

Performansi Pemrosesan *Big Data* pada Virtualisasi Berbasis *Container* dan *Hypervisor*

Muhammad Rashief¹, Sidik Prabowo², Siti Amatullah Karimah³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹rashiefm@student.telkomuniversity.ac.id, ²pakwowo@telkomuniversity.ac.id,

³karimahsiti@telkomuniversity.ac.id

Abstract

Hadoop is a tool that is very helpful in processing large data in a distributed manner. Hadoop MapReduce is a programming model that processes large datasets, because Hadoop deals with large data, this affects the strength of the hardware used. With the virtualization technique, Hadoop's performance can be optimized. In this study comparing two types of virtualization, namely Container and Hypervisor where both have different architectures. Looking at CPU Utilization in Container where ten times higher than the Hypervisor, compares Disk I / O where Container is more optimal than Hypervisor which affects the execution time in running a wordcount job where Container gets four times faster than Hypervisor. The container (Docker) gets better performance than the Hypervisor (VMware) in the test given.

Keywords: hadoop, container, hypervisor, wordcount, CPU utilization, TestDFSIO