

Abstrak

Hadoop merupakan *framework* yang seringkali digunakan untuk pemrosesan terdistribusi terhadap big data. Cluster hadoop melakukan pemrosesan dan penyimpanan data secara terdistribusi pada tiap node. Pembuatan cluster hadoop fisik memerlukan dana dan energi yang besar. Selain itu, fleksibilitas yang terbatas untuk perawatan dan sulit manajemennya.

Teknologi virtualisasi berbasis kontainer seperti *Linux Container* (LXC) dan *Open Virtuozzo* (OpenVZ) memberikan performansi mendekati server fisik. Saat ini pengembangan LXC sebagai pengganti dari OpenVZ pada *container-based virtualization* terus dilakukan. Pada penelitian sebelumnya tanpa menggunakan manajemen mesin virtual dan *I/O scheduler* berbeda, menunjukkan bahwa LXC memberi I/O throughput lebih tinggi dan waktu eksekusi job lebih cepat dari OpenVZ. Pada penelitian ini dilakukan perbandingan OpenVZ dan LXC pada platform Proxmox dengan *I/O scheduler* yang sama.

Pada Proxmox, I/O throughput dan waktu eksekusi MapReduce Job pada kasus parser syslog lebih cepat 5 detik pada OpenVZ. OpenVZ juga memberikan I/O throughput yang lebih besar dibandingkan LXC.

Kata kunci: big data, virtualization, hadoop, distributed computing, performance analysis, container-based