OLAP, OLTP, and NewSQL database systems

In this train, we look at OLAP and OLTP databases and why they are important. We also present NewSQL, a new database type, and discuss how it differs from relational and NoSQL databases.

Learning objectives

In this train, we will review the following concepts:

- OLAP and OLTP databases.
- NewSQL databases.

OLAP and OLTP databases

The increase in the amount, type, and sources of data over the past three decades has led to the birth of two distinct database categories: OLAP and OLTP databases. These categories are suited for analytical, online analytical processing (OLAP), and transactional, online transaction processing (OLTP) workloads.

OLAP databases

Analysts frequently need to group, aggregate, and join data. These operations in relational databases are resource-intensive. With OLAP, data can be pre-calculated and pre-aggregated, making analysis faster. OLAP databases are divided into one or more OLAP cubes. The cubes are designed in such a way that creating and viewing reports become easy. Figure 9 below is an illustration of such an OLAP cube.

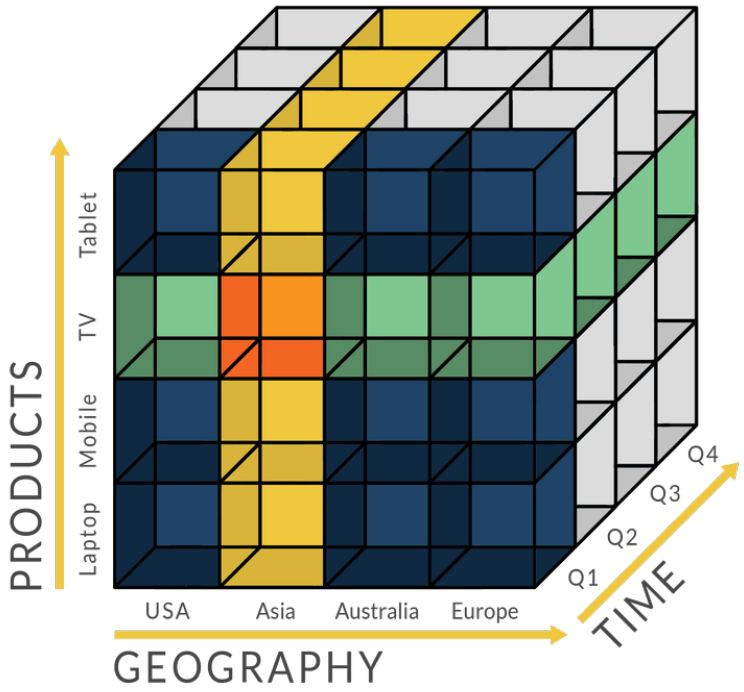

Figure 9: An OLAP cube illustration. A three-dimensional array is built (the cube) to easily query specific product sales by time of year for a specific geographic region.

These cubes can then be navigated using four basic types of analytical operations in OLAP databases:

1. **Roll-up**: View in decreased detail.
2. **Drill-down**: View in increased detail.
3. **Slice and dice**: Focuses on one particular dimension of a cube, and then provides a new sub-cube to analyse. An example would be a cube containing the salaries for men and women in different countries over a few years, and then only selecting the South African "sub-cube" to analyse.
4. **Pivot (rotate)**: Rotate the data to provide an alternative representation thereof. In essence, switch the x, y, and z-axis of the cube.

Usually, the number of users querying an OLAP database is relatively low, primarily consisting of the employees of a company only.

OLTP databases

OLTP databases are concerned with all transactions. These include short-lived, pre-defined queries. OLTP databases are designed with less detail in mind than OLAP databases but need to be built for a much larger number of users.

Examples of OLTP transactions include:

- Search results on an e-commerce website.
- Online banking.
- Online airline ticket booking.
- Sending text messages.
- Grocery purchases via an app or website.
- Purchasing a product on an e-commerce website.

OLTP typically involves:

- Inserting.
- Updating.
- Deleting small amounts of data in a database.

When designing an OLTP system, it is important to ensure that the environment has:

- Short response times.
- Only small transformations are applied.
- The system has high concurrency.
- The system can handle large volumes of continuous data streams.
- The system is highly available.

Below we provide a brief comparison of OLAP and OLTP database systems:

Properties	OLAP	OLTP
Use case	It is used for data analysis	It is used to manage a very large number of online short transactions
Database type	It uses data warehouses	It uses traditional DBMSs
Data modification	It is mainly used for data reading	It manages all insert, update, and delete transactions
Response time	Processing is a little slow	Processes in milliseconds
Normalisation	Tables in OLAP databases are not normalised	Tables in OLTP databases are normalised

The new player, NewSQL

NoSQL is very useful for applications such as social media, where eventual consistency is acceptable. This means that users rarely notice or have their user experience ruined if they see a non-consistent view of the database. The data read from these databases include status updates, tweets, etc. Therefore, strong consistency is not essential. However, NoSQL databases are not easy to use for systems where consistency is critical, such as banking and finance platforms. This is where NewSQL comes into play.

NewSQL systems were born out of the desire to combine the scalability and high availability of NoSQL alongside the relational model, transaction support, and structured query language of traditional relational databases. Most NewSQL databases are consequently a complete redesign focused heavily on OLTP or hybrid workloads. NewSQL databases are built for the cloud era, with a horizontally scalable architecture in mind from the start.

Such systems aim to achieve the scalability of NoSQL systems while still providing the ACID attributes ensured by traditional relational databases. NewSQL databases are primarily intended for companies that handle high-profile data and require scalability but also need more consistency than NoSQL databases can provide, such as financial institutions.

NewSQL utilises the relational data model and runs on SQL.

Feature	SQL (relational databases)	NoSQL	NewSQL
Relational property	Yes, follows relational modelling to a large extent	No, it doesn't follow the relational model	Yes, it follows the relational property
ACID	Yes	No	Yes
SQL	Yes	No	Yes, with enhanced performance

Feature	SQL (relational databases)	NoSQL	NewSQL
OLTP	Full	Limited	Full
Scaling	Vertical	Vertical and horizontal	Vertical and horizontal
Query handling	Best suited for simple queries	Best suited for complex queries	Improved performance for simple queries and capable of complex queries
Distributed databases	No	Yes	Yes