

Express - 10

Introduction to Databases

What is a database ?

- A database is an organized collection of structured information or data, typically stored electronically in a computer system.
- It's designed to efficiently store, retrieve, and manage large amounts of data.
- Databases are controlled by a database management system (DBMS).

Types of Databases

- Relational databases: Use tables with predefined schemas (e.g., MySQL, PostgreSQL)
- Non-relational (NoSQL) databases:
 - Document-oriented (e.g., MongoDB)
 - Key-value stores (e.g., Redis)
 - Column-family stores (e.g., Cassandra)
 - Graph databases (e.g., Neo4j)
- Object-oriented databases
- Time-series databases
- Spatial databases

Importance of Database

Data organization: Efficiently structure and organize large amounts of information

Data integrity: Ensure accuracy and consistency of data

Data security: Control access and protect sensitive information

Scalability: Handle growing amounts of data and users

Performance: Quick data retrieval and processing for applications

Data analysis: Enable complex queries and data mining

Concurrent access: Allow multiple users to access and modify data simultaneously

Data persistence: Provide long-term storage of information

Integration: Facilitate data sharing between different systems and applications

Relational Database Overview

- Based on the relational model of data, introduced by E.F. Codd in 1970
- Data is organized into tables (relations) with rows (tuples) and columns (attributes)
- Use structured query language (SQL) for defining and manipulating the data
- Enforce ACID properties (Atomicity, Consistency, Isolation, Durability)
- Examples: MySQL, PostgreSQL, Oracle, Microsoft SQL Server

Non-Relational Database Overview

- Also known as NoSQL (Not Only SQL) databases
- Designed to handle various data models: document, key-value, wide-column, and graph
- Generally more flexible and scalable than relational databases
- Often sacrifice ACID compliance for performance and scalability
- Examples: MongoDB, Cassandra, Redis, Neo4j

What is MongoDB ?

- MongoDB is a popular open-source, document-oriented NoSQL database.
- Developed by MongoDB Inc., first released in 2009.
- Designed to be scalable, flexible, and high-performance.
- Name comes from "humongous," reflecting its ability to handle large volumes of data.
- Written in C++, with drivers available for many programming languages.

MongoDB Document Model

- Documents are the basic unit of data in MongoDB.
- Similar to JSON objects, but stored as BSON (Binary JSON).
- A document is a set of key-value pairs.
- Fields in documents can hold various data types, including other documents and arrays.
- Example of a MongoDB document:

```
{
  "_id": ObjectId("5099803df3f4948bd2f98391"),
  "name": "John Doe",
  "age": 30,
  "email": "johndoe@example.com",
  "address": {
    "street": "123 Main St",
    "city": "Anytown",
    "country": "USA"
  },
  "interests": ["reading", "hiking", "photography"]
}
```


MongoDB Document Model

Documents are grouped into collections (analogous to tables in relational databases).

Collections are grouped into databases.