

OLAP Operations in DBMS

RDBMS

DBMS

Database

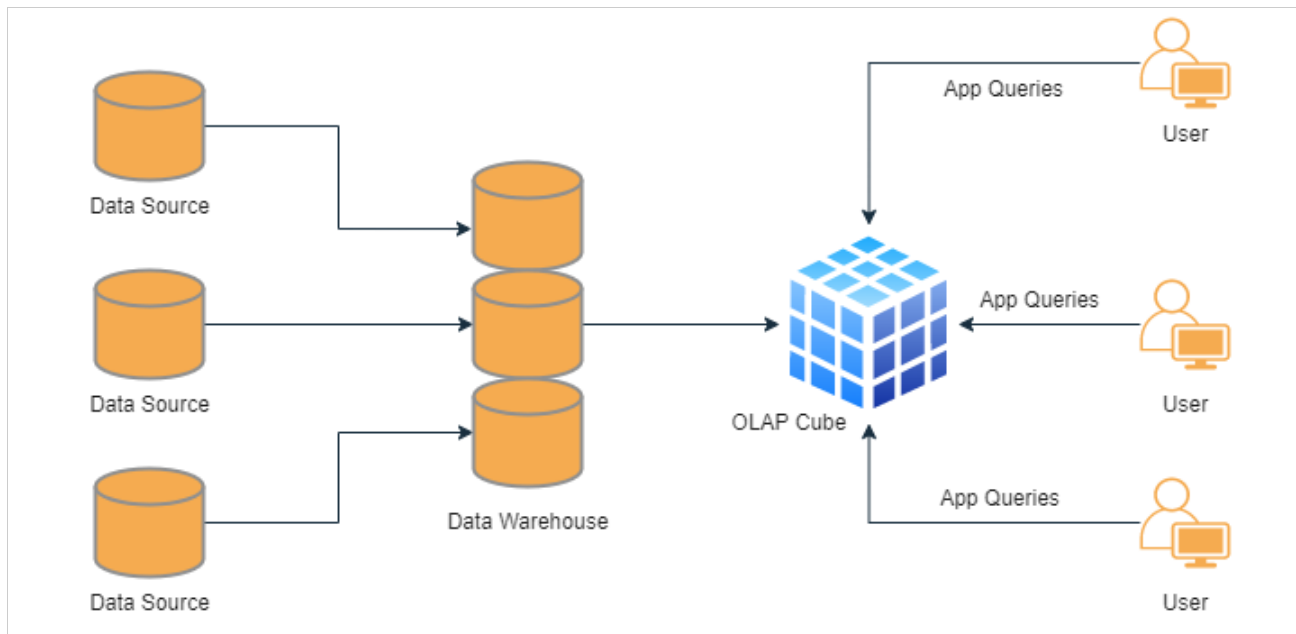
Online Analytical Processing Server (OLAP) is used to enable users to perform analysis on information retrieved from multiple database systems concurrently. It operates on a multidimensional data model. Users can write query data across multiple dimensions, such as location, time, and categories (e.g., Hyderabad-> 2017 -> Sales data).

In OLAP, databases are organized into one or more cubes. These cubes are known as Hyper-cubes and represent a multi-dimensional array of data. Each cell within the array corresponds to specific combination of dimension values. The dimensions define the characteristics or attributes. Data can be sliced and diced for analysis by attributes. Examples of dimensions in a sales database are location, time, product category, and customer segment.

OLAP server are efficient and optimized for querying. It uses techniques such as pre-aggregation and indexing. Users can get valuable insights from large volumes of data stored in diverse database systems. OLAP enables complex analysis operations, including drill-down (going from a higher level of summary data to a more detailed level), roll-up (aggregating data to a higher level of summarization), slicing (selecting a specific subset of data based on dimension values), and dicing (analyzing data from multiple dimensions simultaneously). OLAP servers provide a powerful and flexible environment for performing multidimensional analysis.

OLAP Cube

OLAP cube is also known as a Hypercube. It represents specialized data structure designed for efficient storage and rapid analysis of multidimensional data. Essentially, it captures a snapshot of data at a particular moment in time.



For example, as depicted in the above diagram. It utilizes various OLAP operations. Users can select specific perspective of the hypercube. OLAP cubes provide users to execute analytical queries with multiple dimensions on the available data.

Types of OLAP Servers

There are three types of OLAP (Online Analytical Processing) systems –

Relational OLAP (ROLAP)

These systems directly operate on relational databases. These utilize complex SQL queries to retrieve data from the database. They are capable of handling large volumes of data. But they may take more time.

Multidimensional OLAP (MOLAP)

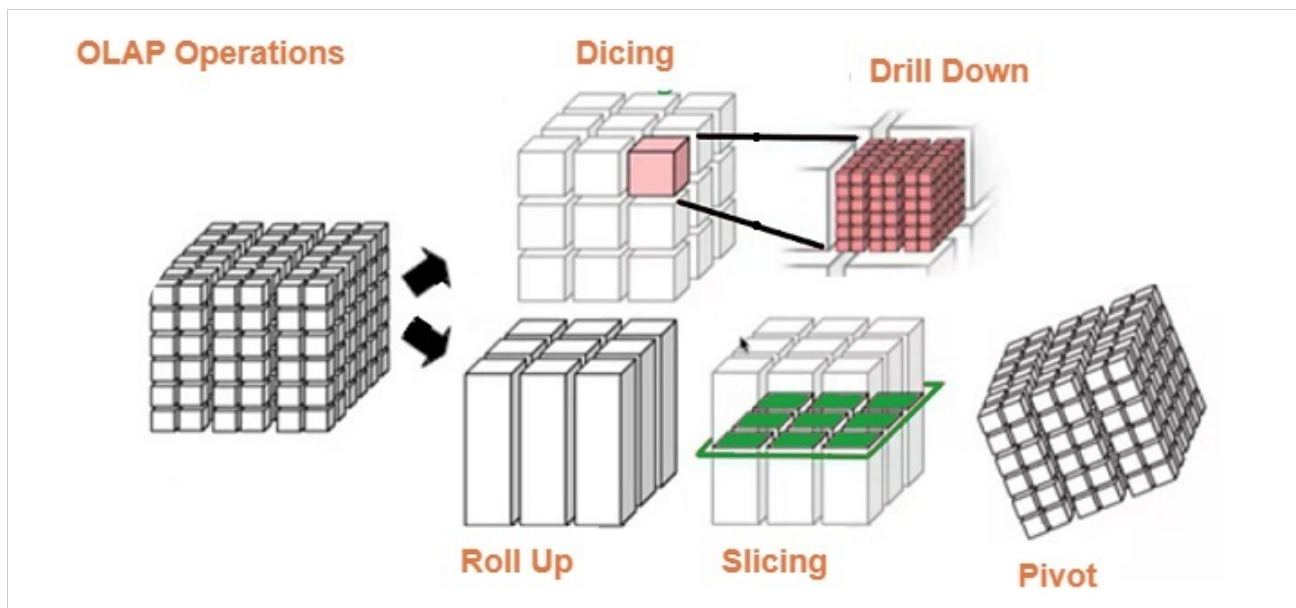
This is considered the traditional form of OLAP. It optimize storage system based on multi-dimensional arrays for data storage. It has positional techniques to access the data physically stored within these multidimensional arrays.

Hybrid OLAP (HOLAP)

It is combination of both ROLAP and MOLAP. It has scalability as advantages of ROLAP systems. It takes less time.

OLAP operations

OLAP operations provide many functionalities to extract valuable insights from data stored in multidimensional hypercubes.



These are explained below.

Drill Down

Drill Down operation enables users to delve deeper into the data cube. It transforms data with few details into highly detailed data. This can be achieved by either descending through hierarchy concepts for a specific dimension or by introducing additional dimensions to the hypercube.

Performing a drill down operation can be achieved through the following techniques:

Moving down in the concept hierarchy

It navigates from higher level of aggregation to lower level of detail within a specific dimension. In the given cube, the drill down operation is performed by moving down in the concept hierarchy of the Time dimension from the Quarter level to the Month level.

Adding a new dimension

This is another approach which introduce new dimension that provides additional granularity. It adds dimension related to a different attribute or aspect of the data. We can analysis and exploration deeply.

For the specific cube given, the drill down operation is moving down in the concept hierarchy of the Time dimension, specifically from the Quarter level to the Month level, to obtain more detailed data about product sales within each city during specific months.

Roll up

Roll up operation is also known as "drill-up" or aggregation. It is technique used to reduce the dimensionality of a data cube by performing aggregation. It is the opposite of the drill-down operation.

Roll up operation aggregates data from a lower level of detail to a higher level by combining similar dimensions across any axis of the cube.

By performing Roll up, the data cube is collapsed along one or more dimensions. It allows for higher-level analysis and a broader view of the data. This aggregation process can be useful for reporting, summarizing data, and extracting key insights from a larger dataset.

Roll-up operation can be achieved through the following techniques:

Climbing up in the concept hierarchy

This involves navigating from a lower level of detail to a higher level of aggregation within a specific dimension. In the given cube, the roll-up operation is performed by climbing up in the concept hierarchy of the Location dimension from the City level to the Country level.

Reducing the dimensions

Another approach to roll-up is by reducing the number of dimensions in the data cube. This can be achieved by combining or eliminating dimensions that provide similar or redundant information, resulting in a higher-level summary of the data.

For the specific cube mentioned in the overview section, the roll-up operation is accomplished by climbing up in the concept hierarchy of the Location dimension, specifically from the City level to the Country level, to obtain a higher-level summary of the data aggregated at the country level.

Dice

Dice operation is technique used to generate new sub-cube from existing hypercube. It selects two or more dimensions from hypercube. It creates new sub-cube that focuses on specific data.

By selecting specific dimensions, the resulting sub-cube contains subset of data that meets the criteria defined by the chosen dimensions. It analysis relevant dimensions and obtain more detailed and specific view of the data.

Slice

Slice operation is technique used to select single dimension from given cube in order to generate new sub-cube. You can view data from different perspective by isolating specific dimension of interest.

By applying the Slice operation, new sub-cube is created. These contain only the data related to the selected dimension. Slice operation is valuable for analyzing data from different perspectives. You can also analysis specific attributes in isolation. Users can have deeper understanding of the data by exploring its behavior and relationships from the standpoint of a chosen dimension.

Pivot

Pivot operation is also known as the Rotate operation. Pivot operation is technique used to provide alternate view of the available data to users. It allows for rotating and reorienting the cube's structure to examine the data from different perspectives. Pivot has flexible data analysis and enhances the understanding of the underlying information by presenting it in alternative orientations or arrangements.

Pivot operation enables users to change the orientation of the cube by rearranging the dimensions, resulting in different arrangement of cells and values. This transformation provides new viewpoint for analyzing and interpreting the data.