UNIT-2
Types of OLAP

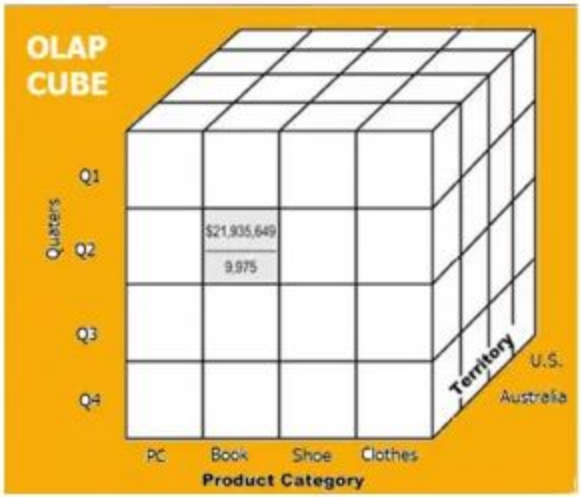
Dr. Shalini Gambhir

What is OLAP?

- **Online Analytical Processing (OLAP)** is a category of software that allows users to analyze information from multiple database systems at the same time. It is a technology that enables analysts to extract and view business data from different points of view.
- 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 cubes. The cubes are designed in such a way that creating and viewing reports become easy. OLAP stands for Online Analytical Processing.

OLAP Cube:

- At the core of the OLAP concept, is an OLAP Cube. The OLAP cube is a data structure optimized for very quick data analysis.
- The OLAP Cube consists of numeric facts called measures which are categorized by dimensions. OLAP Cube is also called the **hypercube**.
- Usually, data operations and analysis are performed using the simple spreadsheet, where data values are arranged in row and column format. This is ideal for two-dimensional data. However, OLAP contains multidimensional data, with data usually obtained from a different and unrelated source. Using a spreadsheet is not an optimal option. The cube can store and analyze multidimensional data in a logical and orderly manner.



How does it work?

- A Data warehouse would extract information from multiple data sources and formats like text files, excel sheet, multimedia files, etc.
- The extracted data is cleaned and transformed. Data is loaded into an OLAP server (or OLAP cube) where information is pre-calculated in advance for further analysis.

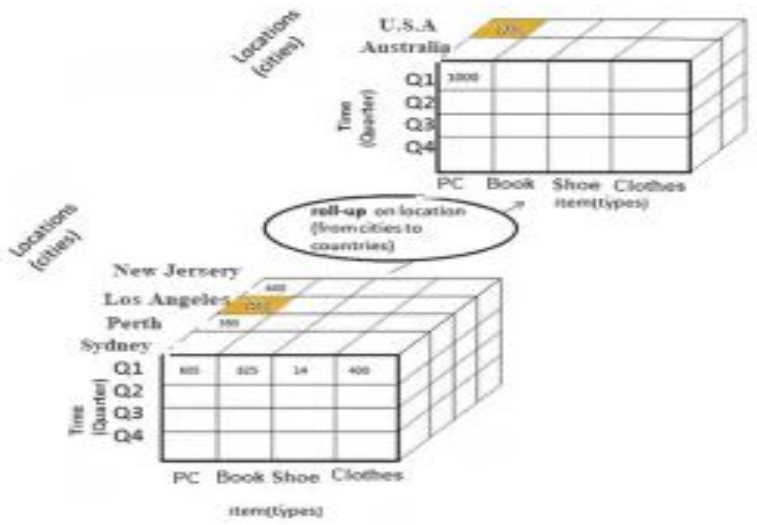
Basic analytical operations of OLAP

- Four types of analytical operations in OLAP are:
- Roll-up
- Drill-down
- Slice and dice
- Pivot (rotate)

1) Roll-up:

- Roll-up is also known as “consolidation” or “aggregation.” The Roll-up operation can be performed in 2 ways
- Reducing dimensions
- Climbing up concept hierarchy. Concept hierarchy is a system of grouping things based on their order or level.

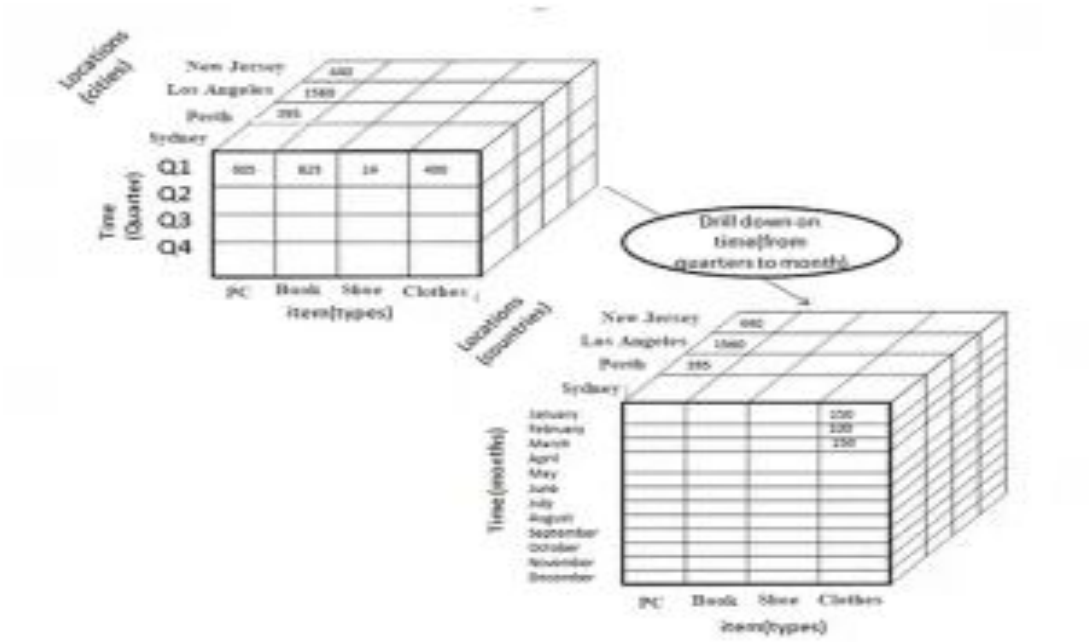
- Consider the following diagram



- In this example, cities New Jersey and Lost Angles and rolled up into country USA
- The sales figure of New Jersey and Los Angeles are 440 and 1560 respectively. They become 2000 after roll-up
- In this aggregation process, data is location hierarchy moves up from city to the country.
- In the roll-up process at least one or more dimensions need to be removed. In this example, Quater dimension is removed.

2) Drill-down

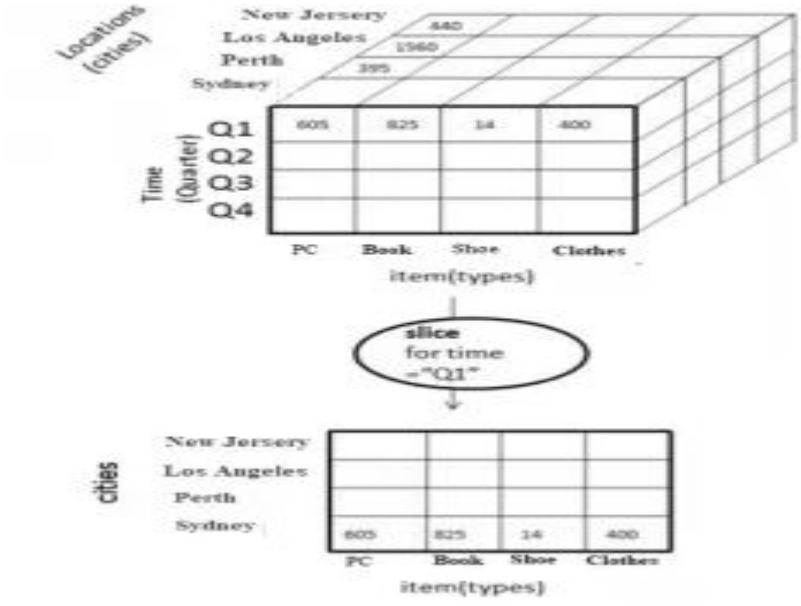
- In drill-down data is fragmented into smaller parts. It is the opposite of the rollup process. It can be done via
- Moving down the concept hierarchy
- Increasing a dimension



Consider the diagram above
Quater Q1 is drilled down to months January, February, and March. Corresponding sales are also registers.
In this example, dimension months are added.

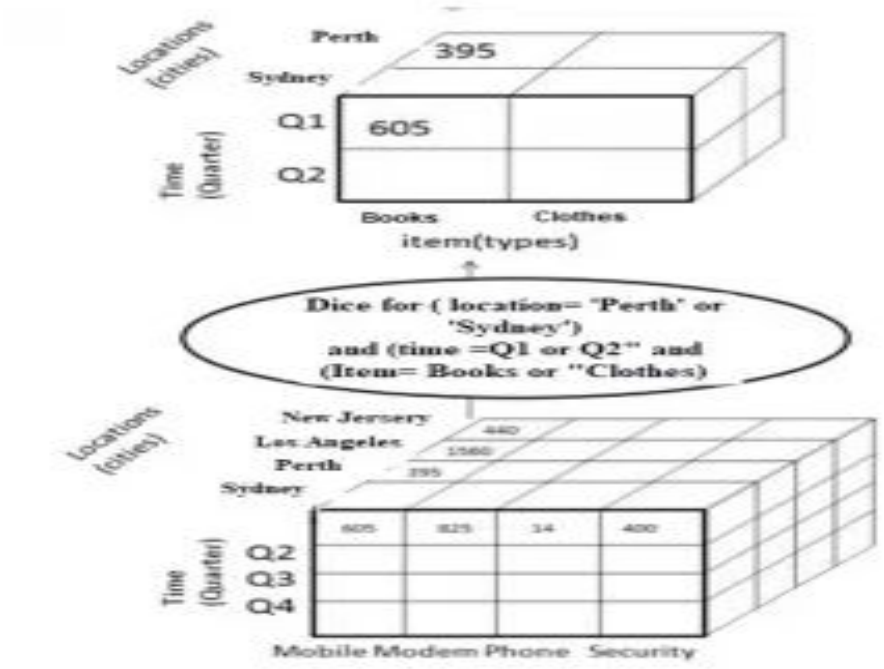
3) Slice:

- Here, one dimension is selected, and a new sub-cube is created.
- Following diagram explain how slice operation performed:
- Dimension Time is Sliced with Q1 as the filter.
- A new cube is created altogether.



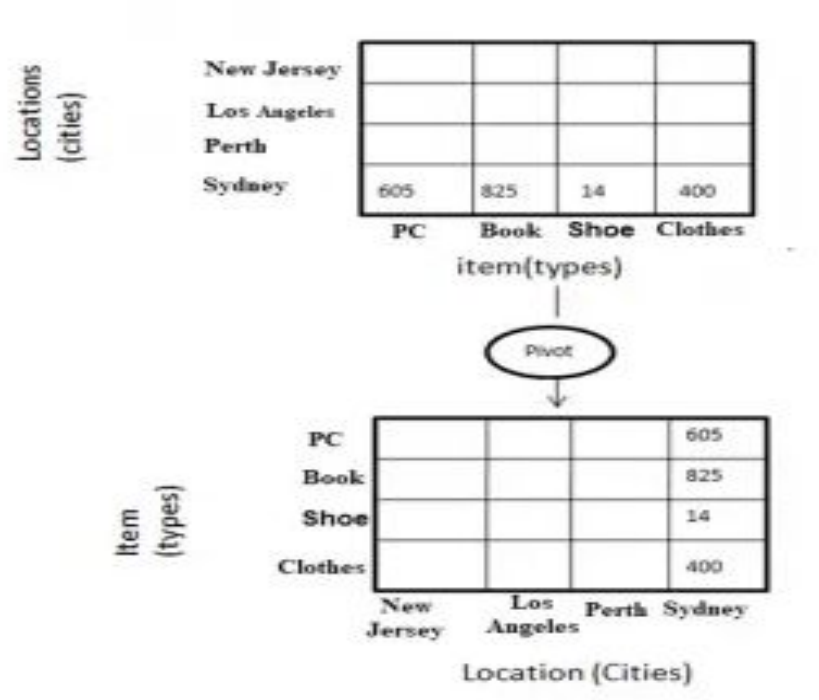
Dice:

- This operation is similar to a slice. The difference in dice is you select 2 or more dimensions that result in the creation of a sub-cube.



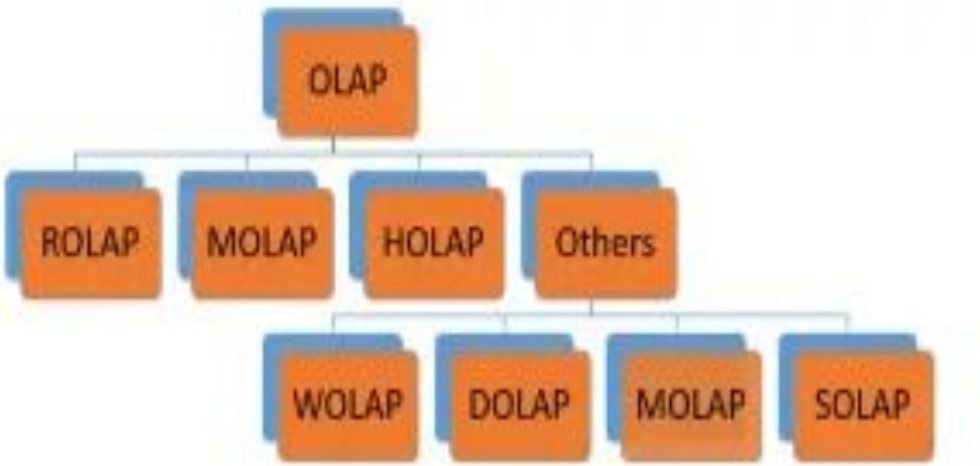
4) Pivot

- In Pivot, you rotate the data axes to provide a substitute presentation of data.
- In the following example, the pivot is based on item types.



Types of OLAP systems

- OLAP Hierarchical Structure



ROLAP

- ROLAP works with data that exist in a relational database. Facts and dimension tables are stored as relational tables. It also allows multidimensional analysis of data and is the fastest growing OLAP.
- **Advantages of ROLAP model:**
- **High data efficiency.** It offers high data efficiency because query performance and access language are optimized particularly for the multidimensional data analysis.
- **Scalability.** This type of OLAP system offers scalability for managing large volumes of data, and even when the data is steadily increasing.

Drawbacks of ROLAP model:

- **Demand for higher resources:** ROLAP needs high utilization of manpower, software, and hardware resources.
- **Aggregately data limitations.** ROLAP tools use SQL for all calculation of aggregate data. However, there are no set limits to the for handling computations.
- **Slow query performance.** Query performance in this model is slow when compared with MOLAP

MOLAP

- MOLAP uses array-based multidimensional storage engines to display multidimensional views of data. Basically, they use an OLAP cube.

What is MOLAP?

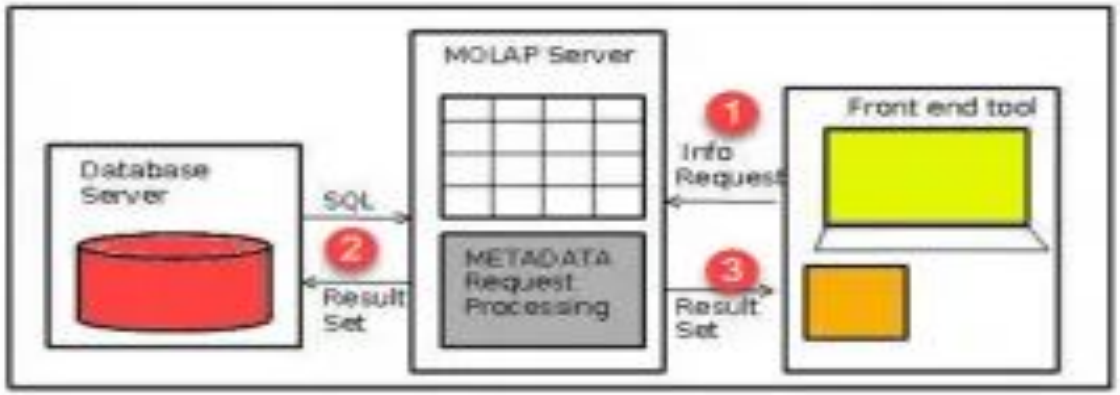
- **Multidimensional OLAP (MOLAP)** is a classical OLAP that facilitates data analysis by using a multidimensional data cube.
- Data is pre-computed, re-summarized, and stored in a MOLAP (a major difference from ROLAP). Using a MOLAP, a user can use multidimensional view data with different facets.
- Multidimensional data analysis is also possible if a relational database is used. But that would require querying data from multiple tables. On the contrary, MOLAP has all possible combinations of data already stored in a multidimensional array. MOLAP can access this data directly. Hence, MOLAP is faster compared to Relational Online Analytical Processing (ROLAP).

Key Points

- In MOLAP, operations are called processing.
- MOLAP tools process information with the same amount of response time irrespective of the level of summarizing.
- MOLAP tools remove complexities of designing a relational database to store data for analysis.
- MOLAP server implements two level of storage representation to manage dense and sparse data sets.
- The storage utilization can be low if the data set is sparse.
- Facts are stored in multi-dimensional array and dimensions used to query them.

MOLAP Architecture

- MOLAP Architecture includes the following components
- Database server.
- MOLAP server.
- Front-end tool.



- Consider above given MOLAP Architectures:
- The user request reports through the interface
- The application logic layer of the MDDB retrieves the stored data from Database
- The application logic layer forwards the result to the client/user.

- MOLAP architecture mainly reads the precompiled data. MOLAP architecture has limited capabilities to dynamically create aggregations or to calculate results that have not been pre-calculated and stored.
- For example, an accounting head can run a report showing the corporate P/L account or P/L account for a specific subsidiary. The MDDB would retrieve precompiled Profit & Loss figures and display that result to the user.

Implementation considerations is MOLAP

- In MOLAP it's essential to consider both maintenance and storage implications to creating strategy for building cubes.
- Proprietary languages used to query MOLAP. However, it involves extensive click and drag support for example MDX by Microsoft.
- Difficult to scale because the number and size of cubes required when dimensions increase.
- API's should provide for probing the cubes.
- Data structure to support multiple subject areas of data analyses which data can be navigated and analyzed. When the navigation changes, the data structure needs to be physically reorganized.
- Need different skill set and tools for Database administrator to build, maintain the database.

MOLAP Advantages

- MOLAP can manage, analyze and store considerable amounts of multidimensional data.
- Fast Query Performance due to optimized storage, indexing, and caching.
- Smaller sizes of data as compared to the relational database.
- Automated computation of higher level of aggregates data.
- Help users to analyze larger, less-defined data.
- MOLAP is easier to the user that’s why It is a suitable model for inexperienced users.
- MOLAP cubes are built for fast data retrieval and are optimal for slicing and dicing operations.
- All calculations are pre-generated when the cube is created.

MOLAP Disadvantages

- One major weakness of MOLAP is that it is less scalable than ROLAP as it handles only a limited amount of data.
- The MOLAP also introduces data redundancy as it is resource intensive
- MOLAP Solutions may be lengthy, particularly on large data volumes.
- MOLAP products may face issues while updating and querying models when dimensions are more than ten.
- MOLAP is not capable of containing detailed data.
- The storage utilization can be low if the data set is highly scattered.
- It can handle the only limited amount of data therefore, it’s impossible to include a large amount of data in the cube itself.

MOLAP Tools

- Essbase — Tools from Oracle that has a multidimensional database.
- Express Server — Web-based environment that runs on Oracle database.
- Yellowfin — Business analytics tools for creating reports and dashboards.
- Clear Analytics — Clear analytics is an Excel-based business solution.
- SAP Business Intelligence — Business analytics solutions from SAP

Summary:

- Multidimensional OLAP (MOLAP) is a classical OLAP that facilitates data analysis by using a multidimensional data cube.
- MOLAP tools process information with the same amount of response time irrespective of the level of summarizing.
- MOLAP server implements two level of storage to manage dense and sparse data sets.
- MOLAP can manage, analyze, and store considerable amounts of multidimensional data.
- It helps to automate computation of higher level of aggregates data
- It is less scalable than ROLAP as it handles only a limited amount of data.

Hybrid OLAP

- Hybrid OLAP is a mixture of both ROLAP and MOLAP. It offers fast computation of MOLAP and higher scalability of ROLAP. HOLAP uses two databases.
- Aggregated or computed data is stored in a multidimensional OLAP cube
- Detailed information is stored in a relational database.

Benefits of Hybrid OLAP:

- This kind of OLAP helps to economize the disk space, and it also remains compact which helps to avoid issues related to access speed and convenience.
- Hybrid HOLAP's uses cube technology which allows faster performance for all types of data.
- ROLAP are instantly updated and HOLAP users have access to this real-time instantly updated data. MOLAP brings cleaning and conversion of data thereby improving data relevance. This brings best of both worlds.

Drawbacks of Hybrid OLAP:

- Greater complexity level**:** The major drawback in HOLAP systems is that it supports both ROLAP and MOLAP tools and applications. Thus, it is very complicated.
- Potential overlaps**:** There are higher chances of overlapping especially into their functionalities.

Advantages of OLAP

- OLAP is a platform for all type of business includes planning, budgeting, reporting, and analysis.
- Information and calculations are consistent in an OLAP cube. This is a crucial benefit.
- Quickly create and analyze “What if” scenarios
- Easily search OLAP database for broad or specific terms.
- OLAP provides the building blocks for business modeling tools, Data mining tools, performance reporting tools.
- Allows users to do slice and dice cube data all by various dimensions, measures, and filters.
- It is good for analyzing time series.
- Finding some clusters and outliers is easy with OLAP.
- It is a powerful visualization online analytical process system which provides faster response times

Disadvantages of OLAP

- OLAP requires organizing data into a star or snowflake schema. These schemas are complicated to implement and administer
- You cannot have large number of dimensions in a single OLAP cube
- Transactional data cannot be accessed with OLAP system.
- Any modification in an OLAP cube needs a full update of the cube. This is a time-consuming process

Summary:

- OLAP is a technology that enables analysts to extract and view business data from different points of view.
- At the core of the OLAP concept, is an OLAP Cube.
- Various business applications and other data operations require the use of OLAP Cube.
- There are primary five types of analytical operations in OLAP 1) Roll-up 2) Drill-down 3) Slice 4) Dice and 5) Pivot
- Three types of widely used OLAP systems are MOLAP, ROLAP, and Hybrid OLAP.
- Desktop OLAP, Web OLAP, and Mobile OLAP are some other types of OLAP systems.