

Ceph: A Scalable, High-Performance Distributed File System

Summary:

The objective of this paper is to introduce Ceph, a distributed file system. As its name suggests, which is an abbreviation of "cephalopod", Ceph employs highly parallel and efficient behavior in data management. Ceph addresses performance, scalability, and reliability issues in storage systems by occupying a unique point in the design space. Ceph boosts the independence between data and metadata management by supplanting assignment tables with the CRUSH mechanism. Here, CRUSH utilizes a pseudo-random information distribution function intended for heterogeneous and dynamic groups of non-stable object storage devices (OSDs) to manifest assumptions on shedding design like allocation lists. RADOS leverages intelligent OSDs to manage data replication, failure detection and recovery, low-level disk allocation, scheduling, and data migration without encumbering any central server(s). EBOFS provides more appropriate semantics and superior performance by addressing the specific workloads and interface requirements present in Ceph. A dynamic distributed metadata cluster provides extremely efficient metadata management and seamlessly adapts to a wide range of general purpose and scientific computing file system workloads.

Strengths, Weaknesses, and Later On:

Through intensive experiments, it is proved that Ceph provides excellent performance, reliability, and scalability, but there are several elements that are not implemented yet. Such elements are, at the time the paper is proposed, planned to be implemented in the future. Later on, Ceph is extended as a complete platform for distributed filesystem management.