Rhodes University

Department of Computer Science

Computer Science Honours Project Proposal

Benchmarking Databases

By Samy Kabangu

Supervisor: Mr. John Ebden

1. Background

Many claims are being made by different vendors in favour of their products in terms of performance as competition increases in the database computing industry. Benchmarking a given database is a process of performing well defined tests on that particular database management system for the purpose of evaluating its performance. The Response time and the throughput are the two main criteria on which the performance of a database can be measured. To achieve an optimum performance, the specified response time should not be exceeded. The throughput refers to the number of transactions a database server can handle during a given period of time. The factors that affect the performance of a database are classified into three general categories which are: The database applications (Design considerations), the database system (engine) and the System resources (Hardware) [4].

Industry Standard Database Benchmarks such as the Transaction Processing Performance Council (TPC) were developed to provide a platform for benchmarking various database products under uniform conditions in order to compare them with respect to their performance. Hence database vendors use the benchmark test results that most reveal the performance of their product. Some of the issues around the industry standard database benchmarks reside on the facts that they simulate real world workload which might not reflect the actual workload of a particular company or organization, they are performed on specific hardware and operating system which makes their adoption a little difficult to individual organization [3]. Some vendors use techniques such as caching the data and the SQL statement into memory (RAM) in order to avoid disk access so as to improve the benchmark performance [1]. So, the design as well as the

implementation of sound custom database benchmarks and tuning techniques appears to be of great importance especially within a particular organization (company).

2. Statement of the problem

Our objective is to come up with a best way of testing a database management system. An investigation into the most critical benchmarking techniques and factors in tuning database management systems will be the main focus of this project.

3. Intended Approach

In order to achieve what we intend to do, we will firstly provide ourselves with a sound understanding of the factors that influence the performance of a database system which will enable us in the investigation of some of the well known industry standard database benchmarks such as the TPC.

Secondly, we will run some benchmark tests on Microsoft SQL Server 2008 (Enterprise edition) using the TPC-H benchmark which is one of the TPC Benchmark suites designed for ad hoc queries and decision support. The reason being, SQL Server 2008 is the latest database management system product being released by Microsoft and spectacular claims are made concerning its performance compared to Oracle 11g [2].

4. Significance of the project

Evaluating the performance of a database (Benchmarking a database) is an operation that needs to be undertaken in order to approximate the capabilities of a database system for current operations or transactions as well as to model the trend of the response time as the workload increases over time or in case of many users accessing the system concurrently. Hence, it appears to be essential for an organization (company) to know the limit of its system in order to

implement adequate techniques about how to benefit from it to the fullness of its performance and predict bottlenecks as the scale of the database increases.

5. Plan of Action

First Semester	
Action	**Proposed Dates**
Study Recommended papers and tutorials related to research project	n/a
Write Research Proposal	2nd March
Oral presentation	25th March
Install and become familiar with Microsoft SQL server	06th April – 06th May
Study of relevant TPC suites and some other Industry Standard benchmarks.	6th April – 6th May
Hand in Literature Review	26th May
Run test Practice	30th April – 29th May

6. Expected results

The results that we expect to obtain are given as follows:

6.1 The first result that we expect to obtain is the implementation of sound practice techniques of tuning a database in order to obtain an optimum performance from it.

6.2 Secondly, we would like to be able at the end of our project to answer the question: "Why the TPC benchmarks are better than other industry standard benchmarks?"

6.3 Knowing the primary rigging database methods used by vendors to increase the performance of their products.

6.4 We will also observe if the performance of new software (Microsoft SQL Server 2008 - 32 and 64 bit versions) running on slow hardware would match the performance of old software (Microsoft SQL Server 2005) running on fast hardware.

6.5 Knowing if Microsoft SQL Server 2008 running on a core 2 Quad processor would match the performance published by the TPC-H benchmark results run on mainframe hardware will also be one of our expected results.

6.6 Finally, we will compare the performance results test of SQL Server 2008 against the one of Oracle 11g in order to know which one of the two database management systems performs better.

7. Future Extension

This project can be extended to the design of database benchmarks.

8. Reference

[1] Burleson D, Burleson Consulting, Database benchmark: What you need to know, http://www.dba-oracle.com/art_db_benchmark.htm [Accessed 08-03-2009]

[2] Microsoft SQL Server 2008, White Paper: SQL Server 2008 Compared to Oracle Database 11g, http://www.microsoft.com/sqlserver/2008/en/us/sql2008-oracle11g.aspx [Accessed 08-03-2009

[3] Oracle, Database Benchmarking, http://wiki.oracle.com/page/Database+Benchmarking?t=anon, [Access 08-09-2009]

[4] Petkovic D., *Microsoft SQL Server 2005, A Beginner's Guide*, The McGraw-Hill Companies, 2006

9. Bibliography

1. Ault M., Burleson D., Fernandaez C., Klein K., Scalzo B., *Database Benchmarking*, Practical Methods for Oracle & SQL Server, Rampant TechPress, April 2007

2. Boral H., DeWitt D.J., A methodology for Database System Performance Evaluation, http://pages.cs.wisc.edu/~dewitt/includes/benchmarking/sigmod84.pdf ,[ACCESS 28-02-2009]

3. Gray J., Dujmovic J., The Benchmarking Handbook for Database and Transaction Processing Systems, ACM Press,

4. Microsoft, SQL Server 2008: Benchmarks, http://www.microsoft.com/sqlserver/2008/en/us/benchmarks.aspx, [ACCESS 01-03-2009]

5. Oracle, Oracle Benchmark Results, http://www.oracle.com/solutions/performance_scalability/benchmark_results.html, [ACCESS 01-03-09]

6. Petkovic D., Microsoft SQL Server 2005, *A Beginner's Guide*, The McGraw-Hill Companies, 2006

7. Rubenstein W., Kubicar M., Cattell R., Benchmarking simple database operations, ACM, New York, USA,1987

8. Shanley K, History and overview of the TPC, Http://www.tpc.org/information/about/history.asp, [ACCESS 24-02-2009]

9. Traeger A., Zadok E., Joukov N., Wright C., A nine year study of file system and storage benchmarking, ACM, New York, 2008