

Performance Comparison of MySQL and PostgreSQL Database Management Systems in Mixed Workload Environments

Student Name: Alwaleed Alissa

Course: Research Methodology in Computer Science (A1800)

Supervisor: Dr. Mohammed Anbar

Submission Date: August 15, 2026

Abstract

Database systems are used in many software applications, and choosing a suitable database can affect the performance of the application. This becomes more important when the system has to deal with different types of work at the same time.

This research compares MySQL and PostgreSQL in a mixed workload environment. The workloads include read operations, write operations, and more complex analytical queries. The experiment is planned using a dataset based on the general idea of the TPC-H benchmark. The performance will be measured using response time, CPU usage, memory usage, and I/O operations.

Three levels of concurrency and three different data sizes will be used in the experiment. Based on the previous studies reviewed in this research, PostgreSQL is expected to have better performance with complex analytical queries, while MySQL may perform better with simple queries and write operations.

The main purpose of this research is to make a practical comparison between MySQL and PostgreSQL and provide some recommendations that can help developers choose the database system that fits their application.

Keywords: MySQL, PostgreSQL, Mixed Workloads, Database Performance, OLTP, OLAP

Chapter One: Introduction

1.1 Introduction

Database systems are an important part of most modern software applications. Many applications do not use only one type of database operation. For example, an application can process normal transactions while users are also running analytical queries on the same database.

Because of this, choosing a database system is not always easy. A database that performs well with simple transactions may not give the same performance when more complicated queries are used. Therefore, testing database performance with different workloads can help in making a better choice.

MySQL and PostgreSQL are two popular open-source relational database management systems. MySQL is commonly used because it is relatively simple and works well with normal transactional queries. PostgreSQL provides many advanced SQL features and can be useful for complex queries and analytical workloads.

Even though both systems are widely used, their performance can change depending on the type of workload, the amount of data, and the number of users working at the same time. For this reason, this research compares MySQL and PostgreSQL using a mixed workload that includes read, write, and analytical operations.

The main purpose of this study is to compare the two database systems through an experiment, understand their strengths and limitations, and provide some practical recommendations for choosing between them.

1.2 Theoretical Background

A relational database management system stores information mainly in tables that contain rows and columns. SQL is used to create and manage the data, as well as to retrieve and modify it.

When the amount of data becomes larger and queries become more complicated, database performance becomes more important. Different factors can affect performance, including the query structure, indexes, hardware resources, number of users, and database configuration.

MySQL

MySQL is an open-source relational database system that is widely used, especially in web applications and systems that depend on transactions. It is generally known for its simple design and good performance with straightforward queries.

MySQL uses the InnoDB storage engine, which supports important database features such as transactions and foreign-key constraints. Because of these features, MySQL can be suitable for many online transaction processing (OLTP) applications.

However, complex analytical queries can require more processing. Queries that use several JOIN operations, aggregation, or window functions may have different performance compared with PostgreSQL.

PostgreSQL

PostgreSQL is an open-source object-relational database management system. It provides many advanced SQL features, including Multi-Version Concurrency Control (MVCC), window functions, recursive queries, and advanced query-processing capabilities.

These features can make PostgreSQL suitable for applications that need complex queries and analytical processing.

On the other hand, using more advanced features can sometimes require more system resources. This may be noticeable when the database is processing many simple queries at the same time.

1.3 Problem Statement

Although MySQL and PostgreSQL are both widely used, deciding which one is better can be difficult when an application has different workloads running together.

One problem is that many database performance studies concentrate on only one type of workload. For example, a study may focus mainly on read operations or write operations. This does not always represent real applications where several types of operations happen at the same time.

Another problem is resource usage. Looking only at response time does not always explain the reason for a performance difference. CPU usage, memory usage, and I/O operations can give more information about how the database handles different workloads.

There is also a need for simple recommendations that developers can use when selecting a database for applications that have mixed workloads. Looking at workload type, concurrency, and data size together can provide more useful information.

For these reasons, this research proposes an experiment to compare MySQL and PostgreSQL under mixed workload conditions.

1.4 Research Objectives

The objectives of this study are:

1. To measure the response time of four query types under three levels of concurrency.
2. To examine CPU, memory, and I/O resource usage during the workload.
3. To study how increasing the data size affects MySQL and PostgreSQL performance.
4. To compare the two database systems using read, write, and analytical workloads.
5. To provide practical recommendations for selecting a suitable database system.

1.5 Research Scope

This research focuses on:

- **Database systems:** MySQL 8.0 and PostgreSQL 16.
- **Operating environment:** A virtual server with 4 vCPUs, 16 GB RAM, and SSD storage.
- **Dataset:** A simplified model inspired by TPC-H and consisting of five tables.
- **Query types:** Simple read, medium read, complex analytical query, and batch write.
- **Concurrency levels:** 50, 100, and 200 concurrent users.
- **Performance measures:** Response time, CPU usage, RAM usage, and I/O operations.
- **Excluded areas:** NoSQL databases, distributed databases, security, backup, and recovery.

1.6 Research Limitations

There are some limitations that should be considered in this research.

First, the experiment uses one server with fixed hardware. The results could therefore be different if the experiment was performed on different hardware, cloud infrastructure, or a distributed system.

Second, the largest dataset in the planned experiment is limited to 10 million records. Larger datasets, such as 50 or 100 million records, could provide more information about scalability, but they are outside the available time and resources.

Another limitation is that the study mainly focuses on performance. Other factors such as administration, maintenance, operating costs, licensing, and ease of deployment are not discussed in detail.

Finally, each test is planned to be repeated five times. However, a laboratory environment cannot represent every situation that may happen in a real production system.

Chapter Two: Literature Review

2.1 Previous Studies

Several previous studies have examined relational database performance and the factors that can affect it.

Cooper et al. (2010)

Cooper et al. introduced the Yahoo! Cloud Serving Benchmark (YCSB), which is a framework used to evaluate the performance of cloud-serving systems. Their work mainly considered read and write workloads.

This study is useful because it shows how standardized performance measurements can be used. However, it does not directly compare MySQL and PostgreSQL under mixed workloads.

Salunke and Ouda (2024)

Salunke and Ouda compared the performance of PostgreSQL and MySQL. Their results showed good performance from PostgreSQL when dealing with complex queries, while MySQL performed well in simple transactional operations.

This study is related to the current research because it provides a useful way to compare the two database systems. However, the workload in that study was not specifically designed around mixed workloads.

Kurniadi et al. (2026)

Kurniadi et al. studied the performance of relational database management systems when processing large amounts of textual data.

Their results showed that the performance of different database systems can become similar under some conditions. However, the study did not specifically focus on complex OLAP queries.

Napradit et al. (2025)

Napradit et al. compared database performance in containerized and traditional environments. Their findings showed that the environment used to run the database can affect its performance.

This is important for the current research because both database systems should be tested in the same environment as much as possible.

Tanujaya and Suryasari (2025)

Tanujaya and Suryasari examined how SQL query structure affects the performance of MySQL, PostgreSQL, and Microsoft SQL Server.

Their study showed that the structure and optimization of a query can have a noticeable effect on response time. This supports the idea of using similar query structures when comparing database systems.

2.2 Critical Analysis of Previous Studies

The previous studies provide several useful ideas for the current research.

Different researchers have used different methods to measure database performance. Some have used benchmark tools and workload-generation tools such as Sysbench and Apache JMeter.

Another important point is that database performance should not be evaluated using response time only. CPU and memory usage can also provide useful information about how the database system works under different conditions.

At the same time, the previous literature has some limitations. Many studies focus on only one workload type. This makes it harder to understand what happens when read, write, and analytical workloads are running together.

Resource usage has also received less attention in some comparisons. Measuring CPU, memory, and I/O usage can help explain why one system performs differently from another.

There is also a need for practical recommendations that developers can use when selecting a database for applications with mixed workloads.

2.3 Research Gaps

Based on the studies reviewed, several research gaps can be identified.

First, there is a need for a comparison that combines different workload types in the same experimental environment.

Second, resource usage needs to be considered along with response time. Measuring CPU, memory, and I/O can give a more complete picture of database performance.

Third, the results of database comparisons should be connected to practical recommendations so developers can use them when choosing a database.

Finally, the relationship between data size and concurrency needs more attention because both factors can influence database performance.

This research tries to address these points by testing different workload types with different data sizes and concurrency levels.

Chapter Three: Research Methodology

3.1 Research Methodology

This research uses an experimental approach. The same testing environment will be prepared for MySQL and PostgreSQL, and the same workload scenarios will be run on both systems.

The purpose of using the same conditions is to make the comparison as fair as possible.

Four types of queries will be tested:

- 1. Simple read query.
- 2. Medium read query.
- 3. Complex analytical query.
- 4. Batch write operation.

The experiments will use three different data sizes and three levels of concurrency.

The experimental setup will include database installation, dataset generation, database configuration, and monitoring. These steps will be kept as similar as possible for both systems.

3.1.1 Research Environment

The same hardware and operating system will be used for both database systems.

Component	Specification
Processor	4 vCPUs at 2.5 GHz
Memory	16 GB DDR4
Storage	100 GB SSD
Operating System	Ubuntu 22.04 LTS

Using the same environment can reduce the effect of hardware differences on the comparison.

3.1.2 Tools Used

The following tools will be used in the experiment:

Tool	Function
Sysbench	Load generation and response-time measurement
pgbench	Initial data generation
Prometheus + Grafana	Real-time resource monitoring
Python + pandas	Statistical data analysis

These tools will help generate workloads, monitor system resources, and analyze the collected data.

3.1.3 Experimental Design

The dataset will contain five tables:

- Customers

- Products
- Categories
- Orders
- Order Details

Three data sizes will be used:

- M1
- M5
- M10

Four query types will be tested:

- **Q1:** Simple read using a primary key.
- **Q2:** Medium read involving a JOIN between two tables.
- **Q3:** Complex analytical query using multiple JOIN operations, GROUP BY, and window functions.
- **Q4:** Batch write operation involving the insertion of 1,000 records.

The experiment will use three concurrency levels:

- 50 concurrent users
- 100 concurrent users
- 200 concurrent users

Each test will be repeated five times.

The total number of planned tests is:

2 database systems × 4 query types × 3 concurrency levels × 3 data sizes × 5 repetitions = 360 tests

3.2 Expected Results

The following results are expected based on the previous studies and the characteristics of MySQL and PostgreSQL.

These numbers are **expected values only**. They are not actual experimental results because the tests have not yet been carried out.

3.2.1 Response Time

Query	MySQL (ms)	PostgreSQL (ms)	Expected Faster System
Q1 Simple	3	4	MySQL

Query	MySQL (ms)	PostgreSQL (ms)	Expected Faster System
Q2 Medium	13	12	PostgreSQL
Q3 Complex	85	63	PostgreSQL
Q4 Write	45	53	MySQL

Based on these expected values, the study expects MySQL to perform well with simple read operations and write-heavy workloads.

PostgreSQL is expected to perform better with medium and complex analytical queries. It is also expected to remain more stable when the dataset becomes larger.

3.2.2 Resource Usage

Resource	MySQL	PostgreSQL
CPU Usage	45%	52%
RAM Usage	6 GB	9 GB

The expected values suggest that PostgreSQL may use more CPU and memory when processing complex queries. In situations where hardware resources are limited, this could make MySQL a more suitable option.

Again, these values are expected values and should not be considered actual measurements until the experiment is completed.

3.2.3 Recommendation Matrix

Application Case	Recommended System
OLTP / Read-intensive	MySQL
OLAP / Analytical Queries	PostgreSQL
Hybrid Applications	PostgreSQL
Large Data Size (> M5)	PostgreSQL
Write-intensive Applications	MySQL
Complex JSON Processing	PostgreSQL

These recommendations are based on the expected behavior of the two systems and the previous studies reviewed in this research.

3.3 Gantt Chart

Activity	Week 1	Week 2	Week 3	Week 4	Week 5
Literature Review	✓				
Environment and Data Setup	✓	✓			
MySQL Tests		✓	✓		
PostgreSQL Tests			✓	✓	
Results Analysis and Report				✓	✓

References

1. Cooper, B. F., Silberstein, A., Tam, E., Ramakrishnan, R., & Sears, R. (2010). Benchmarking cloud serving systems with YCSB. In *Proceedings of the 1st ACM Symposium on Cloud Computing (SoCC)*, 143–154. ACM. <https://doi.org/10.1145/1807128.1807152>
2. Kurniadi, A., Hermawan, I., & Pratama, F. (2026). Performance comparison of relational database management systems for large-scale text data processing. *International Journal of Data Science and Analytics*, 22(1), 45–58.
3. Napradit, A., Masrijun, P., & Rassameeroj, I. (2025). Containerized and host-based RDBMSs performance comparison. In *2025 11th International Conference on Computer and Communications (ICCC)*, 12–16. IEEE. <https://doi.org/10.1109/ICCC68654.2025.11438083>
4. Oracle Corporation. (2024). *MySQL 8.0 Reference Manual – Optimization*. MySQL Documentation.
5. PostgreSQL Global Development Group. (2024). *PostgreSQL 16.6 Documentation – Performance Tips*. PostgreSQL Documentation.
6. Salunke, S. V., & Ouda, A. (2024). A performance benchmark for the PostgreSQL and MySQL databases. *Future Internet*, 16(10), 382. <https://doi.org/10.3390/fi16100382>
7. Tanujaya, D., & Suryasari, D. (2025). The impact of SQL query structure on RDBMS performance: A comparative study. *Journal of Computer Science and Technology*, 40(3), 512–527.
8. Transaction Processing Performance Council (TPC). (2023). *TPC-H Benchmark Specification (Revision 3.0.0)*.