ABSTRACT

Storage requirements in data centers are growing dramatically for most internet and cloud service providers today. Data protection mechanisms are used to protect data from device or system failures. The mechanism used is redundancy with Replication and Erasure Coding. However, Replication is very inefficient in terms of capacity utilization. Erasure Coding has become a viable alternative to Replication as it provides the same level of reliability as Replication with much less storage system usage. Both of these schemes can be done on a Software Defined Storage (SDS) based storage system.

In this final project, the authors conducted a comparative analysis of Erasure Coding and Replication on an opensource storage system based on software defined storage (SDS) to measure the performance of Disk Utilization, IOPS, CPU usage and throughput on reading and writing based on multiple disks from comparisons during Replication and Erasure modes. code on SDS. The number of disks used is 3 disks, 6 disks, and 9 disks with each disk measuring 100 GB and using 3 storage servers which are used as 1 SDS storage server cluster.

Based on the results obtained in this study, Erasure Coding can save three times the storage from Replication. But throughput and IOPS performance in Replication is far superior to Erasure Coding with an average writing and reading on Replication of 31.2 MB / s and 41.4 MB / s, while Erasure Coding is 25.2 MB / s and 27, 3 MB / s. Overall, to get greater storage performance you can use Erasure Coding, for performance that requires speed you can use Replication.

Keywords: *Software Defined Storage, Erasure Coding, Replication, Storage*