SIMATS ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES,

CHENNAI– 602105

TITLE

Fault-tolerant distributed file system

A CAPSTONE PROJECT REPORT

Submitted to

SAVEETHA SCHOOL OF ENGINEERING

Operating System for Process Scheduling

By

MANICHNADAN REDDY A

(192211741)

NEELIMA B

(192111296)

PREM KRISHNA B

(192210064)

Supervisor

Dr G. Mary Valentina

Abstract:
In this report, we delve into the intricate domain of fault-tolerant distributed systems, particularly focusing on their implementation in Python. We address the pressing need for robustness in modern computing infrastructures by designing and evaluating fault-tolerant mechanisms within distributed environments. Our study encompasses a detailed exploration of fault detection, recovery, and resilience strategies, culminating in the development of a fault-tolerant distributed system. Overall, FTDFS provides a robust and resilient platform for storing and accessing data in distributed environments, ensuring high availability and reliability even in the presence of failures and faults.

Introduction:
Operating systems play a pivotal role in managing resources and orchestrating tasks across distributed computing environments. However, ensuring uninterrupted service delivery in the face of hardware failures, network partitions, and software errors poses significant challenges. Our project endeavours to overcome these obstacles by engineering a fault-tolerant distributed system that seamlessly adapts to adverse conditions, thereby bolstering the reliability and availability of critical services. In the world of operating systems, a fault-tolerant distributed file system plays a crucial role in ensuring data reliability and availability across a network of interconnected computers. Unlike traditional file systems, which are often centralized and susceptible to single points of failure, a fault-tolerant distributed file system is designed to withstand various failures, including hardware malfunctions, network disruptions, and even system crashes. At its core, this type of file system leverages redundancy and replication techniques to store multiple copies of data across different nodes within the network. By doing so, it not only enhances data durability but also facilitates seamless access to files, even in the event of node failures or network partitions.

Literature Review:
GFS is a distributed file system developed by Google to provide reliable and scalable storage for large-scale distributed applications. It is designed to handle failures gracefully and ensure high availability and data integrity by Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung.

HDFS is the primary distributed storage system used by the Hadoop framework. It is designed to store large volumes of data reliably across commodity hardware, with built-in fault tolerance mechanisms to handle failures by Doug Cutting and Mike Cafarella.

Ceph is a distributed storage platform that provides object, block, and file storage in a single unified system. It is designed to be fault-tolerant by replicating data across multiple nodes and automatically recovering from hardware failures by Sage Weil, Gregory Farnum.

Gantt Chart:

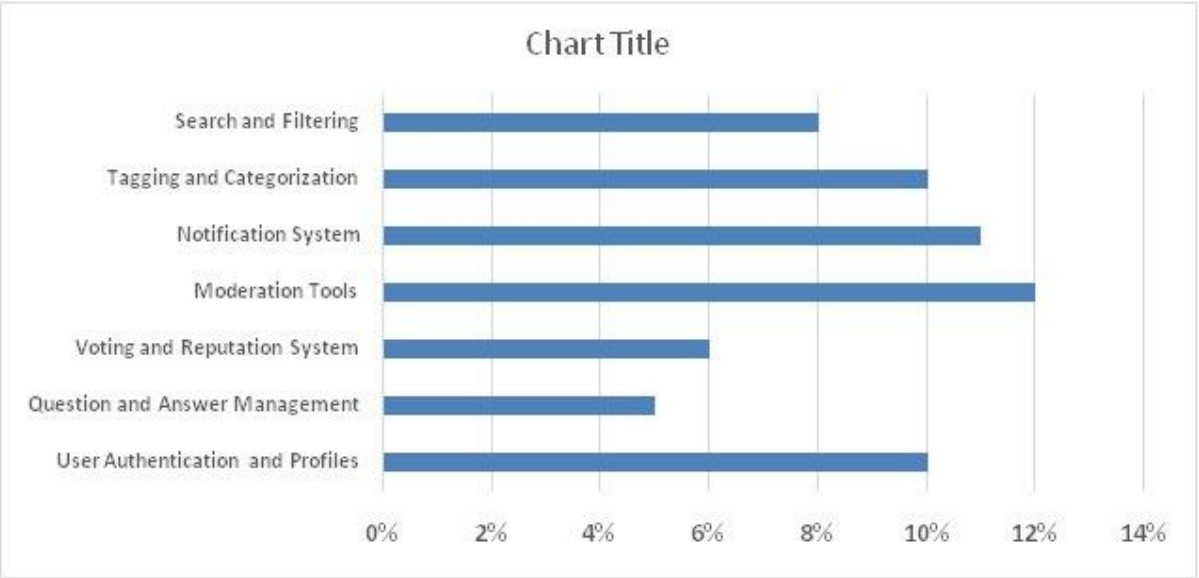
Chart Title
Search and Filtering
Tagging and Categorization
Notification System
Moderation Tools
Voting and Reputation System
Question and Answer Management
User Authentication and Profiles
0%
2%
4%
6%
8%
10%
12%
14%