

Analisis Performansi NoSQL *file system* pada *Hadoop Distribute File System* dan *Cassandra File System*

Juan Yudha¹, Sidik Prabowo², Siti Amatullah Karimah³

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

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹juanyudha@students.telkomuniversity.ac.id, ²karimahsiti@telkomuniversity.ac.id,

³pakwowo@telkomuniversity.ac.id,

Abstract

At this time *technology* is growing rapidly. Based on research carried out by *Big Data* will dominate the technology market. The growth rate of *Big Data* increases every year, and the required *file system* (FS). *Hadoop* is an OS *open source* that is very popular among *Big Data*. But with the development and use of *Big Data*, the specifications on the *file system* also began to increase. A system that only provides to be made from the *bottleneck* category on *Big Data*. But this system starts still lacking, because there is no output in determining the performance of the *file system*. A system is needed to determine the performance of a *file system*. *File System* currently uses structured data based on one example, namely *SQL*. To return unstructured data, it is necessary to do it through *File System* which uses *NoSQL*. For this reason, we must test *file system* based on *load-write* and *load-read* using *Hadoop Distribute File System* (HDFS) and *Cassandra File System* (CFS). The results classification is divided into two, good or not from *file system*. From the research that has been done, the results can consist of *cassandra* having a performance of *throughput* which is better than the difference between HDFS which is equal to 1818.75 seconds for each operation carried out, and for the speed in the process of data obtained results from *Cassandra* better than HDFS the difference is 6028 seconds.

Keywords: *Big Data*, *performance*, *HDFS*, *CFS*, *load-read*, *load-write*.