Samy Benjelloun, Stephen Collins, Aaron Lance, Wesley Rank, Kevin Taweesup
Professor Cynthia Cheng
CIS 3454 Section 1
16 April 2024

Recommendation - Approach to OLAP for a Pharma Manufacturer

There are three main approaches to Online Analytical Processing (OLAP) in data warehousing: Multidimensional OLAP (MOLAP), Relational OLAP (ROLAP), and Hybrid OLAP (HOLAP). Each approach has its strengths and weaknesses, and the choice of which one to use depends on the specific requirements of the data warehouse project. In this detailed exploration of OLAP, we will discuss the key advantages of each approach to help pharmaceutical manufacturers make informed decisions for their data warehouse.

Multidimensional OLAP (MOLAP)

MOLAP utilizes a multidimensional database, often called a cube, to store and analyze data. Data in MOLAP systems is pre-aggregated and stored in a multidimensional format. This format allows MOLAP to respond quickly to queries and manipulate data efficiently. MOLAP is great for business processes such as financial analysis, sales forecasting, and inventory management.

The key advantages of MOLAP include:

- Fast query response times (resulting from pre-aggregated data)
- Efficient data manipulation and analysis capabilities
- Able to manage large volumes of data
- User-friendly interface that is intuitive for end-users

Some of the disadvantages of MOLAP include:

- Significant upfront investment in software and hardware

- Maintenance and data storage can be resource-intensive and complex

- Scalability could be a challenge as the volume of data grows

- It is not a very flexible model to manage changes in schema

Relational OLAP (ROLAP)

ROLAP leverages the existing RDBMS to store and analyze data through direct operation of relational databases. ROLAP does not require pre-aggregated data or a specialized multidimensional storage format, unlike MOLAP. ROLAP is better suited for transaction processing, customer relationship management, and enterprise resource planning systems.

The key advantages of ROLAP include:

- Utilization of existing relational databases (reduces the need for additional infrastructure investment)

- Flexibility in handling changes in schema

- Can leverage scalability features of the RDBMS

- Can manage large volumes of data

Some of the disadvantages of ROLAP include:

- Slower query response times

- Dependence on RDBMS performance

- Complexity in the design and maintenance of the relational database schema

- Increased resource utilization on the database server

Hybrid OLAP (HOLAP)

HOLAP (Hybrid Online Analytical Processing) combines MOLAP and ROLAP, including their strengths, offering flexibility in handling complex data. It also uses multidimensional databases for fast data retrieval and relational databases for managing large volumes of detailed data. This approach supports efficient query performance with scalability, making it suitable for many analytical needs. However, HOLAP systems can be challenging to manage and require a lot of resources for setup and maintenance.

Advantages of HOLAP:

- Faster querying, faster than ROLAP

- Higher processing ability, when related to ROLAP and MOLAP systems

- HOLAP is the best solution when you anticipate increasing data volume over time.

Disadvantages of HOLAP:

- More extensive data volume, as both ROLAP and MOLAP are combined to form HOLAP

- It is hard to maintain as updates are often needed.

- Since HOLAP combines ROLAP and MOLAP, designing the HOLAP system requires balancing storage and performance.

Tables:

·   Sales Orders – Order ID, Year, Week, Order Date, Completion Date

·   Delivery Location – Location ID, City, State, Address

·   Employee – Employee ID, Name, Employee Org, Building

·   Plant – Plant ID, Plant Name

·   Customer – Customer ID, Name

·   Item – Item ID, Item Type, Item Description, Item Cost

·   Fact – Order ID, Location ID, Employee ID, Plant ID, Customer ID, Item ID, Items

Requested, Items Sold, Items Forecasted, Revenue

Granularity:

The granularity of each table would be the ID number. So, Order ID would indicate a

unique record in Sales Orders, Location ID would indicate a unique record in Delivery Location,

and the same would be true for each respective table.

Conclusion:

Based on the tables and granularity outlined above, while also considering the fact that

this will be for a large pharmaceutical manufacturer which deals with a large amount of data,

HOLAP seems to be the most suitable recommendation. HOLAP will enable this manufacturer

to manage querying in a fast and efficient manner for analytical tasks, while also allowing it to

manage large datasets. This approach provides a balanced solution to the organization in terms of scalability, maintenance, and performance.

Work Cited

Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to

Dimensional Modeling. John Wiley & Sons.

Inmon, W. H. (2005). Building the Data Warehouse. John Wiley & Sons.

Chaudhuri, S., & Dayal, U. (1997). An Overview of Data Warehousing and OLAP Technology.

ACM SIGMOD Record, 26(1), 65-75.

Data Warehousing - Relational OLAP. (n.d.). Www.tutorialspoint.com.

https://www.tutorialspoint.com/dwh/dwh_relational_olap.htm

Pedamkar, P., (2023). HOLAP. www.educba.com/holap/

Contribution Sheet:

Samy Benjelloun	HOLAP definition, pros and cons.
Stephen Collins	Introduction, MOLAP, and proofreading
Aaron Lance	Tables and Granularity
Wesley Rank	Conclusion, revision, formatting
Kevin Taweesup	ROLAP definition, pros and cons