

CW

ZY

N

MY

Chengzhang Wu

Zihan Yan

newto

Murad Yassin

Elements of Database Systems

Chapter 6

4/2/2020

Rutgers Business School

1

CW

ZY

N

MY

RK

Chengzhang Wu

Zihan Yan

newto

Murad Yassin

Ronak Kachhia

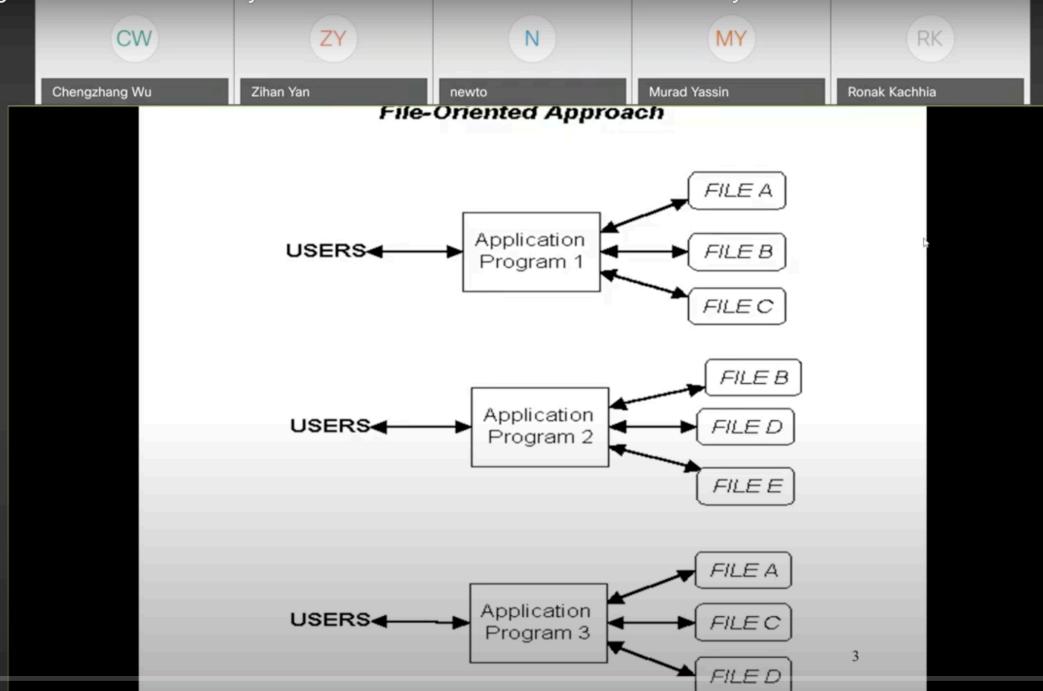
The file-oriented approach

- The file-oriented approach involves creating a set of files, as needed, for each transaction processing application such as sales or purchases.
- What are the problems with such approach?

Rutgers Business School

2

+1



Drawbacks of file-oriented approach

Drawbacks of the File-Oriented Approach	
Drawback	Explanation
Data redundancy	Duplication of data (files) across applications poses a data maintenance problem and can potentially cause problems of data inconsistencies. Excessive duplication also results in high data storage costs.
Proliferation of files	The task of maintaining files can become very complex as the number of applications multiplies, since each application creates its own set of files.
Lengthy application development	Inability to share data in existing files increases the time required to create new applications, since all files needed for the new application must be independently created.
Lack of data independence	Data structures and the procedures that modify the data are both defined within the same program (in the "Data" and "Procedure" divisions). Data structures and/or procedures cannot be independently modified.

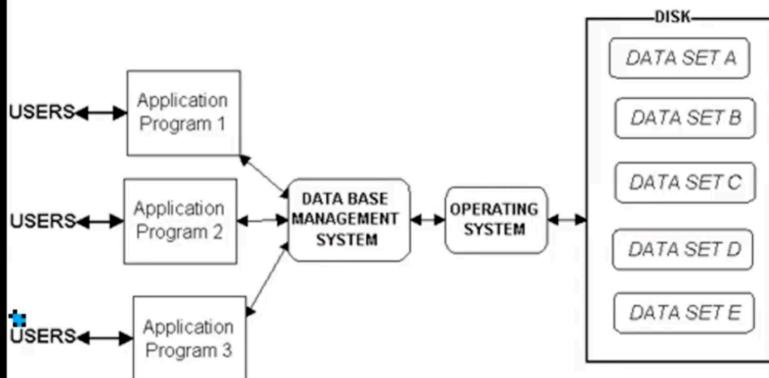
The database approach

- The data base approach centers around creating an organization wide repository of data that all applications and all users can share.
- Each set of data is uniquely stored in the DB.
- Concurrent access to the same data set is handled by the data base management system (DBMS).

Rutgers Business School

5

Database Approach



Rutgers Business School

6

Database

- A data base is an integrated repository of an organization's data containing a series of interrelated data sets.
- The data sets are designed to store data about entities such as customers, employees, and vendors and also events such as sales.
- The repository is integrated in that there is no duplication of data sets
- The data sets are interrelated in that common attributes exist between data sets to signify relationships between entities and events.
- Coordination among the data sets involves ensuring that updates to one data set do not result in data inconsistencies in related data sets.
- The tasks of creating, updating, and managing data sets is handled by the data base management system (DBMS).

Rutgers Business School

7

Management Information System-20200326-Elements of Database System

Advantages of database approach

Advantage	Explanation
Data redundancy virtually eliminated	Each data set is stored just once in the repository, thereby reducing data storage costs.
No data inconsistencies	Since each data item is stored only once, there cannot be multiple versions of that data item.
Data independence	Data structures are defined separately from the application programs. Changes can be made to data structures without having to modify application programs.
Rapid application development	Ability to share data shortens the time required to create new applications.
Centralized backup, control, and security	Backup, control, and security tasks are all handled centrally by the database management system.

Rutgers Business School

8

N ZY CW MY RK
newto Zihan Yan Chengzhang Wu Murad Yassin Ronak Kachhia

Drawbacks of database approach

<i>Disadvantage</i>	<i>Explanation</i>
Complexity of DBMS administration	The administration of large scale database systems requires significant resources and expertise
Data integrity using complex DBMS features	Configuring the database to insure data integrity requires considerable expertise and intricate knowledge of DBMS features. Accountants and auditors must familiarize themselves with control and security concerns for DBMS and how these can be implemented in database environments
Data accessible only through DBMS	Accountants and auditors must be competent in using the DBMS to access data for the purpose of generating useful information and fulfilling audit objectives
Centralized backup, control, and security	Backup, control and security are typically centralized, potentially making the organization vulnerable to a hacker who can break through the central security shield

Rutgers Business School 9 +6

Management Information System-20200326-Elements of Database System

N ZY CW MY RK
newto Zihan Yan Chengzhang Wu Murad Yassin Ronak Kachhia

Data hierarchy

```

graph BT
    Bit[Bit] --> Byte[Byte]
    Byte --> Field[Field]
    Field --> Record[Record]
    Record --> DataSet[Data set (file/table)]
    DataSet --> Database[Database]
  
```

Rutgers Business School 10 +6

|| ▶ 🔍 30:39 / 1:28:48



N newto	ZY Zihan Yan	CW Chengzhang Wu	MY Murad Yassin	RK Ronak Kachhia
------------	-----------------	---------------------	--------------------	---------------------

Fundamental database concepts

- A **primary key** is a unique identifier of a record in a file.
- A **composite key** or a concatenated key are used when multiple fields are required to uniquely identify a record.
- Every field in a data set other than the primary key is referred to as a **non-key attribute**.
- A **foreign key** is a field in a data set which is the primary key in a related data set.

Rutgers Business School

11



N newto	ZY Zihan Yan	CW Chengzhang Wu	MY Murad Yassin	RK Ronak Kachhia
------------	-----------------	---------------------	--------------------	---------------------

Relationships

- Relationships between entities and events, can be of three types:
 - One-to-one (1:1)
 - One-to-many (1:M)
 - Many-to-many (M:M)

Rutgers Business School

12





N newto ZY Zihan Yan CW Chengzhang Wu MY Murad Yassin RK Ronak Kachhia

Relational Model Overview

- A **relation** is simply a two-dimensional table with rows and columns.
- The rows, also referred to as "tuples," are the **records** in the data set
- The columns are fields or **attributes**.
- Many RDBMS are available today, Oracle, Informix, MS SQL Server, DB2, and Sybase.

Rutgers Business School

13

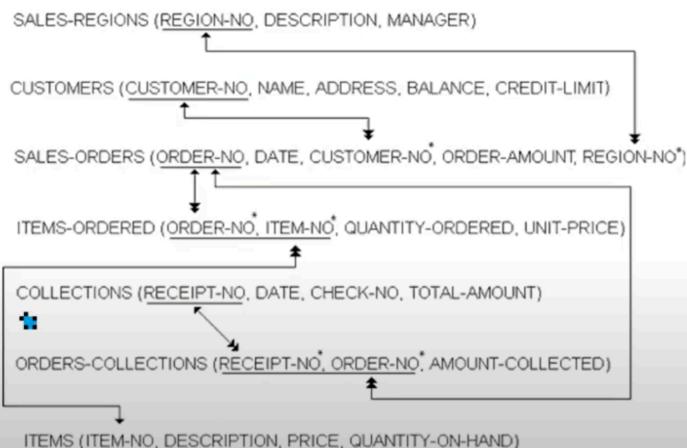
|| ▶ 🔍 46:14 / 1:28:48

CC ⚙️ ⌂ +6



N newto ZY Zihan Yan CW Chengzhang Wu MY Murad Yassin RK Ronak Kachhia

RELATIONAL MODEL



Rutgers Business School

14

|| ▶ 🔍 50:04 / 1:28:48

CC ⚙️ ⌂ +6



N

ZY

CW

MY

RK

newto

Zihan Yan

Chengzhang Wu

Murad Yassin

Ronak Kachhia

Tables rules

- Each table in the data base must have a unique name.
- Duplicate columns and rows are not permitted within a table.
- The sequence of rows and columns is immaterial
- Every table must have a designated primary key - a unique identifier of every row in the table
- A table could have more than one unique identifiers, but only one must be chosen as the primary key
- Relationships between tables are represented using common fields between them

Rutgers Business School

15



1:00:19 / 1:28:48



N

ZY

CW

MY

RK

newto

Zihan Yan

Chengzhang Wu

Murad Yassin

Ronak Kachhia

Entity and referential integrity

- **Entity integrity** means that the primary key field (or fields in case of a composite key) in a table cannot be "Null" and must be unique. (RDBMS are equipped with features to enforce it.)
- **Referential integrity** means that foreign keys must either be "Null" or match an existing value in the "master" table for the foreign key.

Rutgers Business School

16



1:07:08 / 1:28:48





N	ZY	CW	MY	RK
newto	Zihan Yan	Chengzhang Wu	Murad Yassin	Ronak Kachhia

Data validation rules

- The purpose of validation rules is to **prevent** erroneous data from being entered into the table.
- Validation rules can be established for individual fields within a table to restrict the data that can be entered into the field.
- Rules can also be specified to ensure valid relationships between fields

Validation Rules	
<i>Range test</i>	Greater than a minimum and/or less than a maximum value?
<i>Validity test</i>	One of the acceptable values for this field?
<i>Length test</i>	Correct number of digits entered?
<i>Valid combinations test</i>	Correct mathematical or logical relationship between fields in a table?

Rutgers Business School

17

|| ▶ ◀ 1:17:31 / 1:28:48

CC Settings Share


N	ZY	CW	MY	RK
newto	Zihan Yan	Chengzhang Wu	Murad Yassin	Ronak Kachhia

Restricting access to tables

- Access restrictions can be defined at the database level for controlling access to sensitive data.
- Certain tables can be made accessible to all users except for those listed, or only to the listed users.
- Another method of restricting access to sensitive data is to create a **view**.
- Apart from hiding certain columns, rows that contain sensitive data can also be hidden.

Rutgers Business School

18

|| ▶ ◀ 1:24:43 / 1:28:48

CC Settings Share



N

ZY

CW

MY

RK

newto

Zihan Yan

Chengzhang Wu

Murad Yassin

Ronak Kachhia

Data dictionary

- Data dictionary contains a variety of information about the contents of the database.
- Data dictionary may include names of all tables, the columns (attributes) contained in each table along with their data formats, and the privileges held by each user authorized to access the database.

Rutgers Business School

19

6



1:25:31 / 1:28:48



N

ZY

CW

MY

RK

newto

Zihan Yan

Chengzhang Wu

Murad Yassin

Ronak Kachhia

Languages for RDBMS

- The data definition language (**DDL**) is used to program the data base schema.
 - create tables
 - define authorized users of each table
 - specify validation rules for individual tables.
- Data manipulation language (**DML**) is used to allow the programs to access tables in a database whenever required.
- Data query language (**DQL**) is used by end users to perform ad hoc queries on the database.
 - Query By Example" (QBE)
 - Structured Query Language (**SQL**)
- Two additional tools that are provided by RDBMS are report writers and form editors.

Rutgers Business School

20

6



1:27:03 / 1:28:48

