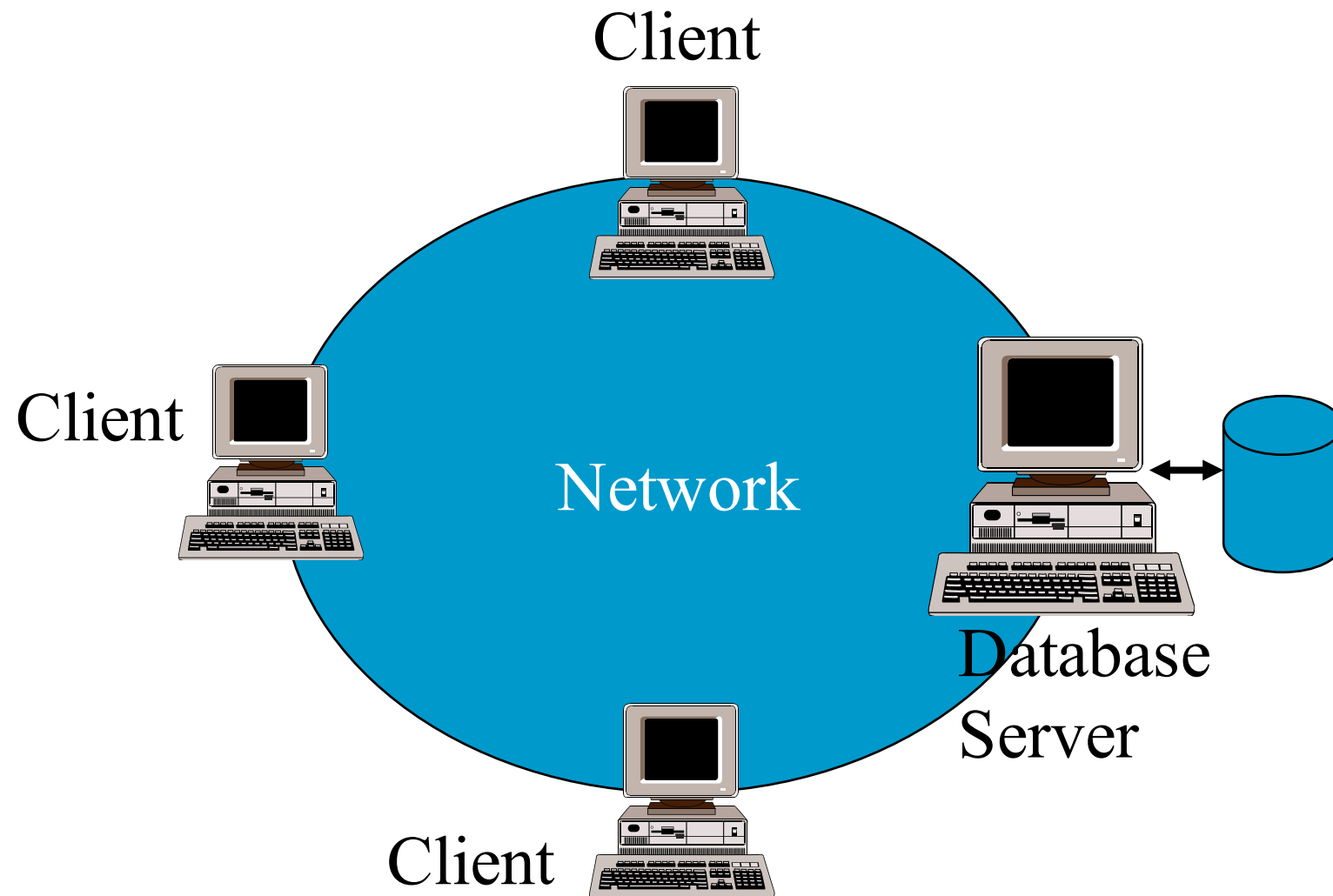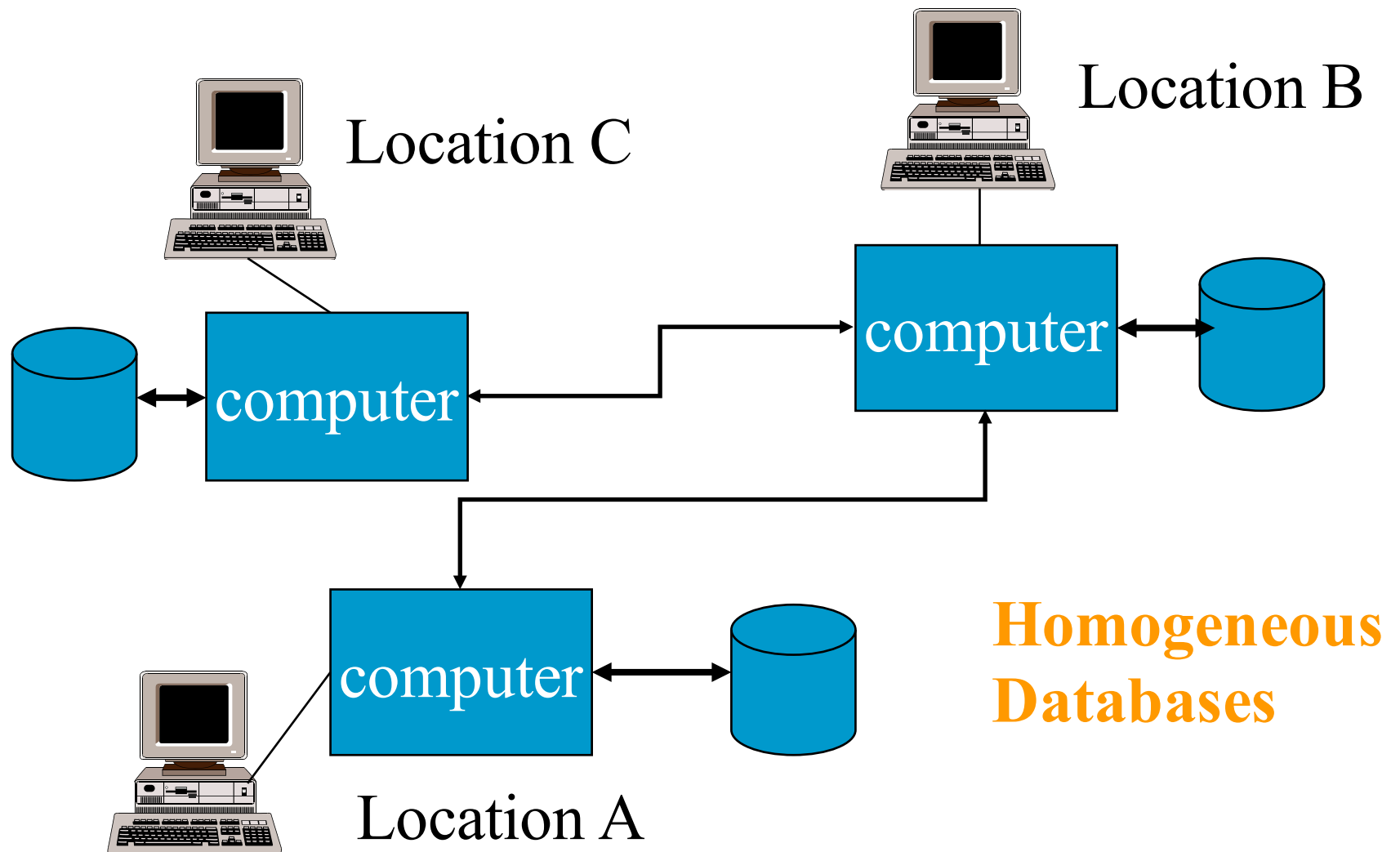
E.G.
Access
FoxPro
Dbase
Etc.

Centralized Databases

Cental
Computer

Client Server Databases

Client

Client

Network

Database
Server

Client

Distributed Databases

Location C

Location B

computer

computer

computer

Location A

Homogeneous Databases

Distributed Databases

Heterogeneous Or Federated Databases

Database Server

Client

Local Network

Comm Server

Client

Remote Comp.

Remote Comp.
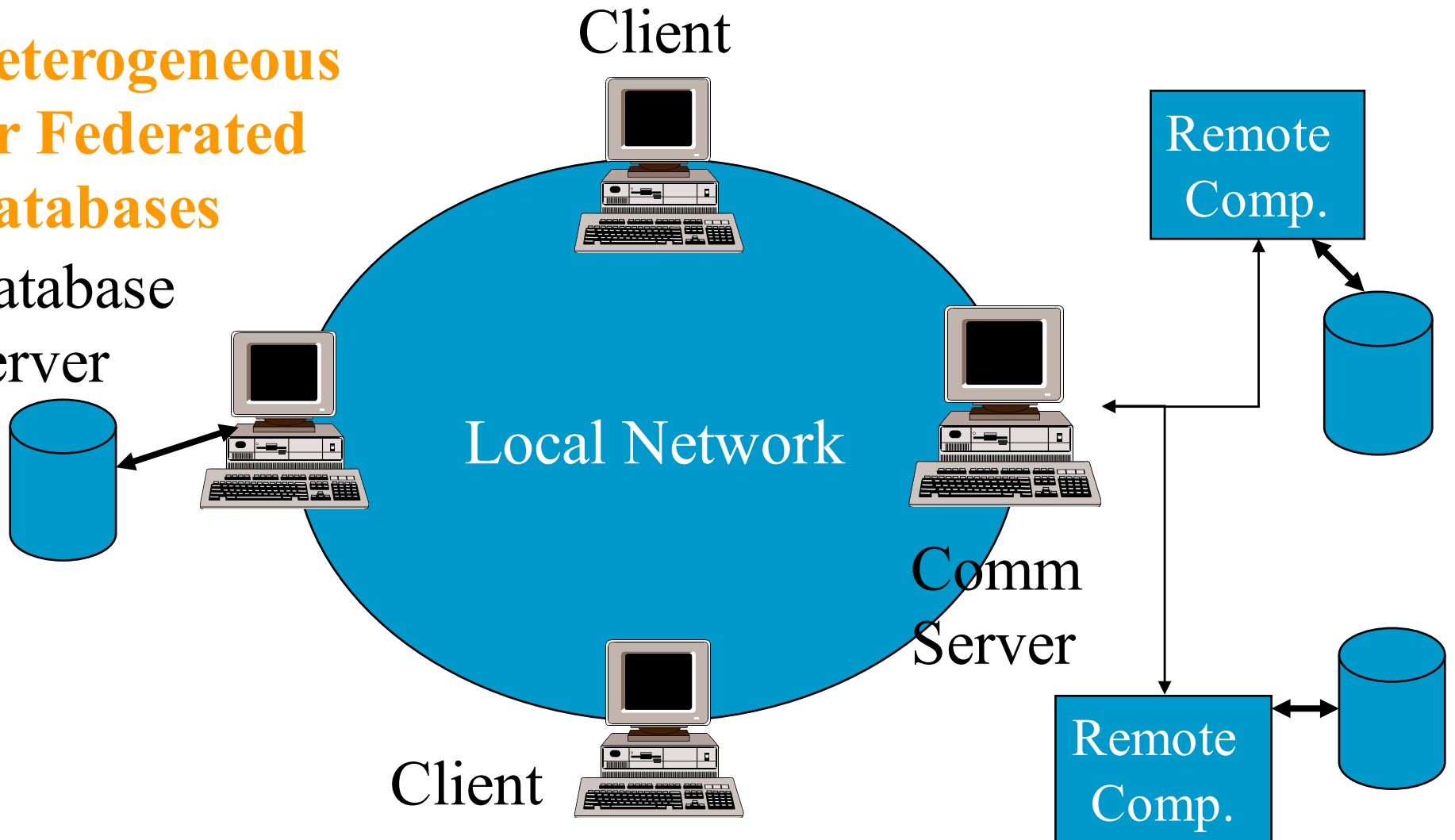
Terms and Concepts

- ***Database Application***
  - An application program (or set of related programs) that is used to perform a series of database activities:
    - Create
    - Read
    - Update
    - Delete
    - On behalf of database users

Range of Database Applications

- PC databases
  - Usually for individual
- WorkGroup databases
  - Small group use where everyone has access to the database over a LAN
- Departmental databases
  - Larger than a workgroup – but similar
- Enterprises databases
  - For the entire organization over an intranet (or sometimes the internet)

Terms and Concepts

- Database activities:
  - *Create*
    - Add new data to the database
  - *Read*
    - Read current data from the database
  - *Update*
    - Update or modify current database data
  - *Delete*
    - Remove current data from the database

Terms and Concepts

- ***Enterprise***
  - Organization
- ***Entity***
  - Person, Place, Thing, Event, Concept...
- ***Attributes***
  - Data elements (facts) about some entity
  - Also sometimes called fields or items or domains
- ***Data values***
  - instances of a particular attribute for a particular entity

Terms and Concepts

- ***Records***
  - The set of values for all attributes of a particular entity
  - AKA "tuples" or "rows" in relational DBMS
- ***File***
  - Collection of records
  - AKA "Relation" or "Table" in relational DBMS

Terms and Concepts

- ***Key***
  - an attribute or set of attributes used to identify or locate records in a file

- ***Primary Key***
  - an attribute or set of attributes that *uniquely* identifies each record in a file

Terms and Concepts

- *DA*
  - Data adminstrator - person responsible for the Data Administration function in an organization
  - Sometimes may be the CIO -- Chief Information Officer
- *DBA*
  - Database Administrator - person responsible for the Database Administration Function

Terms and Concepts

- ***Data Administration***
  - Responsibility for the overall management of data resources within an organization

- ***Database Administration***
  - Responsibility for physical database design and technical issues in database management

- ***Data Steward***
  - Responsibility for some subset of the organization's data, and all of the interactions (applications, user access, etc.) for that data

Terms and Concepts
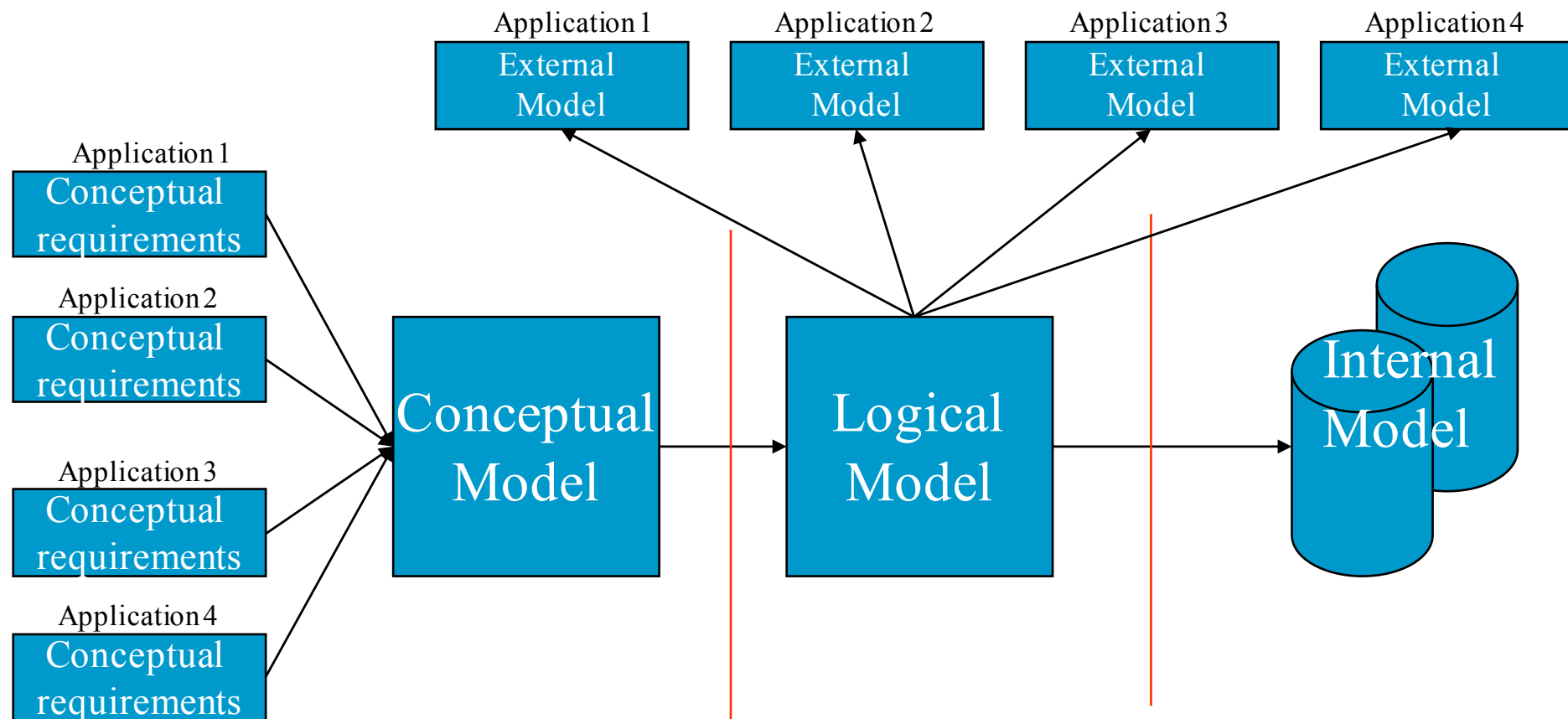
- ***Models***
  - (1) Levels or views of the Database
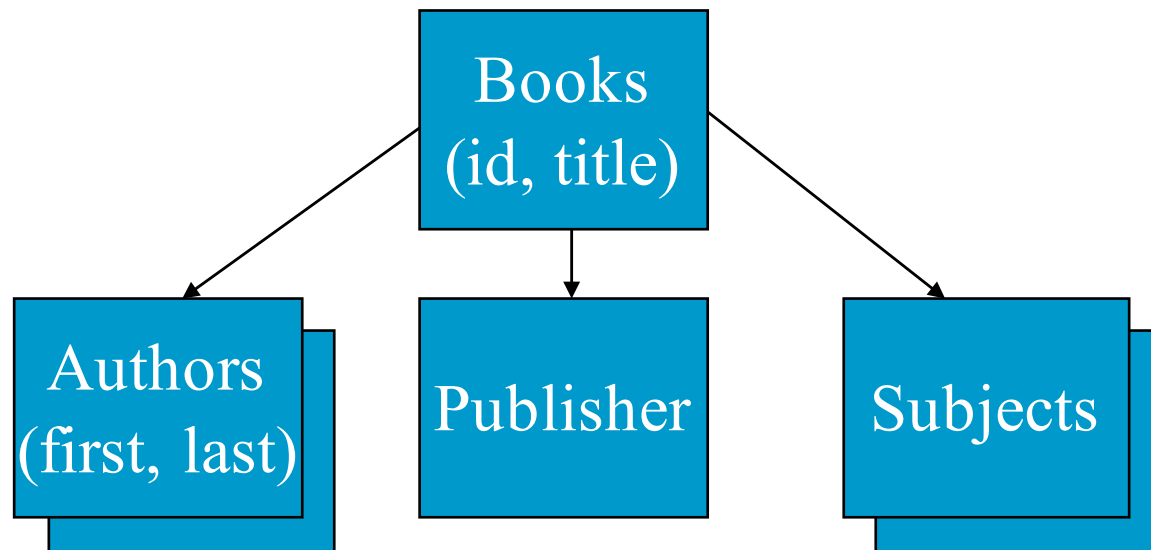   - Conceptual, logical, physical

  - (2) DBMS types
    - Relational, Hierarchic, Network, Object-Oriented, Object-Relational

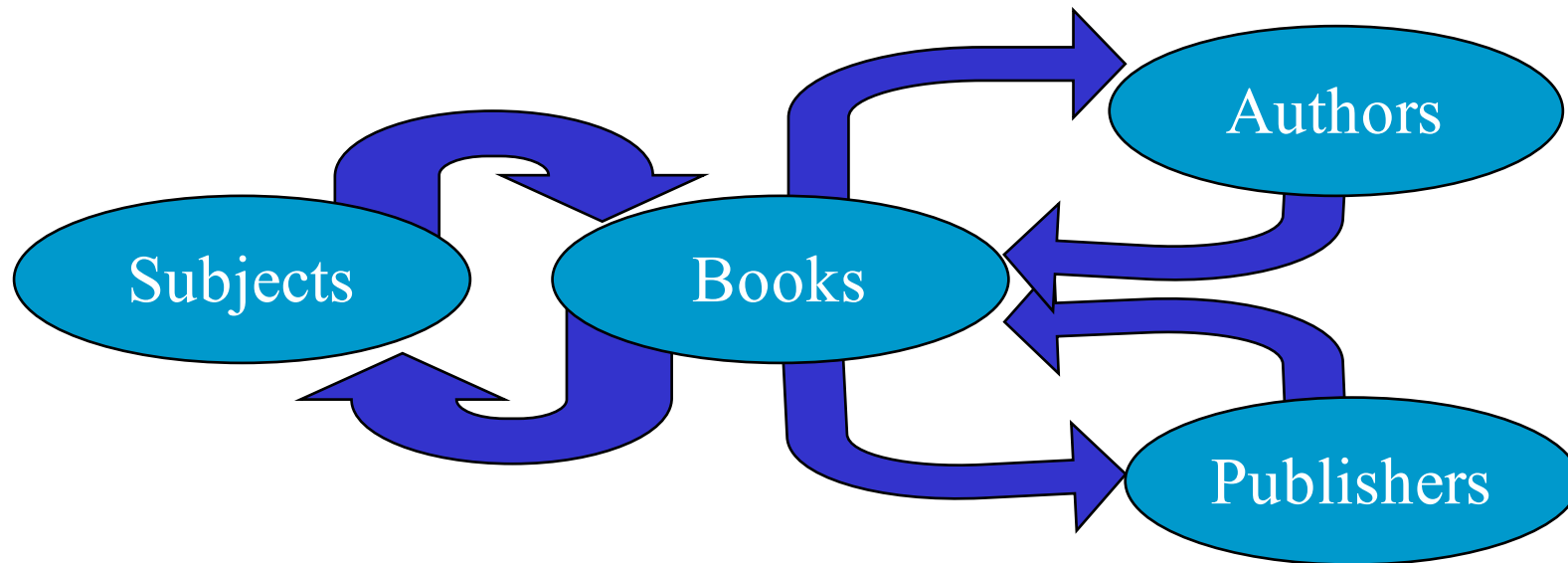# Models (1)

Data Models(2): History

- Hierarchical Model (1960's and 1970's)
  - Similar to data structures in programming languages.

Data Models(2): History

- ## Network Model (1970's)
  - Provides for single entries of data and navigational "links" through chains of data.

Data Models(2): History

- ## Relational Model (1980's)
  - Provides a conceptually simple model for data as relations (typically considered "tables") with all data visible.

Book ID	Title	pubid	Author id
1	Introductio	2	1
2	The history	4	2
3	New stuff a	3	3
4	Another tit	2	4
5	And yet m	1	5

pubid	pubname
1	Harper
2	Addison
3	Oxford
4	Que

Authorid	Author nan
1	Smith
2	Wynar
3	Jones
4	Duncan
5	Applegate

Book ID	Subid
1	2
2	1
3	3
4	2
4	3

| Subid	Subject
1	cataloging
2	history
3	stuff

Data Models(2): History

- Object Oriented Data Model (1990's)
  - Encapsulates data and operations as "Objects"
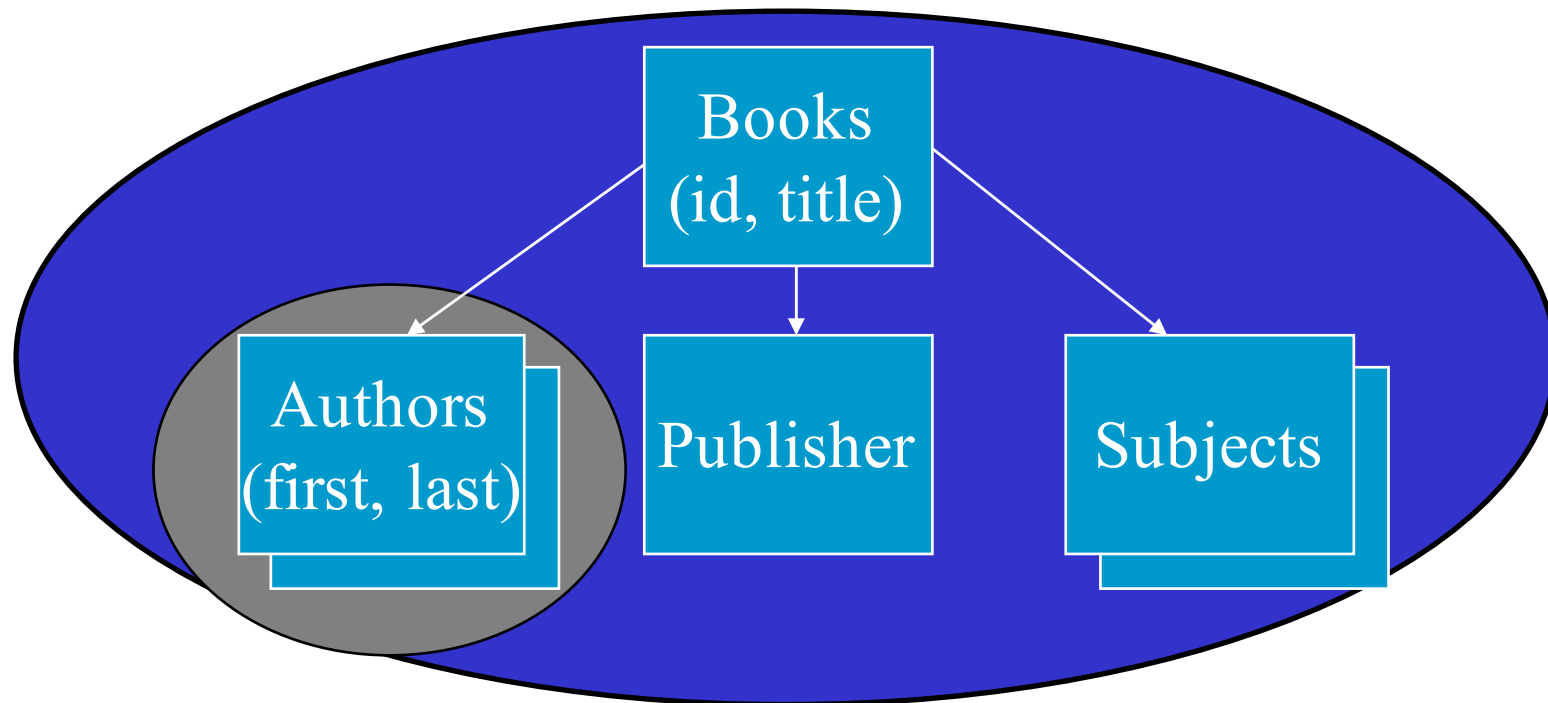
Data Models(2): History

- ## Object-Relational Model (1990's)
  - Combines the well-known properties of the Relational Model with such OO features as:
    - User-defined datatypes
    - User-defined functions
    - Inheritance and sub-classing

Database System Life Cycle
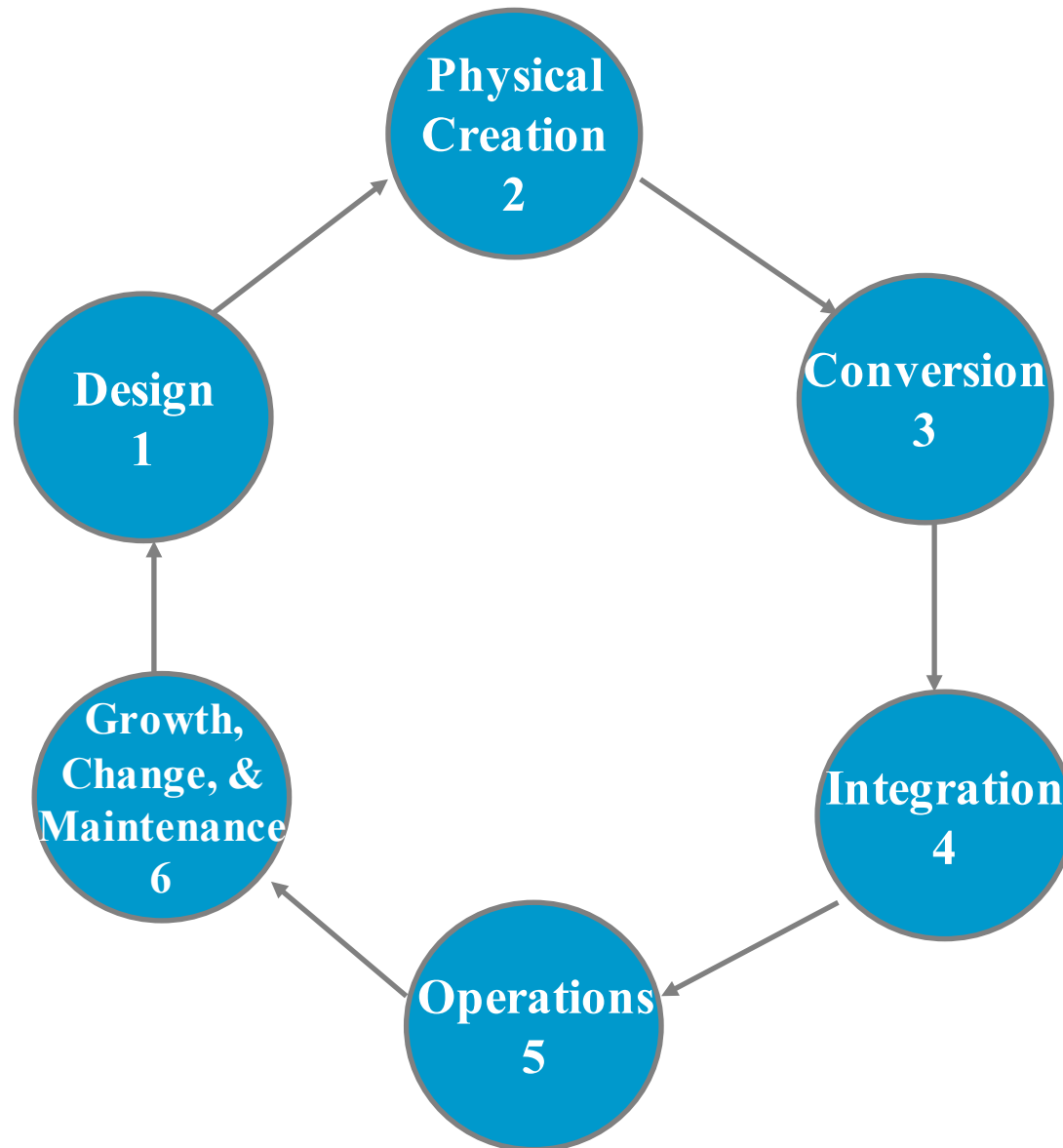
Design

- Determination of the needs of the organization
- Development of the Conceptual Model of the database
  - Typically using Entity-Relationship diagramming techniques
- Construction of a Data Dictionary
- Development of the Logical Model

Physical Creation

- Development of the Physical Model of the Database
  - data formats and types
  - determination of indexes, etc.
- Load a prototype database and test
- Determine and implement security, privacy and access controls
- Determine and implement integrity constraints

Conversion

- Convert existing data sets and applications to use the new database
    - May need programs, conversion utilities to convert old data to new formats.

Integration

- Overlaps with Phase 3
- Integration of converted applications and new applications into the new database
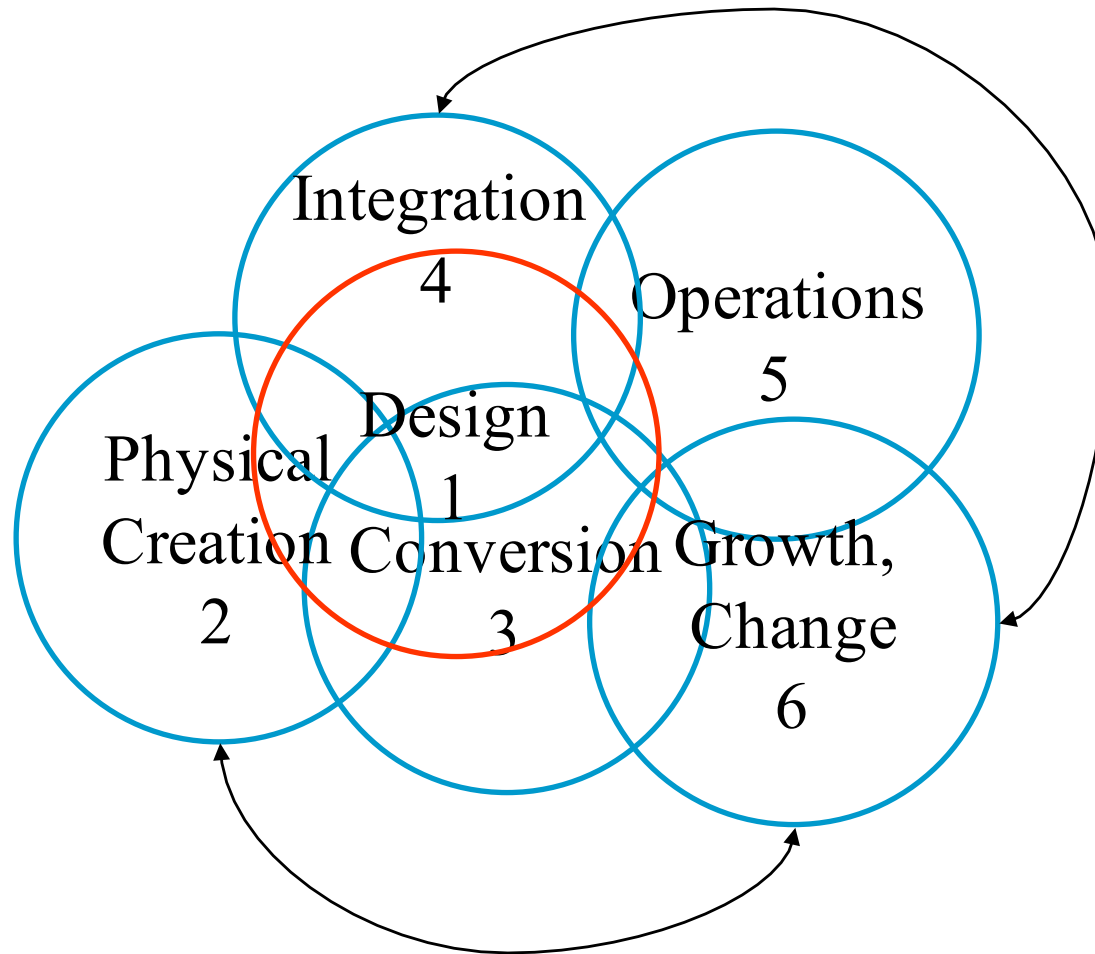
Operations

- All applications run full-scale
- Privacy, security, access control must be in place.
- Recovery and Backup procedures must be established and used

Growth, Change & Maintenance

- Change is a way of life
  - Applications, data requirements, reports, etc. will all change as new needs and requirements are found
  - The Database and applications and will need to be modified to meet the needs of changes

Integration
4

Operations
5

Design
1

Physical
Creation
2

Conversion
3

Growth,
Change
6

Next Time