



**RAMA**  
**UNIVERSITY**

[www.ramauniversity.ac.in](http://www.ramauniversity.ac.in)

**FACULTY OF ENGINEERING**

**DATA MINING & WAREHOUSEING**  
**LECTURE-12**

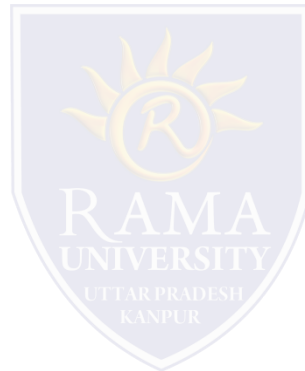
**MR. DHIRENDRA**

ASSISTANT PROFESSOR

RAMA UNIVERSITY

# OUTLINE

- ❖ **OLAP SERVER ARCHITECTURES**
- ❖ **MOLAP**
- ❖ **ROLAP**
- ❖ **HOLAP**
- ❖ **MCQ**
- ❖ **REFERENCES**



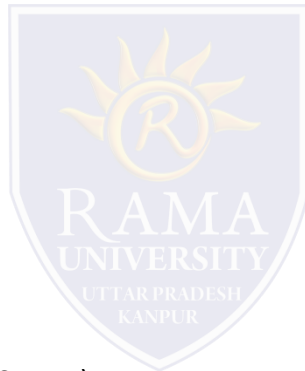
# OLAP Server Architectures

## Relational OLAP (ROLAP)

- ❑ Use relational or extended-relational DBMS to store and manage warehouse data and OLAP middle ware
- ❑ Include optimization of DBMS backend, implementation of aggregation navigation logic, and additional tools and services
- ❑ Greater scalability

## Multidimensional OLAP (MOLAP)

- ❑ Sparse array-based multidimensional storage engine
- ❑ Fast indexing to pre-computed summarized data



## Hybrid OLAP (HOLAP) (e.g., Microsoft SQLServer)

- ❑ Flexibility, e.g., low level: relational, high-level: array

## Specialized SQL servers (e.g., Redbricks)

- ❑ Specialized support for SQL queries over star/snowflake schemas

- In the OLAP world, there are mainly two different types: Multidimensional OLAP (MOLAP) and Relational OLAP (ROLAP).

Hybrid OLAP (HOLAP) refers to technologies that combine MOLAP and ROLAP.

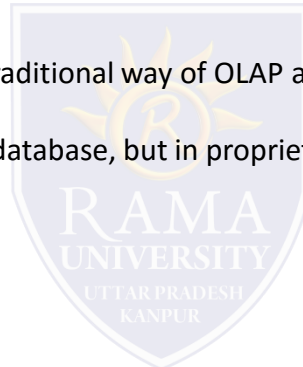
- **MOLAP**

Excellent performance- this is the more traditional way of OLAP analysis. In MOLAP, data is stored in multidimensional cube. The storage is not in the relational database, but in proprietary formats.

- **Products**

Pilot, Arbor Essbase, Gentia

- Pre-calculating or pre-consolidating transactional data improves speed.
- Rolling up and Drilling down through aggregate data.
- With MDDs, application design is essentially the definition of dimensions and calculation rules, while the RDBMS requires that the database schema be a star or snowflake



- **MOLAP Example**

- Analysis and budgeting in a financial department
- Sales analysis

- **MOLAP Tools**

- ORACLE Express Server
- ORACLE Express Clients (C/S and Web)
- MicroStrategy's DSS server
- Platinum Technologies' Plantinum InfoBeacon



- **Advantages:**

- Excellent performance MOLAP cubes are built for fast data retrieval, and are optimal for slicing and dicing operations.
- They can also perform complex calculations. All calculations have been pre-generated when the cube is created. Hence, complex calculations are not only doable, but they return quickly.

- **Disadvantages:**

- It is limited in the amount of data it can handle. Because all calculations are performed when the cube is built, it is not possible to include a large amount of data in the cube itself. This is not to say that the data in the cube cannot be derived from a large amount of data. Indeed, this is possible. But in this case, only summary-level information will be included in the cube itself.
- It requires an additional investment. Cube technology are often proprietary and do not already exist in the organization.

# ROLAP

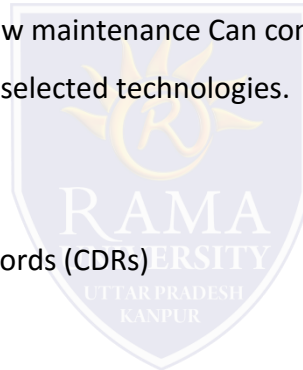
- ❑ This methodology relies on manipulating the data stored in the relational database to give the appearance of traditional OLAP's slicing and dicing functionality. In essence, each action of slicing and dicing is equivalent to adding a "WHERE" clause in the SQL statement.
- ❑ Defines complex, multi-dimensional data with simple model
- ❑ Reduces the number of joins a query has to process
- ❑ Allows the data warehouse to evolve with rel. low maintenance Can contain both detailed and summarized data.
- ❑ ROLAP is based on familiar, proven, and already selected technologies.

## ROLAP Example

- ❑ Telecommunication startup: call data records (CDRs)
- ❑ ECommerce Site
- ❑ Credit Card Company

## ROLAP Tools

- ❑ ORACLE 8i
- ❑ ORACLE Reports; ORACLE Discoverer
- ❑ ORACLE Warehouse Builder
- ❑ Arbors Software's Essbase

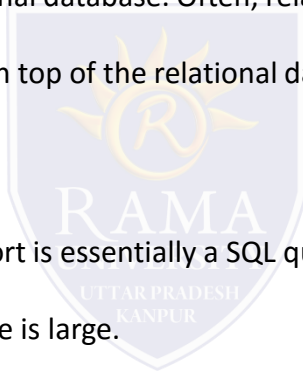


## Advantages:

- ❑ It can handle large amounts of data. The data size limitation of ROLAP technology is the limitation on data size of the underlying relational database. In other words, ROLAP itself places no limitation on data amount.
- ❑ It can leverage functionalities inherent in the relational database. Often, relational database already comes with a host of functionalities. ROLAP technologies, since they sit on top of the relational database, can therefore leverage these functionalities.

## Disadvantages:

- ❑ Performance can be slow. Because each ROLAP report is essentially a SQL query (or multiple SQL queries) in the relational database, the query time can be long if the underlying data size is large.
- ❑ It has limited by SQL functionalities. Because ROLAP technology mainly relies on generating SQL statements to query the relational database, and SQL statements do not fit all needs (for example, it is difficult to perform complex calculations using SQL), ROLAP technologies are therefore traditionally limited by what SQL can do. ROLAP vendors have mitigated this risk by building into the tool out-of-the-box complex functions as well as the ability to allow users to define their own functions.



# HOLAP

- ❑ HOLAP technologies attempt to combine the advantages of MOLAP and ROLAP. For summary-type information, HOLAP leverages cube technology for faster performance. When detail information is needed, HOLAP can "drill through" from the cube into the underlying relational data.

## Examples

- ❑ Sales department of a multi-national company
- ❑ Banks and Financial Service Providers



## ❑HOLAP: Tools

- ❑ ORACLE 8i
- ❑ ORACLE Express Serve
- ❑ ORACLE Relational Access Manager
- ❑ ORACLE Express Clients (C/S and Web)

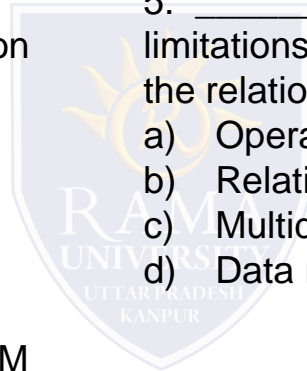
## ❑HOLAP = Hybrid OLAP:

- ❑ Best of both worlds
- ❑ Storing detailed data in RDBMS
- ❑ Storing aggregated data in MDBMS
- ❑ User access via MOLAP tools



# Multiple Choice Question

1. Source data from the warehouse comes from \_\_\_\_\_.
  - a) ODS
  - b) TDS
  - c) MDDDB
  - d) ORDBMS
  
2. \_\_\_\_\_ is a data transformation process.
  - a) Comparison
  - b) Projection
  - c) Selection
  - d) Filtering
  
3. The technology area associated with CRM is \_\_\_\_\_.
  - a) specialization
  - b) generalization
  - c) personalization
  - d) summarization
  
4. SMP stands for \_\_\_\_\_.
  - a) Symmetric Multiprocessor.
  - b) Symmetric Multiprogramming.
  - c) Symmetric Metaprogramming.
  - d) Symmetric Microprogramming.
  
5. \_\_\_\_\_ are designed to overcome any limitations placed on the warehouse by the nature of the relational data model.
  - a) Operational database.
  - b) Relational database.
  - c) Multidimensional database.
  - d) Data repository.



# REFERENCES

- [https://www.tutorialspoint.com/dwh/dwh\\_overview.htm](https://www.tutorialspoint.com/dwh/dwh_overview.htm)
- <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf> DATA MINING BOOK WRITTEN BY Micheline Kamber
- <https://www.javatpoint.com/three-tier-data-warehouse-architecture>
- M.H. Dunham, “ Data Mining: Introductory & Advanced Topics” Pearson Education
- Jiawei Han, Micheline Kamber, “ Data Mining Concepts & Techniques” Elsevier
- Sam Anahory, Denniss Murray,” data warehousing in the Real World: A Practical Guide for Building Decision Support Systems, “ Pearson Education
- Mallach,” Data Warehousing System”, TMH

