© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

Improving Resource Management And Solving Scheduling Problem In Data Ware House Using OLAP And OLTP

[1]Asha Rani, Department of CSE, IITB College, Sonipat, contactme.ashu2@gmail.com
[2]Ritika Saroha, Assistant Prof., Department of CSE, IITB College, Sonipat, saroharitika@gmail.com

ABSTRACT: Data in the warehouse and data marts is stored and managed by one or more warehouse servers, which present multidimensional views of data to a variety of front end tools: query tools, report writers, analysis tools, and data mining tools. Finally, there is a repository for storing and managing metadata, and tools for monitoring and administering the warehousing system. Data warehouses, in contrast, had been targeted for decision support. Historical, summarized & consolidated data would be more important than detailed, individual records. Work load had been query intensive within mostly ad hoc, complex queries that could access millions of records & perform lot of scans, joins, & aggregates. Query throughput & response times had been more important than transaction throughput.

OLAP performs multidimensional analysis of business data & provides capability for complex calculations, trend analysis, & sophisticated data modelling.

[1] Introduction

Meaning of Data Warehouse was firstly coined by Bill Inm on in 1990. This data helps analysts to take informed very important decisions in group. An operational database undergoes frequent changes on a daily basis on account of transactions that taken area. Think a business management wants to analyze previous feedback on any data such as a product, a supplier, or any consumer data, then executive will had no data available to analyze because previous data has been updated due to transactions.

Using Data Warehouse Information

There are decision support technologies that help utilize data on hand in a data warehouse. These technologies help to use warehouse quickly & effectively. They can gather data, analyze it, & take decisions based on information present in warehouse. Information gathered in a warehouse can be used in any of following domains:

1. Tuning Production Strategies or product strategies can be well tuned by repositioning products & managing product portfolios by comparing sales quarterly or yearly.

2. consumer Analysis or consumer analysis are done by analyzing customer's buying preferences, buying time, budget cycles, etc.

3. Operations Analysis or Data warehousing also helps in custumer relation management, & making environmental corrections. information also allows us to analyze business operations.

Architecture of Data Ware House

It includes tools for extracting data from multiple operational databases & external sources; for cleaning, transforming & integrating this data; for loading data into data warehouse; & for periodically refreshing warehouse to reflect updates at sources & to purge data from warehouse, perhaps onto slower archival storage. In addition to main warehouse, there might be several departmental data marts

.

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

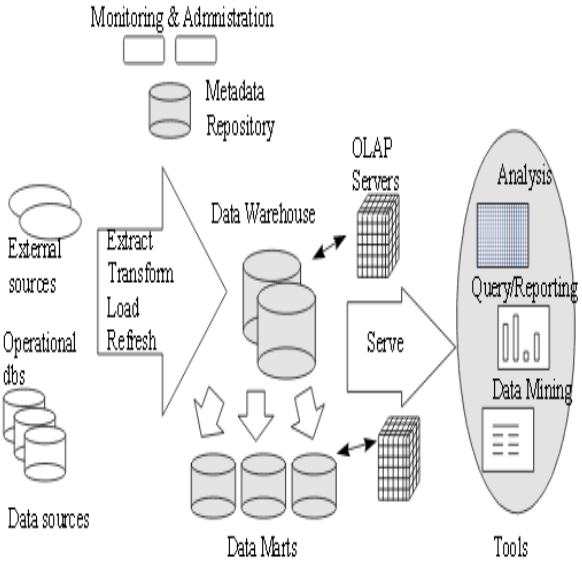


Fig 1. Data Ware house Architecture

[2]Literature Review

A research published on title "An Overview of Data Warehousing and OLAP Technology" by Surajit Chaudhuri in 1997 focused on Data warehousing & on-line analytical processing (OLAP).

According to them these are essential elements of decision support, which has been increasingly become a focus of database industry. Many commercial products & services are now available, & all of principal database management system vendors now had offerings in these areas.

Surajit Chaudhuri wrote on An Overview of Data Warehousing & OLAP Technology (Appears in ACM Sigmod Record, March 1997).

Data warehousing & on-line analytical processing (OLAP) are essential elements of decision support, which has increasingly become a focus of database industry. Many commercial products & services are now available, & all of principal database management system vendors now had offerings in these areas. Decision support places some rather different requirements on database technology compared to traditional on-line transaction processing applications.

Manjunath T. N. wrote on Realistic Analysis of Data Warehousing & Data Mining Application in Education Domain

Data-driven decision support systems, such as data warehouses can serve requirement of extraction of information from more than one subject area. Data warehouses standardize data across organization so as to had a single view of information. Data warehouses can provide information required by decision makers. Developing a data warehouse for educational institute are less focused area since educational institutes are non-profit & service oriented organizations. In present day scenario where education has been privatized & cut throat competition are prevailing, institutes needs to be more organized & need to take better decisions. **Mr. Dishek Mankad wrote on "The Study on Data Warehouse Design & Usage"**

Data ware housing are a booming industry within many interesting research problem. data warehouse are concentrated on only few aspects. Here we are discussing about data warehouse design & usage. Let's look at various approaches to data ware house design & usage process & steps involved. Data warehouse can be built using a top-down approach, bottom –down approach or a combination of both. In this research paper we are discussing about data warehouse design process.

"**Comparative Study of Various Bitmap Indexing Techniques Used in Data Warehouse**" was published in 2012 by Bikramjit Pal1, Anirban Bhattacharjee, Satyajit Ghosh, Rajdeep Chowdhury & Dr. Mallika De. In their paper they explained that for running complex query, performing aggregated function & handling huge no of data in data warehouse bitmap indexing has been become most popular indexing technique recently. they study varies type of bitmap indexing techniques (simple bitmap & encoded bitmap) & perform aggregated operation on query within help of both simple & encoded bitmap indexing & analyses result, which was really interesting.

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

[3] TOOLS & TECHNOLOGY

Online Analytical Processing Server (OLAP) are based on multidimensional data model. It allows managers, & analysts to get an insight of information through fast, consistent, & interactive access to information. This chapter cover types of OLAP, operations on OLAP, difference between OLAP, & statistical databases & OLTP.

Types of OLAP Servers

We had four types of OLAP servers:

1. Relational OLAP (ROLAP)
2. Multidimensional OLAP (MOLAP)
3. Hybrid OLAP (HOLAP)
4. Specialized SQL Servers

Relational OLAP

ROLA P servers are placed between relational back-end server & client front-end tools. To store & manage warehouse data, ROLAP uses relational or extended-relational DBMS.

ROLAP includes following:

1. Implementation of aggregation navigation logic.
2. Optimization for each DBMS back end.
3. Additional tools & services.

Roll-up

Roll-up performs aggregation on a data cube in any of following ways:

1. By climbing up a concept hierarchy for a dimension
2. By dimension reduction

The following diagram illustrates how roll-up works.

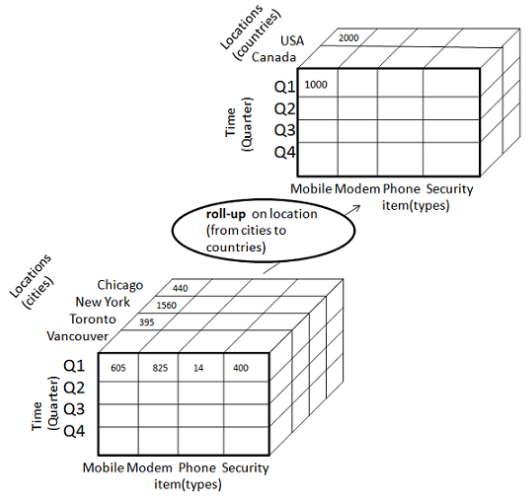


Fig.2 Roll up operations

Drill-down

Drill-down are reverse operation of roll-up. It are performed by either of following ways:

1. By stepping down a concept ladder for a dimension
2. By introducing a new dimension.

The following diagram illustrates how drill-down works:

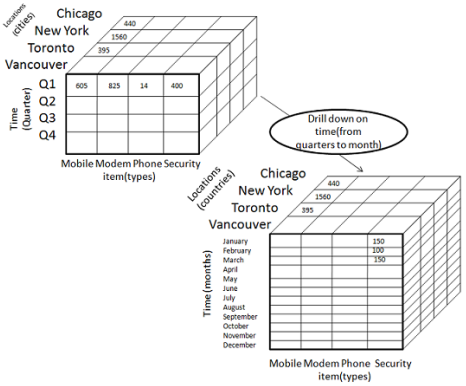


Fig. 3. Drill Down Operation

[4] PROPOSED WORK

1. Investigation of new face in data ware house management.
2. Study of runaway queries problems.

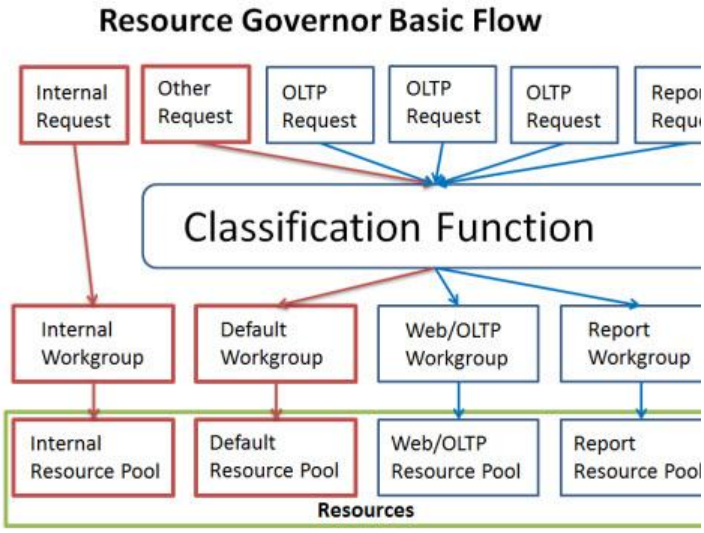
© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

3. Taking corrective measurement to manage Resource
4. Taking corrective steps in order to solve scheduling problem.

Investigation of new Challenges in data ware house management

Data warehousing projects are one of its kinds. All data warehousing projects do not pose same challenges & not all of them are complex but they are always different. Knowing these challenges upfront are your best bet to avoid them.

Data warehousing are different.. For most part of it, these projects are heavily dependent on backend infrastructure in order to support front-end user reporting. But these are not only reasons why doing data warehousing are difficult. In below list we show top 5 reasons which actually make things complex on practical ground.



. **Fig:4 Resource Governor Basic Flow**

Taking corrective step in order to solve scheduling problem

Trigger would be used to scheduling. We will use trigger to schedule a specific tast on specific event. A trigger are a database object that are *attached* to a table. In many aspects it are similar to a stored procedure. As a matter of fact, triggers are often referred to as a "special kind of stored procedure." main difference between a trigger & a stored procedure are that former are attached to a table & are only *fired* when an INSERT, UPDATE or DELETE occurs. We specify modification action(s) that fire trigger when it are created.

[5] Result & discussion

Here we had chosen a huge database of MLM Company. Records of Approximate 8000 people had been maintained here along within their daily payout & regular buying.

Handling Challenges in data ware house management using query optimization

There are several Factors that would effect query processing

1. Data has been extracted from local or Remote server

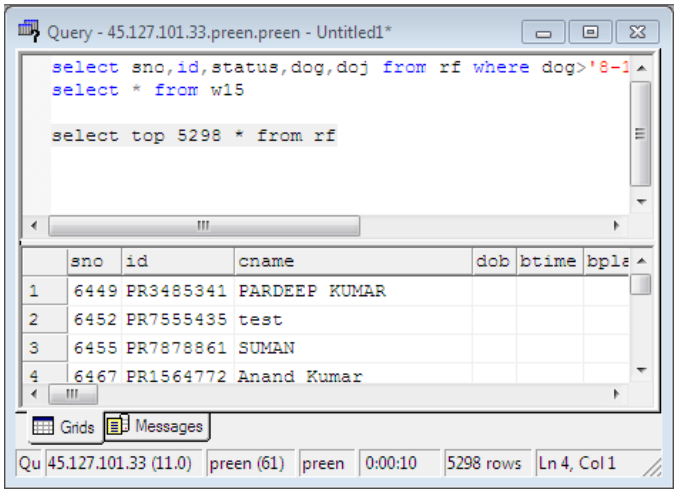


Fig 5 Here in above diagram we had extracted records from remoter server & it took 10 seconds in case of 5298 records.

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

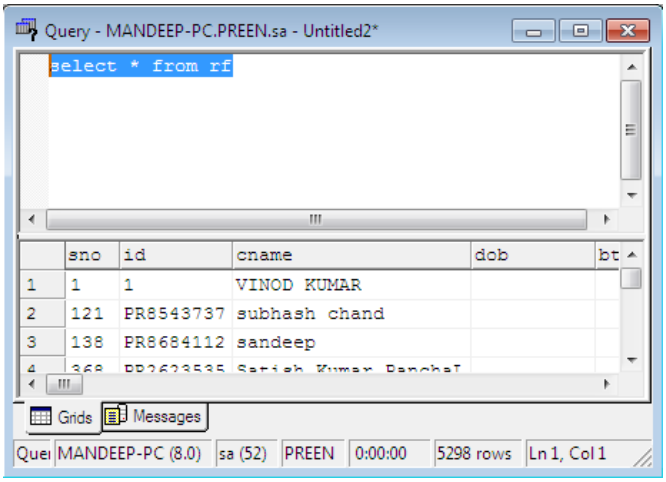


Fig 6 Here in above diagram we had extracted records from local server & it took 0 seconds in case of 5298 records.

2. Number of columns extracted

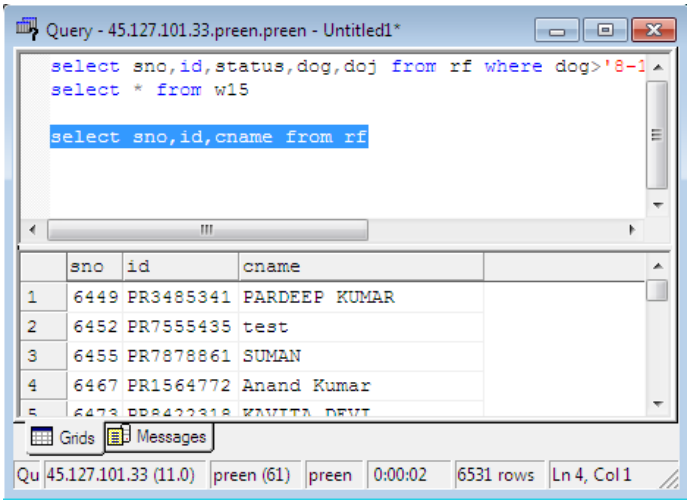


Fig 7 If 3 columns had been retrieved from remote server than it will take about 2 seconds.

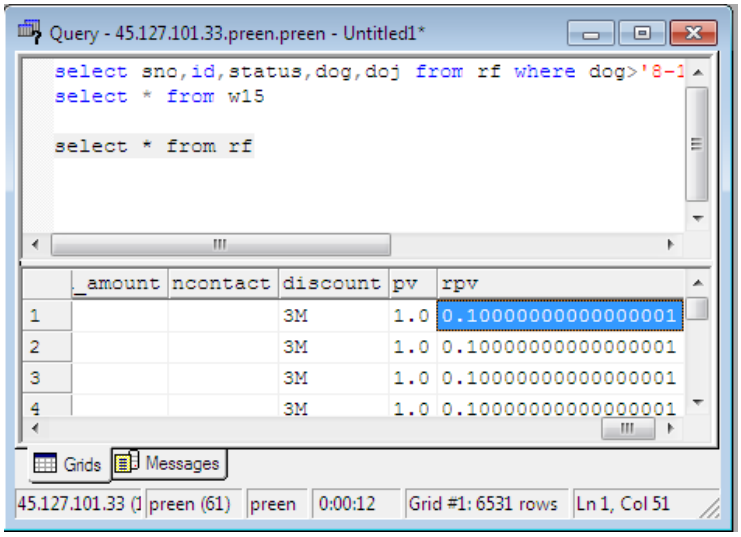


Fig 8 If 51 columns had been retrieved from remote server than it will take about 12 seconds.

[6] CONCLUSION

Relational OLAP servers is placed to relational back-end server & client front-end tools. To store & manage warehouse data, relational OLAP uses relational or extended-relational DBMS. ROLAP servers are highly scalable. ROLAP tools analyze big range of data across multiple dimensions. ROLAP tools store & analyze highly volatile & changeable data.

As per our simulation time taken to read records got reduced if speed of network connection are fast & Query are Optimized. Here we had made comparative analysis of time taken by optimized & normal query. Various factors effect time taken by query. These factor are number of area in table, Complexity of Sql Query, Connection speed.

Achieving performance objectives are not easy. In first place, setting up performance objectives itself are a challenging task. An untrained user can easily drift before setting up some performance goals that are unrealistic for a given data warehousing scenario. Hence for users of data warehouse, it are generally considered safe to set up performance goals in terms of practical usability requirements.

REFERENCES

© INTERNATIONAL JOURNAL FOR RESEARCH PUBLICATION & SEMINAR
ISSN: 2278-6848 | Volume: 07 Issue: 07 | July – September 2016
Paper is available at www.jrps.in | Email : info@jrps.in

[1] Mr. Dishek Mankad "The Study on Data Warehouse Design and Usage" International Journal of Scientific and Research Publications , Volume 3, Issue 3, March 2013 ISSN 2250- 3153

[2] Surajit Chaudhuri wrote on An Overview of Data Warehousing and OLAP Technology (Appears in ACM Sigmod Record, March 1997).

[3] Manjunath T. N. wrote on Realistic Analysis of Data Warehousing and Data Mining Application in Education Domain

[4] Weiss, Sholom M.; and Indurkhya, Nitin (1998); Predictive Data Mining, Morgan Kaufmann

[5] Kimball, R.The Data Warehouse Toolkit. John Wiley, 1996.

[6] Barclay, T., R. Barnes, J. Gray, P. Sundaresan, "Loading Databases using Dataflow Parallelism." SIGMOD Record, Vol.23, No. 4, Dec.1994.

[7] Blakeley, J.A., N. Coburn, P. Larson. "Updating Derived Relations: Detecting Irrelevant and Autonomously ComputableUpdates." ACM TODS, Vol.4, No. 3, 1989.

[8]Gupta, A., I.S. Mumick, "Maintenance of Materialized Views: Problems, Techniques, and Applications." Data Eng. Bulletin, Vol. 18, No. 2, June 1995. 9 Zhuge, Y., H. Garcia-Molina, J. Hammer, J. Widom, "View Maintenance in a Warehousing Environment, Proc. Of SIGMOD Conf., 1995.

[9] Roussopoulos, N., et al., "The Maryland ADMS Project: Views R Us." Data Eng. Bulletin, Vol. 18, No.2, June 1995.[11] O'Neil P., Quass D. "Improved Query Performance withVariant Indices", To appear in Proc. of SIGMOD Conf., 1997.

[10] O'Neil P., Graefe G. "Multi-Table Joins through BitmappedJoin Indices" SIGMOD Record, Sep 1995.

[11] Harinarayan V., Rajaraman A., Ullman J.D. " Implementing Data Cubes Efficiently" Proc. of SIGMOD Conf., 1996.

[12] Chaudhuri S., Krishnamurthy R., Potamianos S., Shim K."Optimizing Queries with Materialized Views" Intl.Conference on Data Engineering, 1995.

[13] Levy A., Mendelzon A., Sagiv Y. "Answering Queries Using Views" Proc. of PODS, 1995. 16 Yang H.Z., Larson P.A. "Query Transformations for PSJ Queries", Proc. of VLDB, 1987

[14] *Witten, Ian H.; Frank, Eibe; Hall, Mark A. (30 January 2011). Data Mining: Practical Machine Learning Tools and Techniques (3 ed.). Elsevier. ISBN 978-0-12-374856-0.*

[15] Ye, Nong (2003); The Handbook of Data Mining, Mahwah, NJ: Lawrence Erlbaum

[16] Cabena, Peter; Hadjnian, Pablo; Stadler, Rolf; Verhees, Jaap; Zanasi, Alessandro (1997); Discovering Data Mining: From Concept to Implementation, Prentice Hall, ISBN 0-13-743980-6

[17] M.S. Chen, J. Han, P.S. Yu (1996) "Data mining: an overview from a database perspective". Knowledge and data Engineering, IEEE Transactions on 8 (6), 866–883

[18] Feldman, Ronen; Sanger, James (2007); The Text Mining Handbook, Cambridge University Press, ISBN 978-0-521-83657-9

[19] Guo, Yike; and Grossman, Robert (editors) (1999); High Performance Data Mining: Scaling Algorithms, Applications and Systems, Kluwer Academic Publishers

[20] Han, Jiawei, Micheline Kamber, and Jian Pei. Data mining: concepts and techniques. Morgan kaufmann, 2006.