YCompany - eClaim System’s Database Engine

DAR Document



Nagarro Software Pvt. Ltd.

Sumit Chauhan

|  |  |  |  |
| --- | --- | --- | --- |
| **Revision History** | | | |
| Version | Date | Author/Contributor | Comments |
| 1.0 | 11/05/2018 | Sanjeet Chauhan | Initial Version |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Contents

[1 Introduction 4](#_Toc529180596)

[1.1 Objective and scope of document 4](#_Toc529180597)

[2 Requirements at a Glance 5](#_Toc529180598)

[3 Available tools 6](#_Toc529180599)

[3.1 MS SQL Server 6](#_Toc529180600)

[3.1.1 Features 6](#_Toc529180601)

[3.1.2 Pricing 6](#_Toc529180602)

[3.2 MySQL Server 7](#_Toc529180603)

[3.2.1 Features 7](#_Toc529180604)

[3.2.2 Pricing 7](#_Toc529180605)

[3.3 Oracle 7](#_Toc529180606)

[3.3.1 Features 7](#_Toc529180607)

[3.3.2 Pricing 8](#_Toc529180608)

[4 Comparison Analysis 8](#_Toc529180609)

[4.1 Point Matrix 8](#_Toc529180610)

[4.2 Comparison 8](#_Toc529180611)

[5 Recommendation 9](#_Toc529180612)

[6 Assumptions 10](#_Toc529180613)

[7 Risks 11](#_Toc529180614)

[8 Appendix 12](#_Toc529180615)

[8.1 References 12](#_Toc529180616)

# Introduction

As per requirement of the YComapany – eClaim design, it is required to choose a database engine that provides the desired level of performance, efficiency at minimum cost.

This document is for analyzing various database engines available in the marketplace and providing a detailed comparison of the available option. Comparison is made based on intended features using a point-based scoring system to identify the best fit.

## Objective and scope of document

The objective of system is to identify all the features required for a database engine to qualify for fittment to the system. Futher, not all features are equally important, some carry a higher weightage other may carry lower weightage.

Also, there are large number of database engines available in the market place but not all available options are compared in this document. Only the databases that are relied & trusted worldwide are cosidered for detailed comparision.

# Requirements at a Glance

The system require a database engine that provides easy integration, low prices and great performance to be integrated with various services. The system design divide the data of the system into multiple databases to divide the load on the system. So, performance is a requirement but unlike many other production systems the data stored is not in a single databse so there is no possibilty of a singel database server becoming a bottleneck for overall system performance.

Comparison of the available database engines will be done on following parameters:

* Performance
* Pricing
* Ease of development & integration
* Security & Compliance
* Replication
* Backup & Recovery
* Support & Online Help
* Manageability
* Scalability

# Available tools

There are many database engines available in the marketplace being used for enterprises application. Most common are : Oracle, MS SQL Server, MySQL, Aurora Db, Maria DB, Postgree SQL etc.

We are considering MS SQL Server, MySQL & Oracle for comparison as these are the most trusted, most widely used and can be integrated easily in the proposed system.

## MS SQL Server

Microsoft SQL Server is a relational database management system developed by Microsoft. As a database server, it is a software product with the primary function of storing and retrieving data as requested by other software applications—which may run either on the same computer or on another computer across a network (including the Internet).

Microsoft markets at many different editions of Microsoft SQL Server, aimed at different audiences and for workloads ranging from small single-machine applications to large Internet-facing applications with many concurrent users.

### Features

* SQL Server is a high-performance database able to perform at massive scale, consistently leading in TPC-E OLTP workload benchmarking.
* Cross platform database engine.
* Scale to petabytes of data for enterprise-grade relational data warehousing—and integrate with non-relational sources like Hadoop.
* Protect data at rest and in motion with a database that has the least vulnerabilities of any major platform for six years running in the NIST vulnerabilities database (National Institute of Standards and Technology, National Vulnerability Database, Jan 17, 2017).
* Gain mission-critical uptime, fast failover, easy setup, and load balancing of readable secondaries with enhanced Always On
* Build intelligent applications with SQL Server Machine Learning Services using R and Python.
* Combine in-memory column store and row store capabilities in SQL Server 2017 for real-time operational analytics—fast analytical processing right on transactional data.
* Develop once and deploy anywhere with consistent experience from on-premises to cloud with support for Windows and Linux as well as Docker containers.
* Point in time restore
* Automated & manual backups
* SQL Server mirroring for high availability & failover solutions

### Pricing

Pricing has been considered with following configuration:

|  |  |
| --- | --- |
| Avg Usage | 70% |
| DB Engine Type | SQL Server 2017 - Standard |
| Deployment Model | Multiple Availability Zone |
| Storage | 5TB |
| CPU | 16vCPU |
| Memory | 64GB |

For the configuration mentioned above, an instance of SQL Server costs **6996 USD/Mo**.

## MySQL Server

MySQL is an open-source relational database management system. MySQL was owned and sponsored by a single for-profit firm, the Swedish company MySQL AB, now owned by Oracle Corporation. For proprietary use, several paid editions are available, and offer additional functionality.

### Features

* Easy backup and restore
* Ability to create Multi – AZ secondaries, Read Replicas
* Compliant with many industry standards
* Enables delivering high-performance and scalable Online Transaction Processing (OLTP) applications.
* Ease of use that has made MySQL famous along with industrial strength performance and reliability.
* Lower total cost of ownership - MySQL enables minimizing database TCO
* MySQL Workbench provides an integrated development, design and administration environment to make developers and DBAs more productive.

### Pricing

Pricing has been considered with following configuration:

|  |  |
| --- | --- |
| Avg Usage | 70% |
| DB Engine Type | MySQL Standard 8.0 |
| Deployment Model | Multiple Availability Zone |
| Storage | 5TB |
| CPU | 16vCPU |
| Memory | 64GB |

For the configuration mentioned above, an instance of MySQL Server costs **2882** **USD/mo**.

## Oracle

Oracle Database is a multi-model database management system produced and marketed by Oracle Corporation. It is a database commonly used for running online transaction processing (OLTP), data warehousing (DW) and mixed (OLTP & DW) database workloads. The latest generation Oracle database is available on-prem, on-Cloud, or in a hybrid-Cloud environment.

### Features

Oracle is a feature rich database providing wide array of feature for wide array or domains, applications, security & performance requirements. Of course, it comes with all the bases features that any RDBMS system has to offer, there is much more on the plate considering the offerings by oracle database.

Some of the most recent features added to oracle are:

* Database Replay
* Referential, Interval & Virtual column partitioning
* Automatic health monitor
* Advanced compression
* Automatic storage management
* Point in time restore, automated & manual backups
* Multi-AZ deployment for high availability & failover solutions

### Pricing

Pricing has been considered with following configuration:

|  |  |
| --- | --- |
| Avg Usage | 70% |
| DB Engine Type | Oracle 12c |
| Deployment Model | Multiple Availability Zone |
| Storage | 5TB |
| CPU | 16vCPU |
| Memory | 64GB |

For the configuration mentioned above, an instance of SQL Server costs **5014 USD/mo**.

# Comparison Analysis

All the three database engines will be compared on various features such as performance, scalability security, pricing and many more listed in the point matrix below, various feature bear different weightage. The final score of each database will be based on sum of the score for various features.

## Point Matrix

Table below list down various features that are under consideration for comparing th three database engines with their weightage.

|  |  |
| --- | --- |
| Feature | Points |
| Performance | 20 |
| Pricing | 20 |
| Ease of development & integration | 15 |
| Security & Compliance | 10 |
| Replication | 10 |
| Backup & Recovery | 10 |
| Support & Online Help | 5 |
| Manageability | 5 |
| Scalability | 5 |

## Comparison

Feature based relative scoring of all the three database engines is mentioned below

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | SQL server | My SQL | Oracle |
| Performance | **19** | **15** | **17** |
| Security & Compliance | **10** | **9** | **10** |
| Pricing | **10** | **19** | **15** |
| Scalability | **5** | **5** | **5** |
| Replication | **10** | **10** | **10** |
| Backup & Recovery | **10** | **9** | **10** |
| Support & Online Help | **5** | **5** | **4** |
| Manageability | **5** | **5** | **5** |
| Ease of development & integration | **15** | **14** | **13** |
| Total Score | **89** | **91** | **89** |

# Recommendation

Based on the score calculation, it is recommended that My SQL be used as database engine. It provides right mix of features on the requirement scale.

# Assumptions

1. Standard version of all the databases has been considered for comparision, based on the proposed design, enterprise version won’t be required.
2. OLTP database engines wont’be used for analytical purpose, so such features haven’t been considered for comparision. Further regarding analytics,
   1. Such use cases are out of scope.
   2. If a requirement comes in future, the prolem would be solved using no sql database engines populated post ETL processes which is already a part of system design.
3. Pricing mentioned for the considered database engines is done at the time of analysis and keep fluctuating by small amount in cloud environment.
4. The prices mentioned are just for comparision purpose & has been done for a single instance on average hardware. The actual number of database server nodes & number of instances will be identified later on.

# Risks

* The current architecture doesn’t make database as the core building block & the overall system performance doesn’t completely depend on the QphH capabilities of the database engine. So, there is no risk involved from performance perspective.
* From management perspective, the current design relies on cloud services to backup, replicate, update & monitor databases, so minor risk is involved if migration to on premises design is considered. All these tasks will be handled by the on premises team, which is a complex task using the standard versions. For such use case either an upgrade to enterprise version is needed or additional tooling & monitoring support will be required on premises.

# Appendix

## References

1. Price calculator for various databases on AWS

<http://calculator.s3.amazonaws.com/index.html#s=RDS>

1. Recent QphH results

<http://www.tpc.org/tpch/results/tpch_perf_results.asp>

1. SQL Server 2017

<https://www.microsoft.com/en-us/sql-server/sql-server-2017>

1. MySQL 8.0

<https://www.mysql.com/why-mysql/white-papers/whats-new-mysql-8-0/>

1. Oracle 12c

http://www.oracle.com/us/corporate/features/database-12c/index.html