Review on Distributed File System

Presented by Jiawei Xu

Abstract - Distributed File Systems (DFS) have emerged as a pivotal component in the realm of distributed computing, serving as a bedrock for data storage and accessibility across networked computers. This review delves into the exploration and analysis of the key facets of Distributed File Systems, their development, characteristics and functionalities. Fundamental elements such as data replication, consistency, fault tolerance, scalability, and security are systematically introduced, shedding light on their unique attributes and significance in DFS. The study also takes into account diverse DFS architectures such as Google's GFS, Hadoop's HDFS and Microsoft's DFS, offering a comprehensive analysis to compare their respective advantages and limitations in different scenarios. Augmenting this discourse, the paper introduces Alluxio—an open source virtual distributed file system (VDFS) that stands as a major milestone in the DFS landscape. Alluxio pioneers a data abstraction layer for compute frameworks, thus capacitating applications to interconnect with multiple storage infrastructures via a consolidated interface. This feature enhances efficiencies by offering a unified view and faster data access capabilities. This review presents an amalgamation of basic principles, diversified architectures, and recent advancements in Distributed File Systems, aiming to provide readers a holistic understanding and appreciation of this rapidly evolving field.