
CS-313 Assignment 1

Shashank P, 200010048

August 13, 2022

1 Introduction

1.1 Scientists

1.1.1 E.F. Codd

Edgar Frank Codd was an English computer scientist who invented the relational model for database management. Edgar Frank Codd was born in Fortuneswell. He was working for IBM and, from 1961 to 1965, pursuing his doctorate in computer science at the University of Michigan. Two years later, he moved to San Jose to work at IBM's Research Laboratory. Initially, IBM refused to implement the relational model. Then IBM included it in its Future Systems project but put it in charge of its developers, who were unfamiliar with Codd's ideas. Codd's theorem, proven in his seminal work on the relational model, equates the expressive power of relational algebra and relational calculus. Codd received the Turing Award in 1981 and, in 1994, was inducted as a Fellow of the Association for Computing Machinery. In 2004, SIGMOD renamed its highest prize to the SIGMOD Edgar F. Codd Innovations Award in his honour.

Source: E.F. Codd

1.1.2 Michael Stonebraker

Michael Ralph Stonebraker is a computer scientist specializing in database systems. He earned his B.S.E. in 1965 and his M.S. and PhD from the University of Michigan in 1967 and 1971. During his time at the University of California, Berkeley, he focused on relational database management systems such as Ingres and Postgres, and at (MIT) he developed more novel data management techniques such as C-Store, H-Store and SciDB. His awards include the IEEE John von Neumann Medal, SIGMOD Edgar F. Codd Innovations Award. He was elected a member of the National Academy of Engineering for the commercialization of relational and object-relational DB systems. In March 2015, he won the 2014 ACM Turing Award. He won the 2015 Commonwealth Award. Stonebraker is currently a Professor at UC Berkeley.

Source: Michael Stonebraker

1.2 Data Model and Database Application

1.2.1 Data model

A data model defines how to model the logical structure of a database. A data model is a basic entity for introducing abstractions into a DBMS. A data model defines how data is connected, processed, and stored in a system. Your first data model might be a flat data model that requires all the data you use to be on the same level. The previous data model was not very scientific and was prone to many duplicates and update anomalies. The Entity-Relationship (ER) model is based on the concept of real-world entities and the relationships between them. When formulating a real-world scenario into a database model, The ER model creates entity sets, relationship sets, common attributes, and constraints. The ER model is ideal for conceptual database design. The most common data model in DBMS is the relational model. This is a more scientific model than others. This model is based on primary predicate logic and defines tables as n-ary relations.

Source: Data model

1.2.2 Database Application

A database application is a program whose purpose is to retrieve information from a database. From here, you can insert, modify or delete information and it will be updated again in the database. A hallmark of modern database applications is allowing concurrent updates and queries from multiple users. Database applications with web interfaces had the advantage of being able to be used on devices of various sizes. An early example of a database application with a web interface was amazon.com using Oracle's relational database management system. Examples of database applications include Amazon, eBay, Google, Youtube, and Facebook. Many of today's computer systems are database applications such as Facebook built on top of MySQL.

Source: Database Application

2 Digital Applications

Databases form an integral part of many applications developed from small scale to large scale. Some of the apps in India that make use of a large database are as follows:

- **Finance:** Google pay, PhonePe, Paytm, BHIM app, etc.
- **Retail:** Shopsy, Ajio, Flipkart, Amazon, etc.
- **Manufacturing:** Odoo, VTScada, Anvyl, nTask, etc.
- **IT:** Infosys, Wipro, HCL Technologies, Tech Mahindra, etc.

Source: Finance, Retail, Manufacturing, IT

3 OLTP & OLAP

3.1 OLTP

Online transaction processing is data processing that executes a series of transactions simultaneously. These transactions are traditionally called economic or financial transactions. It is recorded and protected so the business can access the information at any time for accounting or reporting purposes. The primary definition of a transaction still underlies most OLTP systems. Therefore, online transaction processing typically involves inserting, updating, and deleting small amounts of data in data stores. The most common architecture for OLTP systems with transactional data is a three-tier architecture consisting of a presentation tier: a business logic layer, and a data storage layer. Relational databases were explicitly built for transactional applications.

Source: OLTP

3.2 OLAP

Online Analytical Processing is software for high-speed multidimensional analysis of large amounts of data. A warehouse, data mart, or other consolidated, centralized data store. In a data warehouse, records are stored in tables and can be organized individually, data for only two dimensions at a time. OLAP extracts data from multiple relational datasets and reorganizes them into multidimensional ones, A format that allows for swift processing and insightful analysis. The core of most OLAP systems, the OLAP cube, is an array-based multidimensional A database that can process and analyze multidimensioned data much faster and more efficiently than traditional relational databases.

Source: OLAP

3.3 Differences

Key Difference between **OLTP** and **OLAP** are as follows:

OLTP	OLAP
Enable real-time executions of multiple transactions	Querying multiple records for analysis
Lightning fast response	Magnitudes slower than OLTP
Modify small data frequently	Usually for large data, but read intensive
Frequent data backups	Infrequent data backups
Less storage space	Usually requires a lot of storage
Run queries with few records	Run complex queries spanning multiple records

Table 1: OLTP vs OLAP

Source: OLTP vs OLAP