

Docker-Based and Virtual Machine-Based Apache Cassandra Database Management System Evaluation

Kithulwatta W.M.C.J.T.^{1,2*}, Jayasena K.P.N.³, Kumara B.T.G.S.³ and Rathnayaka R.M.K.T.⁴

¹Faculty of Graduate Studies, Sabaragamuwa University of Sri Lanka

²Department of Information and Communication Technology,
Uva Wellassa University of Sri Lanka

³Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka

⁴Department of Physical Sciences and Technology, Sabaragamuwa University of Sri Lanka
[*chiranthajtk@gmail.com](mailto:chiranthajtk@gmail.com)

Computer systems are a major component in the current complex era. In computer systems, to archive data and information in a well-defined manner, database management systems (DBMSs) are used. Among the plenty of DBMS vendors, Apache Cassandra is a popular no-SQL DBMS. According to traditional computer system infrastructure management, DBMSs are launched on virtual machines. After the arrival of containerization, most software applications and services were launched on container-oriented infrastructure. Docker is the most popular and trending container vendor. This research study aims to evaluate the performance of Apache Cassandra DBMS on both Docker-based infrastructure and virtual machine-based infrastructure. Docker containerized infrastructure was launched on Ubuntu 18.04 Long Term Support (used package with architecture: GNU/Linux 4.15.0-112-generic x86_64) cloud-based operating system. The host computer was with 15 GB memory capacity and 1 Gbps network bandwidth. On that host computer infrastructure, Docker version 19.03.9 was launched. Both Docker client and server engine communities are version 19.03.9. Docker Application Program Interface (API) version was 1.40. The same host infrastructure was used for the virtual machine-based infrastructure. By executing queries with 5, 50, 500, 5000, and 50000 data records on the Apache Cassandra DBMS, a particular performance evaluation was made. According to the query execution time, the increment of performance on Docker infrastructure was 0.244 (SELECT), 0.295 (DELETE), 0.496 (UPDATE), and 0.708 (INSERT) with 95% confidence. The experiment depicts that the Docker-based platform presents a higher and more advanced infrastructure for the Apache Cassandra DBMS.

Keywords: *Apache Cassandra, Containers, Database Management System, Docker, Virtual machine*