Online Analytical Processing Server

Online Analytical Processing Server (OLAP) is based on multidimensional data models. It allows managers and analysts to get a piece of deep information through fast, consistent, and interactive access to information. In this topic, we cover the types of OLAP, operations on OLAP, a difference between OLAP, and statistical databases and OLTP.

Types of OLAP Servers

OLAP had four types of servers they are listed below

- Relational OLAP (ROLAP)
- Multidimensional OLAP (MOLAP)
- Hybrid OLAP (HOLAP)
- Specialized SQL Servers

Relational OLAP

ROLAP is Relational Online Analytical Processing model where the data is stored in a relational database. for eg rows and columns in the data warehouse. In the ROLAP model data was present in front of the user in the multidimensional form. To display the data, in multidimensional views a semantic layer of metadata is created that maps dimension to the relational tables. Metadata also supports the aggregation of the data.

Whenever the ROLAP engine gets analytical server issues a complex query, it fetches data from the main warehouse and dynamically creates a multidimensional view of data for the user and it differs from MOLAP because MOLAP already has a static multidimensional view of data stored in proprietary databases MDDBs.

As the multidimensional view of data is created dynamically it processes slower in comparison to MOLAP. ROLAP engine deals with large volumes of data.

Relational online analytical processing (ROLAP) is a form of online analytical processing (OLAP) that performs dynamic multidimensional analysis of data stored in a relationaldatabase rather than in a multidimensionaldatabase which is usually considered the OLAP standard.ROLAP uses a relational database and it requires more processing time or disk space to perform some of the tasks that multidimensional databases are designed for. However, ROLAP supports larger user groups and greater amounts of data and is often used when these capacities are crucial, such as in a large and complex department of an enterprise.Relational OLAP servers are placed between relational back-end server and client front-end tools to store and manage the warehouse data and the relational OLAP uses or extended-relational Data Base Management System. ROLAP includes

- Implementation of aggregation navigation logic

- Optimisation for each DBMS back-end
- Additional tools and services.

Advantages

- ROLAP servers can be easily used with existing Relational Data Base Management System.
- Data can be stored efficiently, and no zero facts can be stored.
- ROLAP tools do not use pre-calculated data cubes .
- DSS server of micro-strategy adopts the ROLAP approach.

Disadvantages

- Poor query performance.
- Some limitations of scalability depending on the technology architecture that is utilized.

Multidimensional OLAP

MOLAP is a Multidimensional Online Analytical Processing model. The data was used to analyze and is stored in specialized multidimensional databases (MDDBs). The multidimensional database management systems are proprietary software systems.

These multidimensional databases are formed from the large multidimensional array. The cells or data cubes of this multidimensional databases will carry precalculated and prefabricated data. The Proprietary software systems create this precalculated and fabricated data, while the data is loaded to MDDBs from the main databases.

Now, it is the work of MOLAP engine, it will reside there in the application layer and provide the multidimensional view of data from the Multidimensional Database to the user. Thus when a user request for the data, no time is wasted in calculating the data and the system responses fast.It is the more traditional way of OLAP analysis. In MOLAP, data is stored in a multidimensional cube. The storage is not in the relational database but in proprietary formats. Multidimensional OLAP (MOLAP) uses array-based multidimensional storage engines for multidimensional views of data with multidimensional data storage, the storage utilization may be low if the dataset is sparse. Most of the MOLAP servers use two levels of data storage representation to handle dense and sparse datasets.

Advantages

- MOLAP allows the fastest indexing to the pre-computed summarised data.
- MOLAP cubes are built for fast data retrieval and they are optimal for slicing and dicing operations.

- MOLAP Can perform complex calculations: All calculations have been pre-generated when the cube was created. so complex calculations are not only doable, but they return quickly.
- MOLAP Helps the users to connect a network who need to analyze larger and less-defined data.
- It is easier and suitable for inexperienced users

Disadvantages

- MOLAP is not capable of containing details of data.
- The storage of utilization may be low if the data set is sparse.

Key Differences Between ROLAP and MOLAP

1. ROLAP stands for Relational Online Analytical Processing whereas; MOLAP stands for Multidimensional Online Analytical Processing.
2. In both the cases, ROLAP and MOLAP data is stored in the main warehouse. In ROLAP data is directly fetched from the main warehouse whereas, in MOLAP data is fetched from the proprietary databases MDDBs.
3. In ROLAP, data is stored in the form of relational tables but, in MOLAP data is stored in the form of a multidimensional array made of data cubes.
4. ROLAP deals with large volumes of data whereas, MOLAP deals with limited data summaries kept in MDDBs.
5. ROLAP engines use complex SQL to fetch data from the data warehouse. However, MOLAP engine creates prefabricated and precalculated datacubes to present multidimensional view of data to a user and to manage data sparsity in data cubes, MOLAP uses Sparse matrix technology.
6. ROLAP engine creates a multidimensional view of data dynamically whereas, MOLAP statically stores multidimensional view of data in proprietary databases MDDBs for a user to view it from there.
7. As ROLAP creates a multidimensional view of data dynamically, it is slower than MOLAP which do not waste time in creating a multidimensional view of data.

Hybrid OLAP

Hybrid online analytical processing (HOLAP) is a combination of relational OLAP (ROLAP) and multidimensional OLAP (usually referred to simply as OLAP). HOLAP was developed to combine the greater data capacity of ROLAP with the superior processing capability of the OLAP.

HOLAP can use varying combinations of ROLAP and OLAP technology. Typically it stores data in both a relational database (RDB) and a multidimensional database (MDDB) and uses whichever one is best suited to the type of processing desired. The databases are used to store data in the most functional way. For data-heavy processing, the data is more efficiently stored in an RDB, while for speculative processing, the data is more effectively stored in an MDDB.

HOLAP user can choose to store the results of their queries to the MDDB to save the effort of looking for the same data over and over which saves time. This technique is called "materializing cells".It improves performance and it takes a toll on storage. The user has to strike the balance between performance and storage demand to get the most out of HOLAP but it offers the best features of both OLAP and ROLAP, HOLAP is increasingly preferred.

HOLAP technologies were attempting to combine the advantages of MOLAP and ROLAP. For summary type information's HOLAP leverages cube technology for faster performance. When detail information was needed, HOLAP can "drill through" from the cube into the underlying relational data.It offers scalability of ROLAP and faster computation of MOLAP. HOLAP servers allow storing the large data volumes of detailed information. The aggregations are stored separately in the MOLAP store.

Specialized SQL Servers

Specialized SQL servers provide advanced query language and query processing support for SQL queries over star and snowflake schemas in a read-only environment.
OLAP Operations

OLAP servers were based on the multidimensional view of data and we will discuss OLAP operations in multidimensional data.

Here is the list of OLAP operations −

- Roll-up
- Drill-down
- Slice and dice
- Pivot (rotate)

Roll-up

Roll-up performs aggregation on a data cube in any of the following ways −

- By climbing up a concept hierarchy for a dimension
- By dimension reduction

The following diagram illustrates how roll-up works.

- Roll-up is performed by climbing up to the concept hierarchy for the dimension location.
- Initially the concept hierarchy was "street < city < province < country".
- On rolling up, the data is aggregated by ascending the location hierarchy from the level of the city to the level of the country.
- The data is grouped into cities rather than the countries.
- When roll-up is performed by one or more dimensions from the data cube are removed.

Drill-down

Drill-down is the reverse operation of roll-up. It is performed in the following ways −

- By stepping down a concept hierarchy for a dimension
- By introducing a new dimension.

The following diagram illustrates how drill-down works −

- Drill-down is performed by stepping down to the concept hierarchy for the dimension time.
- Initially, the concept hierarchy was "day < month < quarter < year."
- On drilling down, the time dimension is descended from the level of the quarter to the level of the month.
- When drill-down is performed by one or more dimensions from the data cube were added.
- Drill down will navigate the data from less detailed data to highly detailed data.

The slice operation selects one particular dimension from the given cube and provides a new sub-cube.

We Consider the following diagram that shows how slice works.

- Here the Slice is performed by the dimension "time" using the criterion time = "Q1".
- It will form a new sub-cube by selecting one or more dimensions.

Dice

Dice selects two or more dimensions from a given cube and provides a new sub-cube. We Consider the following diagram that shows the dice operation.

The dice operation on the cube based on the following selection criteria involves in three dimensions.

- (location = "Toronto" or "Vancouver")
- (time = "Q1" or "Q2")
- (item =" Mobile" or "Modem")

Pivot

The pivot operation is also known as rotation. It rotates the data axes in order to view and provide an alternative presentation of data. we Consider the following diagram that shows the pivot operation.

OLTP

OLTP (online transaction processing) is a class of software program it can capable of supporting transaction-oriented applications on the Internet.

In OLTP systems are used for order entry, financial transactions, customer relationship management (CRM) and retail sales. Such systems have a large number of users who conduct short transactions. Database queries are simple, it requires sub-second response times and returns relatively few records.

An important attribute of an OLTP system is its ability to maintain concurrency to avoid single points of failure in OLTP systems are often decentralized.

OLAP vs OLTP

Sr.No.	Data Warehouse (OLAP)	Operational Database (OLTP)
1	Involves historical processing of the information.	Involves day-to-day processing .
2	OLAP systems are used by knowledge workers such as executives, managers and analysts.	OLTP systems are used by clerks, DBAs, or database professionals.
3	Useful in analyzing the business.	Useful in running the business.
4	It focuses on Information out.	It focuses on Data in.
5	Based on Star Schema, Snowflake, Schema and Fact Constellation Schema.	Based on Entity Relationship Model.
6	Contains historical data.	Contains current data.

7	Provides summarized and consolidated data.	Provides primitive and highly detailed data.
8	Provides summarized and multidimensional view of data.	Provides detailed and flat relational view of data.
9	Number or users is in hundreds.	Number of users is in thousands.
10	Number of records accessed is in millions.	Number of records accessed is in tens.
11	Database size is from 100 GB to 1 TB	Database size is from 100 MB to 1 GB.
12	Highly flexible.	Provides high performance.