

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

Abstrak

Hadoop merupakan *tools* yang sangat membantu dalam pemrosesan data berukuran besar secara terdistribusi. Hadoop *MapReduce* merupakan model pemrograman yang memproses *dataset* yang besar, karena Hadoop berhubungan dengan data yang besar, hal ini mempengaruhi kekuatan pada *hardware* yang digunakan. Dengan adanya teknik virtualisasi, kinerja Hadoop dapat dioptimalkan. Pada penelitian ini membandingkan dua jenis virtualisasi yaitu *Container* dan *Hypervisor* dimana keduanya memiliki arsitektur yang berbeda. Dengan melihat *CPU Utilization* pada *Container* dimana 6.9 kali lipat lebih tinggi dibanding dengan *Hypervisor*, membandingkan *Disk I/O* yang dimana *Container* lebih optimal dari *Hypervisor* yang mempengaruhi *execution time* dalam menjalankan *wordcount job* yang dimana *Container* mendapatkan waktu 2.48 kali lebih cepat dari *Hypervisor*. *Container*(Docker) mendapatkan performa yang lebih baik ketimbang *Hypervisor*(VMware) pada pengujian yang diberikan.

Kata kunci : *hadoop, container, hypervisor, wordcount, CPU utilization, TestDFSIO*