

Introduction to Database Technologies



Shoaib Ghachi

Technical Trainer | Mentor |SME

Overview



What is a Database?

Types of Databases

Categories of database models

- Relational Databases-RDBMS

- Non-SQL Databases/Non-RDBMS

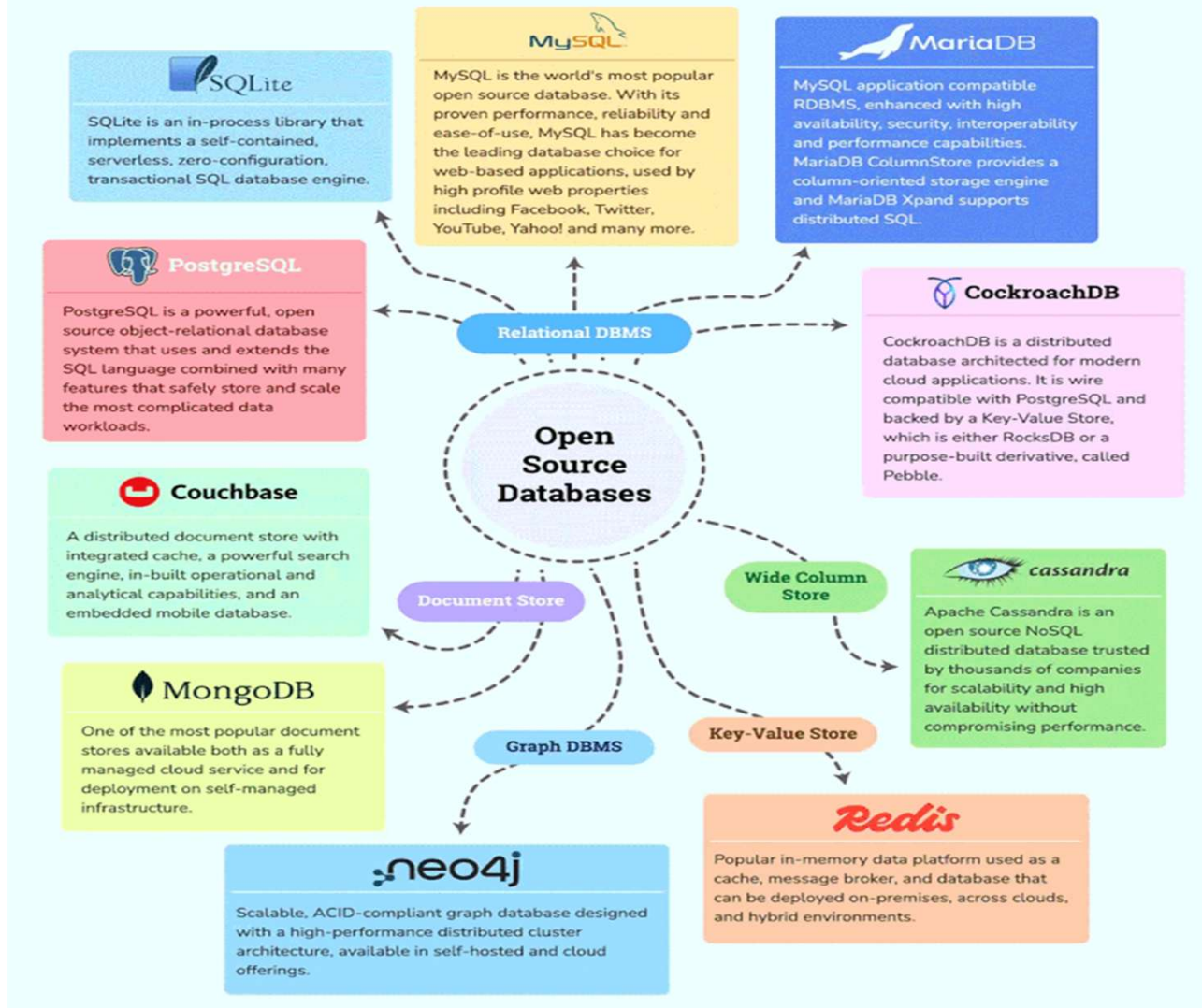
Database Recommendation According to Data

Format

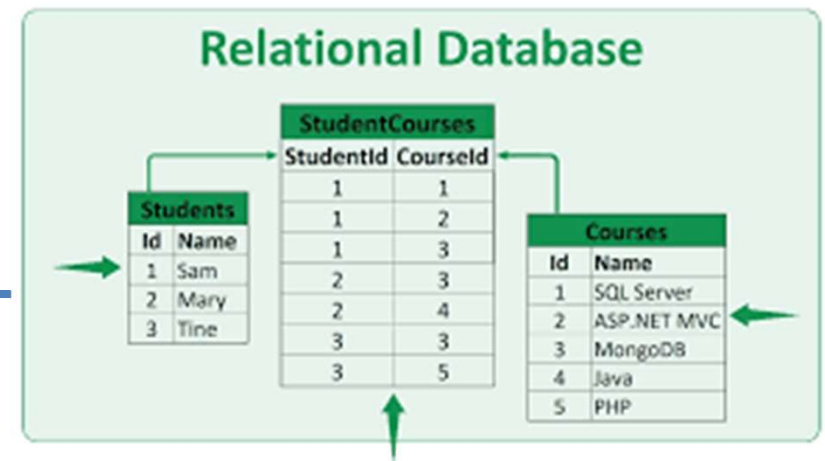
What is a Database?

A database is an organized collection of data that is stored and accessed electronically. Databases are designed to manage, store, retrieve, and update data efficiently. They can handle large volumes of data and allow for data to be accessed, modified, and analyzed in various ways. Databases are critical for a wide range of applications, from simple data storage solutions to complex data management systems in various industries.

Popular Open Source Databases { } DesignGurus.io



RDBMS (SQL) Databases



- Databases can be classified into different types based on their structure, data model, and purpose. Here are some of the main types:.

1.Relational Databases (RDBMS): These databases use a structured query language (SQL) for defining and manipulating data. Data is stored in tables, and relationships are established between these tables.

Examples: MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server.

No-Sql Databases

No-Sql Databases (Non-RDBMS): These databases are designed to handle unstructured data and do not rely on a fixed schema. They are optimized for large-scale data storage and can handle a variety of data types.

- **Types of NoSQL Databases:**

Document-Oriented Databases: Store data in JSON, BSON, or XML documents.

Example: MongoDB.

Key-Value Stores: Store data as key-value pairs.

Example: Redis.

Column-Family Stores: Store data in columns rather than rows.

Example: Apache Cassandra.

Graph Databases: Use graph structures with nodes, edges, and properties to represent and store data.

Example: Neo4j.

Database Recommendation According to Data Format

1. Structured Data: Structured data is highly organized and easily searchable. It resides in fixed fields within a record or file and is typically stored in relational databases.

Recommended Databases:

- MySQL: Open-source, widely used, and suitable for many applications.
- PostgreSQL: Open-source, known for advanced features and compliance with SQL standards.
- Oracle Database: Commercial, highly robust, and used for enterprise-level applications.
- Microsoft SQL Server: Commercial, integrated with Microsoft products and services.

2. Semi-Structured Data: Semi-structured data does not reside in fixed fields or records but contains tags or markers to separate data elements. It includes formats like XML, JSON, and other document-based data.

Recommended Databases:

- MongoDB: Document-oriented NoSQL database that stores data in JSON-like documents.
- CouchDB: Another document-oriented database that uses JSON to store data.
- Couchbase: Combines the capabilities of a document store and key-value store.
- Elasticsearch: Often used for search and analytics, stores data in a semi-structured JSON format.

3.Unstructured Data: Unstructured data does not have a pre-defined data model or organization. It includes text files, multimedia content, emails, etc.

Recommended Databases:

- Apache Hadoop: Framework that allows for distributed storage and processing of large datasets using the Hadoop Distributed File System (HDFS).
- Amazon S3: Object storage service that is highly scalable and used for storing and retrieving large amounts of unstructured data.
- Cassandra: NoSQL database designed to handle large amounts of unstructured data across many commodity servers.
- Elasticsearch: Also suitable for indexing and searching large volumes of unstructured data.

Feature	SQL Databases	NoSQL Databases
Data Model	Relational (tables with rows and columns)	Non-relational (key-value, document, column, graph, etc.)
Schema	Fixed schema (predefined structure)	Dynamic schema (flexible, schema-less)
Scalability	Vertically scalable (increase server capacity)	Horizontally scalable (add more servers/nodes)
Examples	MySQL, PostgreSQL, MS SQL Server, Oracle	MongoDB, Cassandra, Redis, CouchDB, Neo4j
Query Language	Structured Query Language (SQL)	Various: MongoDB uses JSON-like queries, Cassandra uses CQL, etc.
ACID Compliance	Strong ACID compliance (Atomicity, Consistency, Isolation, Durability)	Some are eventually consistent (BASE), but not all NoSQL systems support full ACID
Best Suited For	Complex queries, structured data, transactions	Large-scale unstructured or semi-structured data, high-speed and scalable apps
Data Integrity	Enforced by constraints and relationships	Must be handled manually or by the application
Joins	Supports joins across multiple tables	Typically no joins; joins must be handled in code
Use Cases	Banking, ERP, CRM, traditional enterprise	Real-time analytics, IoT, social networks, big data

References

<https://www.techtarget.com/searchdatamanagement/definition/database>

<https://www.britannica.com/technology/database>

<https://www.simplilearn.com/tutorials/dbms-tutorial/what-is-a-database>

<https://support.microsoft.com/en-us/office/database-basics-a849ac16-07c7-4a31-9948-3c8c94a7c204>

<https://www.matillion.com/blog/the-types-of-databases-with-examples>

<https://www.simplilearn.com/tutorials/dbms-tutorial/what-are-various-types-of-databases>