OLTP vs OLAP/NEW INTERNSHIP PROCESS

Created by Pamela. M., Susan. N. and Mayra. S.

Table of Contents

INTRODUCTION: OLTP vs OLAP

OVERVIEW:

There are two major camps in the relational database development world, OLTP vs OLAP.

An OLTP system is comprised of a set of database tables that are always collecting data. They are designed for business purposes. This is a database that is always running in the background. The data stored in an OLTP database is collected through short and fast data inserts and updates initiated by an end user after a business process. An OLTP database may collect data from websites, Point of Sales systems, Human Resource systems, and Customer Relationship Management platforms. A typical instance of an OLTP system is one that collects data from airline ticket sales. In an OLTP system, the data is stored in separate tables to prevent data redundancy. Eliminating data redundancy in an OLTP system ensures that insert, update, and delete anomalies do not occur.

To eliminate data redundancy and prevent such anomalies, OLTP databases must be normalized. Normalizing a database entails storing descriptive data (varchar data) in separate and distinct dimension tables. In this way, data updates and deletions occur in only one location. This ensures that data is up to date and current always. Normalizing an OLTP database involves going through a refinement process that identifies data objects that should be in a relational database, identifying their relationships and defining tables required and the column within each table. A master transactional table inserts data into these dimension tables. The more a database is normalized, the more dimension tables it will contain. The data collected through the operational system must be backed

up on a repetitive schedule to continue business functions. OLTP databases store current data while historical data is archived in a data warehouse. Relatively simple and standardized queries are used to return records from an OLTP system. For more information on normalization, follow link.

ATTRIBUTES OF OLTP and OLAP

OLTP (On-line Transaction Processing)
- Characterized by many short on-line transactions (**INSERT, UPDATE, DELETE**).
- The main emphasis for OLTP systems is put on very fast query processing, maintaining data integrity in multi-access environments and an effectiveness measured by number of transactions per second.
- In OLTP database there is detailed and current data, and schema used to store transactional databases is the entity model (usually 3NF).

  OLTP Examples.
  - A supermarket server which records every single product purchased at the market.
  - A bank server which records every time a transaction is made for an account
  - A railway reservation server which records the transactions of a passenger

OLAP (On-line Analytical Processing)
- As an analytical database, focus is placed on being able to view the entire data all at once for the purposes of business analysis and intelligence, which includes generation of reports.

- This database is mostly denormalized.
- Designing such a database we must keep in mind the big picture of what metrics we intend to report on to ensure a strategic transfer of information from the OLTP database that provides daily operational data to the data warehouse(DW) from which the OLAP database is built.
- The OLAP database basically consists of a FACTS table that holds information of the business process to be tracked, and various DIMENSIONS tables that consists of elements that contribute to the business process.
- They are usually created in a star schema data model where different DIMENSION tables are related to one FACTS Table for insight into a business process
- Characterized by relatively low volume of transactions.
- Queries are often very complex and involve aggregations.
- For OLAP systems a response time is an effectiveness measure.
- OLAP applications are widely used by Data Mining techniques.
- In OLAP database there is aggregated, historical data, stored in multi-dimensional schemas (usually star schema).

OLAP EXAMPLES

- A Bank manager wants to know how many customers are utilizing the ATM of this branch. Based on this he may take a call whether to continue with the ATM or relocate it.
- An insurance company wants to know the number of policies each agent has sold. This will help in better performance of agents.

OLTP vs OLAP - What's the Benefit.
We will consider the similarities and advantages/benefits of the two systems

First, both OLTP and OLAP are used in business applications in data warehousing and analytics.

OLTP?

Most business applications are OLTP systems. It is hard to imagine a strata of computer engineering that doesn't have some form of OLTP. In fact, a whole branch of computer science grew up around the need to handle transactions in an orderly manner.

Thus, we come to the critical part of the OLTP acronym: The T, which stands for transactional. Transactions are hugely important and keeping them organized is also hugely important.

What does it take to keep transactions organized? It means that database transactions must be stable (or durable, not easily changed), isolated, consistent and Atomic. In simple words, this type of transaction ensures that operations performed by different users do not interfere with each other. For example, if a husband and wife each make a withdrawal from their bank account, atomic transactions make sure that they do not withdraw more than their account holds.

An OLTP system needs each of these properties to ensure that everything is working in an orderly manner. When you bank, you are confident that the money you deposit is credited to your account, not mixed up with your other transactions or, heaven forbid, lost.

Clearly, Transactional processing is very important for our society.

ADVANTAGES OF OLTP
It is easy and the best solution for online shoppers. These systems are most efficient and have excellent response times. Very easy to use; as simple as fill a form and the rest will be taken care of by the web and database servers.

Online banking is completely based on online transaction processing systems. Credit cards are also well-handled by these systems.

You can access anything on the web and choose to buy it because all financial transactions methods are supported by these systems.

OLAP?

OLAP is short for on-line analytical processing. The key letter here is A for analytical which also tells us what the OLAP system does.

It is used to maintain the history of the data and mainly used for deep data analysis and it can also be called a warehouse. This is usually huge Archival data.

Unlike OLTP, OLAP systems work with very large amounts of data, make reporting very easy and quickly across different dimensions. Preserving the accuracy and integrity of transactions is not their purpose, this is up to OLTP. OLAP is here to allow us to find trends, crunch numbers and get the big picture. These systems have a smaller group of users than OLTP systems. For example, you will not interact with your bank's OLAP system, since it is not concerned with recording your account transactions.

As an analytical database, focus is placed on being able to view the entire data all at once for the purposes of business analysis and intelligence, which includes generation of reports.

Data in this environment is mostly clean data where reports can be created across many dimensions using a different dimension.

The OLAP database is mostly created in a star schema data model where the fact tables and the dimensional tables are connected by foreign key relations.

Because of the foreign keys relations in the OLAP environment Data integrity is enforced. Data modifications are not frequent. It has de-normalized tables with indexes highly recommended. There are less users compared to OLTP and is read only.

The OLAP environment is designed for Data Modelling. Today's business environment is always and quickly changing. Today's OLAP environment must support modelling tools like the cube. A cube is a smart data modelling tool or structure that holds multi-dimensional data in an OLAP environment. The cube gets its information from the Schema model that consists of the fact and dimensional tables for reporting purposes. A cube is a very versatile tool much needed for the creation of analytical and business reports. Because the information in the cube is Pre-aggregated, reports can be created very easily and quickly across many dimensions. The cube can process very large volumes of data because they are

optimized for analytical purposes, so they are able to report on millions of records at a time. A cube is a dataset.

OLAP performs multidimensional analysis of business data and provides the capability for complex calculations, trend analysis, and sophisticated data modeling. It is the foundation for many kinds of business applications for Business Performance Management, Planning, Budgeting, Forecasting, Financial Reporting, Analysis, Simulation Models, Knowledge Discovery, and Data Warehouse Reporting.

OLAP enables end-users to perform ad hoc analysis of data in multiple dimensions, thereby providing the insight and understanding they need for better decision making.

ADVANTAGES OF OLAP

Knowledge is the foundation of all successful decisions. Successful businesses continuously plan, analyze and report on sales and operational activities to maximize efficiency, reduce expenditures and gain greater market share.

Statisticians will tell you that the more sample data you have, the more likely the resulting statistics will be true. Naturally, the more data a company can access about a specific activity, the more likely that the plan to improve that activity will be effective. All businesses collect data using many different systems, and the challenge remains: how to get all the data together to create accurate, reliable, fast information about the business.

A company that can take advantage and turn it into shared knowledge, accurately and quickly, will surely be better positioned to make successful business decisions and rise above the competition.

OLAP technology has been defined as the ability to achieve "fast access to shared multidimensional information." Given OLAP technology's ability to create very fast aggregations and calculations of underlying data sets, one can understand its usefulness in helping business leaders make better, quicker "informed" decisions.

A company that can take advantage and turn it into shared knowledge, accurately and quickly, will surely be better positioned to make successful business decisions and rise above the competition.

Data Architecture Model
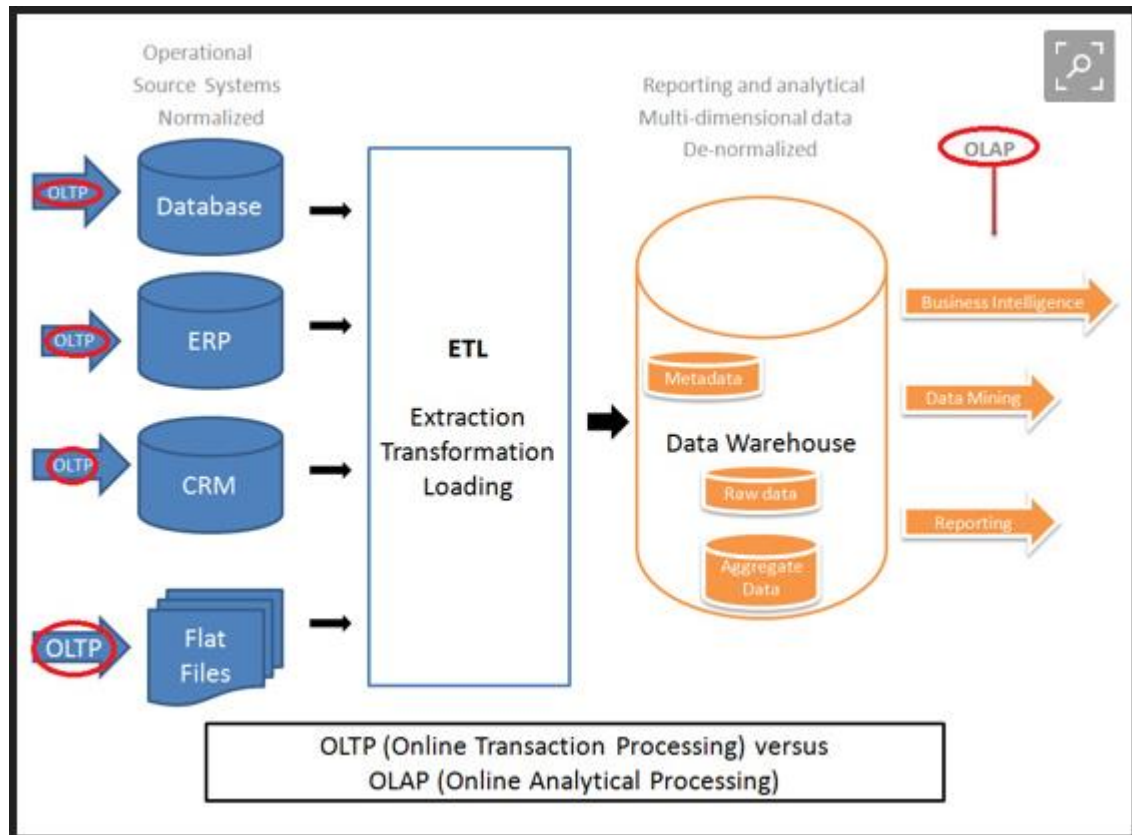
HOW OLTP vs OLAP OPERATE:

	OLTP System Online Transaction Processing (Operational System)	OLAP System Online Analytical Processing (Data Warehouse)
Source Data	Operational data; OLTPs are the original source of the data.	Consolidation data; OLAP data comes from the various OLTP Databases
Purpose	To control and run fundamental business tasks	To help with planning, problem solving, and decision support
What the data does	Reveals a snapshot of ongoing business processes	Multi-dimensional views of various kinds of business activities
Inserts and Updates	Short and fast inserts and updates initiated by end users	Periodic long-running batch jobs refresh the data
Queries	Relatively standardized and simple queries Returning relatively few records	Often complex queries involving aggregations
Processing Speed	Typically, very fast	Depends on the amount of data involved; batch data refreshes and complex queries may take many

		hours; query speed can be improved by creating indexes
Space Requirements	Can be relatively small if historical data is archived	Larger due to the existence of aggregation structures and history data; requires more indexes than OLTP
Database Design	Highly normalized with many tables	Typically, de-normalized with fewer tables; use of star and/or snowflake schemas
Backup and Recovery	Backup religiously; operational data is critical to run the business; data loss is likely to entail significant monetary loss and legal liability	Instead of regular backups, some environments may consider simply reloading the OLTP data as a recovery method

TYPES OF OLAP

Like everything else in technology, there's not a single way to solve the technology problem. There are three basic approaches to OLAP: MOLAP, ROLAP, and HOLAP.

Multidimensional Online Analytical Processing (MOLAP) is the most standard approach to OLAP solutions. It uses a multidimensional database which directly

stores the information contained in the various cubes. This is the best performing solution when using SQL Server Analysis services.

Relational Online Analytical Processing (ROLAP) provides the same solution but uses a relational database for storage of the data. This approach translates native OLAP queries, written in a language called multidimensional expressions (MDX) into the appropriate SQL statements. This is primarily done to prevent the need for another copy of the data. The data created directly by the online transaction processing (OLTP) applications are used. The primary disadvantage to this solution is that it does not perform as well as a MOLAP database.

Hybrid Online Analytical Processing (HOLAP) is a hybrid approach to the solution where the aggregated totals are stored in a multidimensional database while the detail data is stored in the relational database. This is the balance between the data efficiency of the ROLAP model and the performance of the MOLAP model.

If you're developing applications, you ultimately do not care which model is being used because your code will work against any kind of OLAP model (unless you're creating the cubes yourself). It is only a concern when a large-scale project is being rolled out and you must decide which solution will meet its unique needs.

For deciding on whether to use OLTP or OLAP follow Link.