ISSN No: - 2456 – 2165

A Comprehensive Review on OLAP Models and operations

Max Rodrigues

V.E.S Institute of Technology, Collector's colony,

Chembur, Mumbai-74

***Abstract* — In this paper, the Online Analytical Processing system (OLAP) and its logical models have been discussed. By the use of OLAP operations, data analysis is done which can be used to help the user or the person in decision making. The modern data warehouses store or accumulate a large amount of data to perform computations. Here, OLAP operations are used for the computation of complex sub queries in the data warehouse. Decision support systems depend on multidimensional modeling. Therefore, by using OLAP (Online Analytical Processing), decision makers can analyze the data and can arrive at a point where the highest probability of the decision taken will result in success. Data at different aggregation levels, are analyzed using OLAP operators such as roll-up and drill-down. Roll-up operators decrease the details of the measure, aggregating it along the dimension hierarchy and drill-down operators increase the details of the measure.**

***Keywords* —** *OLAP, data warehousing, logical models.*

I. INTRODUCTION

OLAP is an Online Analytical Processing system which performs multidimensional analysis of business data. It has the capability to perform complex calculations, trend analysis and sophisticated data modeling. OLAP is part of the broader category of business intelligence, which also compasses relational database, report writing and data mining [1].OLAP tools enable users to analyze multidimensional data interactively from multiple perspectives. OLAP consists of three basic analytical operations: Roll up, drill-down and slicing and dicing. The three types of OLAP models are: MOLAP (multi-dimensional online analytical processing), ROLAP (Relational online analytical processing), HOLAP (Hybrid online analytical processing)

II. OLAP MODELS

A. MOLAP

MOLAP (multi-dimensional online analytical processing) is the classic form of OLAP and is sometimes referred to as just OLAP. MOLAP stores this data in an optimized multi-dimensional array storage, rather than in a relational database.[2]

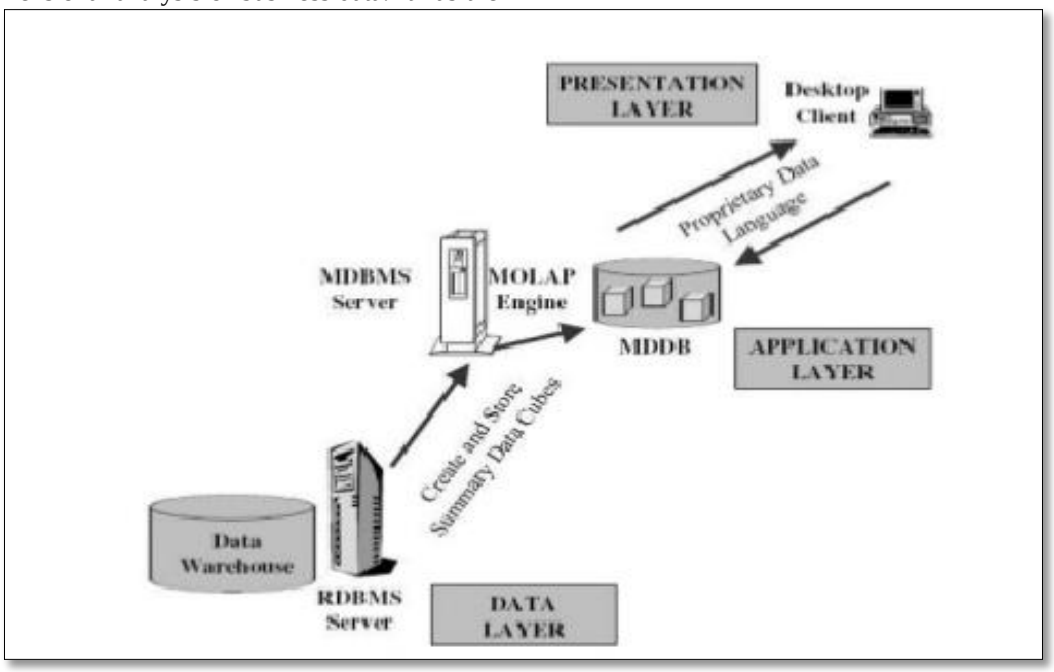

Fig 1 Schematic of a Molap model [3]

ISSN No: - 2456 – 2165

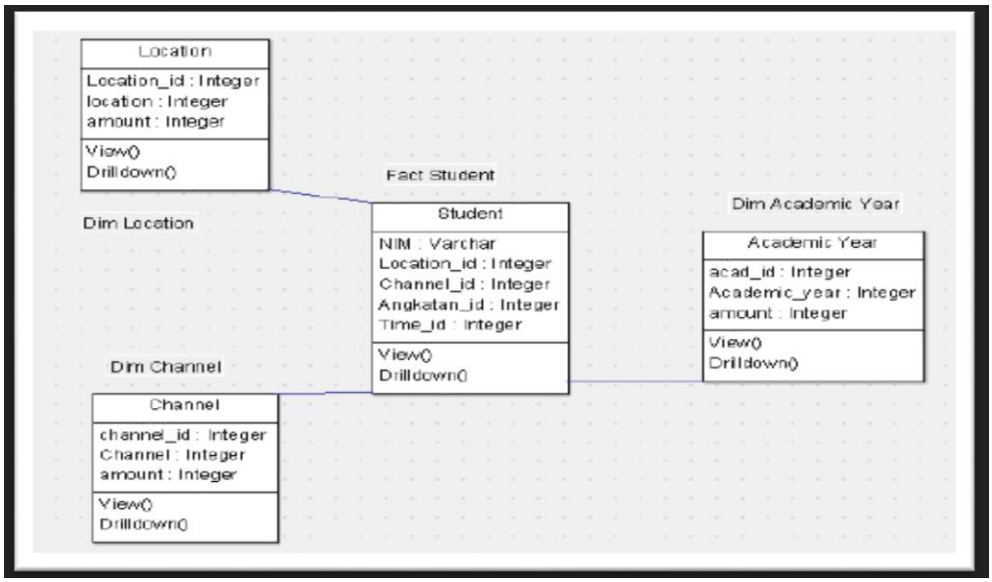

Fig 2 Molap schema [4]

B. Advantages of MOLAP

- The advantage of the MOLAP architecture is, that it provides a direct multidimensional view of the data [5]
- Fast query performance due to optimized storage, multidimensional indexing and caching.[2]
- Smaller on-disk size of data compared to data stored in relational database due to compression techniques. [2]
- Automated computation of higher level aggregates of the data and it is very compact for low dimension data sets. [2]
- Array models provide natural indexing and effective data extraction is achieved through the pre-structuring of aggregated data. [2]

C. *ROLAP*

ROLAP works directly with relational databases and does not require pre-computation. The base data and the dimension tables are stored as relational tables and new tables are created to hold the aggregated information. It depends on a specialized schema design. This methodology relies on manipulating the data stored in the relational database to give the appearance of traditional OLAP's slicing and dicing functionality[2]

D. Advantages of ROLAP

- ROLAP is considered to be more scalable in handling large data volumes, especially models with dimensions with very high cardinality (i.e., millions of members).[2]
- The data are stored in a standard relational database and can be accessed by any SQL reporting tool (the tool does not have to be an OLAP tool).[2]
- ROLAP architecture has two advantages: (a) it can be easily integrated into other existing relational information systems, and (b) relational data can be stored more efficiently than multidimensional data [5]

E. HOLAP Hybrid OLAP

The undesirable trade-off between additional ETL cost and slow query performance has ensured that most commercial OLAP tools now use a "Hybrid OLAP" (HOLAP) approach, which allows the model designer to decide which portion of the data will be stored in MOLAP and which portion in ROLAP. [2]

III. OLAP OPERATIONS

A. ROLL UP [6]

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

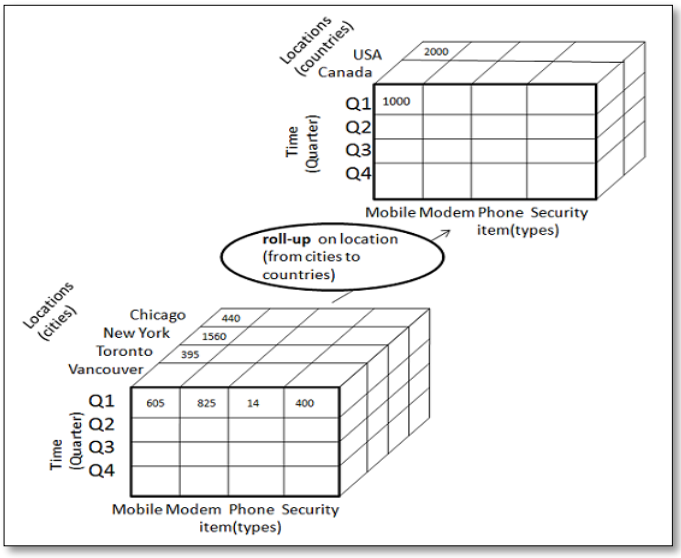

Fig 3 Roll up operation [6]

ISSN No: - 2456 – 2165

- Roll-up is performed by climbing up a 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 city to the level of country.
- The data is grouped into cities rather than countries.
- When roll-up is performed, one or more dimensions from the data cube are removed.

B. DRILL DOWN [6]

Drill-down is the reverse operation of roll-up. It is performed by either of 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:

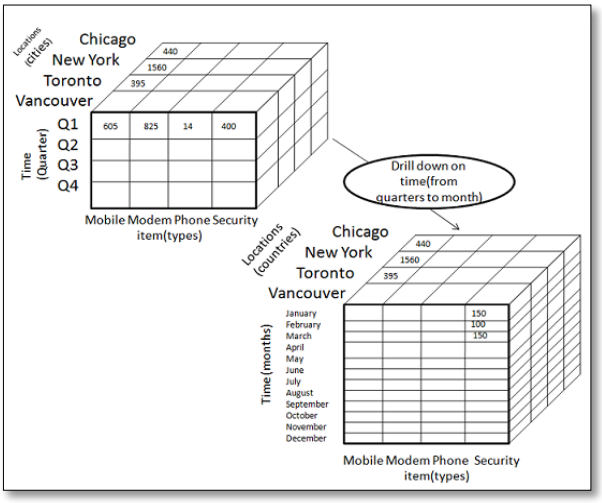

Fig 4 Drill down operation [6]

- Drill-down is performed by stepping down a 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 quarter to the level of month.
- When drill-down is performed, one or more dimensions from the data cube are added.
- It navigates the data from less detailed data to highly detailed data

C. SLICE [6]

The slice operation selects one particular dimension from a given cube and provides a new sub-cube. Consider the following diagram that shows how slice works.

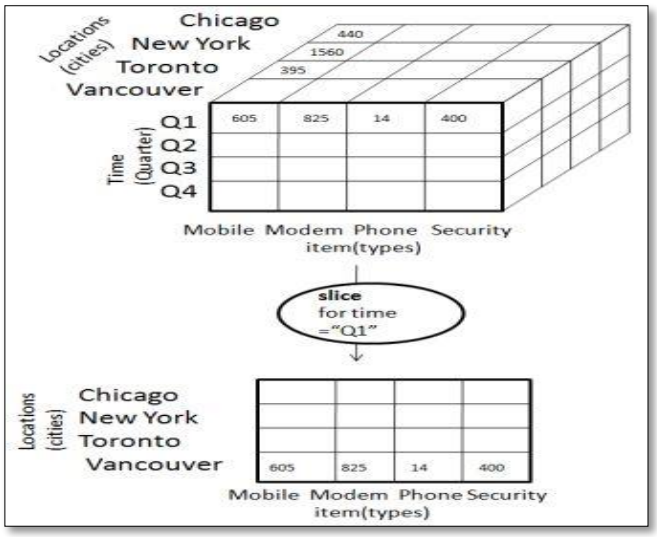

Fig 5 Slice operation [6]

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

D. DICE [6]

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

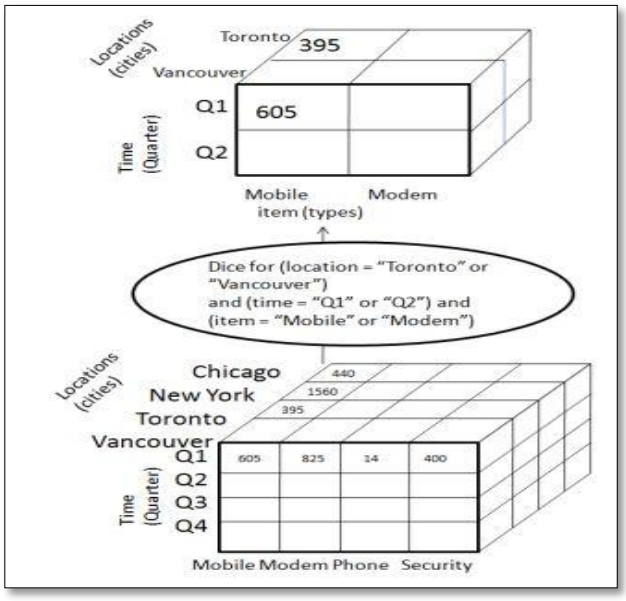

Fig 6 Dice operation [6]

ISSN No: - 2456 – 2165

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

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

IV. CONCLUSIONS

OLAP (Online Analytical Processing) is a fundamental component of a decision making system where a lot of aspects and unknown contents are worth studying. Here, in this study, the various OLAP models and its operations are studied in depth. We have delved into the advantages of ROLAP over MOLAP and how ROLAP is considered to be more scalable in the management of large data volumes.

V. FUTURE SCOPE

Until recently, MOLAP and ROLAP models have been explored and used on a large scale. Therefore, research and investigation should be done for the widespread use of WOLAP – Web-based OLAP, DOLAP – Desktop OLAP, and RTOLAP – Real-Time OLAP. A huge amount of work can be done in the research and development of OLAP Visualization which is still in its preliminary stage. Also, a flexible and efficient system that can support intelligent user interaction with multidimensional data can be explored in the future.

REFERENCES

[1] Deepak Pareek, "Business Intelligence for Telecommunications", CRC Press (2007) ,pp.294.
[2] Nigel Pendse "OLAP architectures", OLAP Report (2006).
[3] Surajit Chaudari, Umeshwar Dayal," An Overview of Data Warehousing and OLAP Technology", VLDB Conference (1996), pp. 1-10.
[4] Dashboard Marketing System for Student's Enrollment. Case Study : UNIS Tangerang - Scientific Figure on ResearchGate.[ONLINE]
[5] Panos Vassiliadis, Timos Sellis," A Survey on logical Models for OLAP Databases, Research Gate (1999), PP. 2-5.
[6] Cristina Ciferri, Ricardo Ciferri Leticia Gómez Markus Schneider Alejandro Vaisman, Esteban Zimányi, "Cube Algebra: A Generic User Centric Model and QueryLanguage for OLAP Cubes, International Journal of Data Warehousing and Mining(2012), pp. 1-23.