

Improving a Hospital Database's Performance by Implementing Materialized Views



Abdulaziz Alhamadani

Department of Computer Science, University of New Hampshire

ABSTRACT

Hospitals rely heavily on databases to store the huge amount of data on a regular daily basis of patients, doctors' reports, test results, operations, visitors of hospitalized patients, and so on. On the other hand, the purpose of storing the data is for retrieval for future use. Hospitals require that data retrieval is fast and reliable. This study is applied on a hospital from Saudi Arabia which has been running their central database on Oracle for the past years. They use views (virtual tables that does not have data in itself) to reduce the complexity of retrieval and other purposes such as security. However, there is a way to improve the performance of their database by distributing the data and creating materialized views instead of having central database and regular views. Distributing the database has several advantages to help increase the performance so are the materialized views. The study applies three different scenarios, and outlines a comparison among the three scenarios by testing them individually.

INTRODUCTION

- The study is based on real database schema of a hospital in Saudi Arabia. The hospital's schema has over 1400 tables but for this project over 120 tables have been borrowed to implement the study.

- The records of patients have been randomly generated by a sophisticated procedure and some temporary tables and random data.

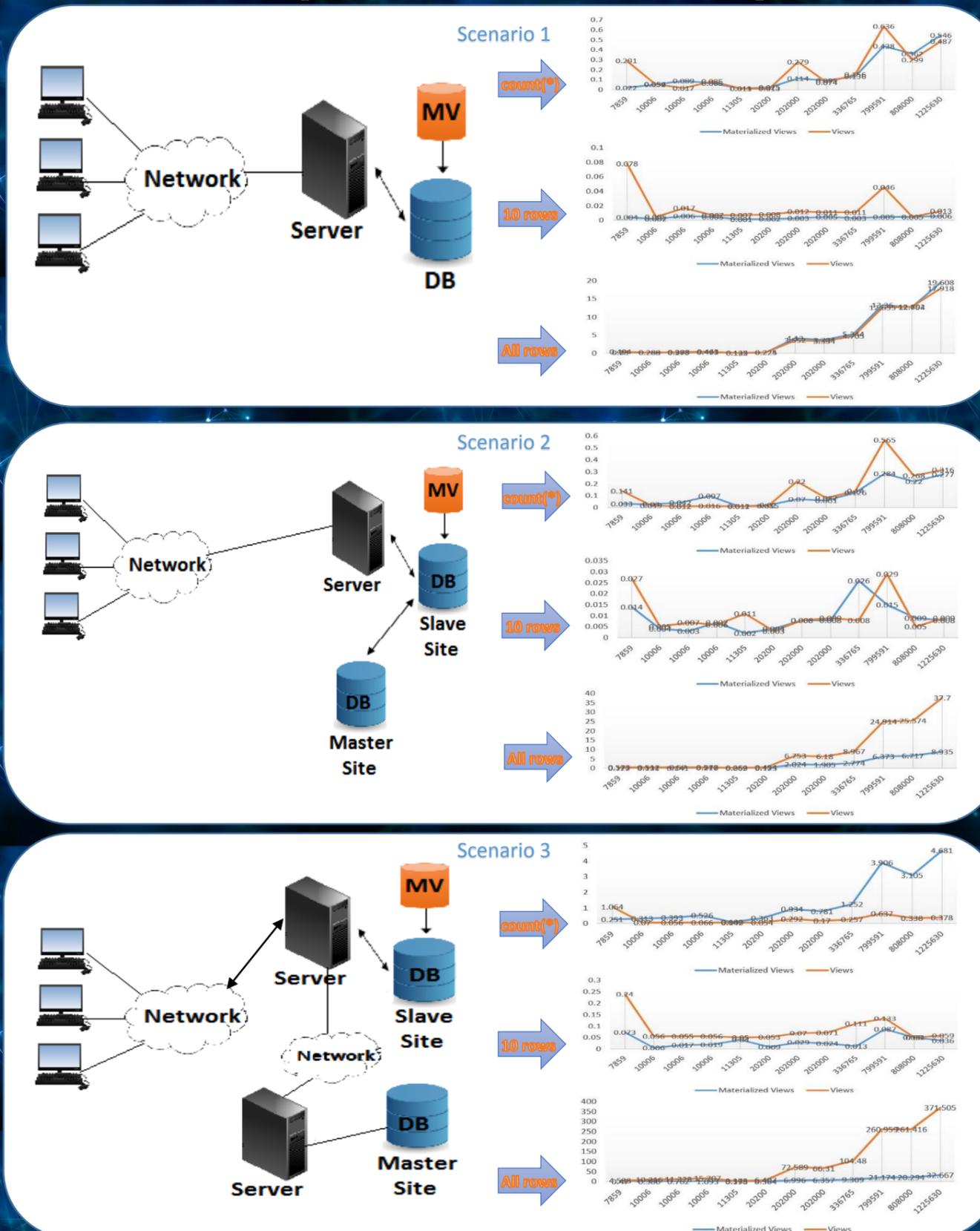
- The initialized records start from 10000 and reaches over a million record by the end of the study.

- The study simulates the workflow of a hospital by admitting more than 600 patients per hour, and more than 16000 patients per day.

- The previous way was using regular views to run the queries, while the suggested way is using materialized views.

- The study implements three scenarios to run the tests; the first is centralized and the other two are distributed as full replicas.

APPROACH → RESULTS



DISCUSSION

- Views: they are virtual tables generated by a SELECT query. The hospital currently uses them entirely for their daily work. They hide critical information and return latest updates but they have performance issues.
- Materialized views: they are actual tables based on the main tables. They have a different technique of updating of regular views. The technique that has been used to update the materialized views is FAST REFRESH.
- Distributing the hospital's database was implemented by fully replicating the data to protect hospital's database from loss and recovery in case of failure.
- A Local Area Network (LAN) has been used to connect the remote databases to ensure there were no network issues during the tests.

CONCLUSION

- **First Scenario:** This scenario does not need materialized views because the performance is similar to the regular views. The hospital could save money by not applying it, but there are limitations of not having remote transactions and not having their data secured.
- **Second Scenario:** the scenario offers data reliability and efficiency when the database grows larger. The costs are moderate but it has limitations such as data loss in case of disasters, and lack of remote transactions. I recommend this scenario over the first scenario for small hospitals that do not expect to grow larger.
- **Third Scenario:** Materialized views are recommended for this scenario. Retrieving complex queries and large data are faster than views. It is reliable, in case of disasters because the slave site could perform on its own. However, it is the most expensive financially, and it relies on the performance of the network. I recommend implementing this scenario for large hospitals to ensure their data availability and high performance.