UNIT 5 INTRODUCTION TO ONLINE ANALYTICAL PROCESSING

Structure

5.0 Introduction
5.1 Objectives
5.2 Online Analytical Processing and its Need
5.3 Characteristics of OLAP
5.4 OLAP and Multidimensional Analysis
5.4.1 Multidimensional Logical Data Modeling and its Users
5.4.2 Multidimensional Structure
5.4.3 Multidimensional Operations
5.5 OLAP Functions
5.6 Data Warehouse and OLAP: Hypercube and Multicubes
5.7 Applications of OLAP
5.8 Steps involved in OLAP Creation
5.9 Advantages of OLAP
5.10 OLAP Architectures
5.11 Online Transaction Processing
5.12 Summary
5.13 Solutions/Answers
5.14 Further Readings

5.0 INTRODUCTION

In the earlier unit, we had studied on Extract, Transform and Loading (ETL) component of data warehousing. Online Analytical Processing is the core component of the data warehouse implementation, providing fast and flexible multi-dimensional data analysis for business intelligence (BI) and decision support applications.

In this unit we will study the need of Online Analytical Processing, its functions, applications and Online Transaction Processing Systems.

5.1 OBJECTIVES

After going through this unit, you should be able to:

- understand the purpose of a OLAP;
- describe the motivation and benefits of OLAP;
- discuss Multidimensional Modeling Structure;.
- describe various OLAP operations;
- list the applications of Multi cube and steps to create OLAP server;
- differentiate between various types of OLAP architectures, and
- understand the online transaction processing, its need, applications and characteristics

5.2 ONLINE ANALYTICAL PROCESSING AND ITS NEED

Online Analytical Processing (OLAP) is the technology to analyze and process data from multiple sources at the same time. It accesses the multiple databases at the same time. It helps the data analysts to collect data from different perspective for developing effective business strategies. The query operations like grouping, join or aggregation can be easily done with OLAP using pre-calculated or pre-aggregated data hence making it much faster than simple relational databases. You can understand OLAP as a multi cubic structure, which has many cubes - each cube is pertaining to some database. The cubes are designed in such a way that generates reports effectively and efficiently.

OLAP is a software used to perform high-speed, multivariate analysis of large amounts of data in data warehouses, data markets, or other unified and centralized data warehouses. The data is broken down for display, monitoring or analysis. For example, sales figures can be related to location (region, country, state/province, company), time (year, month, week, day), product (clothing, male/female/child, brand, type), etc., but In a data warehouse, records are stored in tables, and each table can only sort data on two of the dimensions at a time. Recording and reorganizing them into a multi-dimensional format allows very fast processing and very in-depth analysis.

The primary objective of OLAP is not just data processing. For instance, if a company might compare their sales in the month of January with the month of February then compare those results with another location which may be stored in a separate database. In this case, it needs a multi-view of database design storing all the data categories. Another example is in Amazon it analyzes purchases made by its customers to recommend the customers with a personalized home page of products which are likely to be interested by them. It creates a single platform for all types of business analytical means which includes planning budgeting, forecasting and analysis.

The main benefit of OLAP is the consistency of information and calculations using OLAP systems and we can easily apply security restrictions on users and objects to comply with regulations and protect sensitive data. OLAP assists managers in making decisions by giving multidimensional record views that are efficient to provide, hence enhancing their productivity. Due to the inherent flexibility support provided by organized databases, OLAP functions are self-contained. Through extensive control of analysis-capabilities, it permits simulation of business models and challenges.

Let’s see the need to use OLAP to have better understanding of OLAP over relational databases or online transaction processing (OLTP):

- ***Efficient and Effective methods to improve the sales of an Organization***: In a retail, having multiple products with different number of channels for selling the product across the globe. OLAP makes it effective and efficient to search for a product in a different region within a specified time period(like, excluding weekdays sales or just weekend sales or festival duration sales very specific from a very large data distributed).
- ***It improves the sales of a business***: The data analysis power of OLAP brings effective results in sales. It helps in identifying expenditures which produces a high return of investments (ROI).

Usually, data operations and analysis are performed using the simple spreadsheets, 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.

5.3 CHARACTERISITCS OF OLAP

The main characteristics of OLAP are as follows:

- ***Fast:*** OLAP act as bridge between data warehouse and front-end. Hence helps in the better accessibility of data yielding faster results.
- ***Analysis:*** OLAP data analysis and computational measure and their results are stored in separate data files. OLAP distinguishes better zero and missing values. It should ignore missing values and performs the correct aggregate values. OLAP facilitates interactive query handling and complex analysis for the users.
- ***Shared:*** OLAP operations drill-down or roll-up, it navigates between various dimensions in multidimensional cube making it effective and efficient reporting system.
- ***Multidimensional:*** OLAP has Multidimensional conceptual view and access of data to different users at different levels. The increasing number of dimensions and report generation performance of the OLAP system does not significantly degrade.
- ***Data and Information:*** OLAP has calculation power for complex queries and data. It does data visualization using graphs and charts.

5.4 OLAP AND MULTIDIMENSIONAL ANALYSIS

The multi-dimensional data model stores data in the form of data cube in a data warehouse. Generally, it supports two or three-dimension cubes. It gives the data different views and perspectives. Practically in a retail store the data is maintained month-wise, item-wise, region-wise thus involving many different dimensions.

5.4.1 Multidimensional Logical Data Modeling and its Users

The multidimensional data modeling provides:

- ***Different views and perspectives to the data from different angles:*** The business users have a dimensional and logical view of the data in the data warehouse.
- ***Multidimensional conceptual view***: It allows users to have a dimensional and logical view of the data.
- ***Multidimensional Modeling creates environment for multiuser***: Since the OLAP techniques are shared, the OLAP and database operations, containing retrieval, update, adequacy control, integrity, and security can be easily performed.

For example, in the figure 1, it is shown that the dimensions Time, Regions and Products of a company can be logically saved in a cube. In figure 2, the cross tabular form in every quarter products quantity is shown. In figure 1, Products, Time and

regions these dimensions can be combined into cubes you can imagine what two dimensions would look like by using a spreadsheet metaphor with the time dimension as the columns and the products dimension as the rows if we add data to this view such as units sold that would be a measure. ***Measures*** can be any quantity such as revenues expenses unit's statistics or any text or numerical value if we consider adding the third dimension regions then you can imagine each region being represented as an additional spreadsheet this is how it works when you're limited to a two-dimensional spreadsheet. however, an OLAP cube can represent all three dimensions as a single data set which allows users to fluidly explore all the data from any perspective and despite its name, a cube can hold many more than three dimensions so what's the value of using all that to illustrate this:

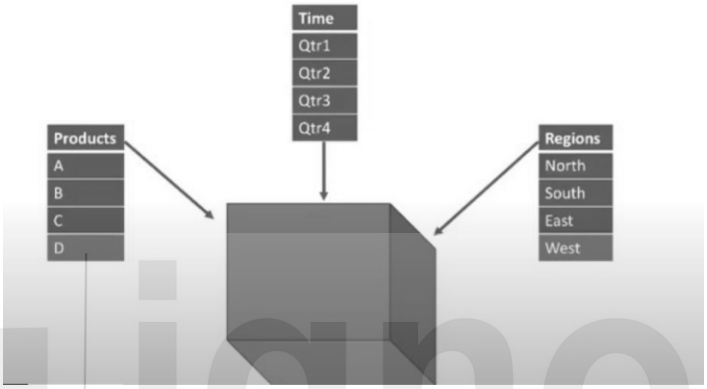

Figure 1: Cube Representation

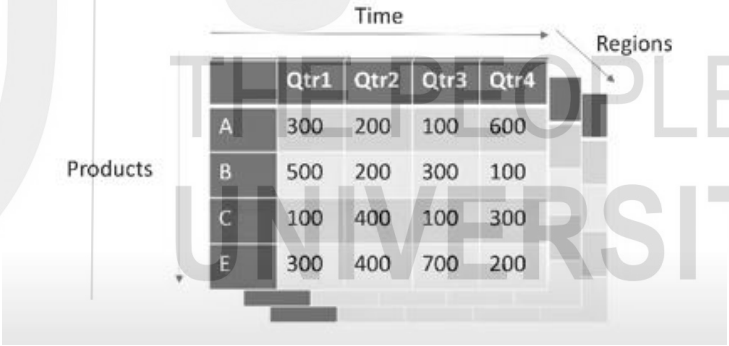

Figure 2: Measurable data shown

Let's say that a manager is tracking sales units with three different spreadsheets with three different dimensions products, quarters and regions from looking at these spreadsheets, it appears that everything is equal as the manager of these stores would probably stock them with the same number of items for each product quarter and region. The manager of a store house makes very different decisions to generate a report with just one or two dimensions or by adding more dimensions and reveal more detail which would allow to make better decisions on managing the inventory of the stores. Hence, you can view OLAP facilitates business oriented multidimensional data having lot of calculations. The data saved in multidimensional structure is very significant in speed thought analysis to companies to take better decisions. OLAP provides the flexibility of data retrieval to generate reports.

5.4.2 Multidimensional Structure

In a multidimensional cube there are dimensions. In each of the dimension of the cube usually we have something called hierarchy which means some of the attributes in a single dimension may form a hierarchy.

For instance let's just use the time dimension we can have years and two quarters into month so kind of we can drill down and look at the more detailed aspects of the dimension. If we use the time dimension as an example, we can have multiple levels of attributes which can help us to form the hierarchy break it down to quarter and break it down to month. Let's use on our grocery database as an example the cells fact table on which has multiple dimensions one of the dimensions is the product dimension and in the product dimension, we can find naturally the following hierarchy. The data has been organized into multiple dimensions and at each level of dimension, contains multiple levels of abstraction defining the Concept Hierarchy. It provides flexibility to view data from different angles.

Concept Hierarchy of a Product

Department → Category → Subcategory→ Brand→ Product

It is important to identify the hierarchy from multi-dimensional cube in terms of query. Then we must look at the performance measure on which attribute or dimension the query is focused on. For instance, department we have grocery and house, so we have two possible values under grocery we have food and then drink obviously based on this relationship.

We know that Department can be broken down into category, so in the grocery department we sell food and drinks and for fresh food if we look at the value, food category can break down into frozen food and fresh food and frozen food can be broken down into two brand codes. So that's kind of the hierarchy you shouldn't count on the sequence of the column within the table to determine the hierarchy you need to really understand your data and sometime study the value stored in the dimensional table that will help you determining hierarchy as well so how this multi-dimensional cube can be used for different queries. Also, during getting the answers, performance measure should be considered by looking at the performance driver for different dimensions.

5.4.3 Multidimensional Operations

Typically, there are four types of OLAP operations are:

i. Roll-up
ii. Drill-down
iii. Slice and Dice
iv. Pivot (rotate)

In practical daily life we come across operations where the manager is interested in knowing the aggregate of data from the concept hierarchy. It can use the concept hierarchy to roll the data up so for instance instead of a daily aggregated data we have monthly aggregate data and quarterly and then annual year.

Concept hierarchy of Time dimension

Year→ Quarter → Month→ Week→ Daily

So, to perform this operation, we can roll-up and store the result of roll-up and subtotal those aggregated data. So, if the manager is interested in going down the concept hierarchy or interested in the minute details to find out the driving attribute responsible for the increase or decrease of sales. For this OLAP operation drill down can be performed.

(i) Roll-up:

- It is an OLAP operation to move from low level of concept hierarchy to higher level known as aggregation.
- It can be performed by reducing the dimensions as well.

In the following example, it is shown a multidimensional cube containing the products of a Home appliances like laptop, furniture, mobile and kitchen appliances, if the manager wants to view the sales of all the products quarterly. The Roll-up operation can be performed on these categories. In this aggregation process, hierarchy moves up and becomes Home Appliance Electronics, as shown in figure 3 below.

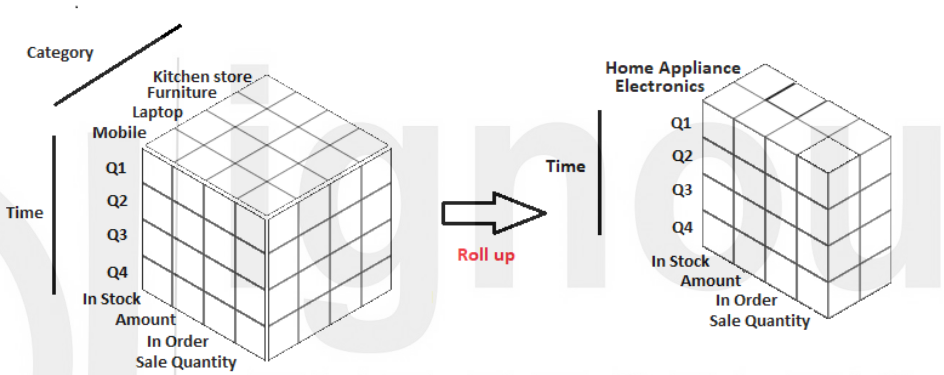

Figure 3: Roll-up on (Category from Home Appliances and Electronics)

(ii) Drill-down:

In this process of scaling down from higher level to the lower level the data is broken into smaller parts.

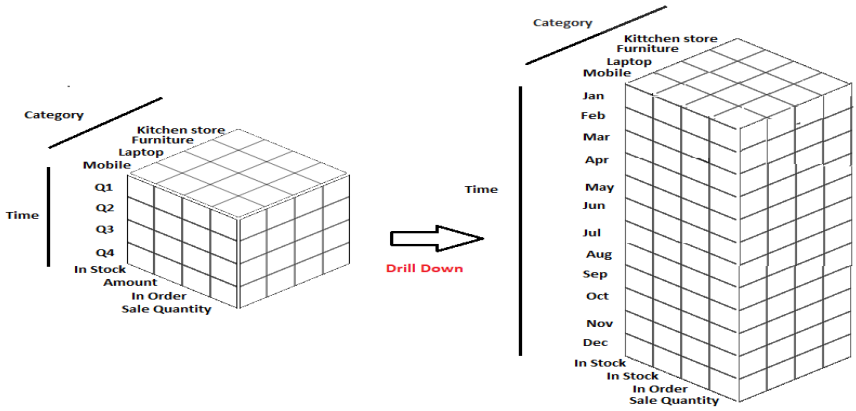

Figure 4: Drill down from Time to Months

You will observe in the above example as shown in figure 4, a multidimensional cube containing products and time. The Time dimension has been drilled-down from

Quarter → Months to observe the sales month-wise.

(iii) Slice:

It is another OLAP operation to fetch the data. In this the query on one dimension is triggered in the database and a new sub cube is created.

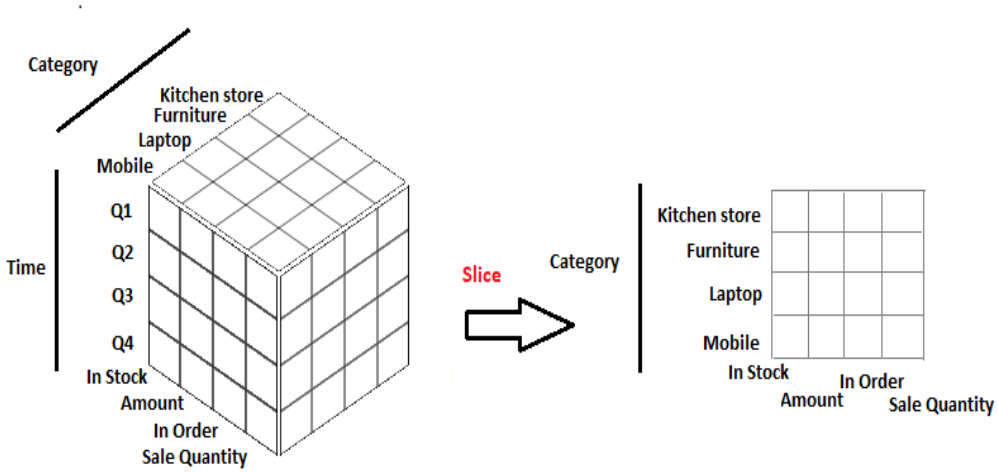

Figure 5: Slice OLAP Operation

In the above figure 5, it can be observed that slice operation is performed on "Time" dimension and a new sub cube is created to retrieve the results.

Slice for Time = "Q1"

(iv) Dice:

This OLAP Dice operation as shown in figure 6 is just like the Projection relational query you have read in RDBMS. In this technique, you select two or more dimensions that results in the creation of a sub cube as shown in the figure.

Dice for (Category= "Laptop" or "Mobile") and (Time = "Q1" or "Q2") and (Stock = "Amount" or "Sale Quantity")

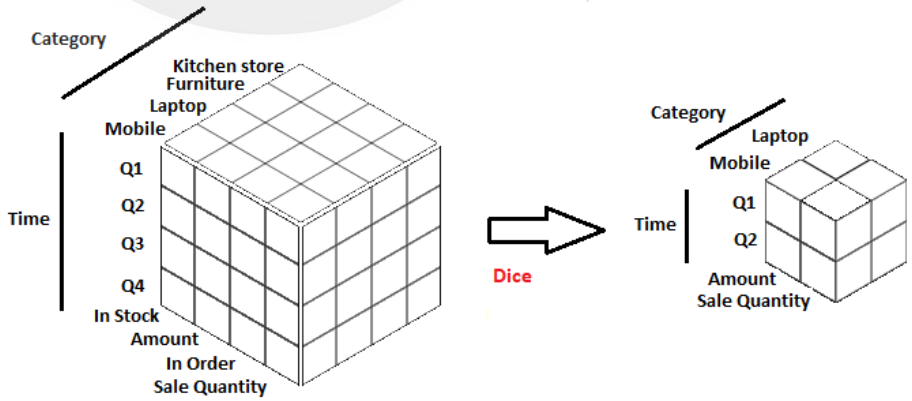

Figure 6: Dice OLAP Operation

(v) Pivot:

This OLAP operation fixes one attribute as a Pivot and rotate the cube to fetch the results, just like an inverted spreadsheet, giving it a different perspective. You can observe in the figure 7, the presentation of the dimensions has been changed to impart a different perspective of the data cube for data analysis.

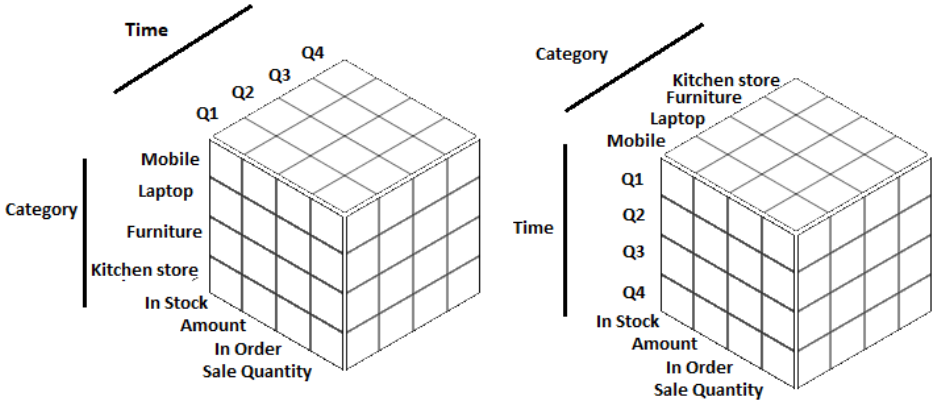

Figure 7: Pivot OLAP operation

☞ Check Your Progress 1:

1) Who are the users of the Multidimensional data modeling?

……

……

……

2) What are the five categories of decision support tool?

……

……

……

5.5 OLAP FUNCTIONS

Online Analytical Processing functions can return the ranking, row numbering. It is very similar to the SQL aggregate functions, but aggregate functions return atomic value.

- The OLAP function returns a scalar value of a query. OLAP functions can be performed at the individual row levels too.
- An OLAP function provides data mining functionalities and data analysis. The detailed data analysis and values are supported with OLAP functions.
- The exhaustive and comprehensive data analysis can be achieved row wise unlike simple SQL functions which produces results in the form of reports like WITH. OLAP runs on rows of the data warehouse.
- OLAP functions uses SQL commands like INSERT/SELECT/ POPULATE on tables or Views.

5.6 DATA WAREHOUSE AND OLAP: HYPERCUBE AND MULTICUBES

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. So, we can say that multidimensional databases can be seen as hypercube and multi cube. Multidimensional cubes have

smaller multiple cubes and in hypercube it seems there is one cube as logically all the data seems to be as one unit of cube.

Table 1: Differences between Multi cube and Hyper cube

	Multi Cube	Hyper Cube
Metadata	Each dimension can belong to many cubes	Each dimension belongs to one cube only
Dimension	Not necessary all the dimensions should belong to some cube	Every dimension owned by a hypercube
Measure Computation	Complex, data can be retrieved from the all the cubes	Simple, as all the numerical facts are available at one place
Multiple	Multicube system, if there are two rows in the DIMENSIONS rowset for which the DIMENSION_NAME value is the same (and the CUBE_NAME value is different), these two rows represent the same dimension. As, sub cubes are built from the same pool of available dimensions.	in a multiple hypercube scenario, it is possible for two hypercubes to have a dimension of the same name, each of which has different characteristics. In this case, the DIMENSION_UNIQUE_NAME value is guaranteed to be different.

5.7 APPLICATIONS OF OLAP

OLAP reporting system is widely used in business applications like:

- Sales and Marketing
- Retail Industry
- Financial Organizations – Budgeting
- Agriculture
- People Management
- Process Management

Examples are Essbase from Hyperion Solution and Express Server from Oracle.

☞ Check Your Progress 2:

1) Explain the OLAP application reporting system in Marketing?

……

……

……

2) What is the purpose of hyper cube. Show slice and dice operation on the sub-cube/hypercube?

……

……

……

3) List the features of an OLAP

……

……

……

5.8 STEPS INVOLVED IN OLAP CREATION

The basic unit of OLAP is an OLAP cube. It is a data structure designed for better and faster retrieval of results from the data analysis. It has dimensions with numeric facts. The data arrangement in rows and columns in multidimensional is the logical view not the physical view. The building of an OLAP cube uses multidimensional array so that the data can be viewed in all the directions, analysis of data and respond to queries can become efficient. For example, the dimensions of cube are customer, time and products and measure count and total sales. The steps involved in the creation of OLAP are as follows:

Step 1: *Extract data from variety of sources* like text, excel sheets, multimedia files, Online Transaction Processing data in flat files.
Step 2*: Transformation and Standardization of data*: Since, the data is distributed and incompatible to each other. It involves the data preprocessing or cleaning part where the semantics of databases are changed into a standard form.
Step 3*: Loading of data*: After all the database nomenclature has been followed then the data is loaded onto the OLAP server or OLAP multidimensional cube.
Step 4*: Building of a Cube for data analysis* • Select the dimensions means set of subsets of significant attributes. • Select the concept hierarchies. • Populate the cube with the relevant data • Select the numeric attribute to apply aggregate function.
Step 5*: Report Generation*

The steps to create OLAP shown in the below figure 8:

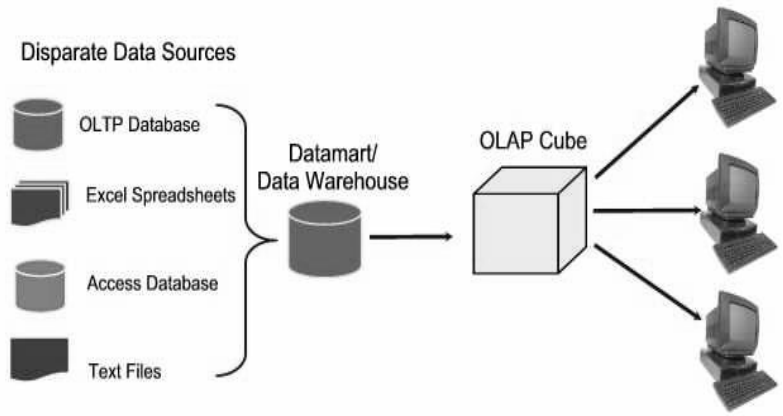

Figure 8 : Steps to create OLAP Cube

5.9 ADVANTAGES OF OLAP

The SQL functions like Group By, Aggregating functions are quite complex to operate in relational databases as compared to multidimensional databases. OLAP can pre-compute the queries can save in sub cubes. The hypercubes also make the computation task faster and saves time. OLAP has proved to an extremely scalable and user – friendly method which is able to perfectly cater to its entire customer needs ranging from small to large companies.

Some listed benefits of using OLAP are as follows:

- ***Data Processing at a faster speed***

The speed of query execution has been tremendous since the use of OLAP technology and is now counted as one of the primary benefits for it. This prevents the customers from spending a lot of time and money on heavy calculations and creating complex reports.

- ***Accessibility***

The cube enables the various kinds of data like – transactional data from various resources, information about every supplier and consumer, etc. all is saved in a concise one location which is easy to operate.

- ***Concise and Fine Data***

OLAP works on the principle of combining multiple and similar records together, which are saved in multiple tables forming a schema between them as a source of connection. Theses tables combine to form the cube to make the massive information concise and yet finely available to the user. Records can be elemental right down to a single element by "drill down" and back to the cube by "drill up" operations.

- ***Data Representation in Multi-Dimension***

OLAP cube is the center of all the data. Each element of the cube contains various attributes and the number of processes performed on it. The cube axes are outlined by the measure and dimension of the cube which is mostly three - dimensional system. This allows the user to take the information from various slices of the cube. A cube slice is a two – dimensional in nature which gives a clear image of the knowledge trying to be represented.

- ***Business Expressions commonly used***

The size of an OLAP cube consisting of data portrays the company's economic and financial conditions. The end user does not manipulate the database files; they deal with end processes like products, salesmen, employees, customers, etc. This gives a reason to even user with less to zero technical background to use OLAP technology.

- ***Situational Scenarios***

The way the cube can cover almost all parts of a data item is through creating various what – if situations; these what – if situations help in extraction of cube information without tampering the original information on the cube. This feature of OLAP technology is responsible for providing the customers the ability to update the values to look at the consequences brought in the cube's situation. Through this feature business intelligence can deeply examine the possible factors of driving a situation in a company and prevent them if necessary.

- ***Easily Understood Technology***

Most of the users or customers working on OLAP technology come from a background of less to minimum technology skills. They mostly do not need any unique training to use this technology, which in return helps the company save some

money. Moreover, OLAP technology providers provide their end users with enough tutorial, documents and some start off technical assistance particularly in case of web based OLAP operations. The end customers are given sessions to continuously work with a group of technical experts so that they do not have to solve all the OLAP issues by themselves.

5.10 OLAP ARCHITECTURE: MOLAP, ROLAP, HOLAP AND DOLAP

There are various types of OLAP architecture: ROLAP, MOLAP, HOLAP and others as shown in the below figure 9.

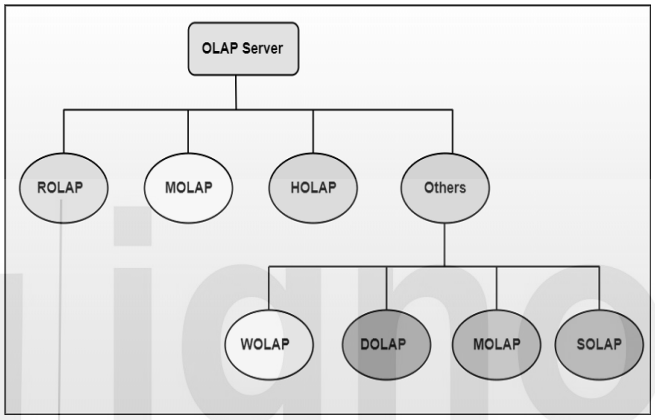

Figure 9: Types of OLAP Architecture

5.10.1 ROLAP Architecture

ROLAP implies Relational OLAP, an application based on relational DBMSs. It performs dynamic multidimensional analysis of data stored in a relational database. The architecture is like three-tier. It has three components viz. front end (User Interface), ROLAP server (Metadata request processing engine) and the back end (Database Server) as shown in the figure 10.

- Database server
- ROLAP server
- Front-End tool

In this three-tier architecture the user submits the request and ROLAP engine converts the request into SQL and submits to the backend database. After the processing of request by the engine, it presents the resulting data into multidimensional format to make the task easier for the client to view it.

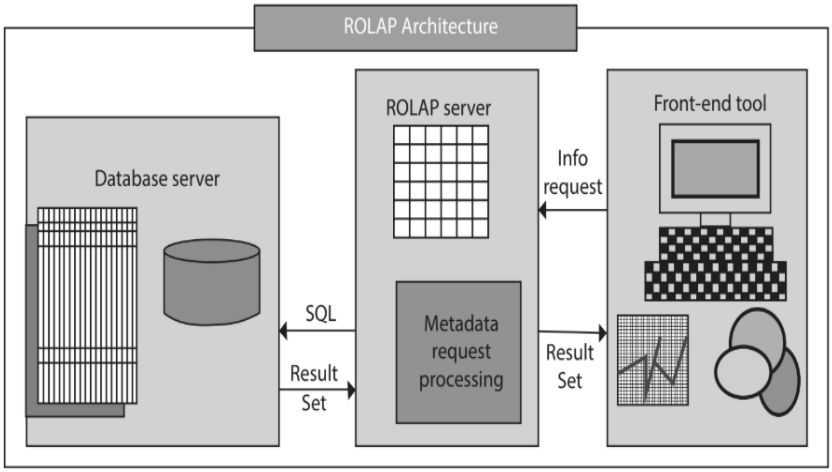

Figure 10 : ROLAP Architechture

The characteristics of ROLAP are:

- ROLAP utilizes the more processing time and disk space.
- ROLAP enables and supports larger user group in the distributed environment.
- ROLAP processes complex queries utilizing the greater amounts of data.

Popular ROLAP products include Metacube by Stanford Technology Group, Red Brick Warehouse by Red Brick Systems.

5.10.2 MOLAP Architecture

MOLAP stands for Multidimensional Online Analytical Processing. It processes the data using the multidimensional cube using various combinations. Since, the data is stored in multidimensional structure the MOLAP engine uses the pre-computed or pre-stored information. The architecture has three components:

- Database server
- MOLAP server
- Front-End tool

As shown in Figure 11, MOLAP engine processes pre-compiled information. It has dynamic abilities to perform aggregation of concept hierarchy. MOLAP is very useful in time-series data analysis and economic evaluation.

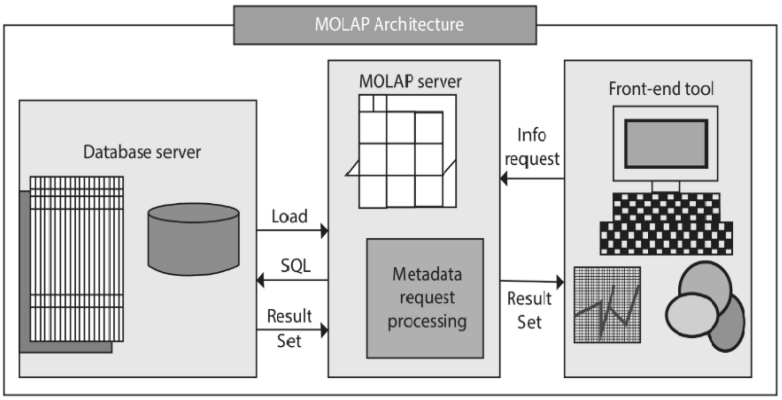

Figure 11 : MOLAP Architechture

The characteristics of MOLAP

- It is a user-friendly architecture, easy to use.
- The OLAP operations slice and dice speeds up the data retrieval.
- It has small pre-computed hypercubes.

Tools that incorporate MOLAP include Oracle Essbase, IBM Cognos, and Apache Kylin.

5.10.3 HOLAP Architecture

It defines Hybrid Online Analytical Processing. It is the hybrid of ROLAP and MOLAP technologies. It connects both the dimensions together in one architecture. It stores the intermediate or part of the data in ROLAP and MOLAP. Depending on the query request it accesses the databases. It stores the relational tables in ROLAP structure, and the data requires multidimensional views, stored and processed using MOLAP architecture as shown in figure 12.

It has the following components:

- Database server.
- ROLAP and MOLAP server.
- Front-end tool

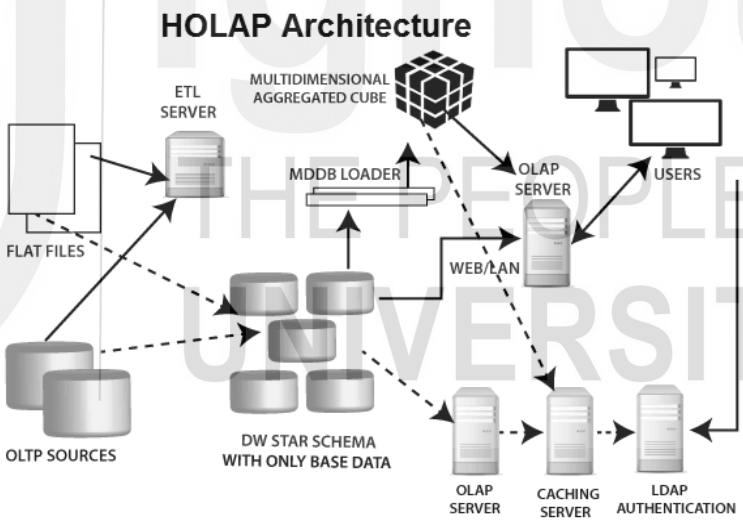

Figure 12 : HOLAP Architecture

The characteristics of HOLAP are:

- Flexible handling of data.
- Faster aggregation of data.
- HOLAP can drill down the hierarchy of data and can access to relational database for any relevant and stored information in it.

Popular HOLAP products are Microsoft SQL Server 2000 presents a hybrid OLAP server.

5.10.4 DOLAP Architecture

Desktop Online Analytical Processing (DOLAP) architecture is most suitable for local multidimensional analysis. It is like a miniature of multidimensional database or it's like a sub cube or any business data cube. The components are:

- Database Server
- DOLAP server
- Front End

The characteristics of DOLAP are:

- The three-tier architecture is designed for low-end, standalone user like a small shop owner in the locality.
- The data cube is locally stored in the system so, retrieval of results is faster.
- No load on the backend or at the server end.
- DOLAP is relatively cheaper to deploy.

5.11 ONLINE TRANSACTION PROCESSING

Online Transactional Processing (OLTP) enables the real-time execution of large numbers of database transactions by large numbers of people, typically over the internet.

A database transaction is a change, insertion, deletion, or query of data in a database. OLTP systems (and the database transactions they enable) drive many of the financial transactions we make every day, including online banking and ATM transactions, eCommerce and in-store purchases, and hotel and airline bookings, to name a very few. In each of these cases, the database transaction also remains as a record of the corresponding financial transaction. OLTP can also drive non-financial database exchanges, including password changes and text messages.

In OLTP, the common, defining characteristic of any database transaction is its *atomicity* (or indivisibility) – a transaction either succeeds as a whole or fails (or is canceled). It cannot remain in a pending or intermediate state.

Examples:

Since the inception of the internet and the e-commerce era, OLTP systems have grown ubiquitous. They're found in nearly every industry or vertical market and in many consumer-facing systems. Everyday examples of OLTP systems include the following:

- ATM machines (this is the classic, most often-cited example) and online banking applications
- Credit card payment processing (both online and in-store)
- Order entry (retail and back-office)
- Online bookings (ticketing, reservation systems, etc.)
- Record keeping (including health records, inventory control, production scheduling, claims processing, customer service ticketing, and many other applications)

5.11.1 Characteristics of OLTP Systems

The most common architecture of an OLTP system that uses transactional data is a three-tier architecture that typically consists of a presentation tier, a business logic tier, and a data store tier. The presentation tier is the front end, where the transaction originates via a human interaction or is system-generated. The logic tier consists of rules that verify the transaction and ensure all the data required to complete the transaction is available. The data store tier stores the transaction and all the data related to it.

The main characteristics of an online transaction processing system are the following:

- **ACID compliance:** OLTP systems must ensure that the entire transaction is recorded correctly. A transaction is usually an execution of a program that may require the execution of multiple steps or operations. It may be complete when all parties involved acknowledge the transaction, or when the product/service is delivered, or when a certain number of updates are made to the specific tables in the database. A transaction is recorded correctly only if all the steps involved are executed and recorded. If there is any error in any one of the steps, the entire transaction must be aborted and all the steps must be deleted from the system. Thus OLTP systems must comply with atomic, consistent, isolated, and durable (ACID) properties to ensure the accuracy of the data in the system.
    - ***Atomic:*** Atomicity controls guarantee that all the steps in a transaction are completed successfully as a group. That is, if any steps between the transactions fail; all other steps must also fail or be reverted. The successful completion of a transaction is called commit. The failure of a transaction is called abort.
    - ***Consistent:*** The transaction preserves the internal consistency of the database. If you execute the transaction all by itself on a database that's initially consistent, then when the transaction finishes executing the database is again consistent.
    - ***Isolated:*** The transaction executes as if it were running alone, with no other transactions. That is, the effect of running a set of transactions is the same as running them one at a time. This behavior is called serializability and is usually implemented by locking the specific rows in the table.
    - ***Durable:*** The transaction's results will not be lost in a failure.

- **Concurrency:** OLTP systems can have enormously large user populations, with many users trying to access the same data at the same time. The system must ensure that all these users trying to read or write into the system can do so concurrently. Concurrency controls guarantee that two users accessing the same data in the database system at the same time will not be able to change that data, or that one user has to wait until the other user has finished processing before changing that piece of data.
- **Scale:** OLTP systems must be able to scale up and down instantly to manage the transaction volume in real time and execute transactions concurrently, irrespective of the number of users trying to access the system.
- **Availability:** An OLTP system must be always available and always ready to accept transactions. Loss of a transaction can lead to loss of revenue or may have legal implications. Because transactions can be executed from anywhere in the world and at any time, the system must be available 24/7.
- **High throughput and short response time:** OLTP systems require nanosecond or even shorter response times to keep enterprise users productive and to meet the growing expectations of customers.
- **Reliability:** OLTP systems typically read and manipulate highly selective, small amounts of data. It is paramount that at any given point of time the data

in the database is reliable and trustworthy for the users and applications accessing that data.

- **Security:** Because these systems store highly sensitive customer transaction data, data security is critical. Any breach can be very costly for the company.
- **Recoverability:** OLTP systems must have the ability to recover in case of any hardware or software failure.

☞ Check Your Progress 3:

1) Compare ROLAP, MOLAP and HOLAP.

……

……

……

2) Write limitations of OLAP cube.

……

……

……

5.12 SUMMARY

OLAP has proven to be an asset in the field of Business Intelligence as it helps in relieving the large amount of data handling along adding the cost benefits of working with this very technique. Furthermore, OLAP providers normally offer their clients with significant documentation, tutorials, and spark off technical assistance in terms of web-primarily based totally OLAP clients. The customers are continuously loose to deal with the group of tech experts while not having to control all the troubles tied to the software program themselves. The concept hierarchies help to organize the dimensions into logical levels. The various OLAP operations help to extract information across sub cubes. The creation of cube and types of OLAPs helps to understand the architecture and usage of various applications of OLAP.

5.13 SOLUTIONS/ANSWERS

Check Your Progress 1

1) Knowledge workers such as data analysts, business analysts, and Executives are the users of OLAP.

2) Decision making Tool features are:
   - Report Generation
   - Query Handling
   - EIS (Executive Information System)
   - OLAP (Online Analytical Processing)
   - Data Mining

Check Your Progress 2

1) In Marketing, OLAP can be used for various purposes as it helps like planning, budgeting, financial marketing, sales data analysis and forecasting. The customer experience is very important to all the companies. So, OLAP works very efficiently in analyzing the data of customers, market research analysis, cost-benefit analysis of any project considering all the dimensions.

There are various OLAP tools available. The OLAP tool should have the ability to analyze large amounts of data, data analysis, fast response to the queries and data visualization. For example, IBM Cognos is a very powerful OLAP marketing tool.

2) Purpose of Hypercube in OLAP: The cube is basically used to represent data with some meaningful measure to compute. Hypercube logically has all the data at one place as a single unit or spreadsheet which makes the computation of queries faster. Each dimension logically belongs to one cube.

For example, a multidimensional cube contains data of the cities of India, Product, Sales and Time with conceptual hierarchy (Delhi→2018→Sales). As, shown in below figures.

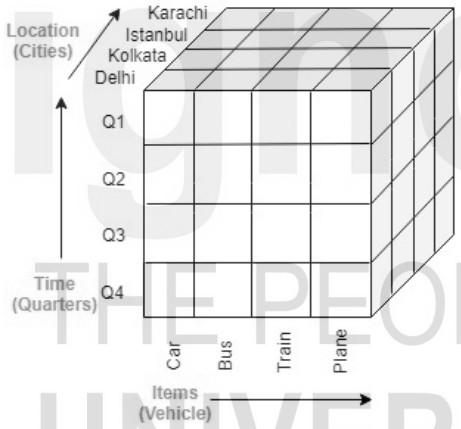

Figure 13: Multidimensional Cube

In the cube given in the overview section, a sub-cube (hypercube) is selected with the following conditions:

Location = “Delhi” or “Kolkata” Time = “Q1” or “Q2”, Item = “Car” or “Bus”

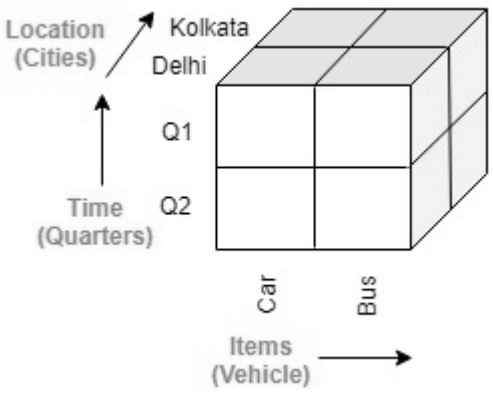

Figure 14: Hypercube or sub-cube

Slice is performed on the dimension Time = "Q1".

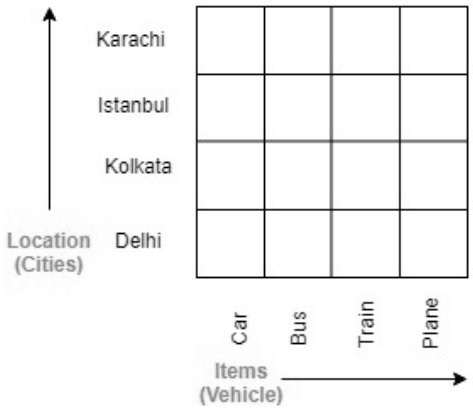

Figure 15: Slice

On the sub-cube, pivot operation is performed.

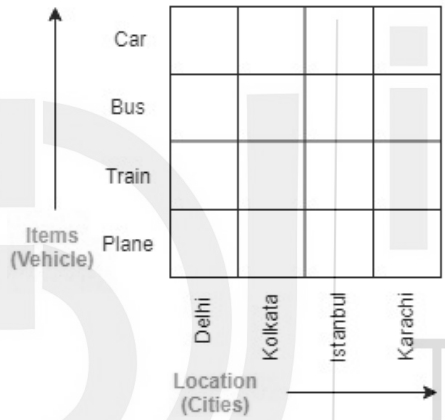

Figure 16: Pivot operation

3) The features of OLAP are:

- Conceptual multidimensional view
- Accessibility of data
- Efficient and flexible Reporting system
- Client/Server architecture
- Supports unrestricted dimensions and aggregation levels
- Uses dynamic sparse matrix handling for faster query results
- Multiuser support

Check Your Progress 3

1) Limitations of OLAP cube are:
   - OLAP requires a star/snowflake schema
   - There is a limited number of dimensions (fields) in a single OLAP cube
   - It is nearly impossible to access transactional data in the OLAP cube
   - Changes to an OLAP cube requires a full update of the cube – a lengthy process

2) Following table 2 summarizes the comparative analysis of ROLAP, MOLAP and HOLAP:

Table 2: Differences between ROLAP, MOLAP and HOLAP

FEATURES	ROLAP	MOLAP	HOLAP
Accessibility of data and Processing time	Very slow because of join operation between tables. The data is fetched from data warehouse.	Fast because of multidimensional storage. The data is fetched from multidimensional data cube.	Fast.
Storage space requirement	Data is stored in relational tables. Comparatively Large storage space requirement	Data is stored in multidimensional tables. Medium storage space requirements	It uses both ROLAP, MOLAP. Small storage space requirements. No duplicate data.
Latency	Low latency	High latency	Medium latency
Query response time	Slow query response time	Fast query response time.	Medium query response time
Volume of data	Used for large volumes of data	Limited volume of data	Can be used in both scenarios
Retrieval of data	Complex SQL queries are used	Sparse Matrix is used	Both
Data View	Static view of data	Dynamic view of data	Both static and dynamic view of data

5.14 FURTHER READINGS

1. William H. Inmon, Building the Data Warehouse, Wiley, 4th Edition, 2005.
2. Data Warehousing Fundamentals, Paulraj Ponnaiah, Wiley Student Edition
3. Data Warehousing, Reema Thareja, Oxford University Press.
4. Alex Berson and Stephen J.Smith, ―Data Warehousing, Data Mining & OLAP‖, Tata McGraw – Hill Edition, 35th Reprint 2016.
5. K.P. Soman, Shyam Diwakar and V. Ajay, ―Insight into Data Mining Theory and Practice, Eastern Economy Edition, Prentice Hall of India, 2006.
6. Ian H.Witten and Eibe Frank, ―Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, Second Edition.