Mod-5

OLAP and its need?

OLAP, which stands for Online Analytical Processing, is a technology used for analyzing and reporting data in a multidimensional way. It allows users to perform complex calculations, create interactive visualizations, and explore data from different angles.

OLAP enables users to gain a deeper understanding and knowledge about various aspects of their corporate data through fast, consistent, interactive access to a variety of possible views of data.

In OLAP, data is organized into dimensions and measures. Dimensions represent different attributes or perspectives of the data, such as time, geography, products, or customers. Measures, on the other hand, are the numerical values or metrics that you want to analyze, such as sales revenue, quantity sold, or profit.

Need of OLAP

OLAP (Online Analytical Processing) is needed for several reasons:

1. Analyzing Large Volumes of Data: OLAP is designed to handle large volumes of data efficiently. It allows users to analyze massive datasets and perform complex calculations, aggregations, and comparisons across multiple dimensions and hierarchies.
2. Multidimensional Analysis: OLAP enables users to analyze data from different dimensions simultaneously. By slicing, dicing, drilling down, and rolling up data, users can explore patterns, relationships, and trends that might not be apparent in a traditional row-based data representation.
3. Faster Query Performance: OLAP databases are optimized for query performance. They use multidimensional structures, such as cubes or star schemas, and employ indexing and caching techniques to enable fast retrieval and analysis of data.
4. Self-Service Business Intelligence: OLAP empowers business users to perform ad-hoc analysis and create customized reports without relying on IT or data experts. Its intuitive interfaces and interactive visualizations make it easier for non-technical users to explore data and derive insights.
6. Flexibility and Scalability: OLAP systems are flexible and adaptable to changing business requirements. They can accommodate new dimensions, measures, or hierarchies without disrupting existing data structures. Additionally, OLAP databases can scale horizontally or vertically to handle increasing data volumes and user concurrency.
7. Data Consistency and Integrity: OLAP ensures data consistency by integrating data from multiple sources into a single, unified view. It helps maintain data integrity by enforcing relationships and constraints between dimensions and measures.

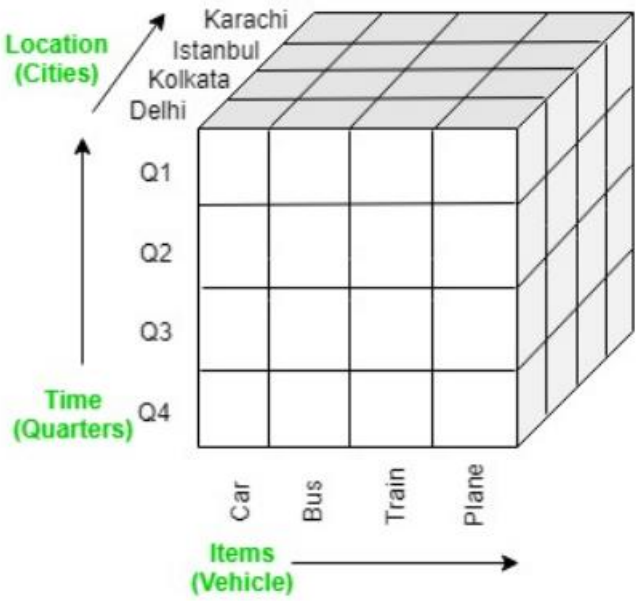
OLTP VS OLAP

Category	OLAP (Online Analytical Processing)	OLTP (Online Transaction Processing)
Definition	It is well-known as an online database query management system.	It is well-known as an online database modifying system.
Data source	Consists of historical data from various Databases.	Consists of only operational current data.
Method used	It makes use of a data warehouse.	It makes use of a standard database management system (DBMS).
Application	It is subject-oriented. Used for Data Mining, Analytics, Decisions making, etc.	It is application-oriented. Used for business tasks.
Normalized	In an OLAP database, tables are not normalized.	In an OLTP database, tables are normalized (3NF).
Usage of data	The data is used in planning, problem-solving, and decision-making.	The data is used to perform day-to-day fundamental operations.

Multidimensional Analysis: Multidimensional analysis is a fundamental concept in OLAP. It involves analyzing data along multiple dimensions simultaneously to gain insights into relationships, patterns, and trends. Multidimensional analysis allows users to examine data from various perspectives, slice and dice data based on different dimensions, and drill down or roll up data to different levels of detail or aggregation.

Some common techniques used in multidimensional analysis include:

1. Slice and Dice: Users can select specific subsets of data by choosing values from one or more dimensions. For example, analyzing sales data for a particular region or a specific time period.
2. Drill Down and Roll Up: Users can navigate through different levels of data hierarchy, drilling down to more detailed levels or rolling up to higher-level summaries. For example, starting with sales data by year and drilling down to quarters, months, or even individual days.
3. Pivot: Users can reorient the dimensions to view the data from different angles or perspectives. This flexibility allows for dynamic analysis and comparison across various dimensions.
4. Calculations and Aggregations: Users can perform calculations, aggregations, and mathematical operations on measures across dimensions. This helps in deriving meaningful insights and supporting decision-making processes.

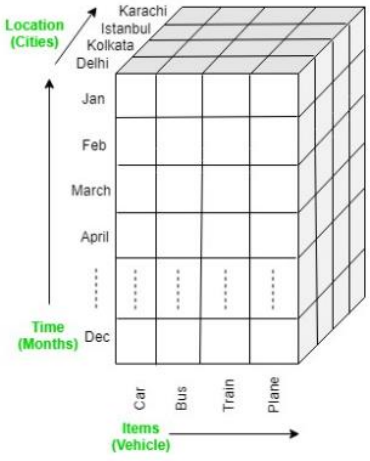


OLAP operations:

There are five basic analytical operations that can be performed on an OLAP cube:

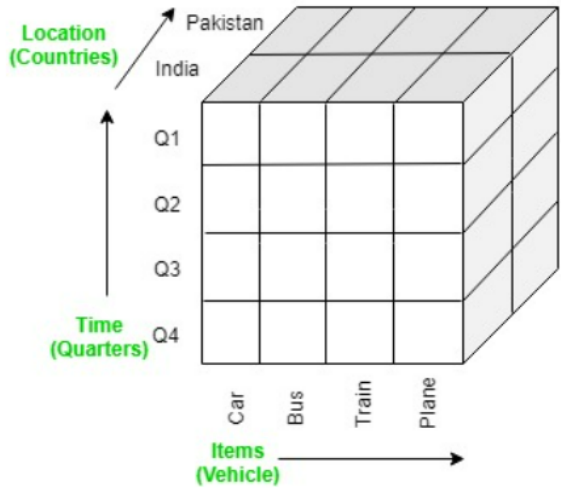
1. **Drill down:** In drill-down operation, the less detailed data is converted into highly detailed data. It can be done by:
   - Moving down in the concept hierarchy
   - Adding a new dimension

   In the cube given in overview section, the drill down operation is performed by moving down in the concept hierarchy of *Time* dimension (Quarter -> Month).

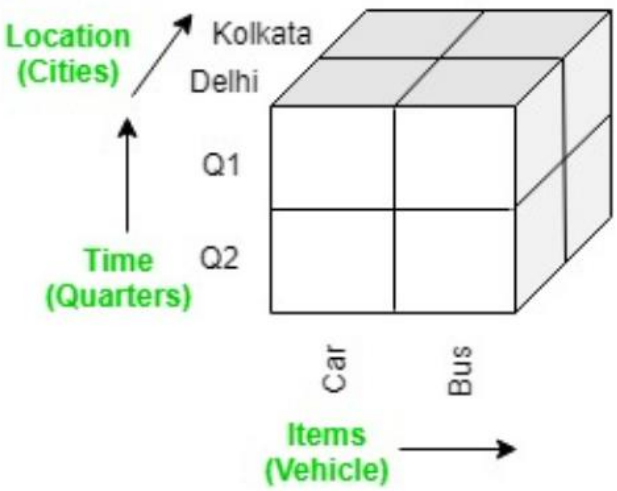


2. **Roll up:** It is just opposite of the drill-down operation. It performs aggregation on the OLAP cube. It can be done by:
   - Climbing up in the concept hierarchy
   - Reducing the dimensions

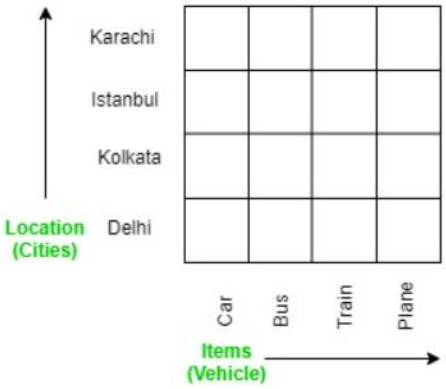
   In the cube given in the overview section, the roll-up operation is performed by climbing up in the concept hierarchy of *Location* dimension (City -> Country).



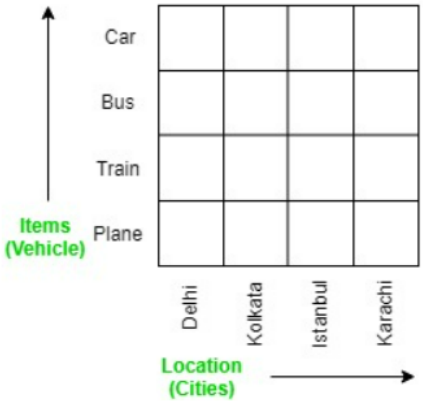
3. **Dice:** It selects a sub-cube from the OLAP cube by selecting two or more dimensions. In the cube given in the overview section, a sub-cube is selected by selecting following dimensions with criteria:
   - Location = “Delhi” or “Kolkata”
   - Time = “Q1” or “Q2”
   - Item = “Car” or “Bus”



4. **Slice:** It selects a single dimension from the OLAP cube which results in a new sub-cube creation. In the cube given in the overview section, Slice is performed on the dimension Time = “Q1”.



5. **Pivot:** It is also known as *rotation* operation as it rotates the current view to get a new view of the representation. In the sub-cube obtained after the slice operation, performing pivot operation gives a new view of it.

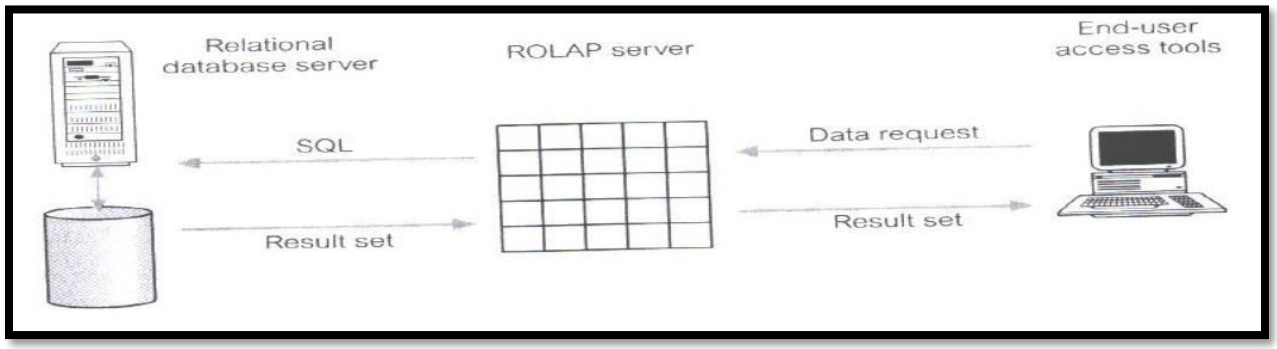


Models of OLAP

ROLAP

ROLAP (Relational Online Analytical Processing) is a model used for organizing and analyzing data in a relational database system. Here are some key points about ROLAP in layman's terms:

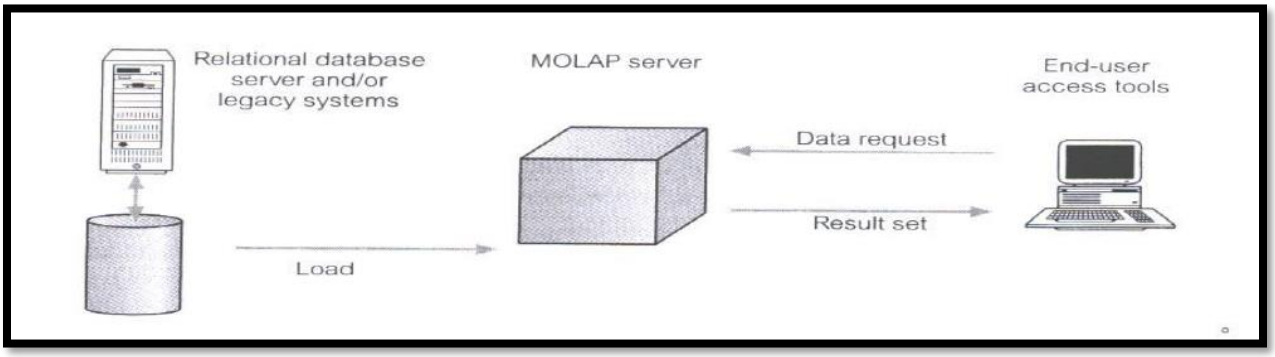
1. Relational Database Foundation: ROLAP is built on top of a relational database system, which is like a structured way of storing and organizing data, similar to a well-organized spreadsheet.
2. Data Organization: ROLAP organizes data in tables, where each table represents a specific aspect of the data. For example, there may be a table for sales transactions, another for customer details, and so on.
3. Flexible Analysis: ROLAP allows users to perform analysis on this structured data by writing queries or using visualization tools. It enables users to ask questions about the data and get meaningful answers.
4. Aggregation and Summarization: ROLAP can aggregate and summarize data across different dimensions. This means you can analyze data at different levels of detail, such as by month, quarter, or year, and perform calculations on the aggregated data.



MOLAP

MOLAP (Multidimensional Online Analytical Processing) is a model used for organizing and analyzing data in a multidimensional way. Here are some key points about MOLAP in layman's terms:

1. Multidimensional Data Storage: MOLAP organizes data in a specialized format called a multidimensional cube. Think of it as a three-dimensional structure where data is stored in cells, with each axis representing a different dimension. For example, you may have dimensions like time, geography, and product categories.
2. Fast Query Performance: MOLAP is designed for fast query performance. By storing data in a pre-aggregated and compressed format, it can quickly retrieve and analyze data from the multidimensional cube. This allows users to interactively explore data and get answers to their analytical questions without waiting for long processing times.
3. Slice, Dice, and Drill: MOLAP allows users to slice, dice, and drill down into data along different dimensions. Slicing means focusing on a specific subset of data, like analyzing sales for a particular time period or region. Dicing involves further refining the analysis by selecting specific values from multiple dimensions. And drilling down allows users to navigate from higher-level summaries to more detailed levels of data.
4. Calculations and Aggregations: MOLAP supports calculations and aggregations within the multidimensional cube. This means you can perform calculations on measures, such as calculating profit margins or growth rates, and aggregate data at different levels of granularity, like aggregating sales from daily to monthly or yearly totals.



HOLAP

HYBRID OLAP (HOLAP)

- HOLAP tools provide limited analysis capability, either directly against RDBMS products, or by using an intermediate MOLAP server.
- HOLAP tools deliver selected data directly from DBMS or via MOLAP server to the desktop (or local server) in the form of data cube, where it is stored, analyzed, and maintained locally is the fastest-growing type of OLAP tools.
- The issues associated with HOLAP tools:
  - The architecture results in significant data redundancy and may cause problems for networks that support many users.
  - Ability of each user to build a custom data cube may cause a lack