International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.2, Issue.2, Mar-Apr 2012 pp-067-072 ISSN: 2249-6645

# PERFORMANCE EVALUATION OF DATABASE MANAGEMENT SYSTEMS BY THE ANALYSIS OF DBMS TIME AND CAPACITY

**Aparna Kaladi[1] and Priya Ponnusamy[2]**
[1]M.E Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India
[2]Asst.Professor, Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India

***ABSTRACT***
*Performance evaluation of database systems are made difficult by the existence of different database management systems, which serves different requirements. As DBMSes are limited in speed, they need to be highly efficient through its performance. A DBMS provides facilities for controlling data access, enforcing data integrity, managing concurrency control, recovering the database after failures and restoring it from backup files, as well as maintaining database security. It is a set of software programs that controls the system organization, storage, management, and retrieval of data in a database. That is why they need to be efficient in storage and speed. This is the same reason why DBMSes need to be evaluated based on their performances. Different DBMSs performs differently in different environment. This paper, evaluate the performance of Three different database management systems(MySql, Sqlite, PostgreSQL). Based on that, comparison of different DBMSs in the same environment is possible. This performance evaluation is done based on responce time and memory capacity.*

***KEYWORDS***
*PostgreSQL ,SQLite, MySQL, Response time, Memory*

## 1. INTRODUCTION

A database management system (dbms) is a complicated set of software programs which controls the organization, storage, management, and retrieval of data in a database. Database management is no longer limited to "monolithic entities". Many solutions have been developed to satisfy the individual needs of users. The development of numerous database options has created flexibility in database management. There are several ways database management has affected the field of technology. Because organizations' demand for directory services has grown as they expand in size, businesses use directory services that provide prompted searches for company information. Mobile devices are able to store more than just the contact information of users, and can cache and display a large amount of information on smaller displays. Search engine queries are able to locate data within the World Wide Web. Online transactions have become tremendously popular for e-business
The last few decades has seen huge transformations in the way businesses are conducted. The growth and diversity of the market has greatly profited consumers through higher availability, better quality and lower prices. The same factors however has made it more difficult for businesses to maintain their competitive edge over one another and hence has forced them to think beyond their product portfolio and look at other means to gain higher visibility and customer satisfaction, maintaining all the while their core advantages on pricing and product through improved and more efficient methods of manufacturing and distribution. Database management systems now form the core of almost all enterprise logic and business intelligence solutions.

Database Systems are one of the key enabling forces behind every business transformations. Apart from supporting enterprise logic they also enable business intelligence. Information is the key to success in today's businesses. However, maintaining information in logically consistent and feasibly retrievable format is a daunting task. More so with the added complications of transaction consistency management, synchronization across multiple repositories spread geographically across the globe, failover management and redundancy management, today's database systems are truly state-of-the-art high performance software systems.

Database management systems need to be efficient in terms of storage and speed. Dynamic addition and deletion of data from the database poses a challenge to maintaining an efficient data retrieval mechanism. Though, limited in speed, database systems need to achieve full throttle through efficient storage and retrieval techniques.

Reliability, availability and fault-tolerance is a huge concern for database systems. Reliability of a system is generally improved through redundancy. Modern businesses cannot afford to lose data or present wrong data. Modern business activities are highly cantered around and dependant on electronic data. Modern database systems thus need to build in high reliability mechanisms in their designs. High level view of database management systems

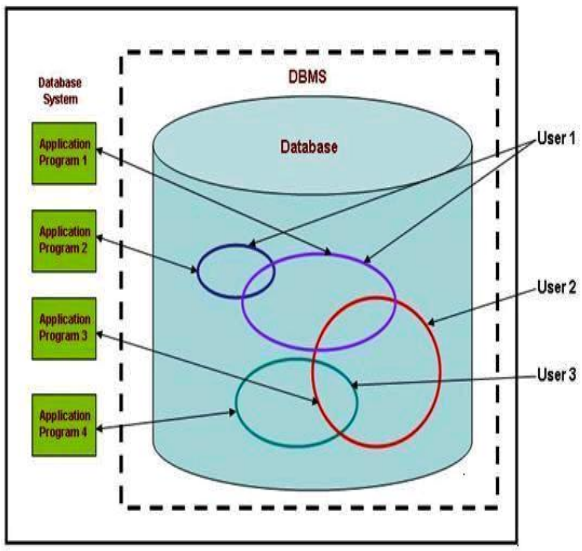


Figure 1 - a generic high level view of a database management system

Performance evaluation of database systems is thus an important concern. However, easier said than done, performance evaluation of database system is a non-trivial activity, made more complicated by the existence of different flavours of database systems fine tuned for serving specific requirements.

## 2. RELATED WORKS

### *2.1* General approaches to database management system performance evaluation*:*

The word "general" means that the approaches mentioned here are generally true for "systems" with a special focus on database systems. Performance analysis of database systems serves two basic purposes:

**1**. For the evaluation of the best configuration and operating environment of a single database system, and

**2**. Studying two or more database systems and providing a systematic comparison of the systems.

Accordingly, some of the analytical modelling methods for evaluating systems that are applicable for database systems too are:

**2.1.1Queuing Models**:
These models are effective to study the dynamics of a database system when it is modelled as a multi-component system with resource allocation constraints and jobs moving around from one component to another. Examples of such dynamic studies are concurrent transaction control algorithms, data allocation and management in distributed database systems etc.

**2.1.2 Cost Models**:
Cost Models are useful in studying the cost in terms of Physical storage and query processing time. The cost model gives some real insight into the actual physical structure and performance of a database system.

**2.1.3 Simulation Modelling**:
A simulation modelling is more effective for obtaining better estimates since it not only analyses the database system in isolation but can effectively analyze the database system with the application program running on top of it and the database system itself operating within the constrained environment of an operating system on a real physical hardware.

**2.1.4 Benchmarking**:
Benchmarking is the best method when multiple database systems need to be evaluated against each other but suffer the inherent setback that it assumes all systems to be fully installed and operational. Benchmarking relies on the effectiveness of the synthetic workloads. Real workloads are non repeatable and hence not good for effective benchmarking.

A set of methods for database performance evaluation assuming the system to be operating in a multi-user environment. Accordingly the three factors that affect the performance of a database system in a multi-user environment are:

1. Multi-programming level
2. Query Mix
3. Extent of data sharing

Data sharing is the condition of concurrent access of a data object by multiple processes. The interesting factor here is that of the query mix. A proper query mix needs to test the appropriate levels of CPU and disk utilization required to serve a particular query.

The query mix needs to properly represent a true multi-user environment. Also, the query mix may be designed to represent a certain percentage of data sharing. Once these have been figured out, the query-mix forms a representative benchmark program and multiple copies of the bench-mark program are issued concurrently to simulated multi-programming effects. Also, different query-mixes allow diversity in the experimental design conditions. The response variable studied is system throughput and response time.

## 3. PERFORMANCE EVALUATION CONCEPT

Two approaches have been taken to early performance analysis of dbms. The first approach is a qualitative analysis in terms of possible impacts of system aspects on product qualities such as performance.. The second approach is quantitative, using models. Smith developed an approach to building a queuing model based on scenario-like "execution graphs" that are specially built for performance analysis.
difficulties posed by early analysis, are:

- incompleteness of the specification, because it is so abstract. This includes open choices of design approaches, algorithms and components to be used, and lack of definition of the final execution environment
- lack of knowledge of the computational effort required
- other aspects such as ignorance of the workload intensity

Factors affecting performance: There are mainly 4 factors affecting the performance of database management systems.

### 3.1. Response time

Response time is defined as the time interval from the instant a command is input to the system to the instant the corresponding reply begins to appear at the terminal.

### 3.2Throughput

Throughput is the overall capability of the computer to process data. It is the combination of IO speed, CPU speed, parallel capabilities of the machine and efficiency of the operating system and system software.

### 3.3. Resources

Hardware and software tools, which include memory, disk speed, cache controller etc.

### 3.4. Memory

The total space required by a query to complete its execution. The values it takes for the completion of its execution

The workload parameters are partly specifications that must be taken as assumptions (such as the number of clients, types of query, query response time, execution time etc.) and partly predictions (such as CPU time for the operations to be implemented, the number of disk accesses and their cache hit ratio).

## 4. PERFORMANCE ANALYSIS OF DATABASE MANAGEMENT SYSTEMS

The development of systems that could adapt gracefully and opaquely to changing data and environment are the need of time. Query optimization, is one of the most happening fields in different areas of computer systems. In the last few years, researchers have been exploring the design of systems that are adaptive to operate in unpredictable and changeable environments.
Nowadays, different types of database are available, like ORACLE dbms, MS Access and SQL server from Microsoft, DB2 from IBM,open Source DBMS MySQL etc.
Im my paper, I am analyzing the Perfomance of three databases management systems. They are,

### 4.1 mysql

MySQL is a relational database management system (RDBMS) that runs as a server providing multi-user access to a number of databases. MySQL database is used for this project because it supports Linux and Java. For administrative purpose MySQL administrator is used. MySQL Administrator is a program for performing administrative operations, such as configuring your MySQL server, monitoring its status and performance, starting and stopping it, managing users and connections, performing backups, and a number of other administrative tasks.You can perform most of those tasks using a command line interface such as that provided by mysqladmin or mysql

### 4.2. Postgresql

PostgreSQL, often simply Postgres, is an object-relational database management system (ORDBMS) available for many platforms including Linux, FreeBSD, Solaris, Windows and Mac OS X.[6] It is released under the PostgreSQL License, which

is an MIT-style license, and is thus free and open source software. As with many other open-source programs, PostgreSQL is not controlled by any single company — a global community of developers and companies develops the system.

**4.3. sqlite**

**S**QLite implements most of the SQL-92 standard for SQL but it lacks some features. For example it has partial support for triggers, and it can't write to views (however it supports INSTEAD OF triggers that provide this functionality). While it supports complex queries, it still has limited ALTER TABLE support, as it can't modify or delete columns.

The time and memory capacity taken by different databases is different, because each and every database have its own specifications..Applications can make use of any of these databases according to users choice. This project aims to identify, time and memory capacity taken by different databases. That is instead of randomly selecting any one of the databases for each applications, it is possible to give provision to the user to select a database which take less time and less memory capacity . Evaluation of the response time of each query must be done. The result can be formulated by combining them. By comparing the results formed from different databases, it will be possible to identify which database take less time and less memory capacity.

Performance evaluation of database systems is an important concern. This is the same reason why Database management systems need to be evaluated based on their performances. Different Database management systems perform differently even in the same environment. There are multiple ways to measure the performance of a system. The most commonly used performance metrics are response time (R) and throughput (X). The response time is defined as the time interval from the instant a command is input to the system to the instant the corresponding reply begins to appear at the terminal. The throughput is generally considered as a measure of the system's productivity, that is, the number of programs or transactions processed during the measurement period. Essentially, they are the same aspect of the system viewed from different perspectives.

Users are interested in the response time while system managers are interested in throughput. Given infinite resources, the expected quality of service can always be provided, but with limited resources, capacity planning and resource management is needed. This project, evaluate the performance of different database management system. Based on that comparison of different DBMSs in the same environment is possible. This performance evaluation is done based on various parameters such as execution time, throughput etc.

## 5. ANALYSIS

Performance evaluations of database systems are made difficult by the existence of different database management systems (DBMS), which serves different requirements. Even though database management systems are limited in speed due to many limitations, they need to achieve full throttle through its efficient performance. Database management systems perform many complicated jobs. That is why they need to be efficient in storage and speed. This is the same reason why database management systems need to be evaluated based on their performances. Different database management systems performs differently even in the same environment.

**5.1. Response time calculation*:***

Query response time is the total time a query takes to complete its execution. It depends on the complexity of the query type. Database time is calculated by aggregating the CPU time and wait time of all user sessions not waiting for idle wait events. The figure shows the number of queries used, type of query, the checkout time across the database time.

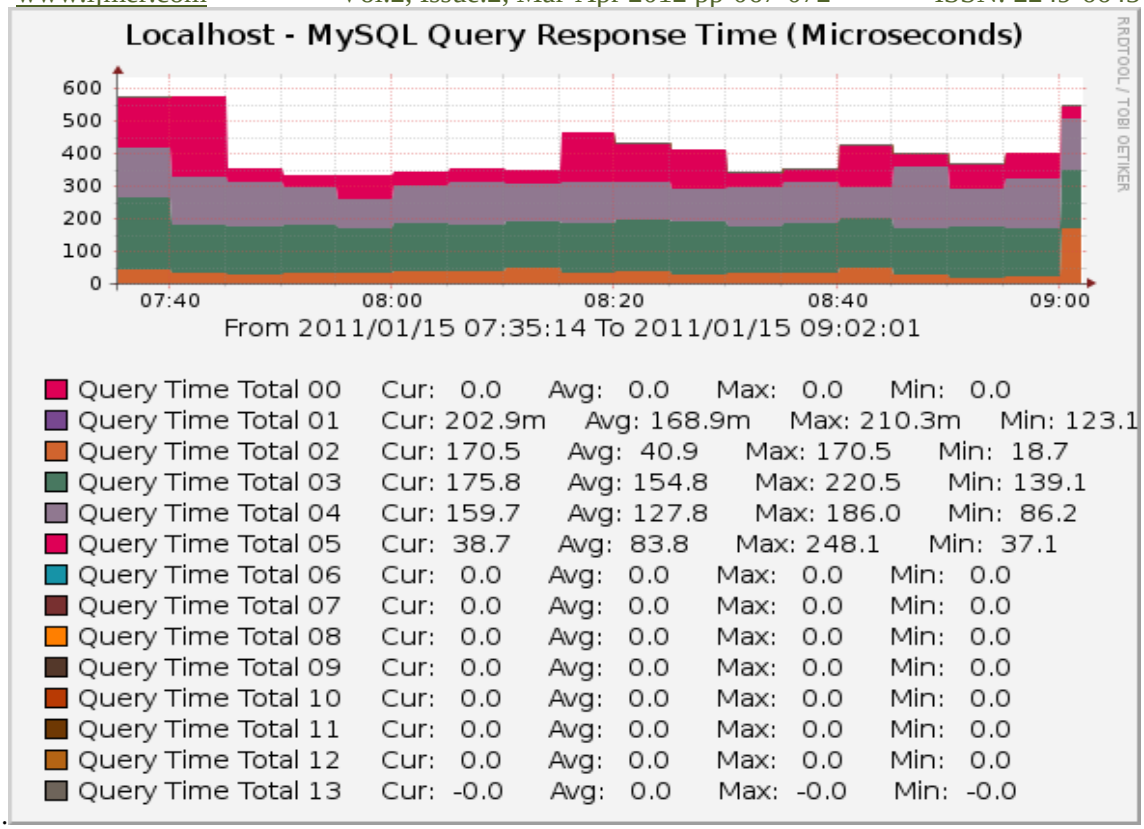


Figure 2: Response time calculate inaccordance with the queries

Figure 2 displays a histogram of the query response time distribution. Because the time units are user-configurable, exact unit labels are not displayed; rather. There are 13 entries on the graph. The graph actually displays the amount of response time spent on queries of various lengths. The units are in microseconds on the graph. When the processor speed changes from 1 GHz to 2 GHz, the modeled response time also increases.

**5.2. Memory capacity calculation:**

The total space required by a query to complete its execution. The values it takes for the completion of its execution. By combining the space, each query takes for execution together gives the total memory space required by that application

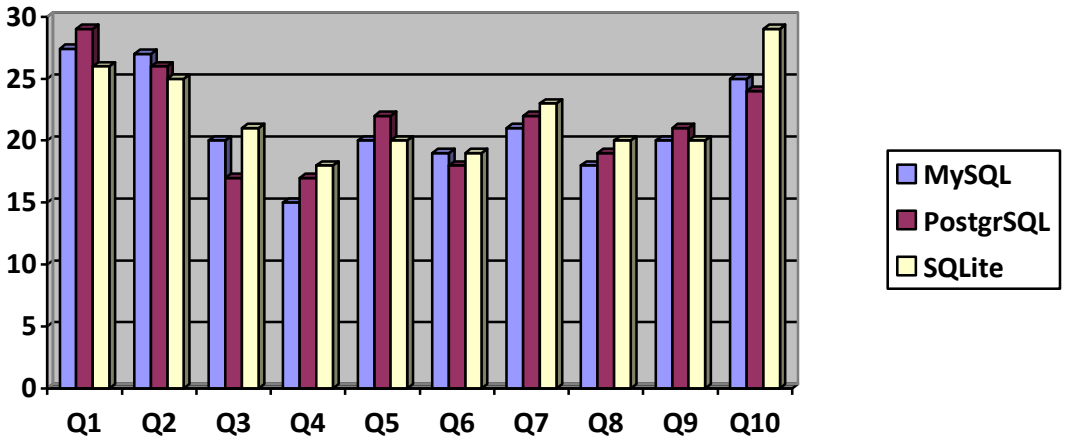


Figure 3: Performance comparison of MySQL,PostgreSQL and Sqlite against different query types

The response time and memory for executing every single operation may be different. It depends upon the current system conditions also. That is whether the system is in ideal state, or multiple processes are running concurrently etc. The analysis is performed against each query type and ,the response time and the memory to execute those query is calculated. By combining the result obtained by these calculations analysis can be done. The analysis is performed against various system conditions also. The graph shows response time of these database management systems which is taken for different queries. Memory capacity taken by different dbms can also be shown in the same way the response time is calculated in milli seconds.

## 6. CONCLUSION

This paper evaluates the performance of three different database management systems They are MySQL, SQLite, postgreSQL. Based on that, comparison of different database management systems in the same environment is possible. This performance evaluation is done based on parameters such as response time, memory capacity etc. The results obtained by evaluating time and memory capacity of different database management systems are stored within a DBMS. These values are compared with each other, to get the specific database management system which gives better performance than the others. Different database management systems works differently even in a unique environment. That is why analysis of the database management system which shows better performance is an important issue. . By comparing the performance of different database management system, selection of the database management system which took less time and less memory in that environment is possible

## REFERENCES

[1] Amol deshpande, Zachary ives and Vijayshankar raman. "Database Performance Evaluation Techniques for Specialized Databases" vol. 1, no. 1 (2007) 1–140c2007

[2] Xilin cui, "A capacity planning study of database management systems with Olap workloads", tcset'2008, february 19-23, 2009, lviv-slavsko, Ukrain

[3] Jian pei, Runying mao, Kkan hu and Hua zhu," Multiple approaches for Query Analysis", sigmod '00: proceedings of the 2000 ACM sigmod international conference on Management of Data, pp 592,2000

[4] Nurul husna mohd saad and Hamidah Ibrahim,"Multilingual Database Management System: A performance evaluation" , 43400 upm, serdang, selangor, Malaysia-2011

[5] Kyoung-Don Kang and Phillip H. Sin and Jisu Oh, A Real-Time Database Testbed and Performance Evaluation, Proceedings of the 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, Pages 319-326 , 2007

[6] Amol Deshpande, Zachary Ives and Vijayshankar Raman. "Adaptive Query Processing" -Foundations and Trends R in Databases Vol. 1, No. 1 (2007) 1–140c2007

[7] N. Bruno and S. Chaudhuri, "Exploiting statistics on query expressions for optimization," in SIGMOD '02: Proceedings of the 2002 ACM SIGMOD international conference on Management of data, (New York, NY, USA), pp. 263– 274, ACM Press, 2002. international conference on Management of data, pp 592,2000

[8] Piotr Mazur,Jan Marlewki,Marek Kamiski. "Universal database repository model for use in near real-time database warehouse environment":TCSET'2008, February 19-23, 2009, Lviv-Slavsko, Ukraine

## Authors

**APARNA KALADI** received B.Tech degree in Computer Science and Engineering from the Mahathma Gandhi University, Kerala in 2010 and pursuing ME degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology under Anna University, Coimbatore. Her research interests include Data mining and computer networks.

**MRS. PRIYA PONNUSAMY** is working as an Assistant Professor in Sri shakthi institute of engineering and technology. She is a research scholar under Anna University of Technology- Coimbatore. She did her ME in Anna University of Technology-Coimbatore and her BE in Tamilnadu College of Engineering, Coimbatore. Her area of interest includes networking and hardware.