

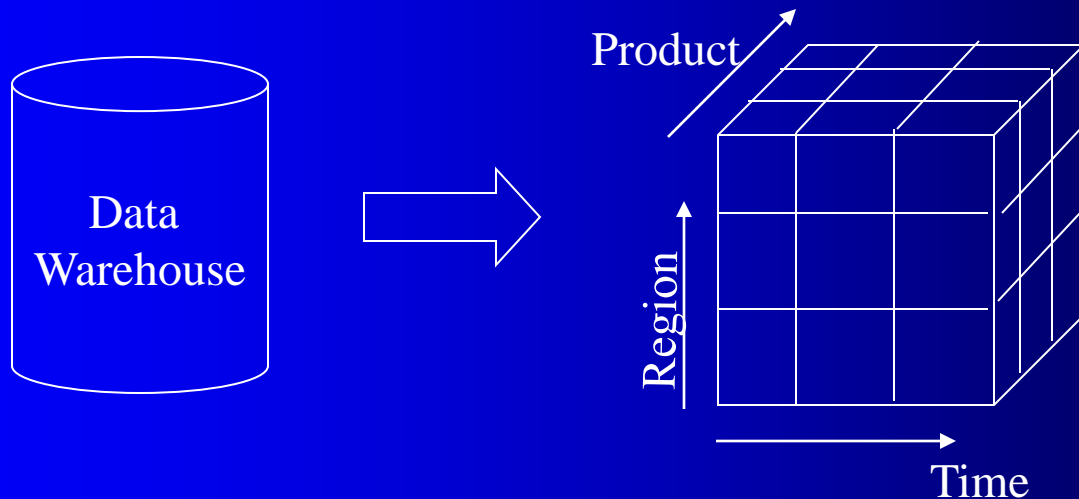
Online Analysis Processing (OLAP)

OLTP and OLAP

- OLTP (on-line transaction processing)
 - Major task of traditional relational DBMS
 - Day-to-day operations: purchasing, inventory, banking, manufacturing, payroll, registration, accounting, etc.
- OLAP (on-line analytical processing)
 - Major task of data warehouse system
 - Data analysis and decision making

Online Analysis Processing(OLAP)

- It enables analysts, managers and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user.



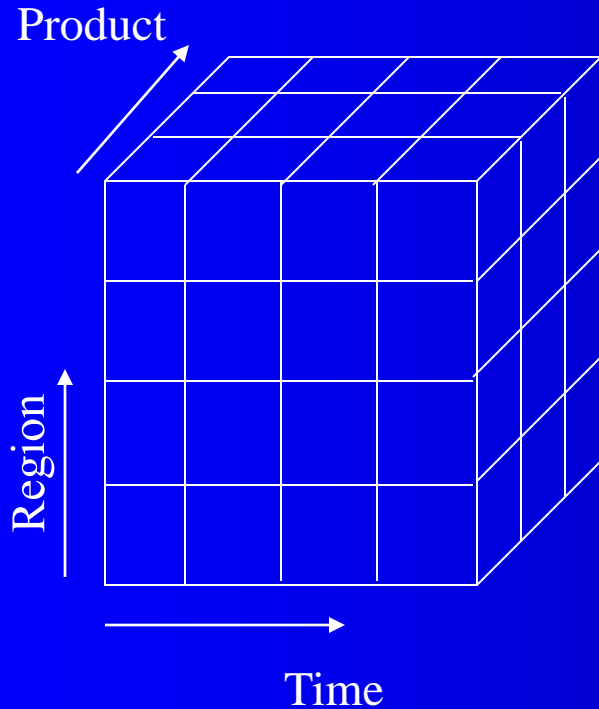
- On-line analytical processing (OLAP) can be performed in data warehouses/marts using the multidimensional data model.
- Typical OLAP operations include rollup, drill-down, slice-and-dice, pivot (rotate).
- OLAP operations can be implemented efficiently using the data cube structure.

OLAP Cube

City	Product	Time	Units	Dollars
All	All	All	113	251.26
Mumbai	All	All	64	146.07
Mumbai	White Bread	All	38	98.49
Mumbai	Wheat Bread	All	13	32.24
Mumbai	Wheat Bread	Qtr1	3	7.44
Mumbai	Wheat Bread	March	3	7.44

OLAP Operations

Drill Down



Category e.g Electrical Appliance



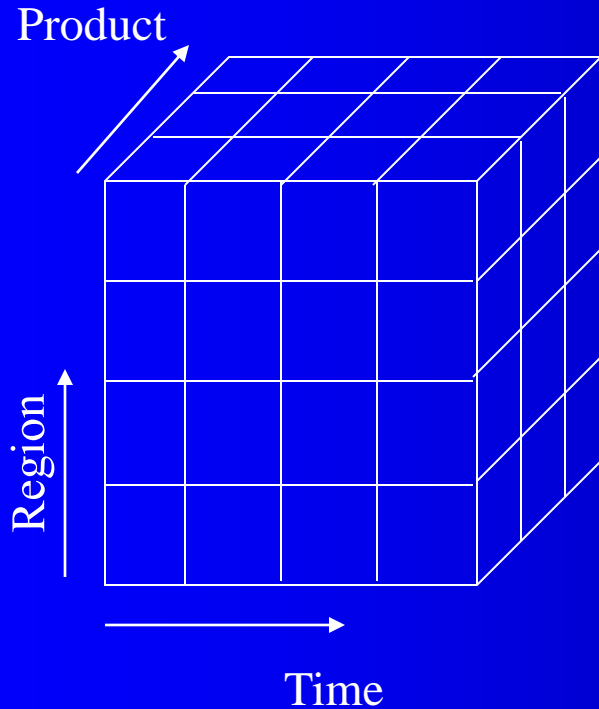
Sub Category e.g Kitchen



Product e.g Toaster

OLAP Operations

Drill Up



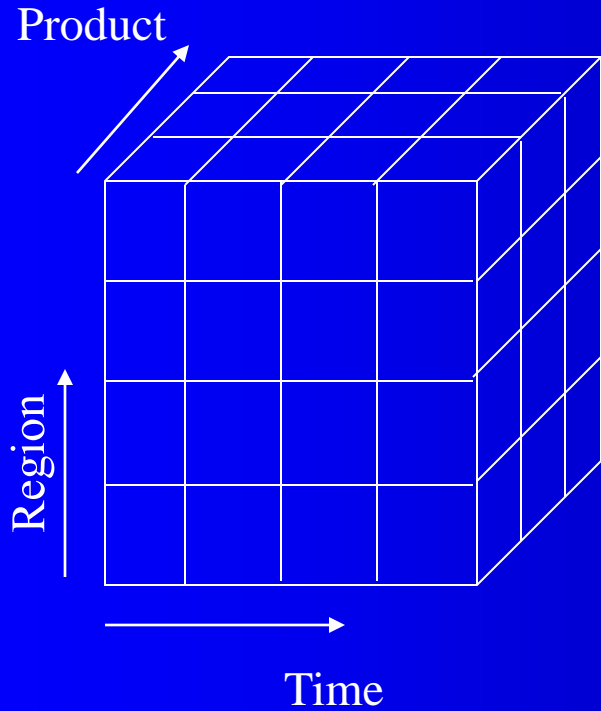
Category e.g Electrical Appliance

Sub Category e.g Kitchen 

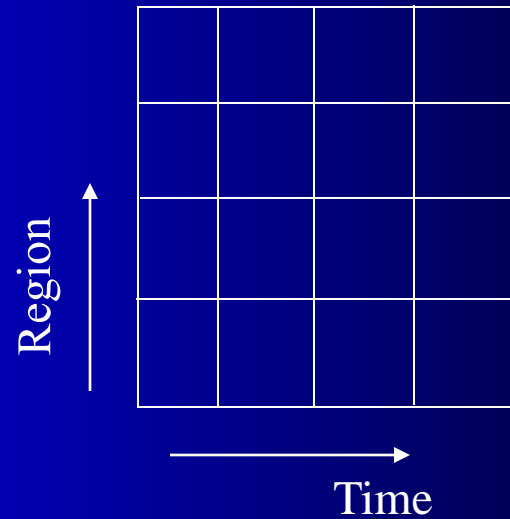
Product e.g Toaster 

OLAP Operations

Slice and Dice

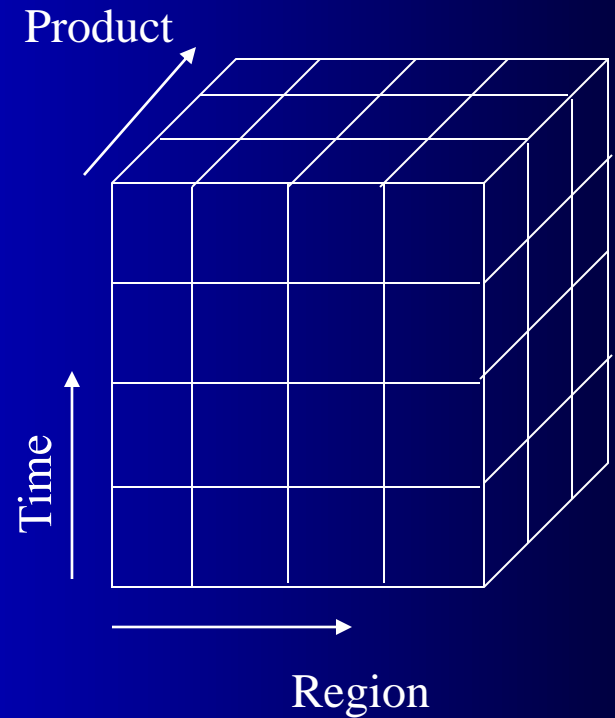
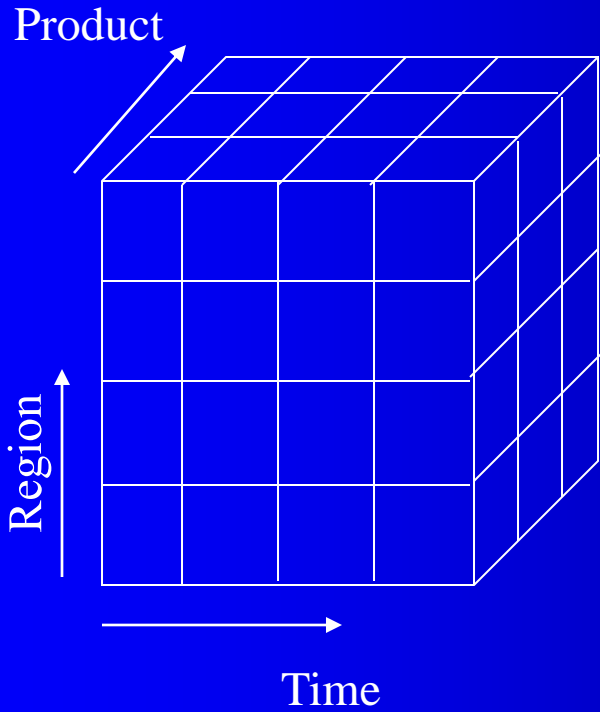


Product=Toaster



OLAP Operations

Pivot



OLAP Server

- An OLAP Server is a high capacity, multi user data manipulation engine specifically designed to support and operate on multi-dimensional data structure.
- OLAP server available are
 - ROLAP server
 - MOLAP server
 - HOLAP server

ROLAP Server

- Relational OLAP (ROLAP) servers: These are the intermediate servers that stand in between a relational back-end server and client front-end tools. They use a relational or extended-relational DBMS to store and manage warehouse data, and OLAP middleware to support missing pieces. ROLAP servers include optimization for each DBMS back end, implementation of aggregation navigation logic, and additional tools and services. ROLAP technology tends to have greater scalability than MOLAP technology.
- The DSS server of Microstrategy, for example, adopts the ROLAP approach.

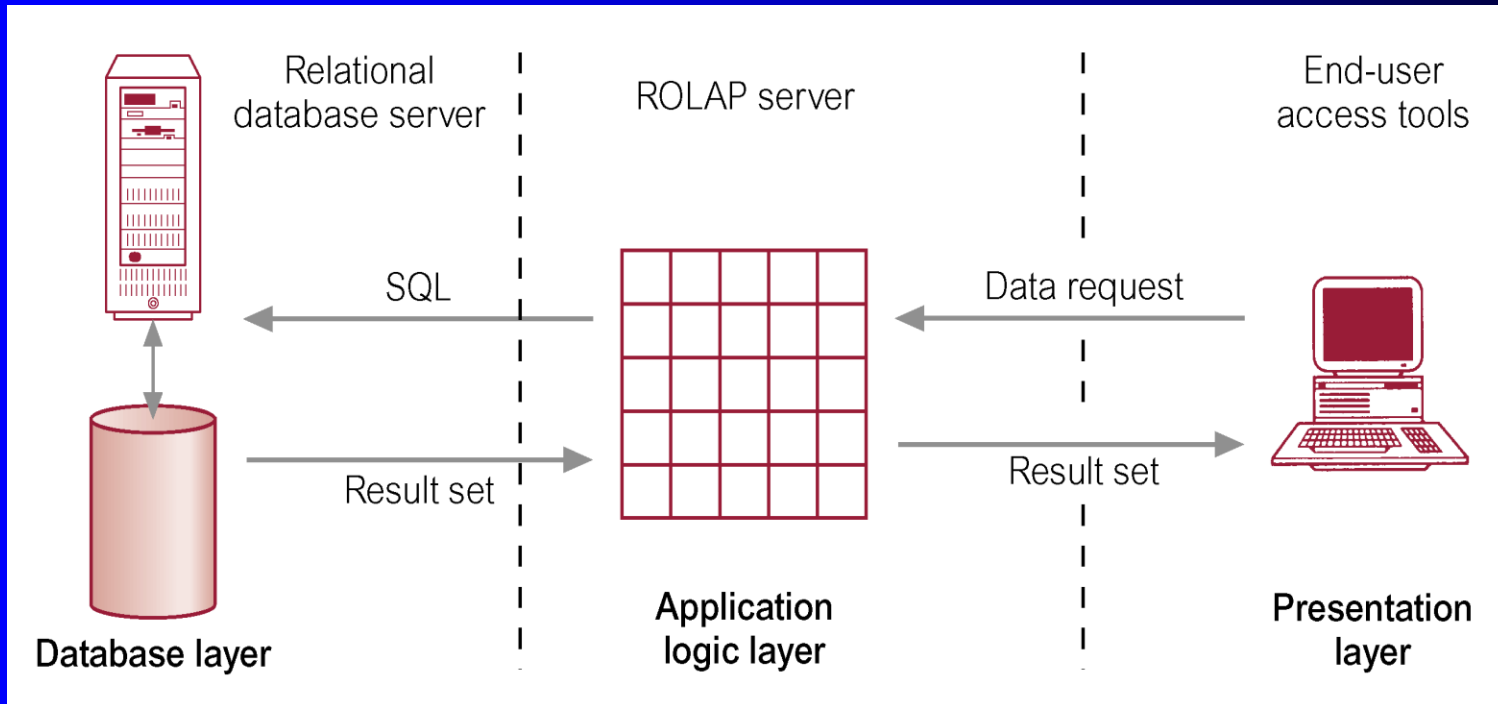
Relational OLAP (ROLAP)

- **Fastest growing style of OLAP technology.**
- **Supports RDBMS products using a metadata layer - avoids need to create a static multi-dimensional data structure - facilitates the creation of multiple multi-dimensional views of the two-dimensional relation.**

Relational OLAP (ROLAP)

- To improve performance, some products use SQL engines to support complexity of multi-dimensional analysis, while others recommend, or require, the use of highly denormalized database designs such as the star schema.

Typical Architecture for ROLAP Tools



ROLAP Tools - Development Issues

- **Middleware to facilitate the development of multi-dimensional applications. (Software that converts the two-dimensional relation into a multi-dimensional structure).**
- **Development of an option to create persistent, multi-dimensional structures with facilities to assist in the administration of these structures.**

MOLAP Server

- Multidimensional OLAP (MOLAP) servers: These servers support multidimensional views of data through *array-based multidimensional storage engines*. They map *multidimensional* views directly to data cube array structures.

Multi-Dimensional OLAP (MOLAP)

- **Uses specialized data structures and multi-dimensional Database Management Systems (MDDDBMSs) to organize, navigate, and analyze data.**
- **Data is typically aggregated and stored according to predicted usage to enhance query performance.**

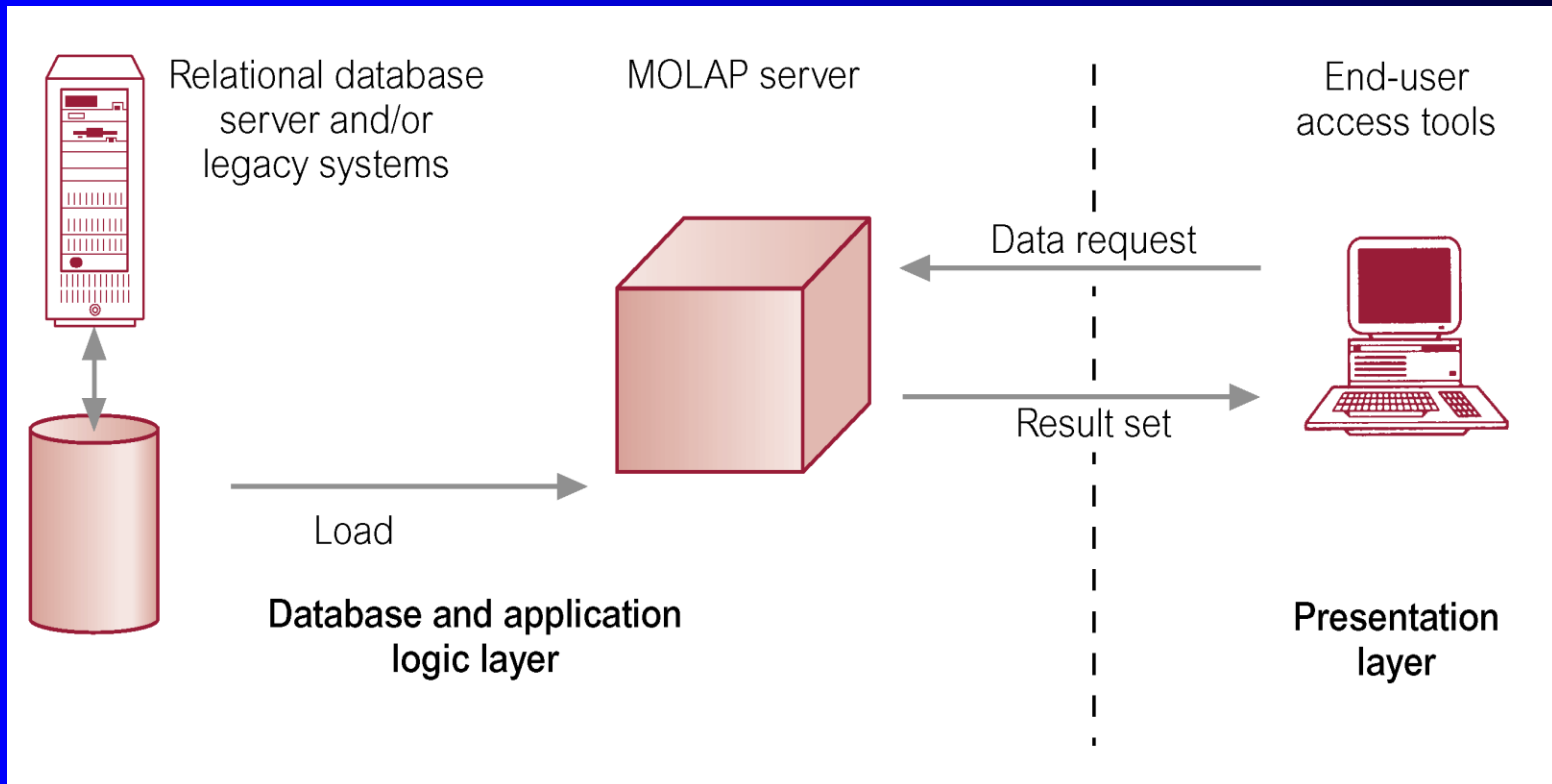
Multi-Dimensional OLAP (MOLAP)

- Use array technology and efficient storage techniques that minimize the disk space requirements through sparse data management.
- Provides excellent performance when data is used as designed, and the focus is on data for a specific decision-support application.

Multi-Dimensional OLAP (MOLAP)

- **Traditionally, require a tight coupling with the application layer and presentation layer.**
- **Recent trends segregate the OLAP from the data structures through the use of published application programming interfaces (APIs).**

Typical Architecture for MOLAP Tools



MOLAP Tools - Development Issues

- **Underlying data structures are limited in their ability to support multiple subject areas and to provide access to detailed data.**
- **Navigation and analysis of data is limited because the data is designed according to previously determined requirements.**

MOLAP Tools - Development Issues

- **MOLAP products require a different set of skills and tools to build and maintain the database, thus increasing the cost and complexity of support.**

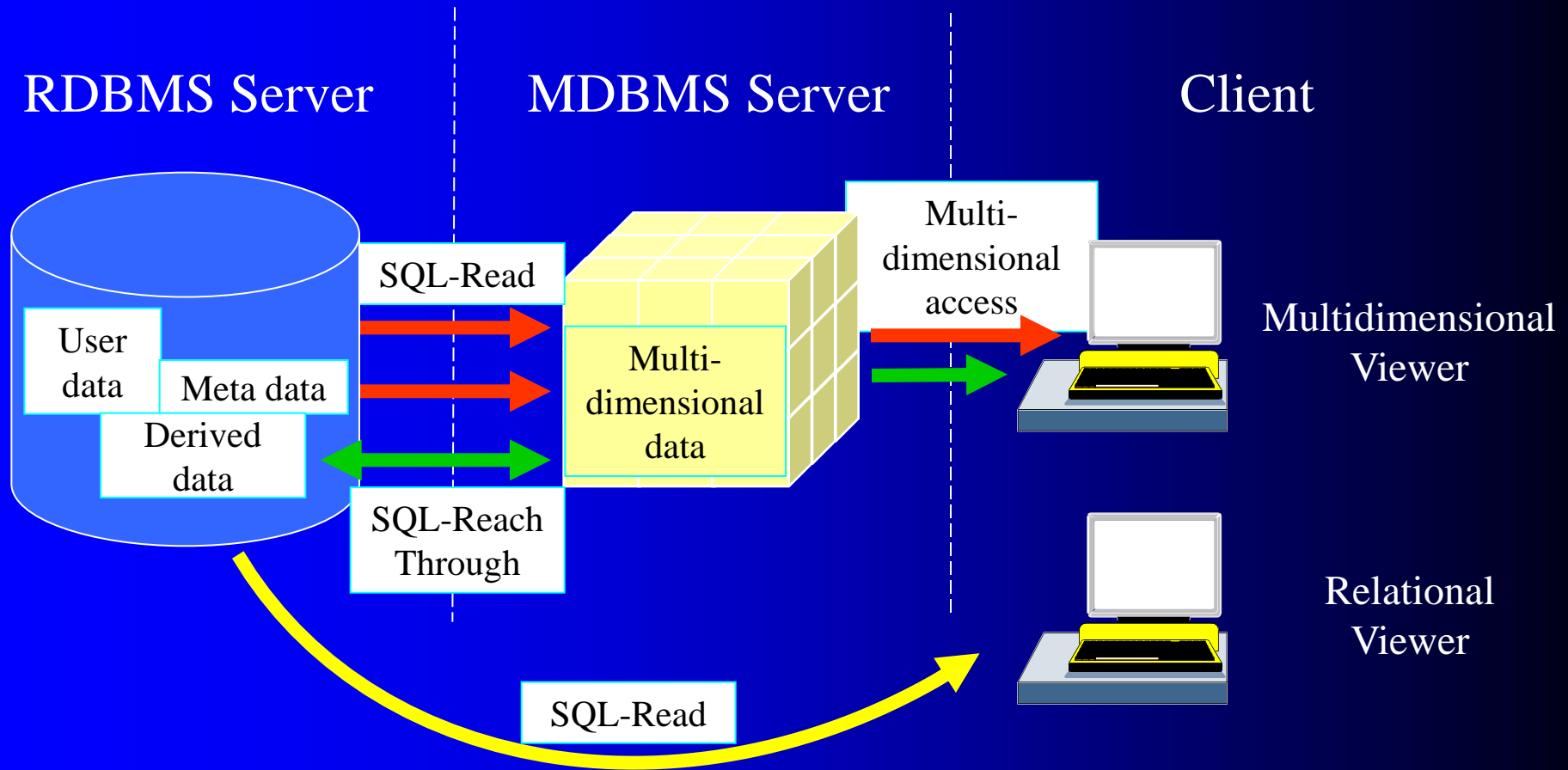
HOLAP Server

- Hybrid OLAP (HOLAP) servers: The hybrid OLAP approach combines ROLAP and MOLAP technology, benefiting from the greater scalability of ROLAP and the faster computation of MOLAP.
- For example, a HOLAP server may allow large volumes of detail data to be stored in a relational database, while aggregations are kept in a separate MOLAP store. The Microsoft SQL Server 2000 supports a hybrid OLAP server.

Hybrid OLAP (HOLAP)

- HOLAP = Hybrid OLAP:
 - Best of both worlds
 - Storing detailed data in RDBMS
 - Storing aggregated data in MDBMS
 - User access via MOLAP tools

Data Flow in HOLAP



Examples

- ROLAP
 - Telecommunication startup: call data records (CDRs)
 - ECommerce Site
 - Credit Card Company
- MOLAP
 - Analysis and budgeting in a financial department
 - Sales analysis
- HOLAP
 - Sales department of a multi-national company
 - Banks and Financial Service Providers

References

- Building Data Warehouse by Inmon
- Data Mining: Concepts and Techniques by Han, Kamber.
- www.datawarehousingonline.com
- www.billinmon.com