NEPAL COLLEGE OF INFORMATION TECHNOLOGY

BALKUMARI LALITPUR

(Affiliated To Pokhara University)

ASSIGNMENT – 1

SUBJECT : Database Management System

Submitted By :
Name : Pradip Dhungana
Roll No : 201751
Semester : 4th

Submitted To :
Instructor : Amit Shrivastava
Department of Software

1. What do you mean DBMS? List some significant differences between a file-processing system and a DBMS.

Answer :

A data-base management system (DBMS) is a generalized tool for manipulating large databases. It is made available through special software for the interrogation, maintenance, and analysis of data. Its interface generally provide a broad range of language to aid all users from clerk to data administrator.

FILE SYSTEM	DBMS
• The file system is software that manages and organizes the files in a storage medium within a computer.	• DBMS is software for managing the database.
• Redundant data can be present in a file system.	• In DBMS there is no redundant data.
• It doesn't provide backup and recovery of data if it is lost.	• It provides backup and recovery of data even if it is lost.
• There is no efficient query processing in the file system.	• Efficient query processing is there in DBMS.
• There is less data consistency in the file system.	• There is more data consistency because of the process of normalization.
• It is less complex as compared to DBMS.	• It has more complexity in handling as compared to the file system.
• File systems provide less security in comparison to DBMS.	• DBMS has more security mechanisms as compared to file systems.
• It is less expensive than DBMS.	• It has a comparatively higher cost than a file system.
• There is no data independence.	• In DBMS data independence exists.
• Only one user can access data at a time.	• Multiple users can access data at a time.
• The user has to write procedures for managing databases.	• The user not required to write procedures.
• Data is distributed in many files. So, not easy to share data.	• Due to centralized nature sharing is easy.
• It give details of storage and representation of data.	• It hides the internal details of Database.
• Integrity Constraints are difficult to implement.	• Integrity constraints are easy to implement.
• Example : Cobol, C++	• Example : Oracle, SQL Server

2. List five responsibilities of a database management system. For each responsibility, explain the problems that would arise if the responsibility were not discharged.

Answer :

Responsibilities of a Database Management System (DBMS) and the problems that would arise if these responsibilities were not discharged are as follows:

I. **Data Security and Access Control :**

Responsibility: A DBMS ensures the security and integrity of data by implementing access control mechanisms, such as user authentication, authorization, and encryption.

Problems if not discharged: Without proper data security and access control, unauthorized users could gain access to sensitive information, leading to data breaches, unauthorized modifications, and compromised data integrity.

II. **Data Integrity and Consistency :**

Responsibility: A DBMS enforces integrity constraints, such as primary key constraints and referential integrity, to maintain the accuracy and consistency of data. Problems if not discharged: Without proper enforcement of integrity constraints, data integrity and consistency may be compromised. Inaccurate or inconsistent data could lead to incorrect results, data corruption, and unreliable decision-making.

III. **Data Backup and Recovery :**

Responsibility: A DBMS performs regular data backups and provides mechanisms for data recovery in case of failures or disasters.

Problems if not discharged: Without regular backups and proper recovery mechanisms, data loss becomes a significant risk. Any accidental or unforeseen event, such as hardware failure, software bugs, or natural disasters, could result in permanent data loss, making it difficult or impossible to restore the database to its previous state.

IV. **Data Concurrency and Transaction Management :**

Responsibility: A DBMS manages concurrent access to the database by multiple users or applications, ensuring data consistency and preventing conflicts using transaction management techniques.

Problems if not discharged: Without effective concurrency control and transaction management, data inconsistencies and conflicts can occur. Concurrent updates from multiple users may result in data inconsistencies and loss of data integrity. Incomplete or incorrect transactions could lead to financial losses, incorrect reporting, and customer dissatisfaction.

V. Database Performance Optimization :

Responsibility: A DBMS optimizes database performance by managing storage structures, query optimization, indexing, and caching mechanisms.

Problems if not discharged: Without performance optimization, a database may suffer from slow response times, inefficient resource utilization, and degraded user experience. Queries and operations could take longer to execute, leading to decreased productivity, scalability issues, and increased hardware costs.

3. What are main functions of a database administrator (DBA)? List six major steps that you would take in setting up a database for a particular enterprise.

Answer :

Database Administrator is a person having central control over data and programs acessing that data. The database administrator is a manager whose responsibilities are focused on management of technical aspects of the database system. The functions of the database administrator are summarized as follows :

- Authorizing access to the database.
- Coordinating and monitoring its use.
- Acquiring hardware and software resources as needed.
- Backup and recovery. DBA has to ensure regular backup of the database, in-case of damage, suitable recovery procedure are used to bring the database up with little downtime as possible.

The major steps to design a database for a particular enterprise are :

1) First find out the requirements of the user :
   - Identify and gather the specific needs and expectations of the users or stakeholders.
   - Conduct interviews, surveys, and meetings to understand their data management requirements.

2) Design a view for each important application :
   - Analyze the requirements gathered in the previous step and identify key applications that rely on the database.
   - Design views or interfaces that cater to the specific needs of each application, ensuring data accessibility and usability.

3) Integrate the views giving the conceptual schema, which is the union of all views :
   - Combine and integrate the individual views designed for each application into a unified conceptual schema.
   - Ensure that the conceptual schema represents the collective requirements of all applications and provides a holistic view of the database.

4) Map to the data model provided by the DBMS (usually relational).
   - Choose a suitable data model provided by the selected database management system (usually relational).
   - Map the conceptual schema designed in the previous step to the chosen data model, identifying tables, relationships, and attributes.

5) Design external views.
   - ◆ Determine additional external views or interfaces required for users or applications that need customized or filtered access to the data.
   - ◆ Design these external views to provide the desired data representation, security measures, and user-friendly functionalities.

6) Choose physical structures (indexes, etc.)
   - ◆ Determine the physical structures required for efficient data storage and retrieval, such as indexes, partitions, or clustering.
   - ◆ Consider performance requirements, data volume, and query patterns to select appropriate physical structures and optimize database performance.

4. Explain the distinctions among the terms primary key, candidate key, and superkey.

Answer :

KEYS:

Keys are attributes or group of attributes which are used for uniquely identifying a row in a relation . Keys are broadly classified into :

- Super key
- Candidate key
- Primary key

I. SUPER KEY :
Super keys are mainly constraints on relations. It doesn't allow two entities to have same values for thatparticular subset of attributes.

II. CANDIDATE KEY :
Candidate key is minimal super key. In a relational schema candidate key is that minimal set of attributes whicuniquely identify tuples of agreeing relation.

III. PRIMARY KEY :
A primary key is a column or a set of columns in a table that uniquely identifies each record (row) in the table. It serves as a unique identifier and provides a way to distinguish one record from another.

Distinctions :

Let’s take a example table “STUDENTS” with attributes and datatypes :

Attribute	Data Type
StudentID	Integer
StudentName	String
Email	String
Phone	String
RollNumber	String
Age	Integer

I. **Superkey:** {StudentID, StudentName, Email, Phone, RollNumber, Age}

The superkey includes all attributes of the table and can uniquely identify each student. However, it contains redundant attributes.

II. **Candidate Key:** {StudentID}

The candidate key is a minimal superkey without redundant attributes. In this example, StudentID is the only attribute that is required to uniquely identify each student.

III. **Primary Key:** {StudentID}

The primary key is a specific candidate key chosen as the main identifier for the table. In this example, StudentID is selected as the primary key, uniquely identifying each student and serving as the primary means of identification.

5. A university registrar's office maintains data about the following entities:

(a) Courses, including number, title, credits, syllabus, and prerequisites;

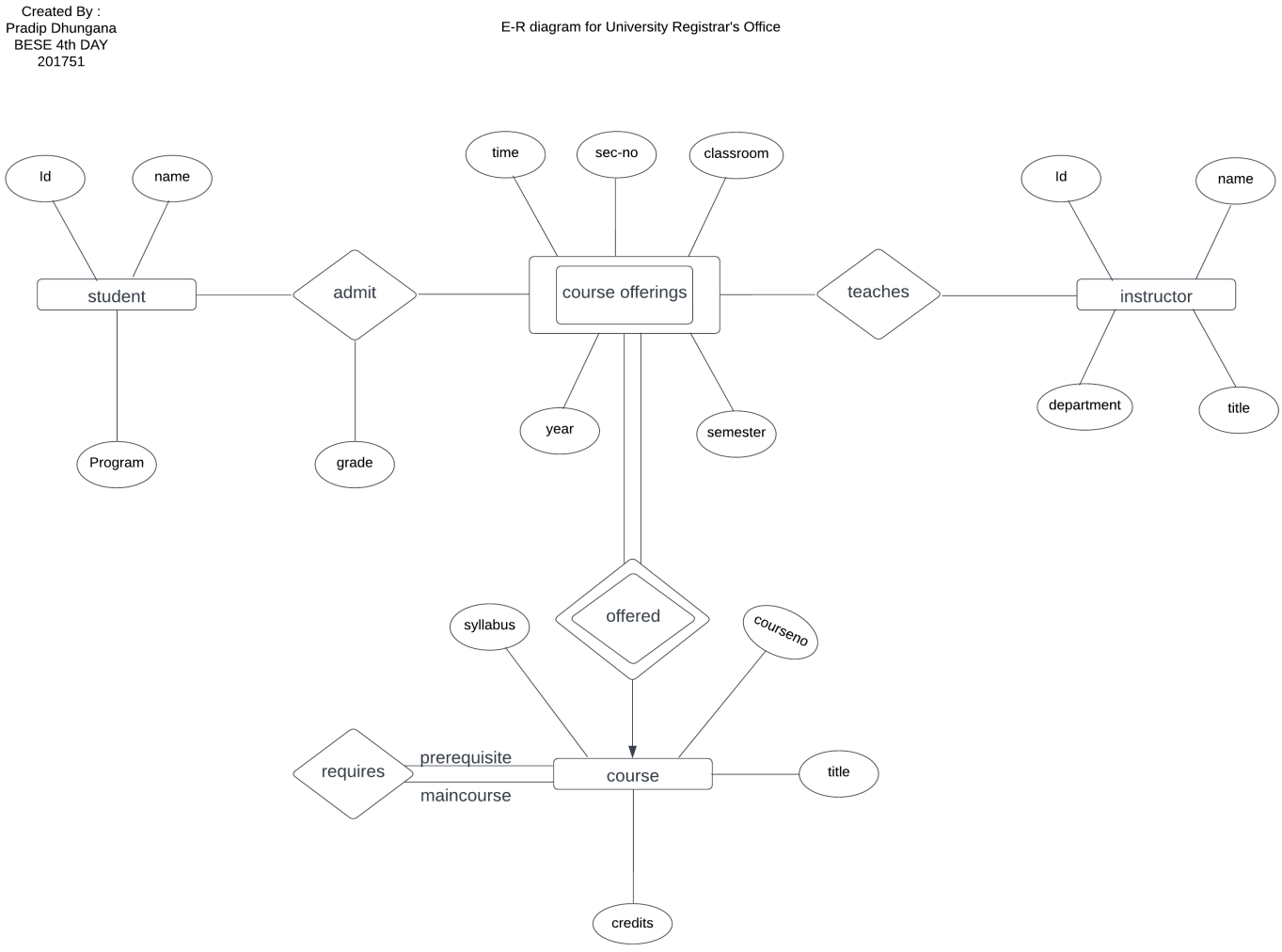
(b) Course offerings, including course number, year,semester, section number, instructor(s), timings, and classroom;

(c) Students, including student-id, name, and program; and

(d) Instructors, including identification number, name, department, and title.

Further, the enrollment of students in courses and grades awarded to students in each course they are enrolled for must be appropriately modeled. Construct an E-R diagram for the registrar's office.Document all assumptions that you make about the mapping constraints.

Answer :



6. Explain the difference between a weak and a strong entity set.

Answer :

Weak Entity set :

A weak entity set is an entity set that cannot exist without being associated with a corresponding strong entity set.

Strong Entity set :

A strong entity set is an entity set that can exist independently without any dependencies on other entity sets.

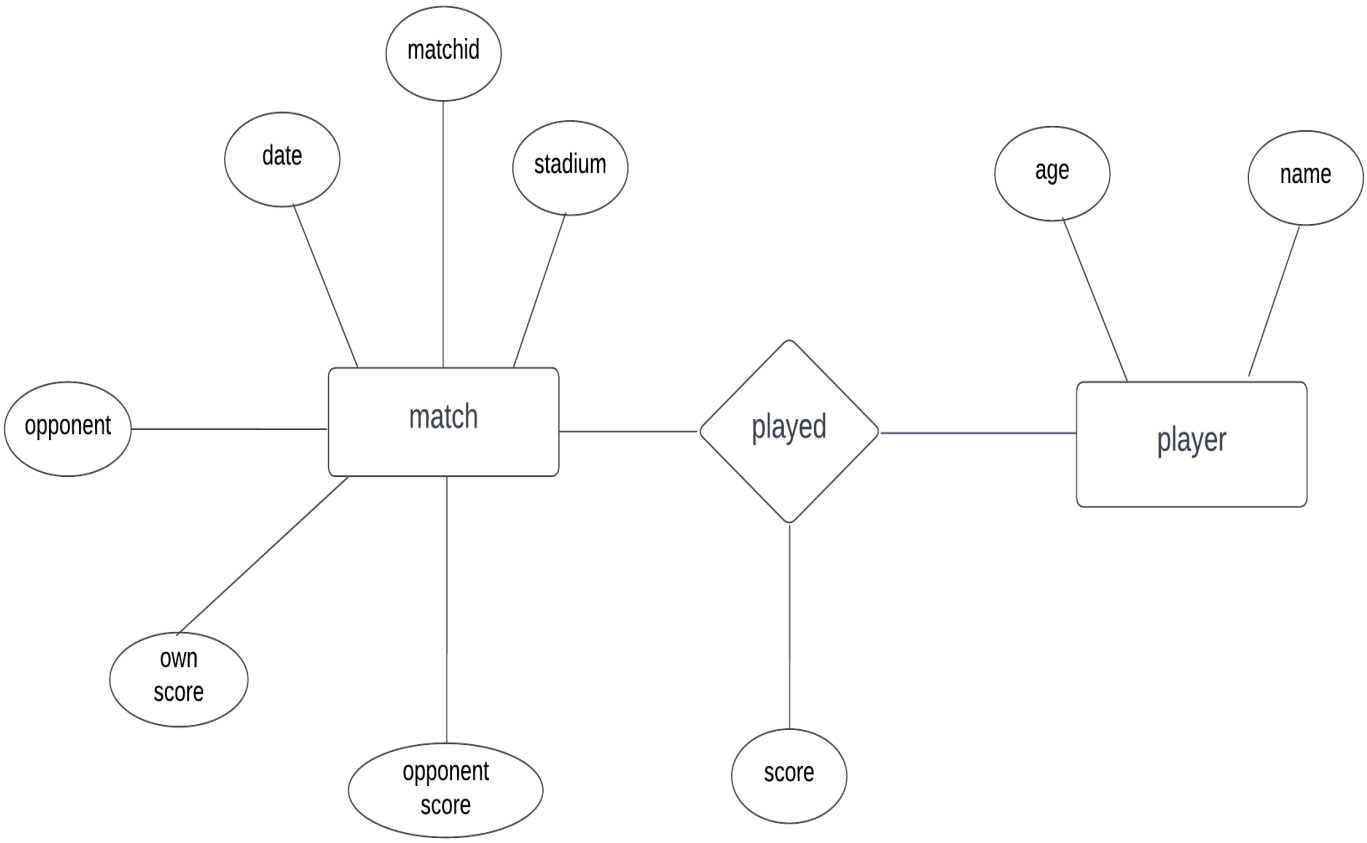
Weak Entity Set	Strong Entity Set
◆ Existence and identifier dependent	◆ Independent of other entities
◆ Double-lined rectangle	◆ Single-lined rectangle
◆ Identifier includes relationship to strong entity	◆ Has its own unique identifier
◆ Always (1,1) with respect to strong entity types	◆ Varies based on the relationship and business context
◆ Requires existence of corresponding strong entity	◆ Can exist independently without dependencies
◆ Cannot be identified without the strong entity	◆ Can be identified without dependencies
◆ Subset of attributes of the strong entity set	◆ Own set of attributes
◆ Directed edge from weak entity to strong entity	◆ No directed edge

7. Design an E-R diagram for keeping track of the exploits of your favorite sports team. You should store the matches played, the scores in each match, the players in each match and individual player statistics for each match. Summary statistics should be modeled as derived attributes.

Answer :

Created By :
Pradip Dhungana
BESE 4th DAY
201751

E-R diagram for Sports Team

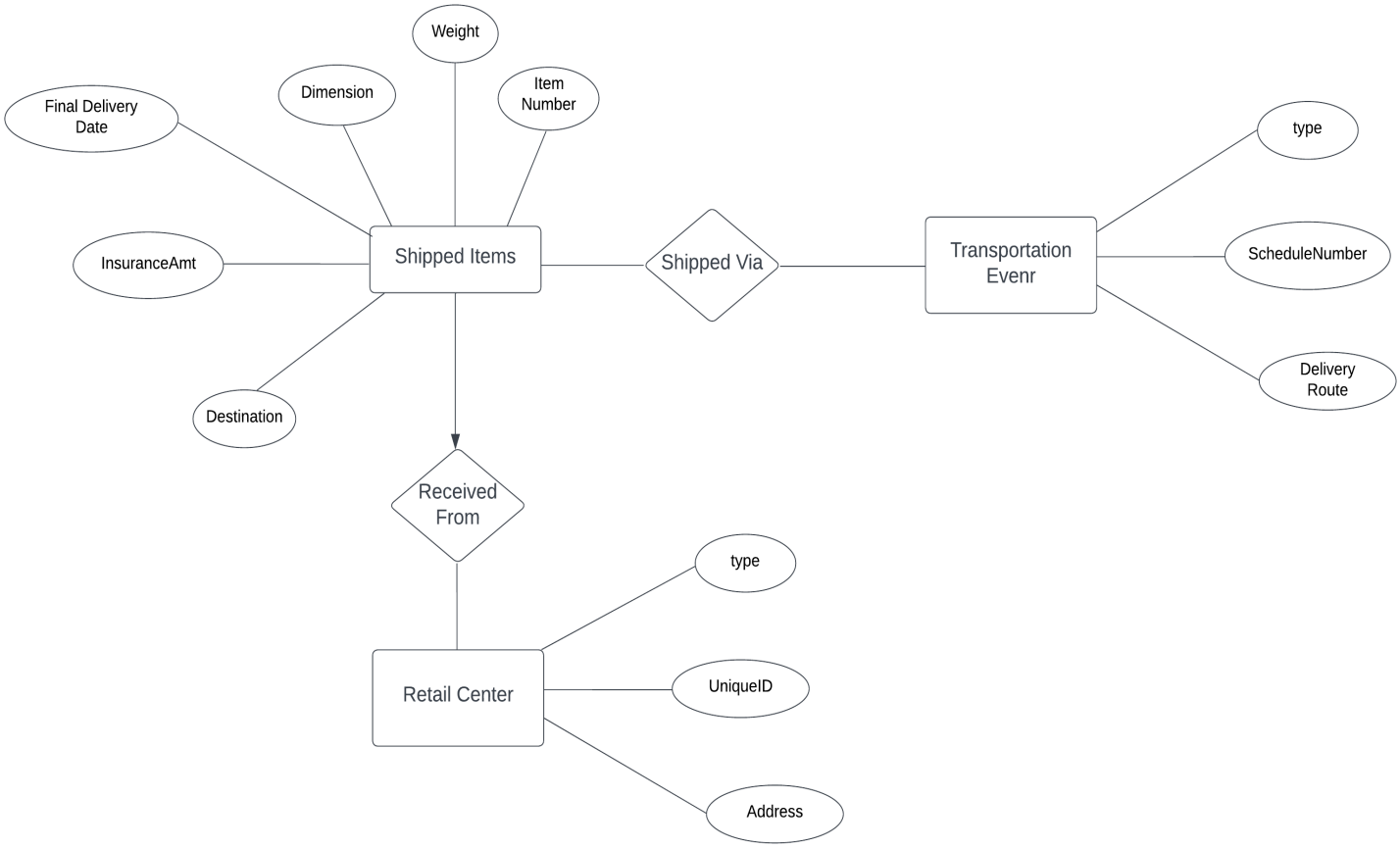


8. UPS prides itself on having up-to-date information on the processing and current location of each shipped item. To do this, UPS relies on a company-wide information system. Shipped items are the heart of the UPS product tracking information system. Shipped items can be characterized by item number (unique), weight, dimensions, insurance amount, destination, and final delivery date. Shipped items are received into the UPS system at a single retail center. Retail centers are characterized by their type, uniqueID, and address. Shipped items make their way to their destination via one or more standard UPS transportation events (i.e., flights, truck deliveries). These transportation events are characterized by a unique scheduleNumber, a type (e.g, flight, truck), and a deliveryRoute. Please create an Entity Relationship diagram that captures this information about the UPS system. Be certain to indicate identifiers and cardinality constraints.

Answer :

Created By :
Pradip Dhungana
BESE 4th DAY
201751

E-R diagram for UPS Shipping System



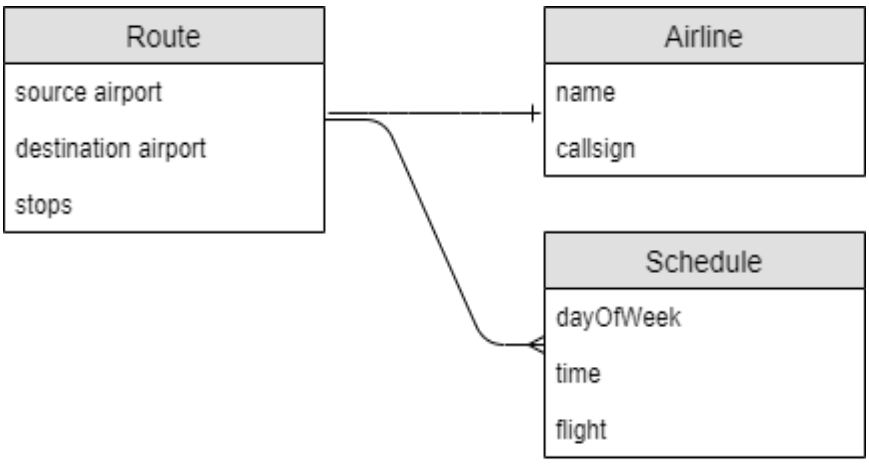
9. Compare between Conceptual, Logical and Physical Data models.

Answer :

Conceptual, logical, and physical data models are three different levels of abstraction in database modeling, each serving a specific purpose in the database design process. Here is a brief comparison between these three types of data models:

A) **Conceptual data model :**

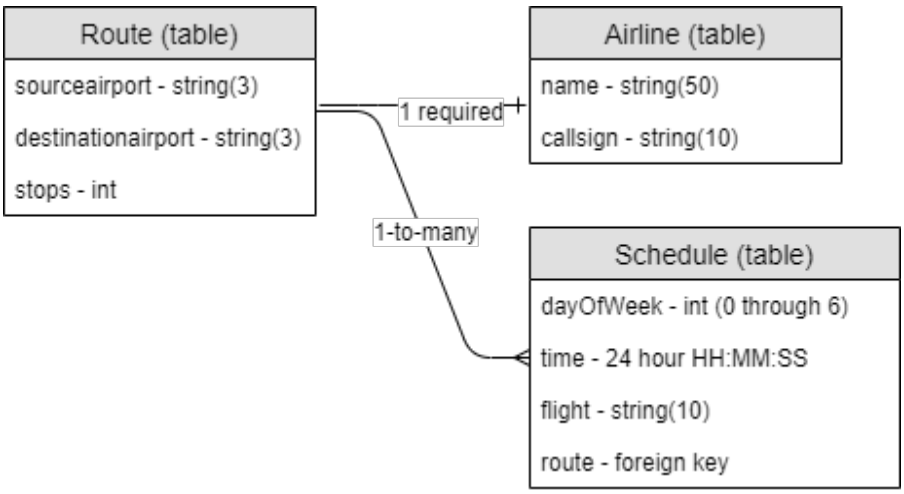
A conceptual data model provides an overview of the entire database. It defines the entities, properties, and relationships in the database without going into implementation details.



It is typically created during the requirements gathering phase of a project and is used to communicate with stakeholders in non-technical language.

B) Logical data model :

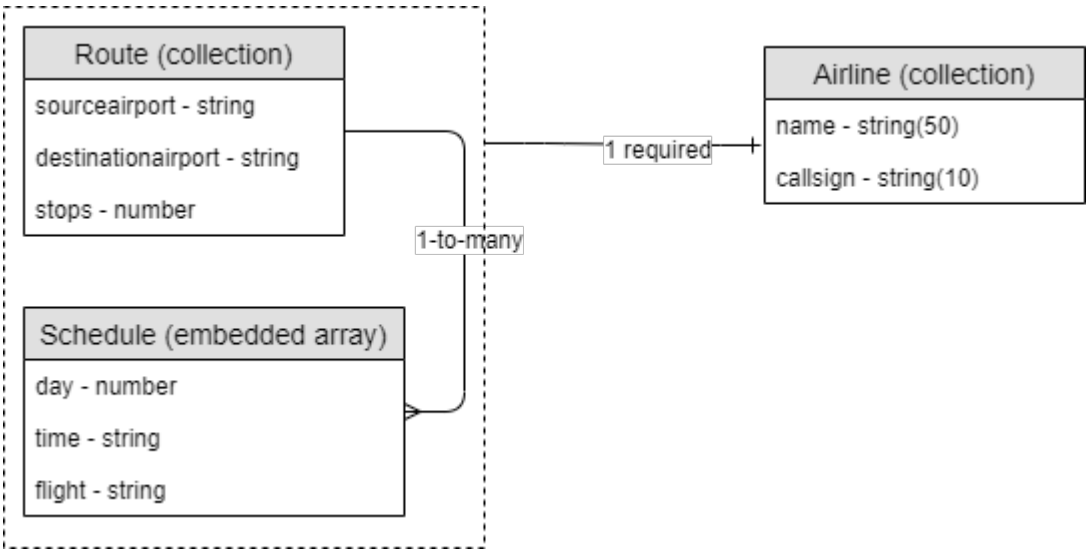
Logical data models provide more detail than conceptual models. It defines the structure of the database in terms of entities, properties, and relationships, but it does not specify how the database will be implemented.



A logical data model is used to represent data requirements and is often used as the basis for physical database design.

C) Physical data model :

A physical data model is the most detailed level of abstraction. It defines how data will be stored in the database, including tables, columns, indexes, and other database objects.



A physical data model is used by database administrators to perform database design. It provides a blueprint for creating a database and is commonly used to create a database schema.

References

1. *Fry, J.P. and Sibley, E.H., 1976. Evolution of data-base management systems. ACM Computing Surveys (CSUR), 8(1), pp.7-42.*
2. *GeeksforGeeks. (2020). Difference between File System and DBMS. [online] Available at: https://www.geeksforgeeks.org/difference-between-file-system-and-dbms/.*
3. *Jain, M.P., Comparative Study Of Relational Database Model and Resource Description Framework (RDF) Model. International Journal on Recent and Innovation Trends in Computing and Communication, 2(7), pp.1879-1882.*
4. *Jaeschke, P., Oberweis, A. and Stucky, W., 1994. Extending ER model clustering by relationship clustering. In Entity-Relationship Approach—ER'93: 12th International Conference on the Entity-Relationship Approach Arlington, Texas, USA, December 15–17, 1993 Proceedings 12 (pp. 451-462). Springer Berlin Heidelberg.*
5. *Sumathi, S. and Esakkirajan, S., 2007. Fundamentals of relational database management systems (Vol. 47). Springer.*
6. *Zachman, J.A., 2000. Conceptual, logical, physical: It is simple. Zachman International, Co, 36.*