****

*Paper Review:*

*“Performance Analysis of Distributed File System Based on RAID Storage for Tapeless Storage”*

Course Name**:**  Distributed computer systems

Course ID**: CSE707**

Section: 01

Submitted To

Annajiat Alim Rasel

Senior Lecturer

Department of Computer Science and Engineering, BRAC University

Date of Submission

1st november 2023

# Summary

This Report is an Analytical study of the Research Article titled “***Performance Analysis of Distributed File System Based on RAID Storage for Tapeless Storage***” published on 16 October 2023. Researchers working on this article are JUNGBIN KIM 1 , HYEON-JIN YU 1 , HYEONGBIN KANG 1 , JAE-HYUCK SHIN1 , HEESEOK JEONG 2 , AND SEO-YOUNG NOH 1 respectively. The report will showcase the summary and findings of the articles as well as the motivation, contribution and and Limitations found whilst analyzing the article and will showcase any possible scopes or applications by adapting the article. To begin with, The Article discusses the design and evaluation of a distributed file system (DFS) based on RAID storage as an alternative to tape storage. It showcases experimental results for read and write operations, comparing different layouts and file systems. The study emphasizes the importance of metrics such as bandwidth, IOPS, and latency in understanding storage performance. It also highlights the significance of workload patterns in tape storage scenarios and the need for individual assessment when configuring large-scale storage systems.

# Motivation/Purpose/Aims/Hypothesis

The motivation behind the Article is to propose a disk-based archival storage solution as an alternative to tape storage. The main objective is to evaluate the performance of a RAID storage-based DFS system and compare it with existing file systems. The assumption is that the RAID storage-based DFS system will exhibit greater performance in terms of bandwidth, IOPS, and latency compared to tape storage.

# Contribution

The article contributes to the field by providing experimental data that illustrate the performance metrics of the RAID storage-based DFS system. It displays the benefits and drawbacks of various layouts and offers insights into the read and write benchmarks for each. The significance of workload patterns and individual evaluation in setting up large-scale storage systems is also emphasized in the article.

# Methodology

To evaluate the performance of the RAID storage-based DFS system, the methodology used was conducting experiments on read and write operations comparing different layouts such as RAID 6 RAIN, replication, and distribution. The researchers measured performance metrics such as bandwidth, IOPS, and latency to assess the system's performance and also considered workload patterns in tape storage scenarios to provide a comprehensive evaluation.In simpler terms, they conducted tests to evaluate the system's performance in reading and writing data by experimenting with various data organization methods and assessed the system's speed, data handling capacity, and compatibility with different types of data, such as data stored on tape.

# Conclusion

The paper concludes that the RAID storage-based DFS system offers superior performance in terms of bandwidth and latency compared to tape storage. It highlights the advantages of the GlusterFS file system in distributed read items and the distinct bandwidth differences across different layouts in write benchmarks. The study emphasizes the need for individual assessment and workload analysis when configuring large-scale storage systems.

# Limitations

In any comprehensive research in any particular field there Might come across various limitations which later on will be Reviewed by many other researchers in this particular field. By analyzing the article Some Possible Limitations were observed.

## 1st Limitation

One limitation observed in the Article is that it focuses on specific file systems and layouts, such as GlusterFS and RAID 6 RAIN. This means that the results may not be generalizable to other file systems or configurations. Additionally, the experiments do not consider all possible workload patterns and scenarios, which may limit the applicability of the findings.

## 

## 2nd Limitation

One limitation of the paper is that it does not compare the RAID storage-based DFS system with other alternative storage solutions. The paper focuses on tape storage as the main point of comparison. A broader comparison with other disk-based archival storage solutions or cloud-based storage systems could provide a more comprehensive understanding of the system's performance.

# Synthesis

After analyzing the article, it appears that the ideas presented could be beneficial to organizations or businesses that require large-scale storage systems. The RAID storage-based DFS system has some advantages in terms of bandwidth, IOPS, and latency, making it ideal for situations where high-performance storage is required. The study also emphasizes the importance of workload analysis and individual assessment when configuring optimal storage setups. In the future, there may be more research on optimizing the system's performance, exploring different file systems and layouts, and conducting comparative studies with other systems.

# Future Scope of the Article

The article lays the foundation for further research and development in the field of distributed file systems. Future studies could focus on optimizing the performance of DFS in different storage environments, exploring new technologies and architectures for distributed file systems, and addressing the challenges associated with scalability, latency, and cost-effectiveness.