***Does Varying BeeGFS Configuration Affect the I/O Performance of HPC Workloads?-Review***

BeeGFS as a *parallel* file system indicates that files are striped over multiple server nodes to maximize read/write performance and scalability of the file system.

Abstract-The increasing gap between processor and disk performance leads to high performance computing (HPC) applications facing I/O bottlenecks. This makes parallel file systems one of the most important components in an HPC cluster. This work analyzes the I/O performance of different workloads for various BeeGFS configurations.

**Introduction-**

The Parallel File System (PFS) is a critical component of the I/O subsystem, enabling multiple clients to store and access data spread across multiple storage nodes through parallel I/O paths. BeeGFS, a POSIX-compatible PFS, is widely favored for its ease of installation and management.

This paper investigates the impact of altering the striping configuration and OST allocation strategy of BeeGFS on I/O performance. Specifically, our work tries to answer the following research questions.

• RQ1: How does the OST allocation strategy affect load imbalance across OSSs and OSTs?

• RQ2: How does varying chunk size and number of targets affect I/O performance?

**Methodology-**

The cluster has a total of 10 nodes with 4 clients, 2 metadata servers (MDS), 1 management server (MGS), and 3 object storage servers (OSS), each with 3 object storage targets (OST). It is equipped with BeeGFS-7.3.2 with 4.5TB of total capacity. Each client node has 8GB of memory. All nodes in the cluster are connected via a 1Gbps Ethernet link. The OSSs are each equipped with 32GB of DRAM. Fig. 3 shows the cluster setup with corresponding the BeeGFS services. Sx Ty represents the yth target in the xth server.

IOR is used to answer the research questions RQ1 and RQ2. The CPU, memory and network utilization of the cluster are monitored using the ELK stack [9]. The benchmark tools measure the I/O throughput and latency of the workloads. The beegfs-df command is used to check the load distribution of the chunks across the OSTs.

**CONCLUSION**-

BeeGFS is gaining popularity among small-scale HPC clusters. Our work studied the effects of varying striping configurations for various I/O workloads. We observed that the default OST allocation strategies lead to load imbalance on OST and OSS levels. The default striping configuration gives sub-optimal performance for various workloads. These results highlight the need for a tool for adaptive striping