Database Query and Its Optimization: A Conclusive Report

Numptik Voratas
Asian College of Engineering, Thailand.

Query is a statement or group of statement that adequately execute some basic database operations viz. 'Read', 'Write','Delete', and 'Update'. It plays a consequential role in managing and retrieving data. In general, distributed queries are more complex and complicated as compared to centralized queries. Queries can be categorized as data creation and data destruction, Data management queries, Data control quarry, OLTP and DSS quarries. In data creation and data destruction quarries create, insert and drop quarries are used. In data management quarry data is managed and manipulate, data can be insert, delete and update. In data control query, one can save data using commit command; permission can be granted using grant command [1][2][3].

In online transaction processing (OLTP) the work analysis and query optimization is done. In decision support system (DSS) queries used to retrieve data from large database. The execution time is not predictable in DSS query. Decision support system (DSS) queries are more complex as compare to online transaction processing queries (OLTP). The running time of DSS queries are unpredictable as compare to OLTP. The process of optimization in Decision support system (DSS) queries is complex as compare to OLTP queries. A distributed DSS query is used to retrieve data from multiple sites. In online transaction processing system (OLTP); real updates are performed. However, DSS queries execute batches as compared to real time updates. Online transaction processing (OLTP) database applications are optimal for managing changing data; these applications typically have many users who are performing transaction at the same time that change real time data, in other words OLTP is a live database. On other side the tables in a decision support database are heavily indexed and the raw data is frequently preprocessed and organized to support various types of queries to be used. The OLTP and DSS queries can be differentiated on the basis of different parameters as mentioned below [1][4][5][6]:

S.No.	Parameter	OLTP	DSS
1	Abbreviation	Online transaction processing.	Decision support system.
2	Nature of Data	Common routine work.	Aggregate or subroutine information.
3	Execution Time	Less	More
4	Amount of Time	Predictable	Unpredictable
5	Resource consumption	Less	More

A number of heuristics have been applied in recent times, which proposed new algorithms for substantially improving the performance of a query[1][2][3]. As stated by Manik Sharma et al. (2015) there are two major types of database queries called DSS and OLTP queries. To optimize a *DSS* query on the basis of usage of system resources, one has to find an optimal query execution plan which minimizes the *Total Costs* of a query. For finding the optimal query execution plan, the costs of

different performance metrics *viz. Input-Output Costs, Processing Costs and Communication Costs, Total Costs* and *Runtime* of *DSS* query should be computed and analyzed. Here, *Total Costs* represents the usage of the system resources required to execute a query. It is also defined as the sum of *Local Processing Costs (LPC)* and *Communication Costs (CMCT).* Several authors have used different techniques like exhaustive enumeration, dynamic programming, genetic algorithm, restricted genetic algorithm; entropy based restricted genetic algorithm and other hybrid evolutionary algorithms to optimize the distributed queries. Authors designed DSS queries using one of the DSS benchmark. The designed queries were then examined and analyzed on the basis of execution time, total costs and the quality of the solution[6][7][8][9][10].

Authors have analyzed the effect of varying Input output to communication costs ratio over the total costs of the DSS query. Authors found that that by reducing the I/O to communication costs from 1:1.6 to 1:1, one is able to drastically reduce the communication costs[11]. Some of the authors have also statistically analyzed the performance of the query optimizer [12].

It is found that both OLTP and DSS need different types of query optimizer. Moreover, several techniques have been used to optimize the query. However, little work has been done using emerging evolutionary techniques. One should employ latest evolutionary techniques to optimize the OLTP and DSS queries.

References

1. Komal Sharma, Ankur Bhardwaj. “Types of Queries in Database System”. International Journal of Computer Applications and Information Technology. 7 (2) : 149-152.
2. Sharma M, Singh G, Singh R. Design and analysis of stochastic DSS query optimizers in a distributed database system. Egyptian informatics journal. 2016 Jul 1;17(2):161-73. doi:10.1016/j.eij.2015.10.003
3. Hevener AR, Yao SB. Query processing in distributed database systems. IEEE Transaction Software Engineering. 1979;5(3):177–87.
4. Manik Sharma, Gurvinder Singh, Gurdev Singh, Rajinder Singh. “Analysis of DSS Queries in Distributed Database System Using Exhaustive and Genetic Approach”. International Journal of Advanced Computing, Vol.36, Issue.2, April 2013
5. Mor Jyoti, Kashyap Indu, Rathy RK. Analysis of query optimization techniques in databases. Int. J. Comp. Appl. 2012;47(15):5–9.
6. Sharma M, Singh G, Singh R, Singh J. Design and Analysis of Stochastic Query Optimizer for Biobank Databases. In Computational Science and Its Applications (ICCSA), 2015 15th International Conference on 2015 Jun 22 (pp. 47-51). IEEE.
7. Sevinc Ender, Cosar Ahmat. An evolutionary genetic algorithm for optimization of distributed database queries. Comp. J. 2011;54:717–25.
8. Manik Sharma, Gurvinder Singh, Rajinder Singh. 2013. “Design and comparative analysis of DSS queries in distributed environment”. IEEE International Conference on Computer Science and Engineering (ICSEC 2013). Doi: 10.1109/ICSEC.2013.6694756
9. Manik Sharma, Gurdev Singh, Rajinder Virk,” Analysis of DSS queries in distributed database system “, International Journal of Networking and Parallel Computing. Vol.1, Issue 3, Dec 2012-Jan2013.
10. T.V.Vijay Kumar, Vikram Singh, “Distributed query processing plans generations using GA”, IJCTE, Vol 3.No.1, 2011.

11. M. Sharma, G. Singh, R. Singh, "*DSS Query Optimization and Effect of Input Output / Communication Cost Metrics*", International Journal of Computer Sciences and Engineering, Vol.5, Issue.11, pp.198-203, 2017.
12. Sharma M, Singh G, Singh R, Singh S. Statistical Analysis of DSS Query Optimizer for a Five Join DSS Query. International Journal of Computer Applications. 2016;141(6):1-4.