

Problem Statement

Data modeling is the process of creating a data model by using some techniques.[2] The techniques used in this process can be various including IE¹, IDEF1X², UML³, and E-R Diagram⁴.

This article will evaluate these different methods of data modeling and then try to state the differences between them.

References

- [1] datawarehouse4u. Oltp vs. olap. <http://datawarehouse4u.info/OLTP-vs-OLAP.html>. Accessed March 28, 2018.
- [2] Wikipedia. Data modeling - wikipedia. https://en.wikipedia.org/wiki/Data_modeling. Accessed April 3, 2018.

Overview of Data Modeling

As stated before, data modeling is the process of creating a data model. Data modeling is a very helpful scratch when people from different fields want to understand or share the design and concepts of the database. A data model can be conceptual, logical, or physical. Conceptual data models are more about concepts, while physical model pays more attention to implementation.

There are various techniques that can be used in data modeling. And IE, IDEF1X, UML, and E-R Diagram are the most commonly used among those techniques.

Critical Thinking

Table 1: The Differences Between OLTP and OLAP

	OLTP	OLAP
IE?	traditional relational database	data warehouse system
IDEF1X	consolidation data	
UML	daily transaction processing	mining useful information
E-R Diagram	short and piecemeal	time-consuming and acting on lots of data

¹Information Engineering

²Integration DEFinition for information modeling

³Unified Modeling Language

⁴Entiy-Relation Diagram

Question

When I looked up for the information of OLTP and OLAP, I saw some implementation details of OLAP. OLAP is implemented in a variety of ways, such as MOLAP, ROLAP, and HOLAP. So what are MOLAP, ROLAP, and HOLAP? What is the difference between them? Which implementation method to use?

Method

Describe how you are going to answer your own question stated above.

Analysis and Discussion

MOLAP is Multidimensional OLAP. In MOLAP, data is indexed directly into a multidimensional database. So what is a multidimensional database? A multidimensional database is a database in which all data is stored in an array for the user. In this way, users can obtain a high-dimensional cube in this multidimensional data by adding multidimensional constraints. For example, the user queries cities where annual precipitation is between 500 and 700, latitude is between 30 degrees and 50 degrees north latitude, and longitude is between 90 degrees and 110 degrees east longitude. Due to the use of fast indexing, MOLAP's time efficiency is very high. However, because the data in high-dimensional cubes may be sparse, the space efficiency is relatively poor.

The full spelling of ROLAP is Relational Online Analytical Processing. As you can guess from the name, ROLAP is an OLAP built on a relational database. When using ROLAP, the system converts the query to a SQL statement and executes it on a relational database. For the problem of city selection just now, ROLAP adds a few "WHERE". Because it is built on a relational database, ROLAP is very space efficient, but time efficiency is not as high as MOLAP.

HOLAP is Hybrid OLAP, which is a strategy that combines the previously mentioned MOLAP and ROLAP. For a rough query (for example, the number of cities that meet the requirements), HOLAP uses MOLAP to query high-dimensional cubes. However, when you need to query more detailed information (such as the name of each city), you need to use the underlying relational database to query.