Software Defined Storage

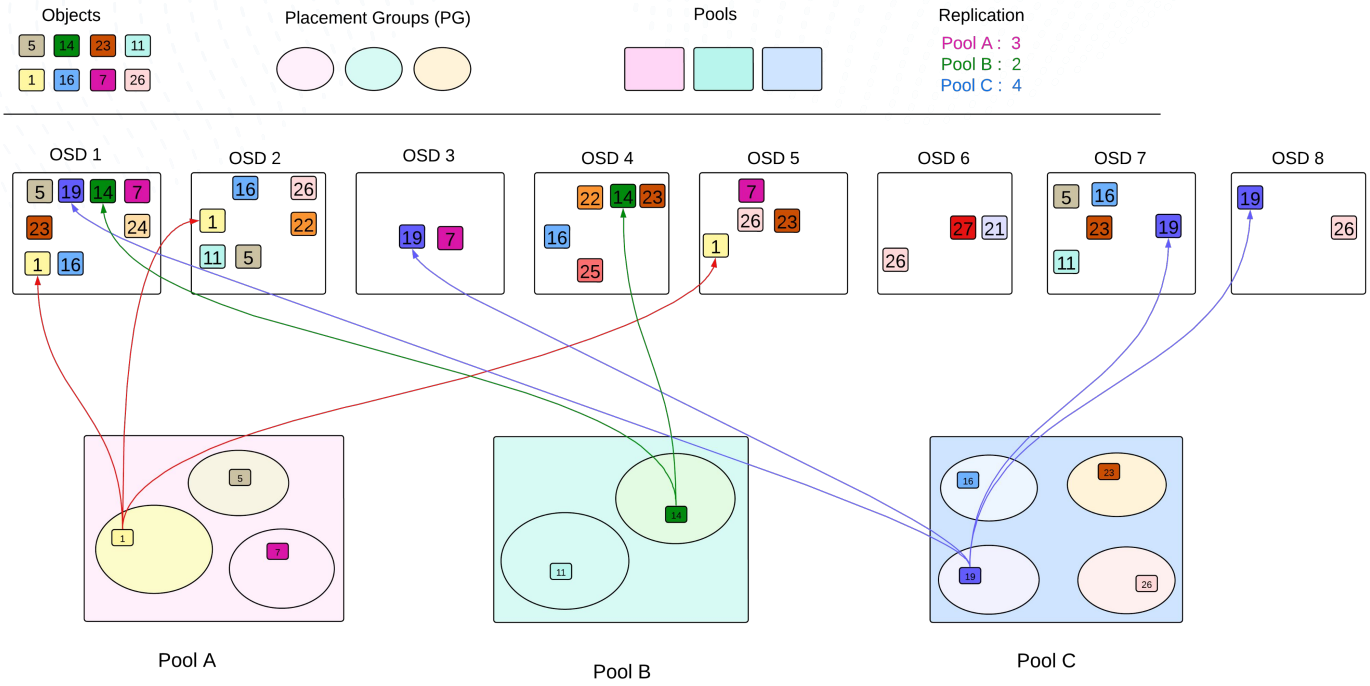
XLOUD TECHNOLOGIES

Ceph is an open-source storage technology designed to provide file, block, and object storage in a single system. Its primary objective is to achieve self-healing and self-management capabilities, thereby minimizing administrative overheads. It also aims for completely distributed operations without a single point of failure and is scalable to the Zetabyte level. Ceph software also runs on commodity hardware and replicates data to make it fault-tolerant.

ceph

ZB-Scalability **Asynchronous Replication** **High Availability** **Unified Storage** **Disaster Recovery**

Mitigating Risks **Data Redundancy** **Fault Tolerance** **Software Defined Storage** **Hardware Independent**

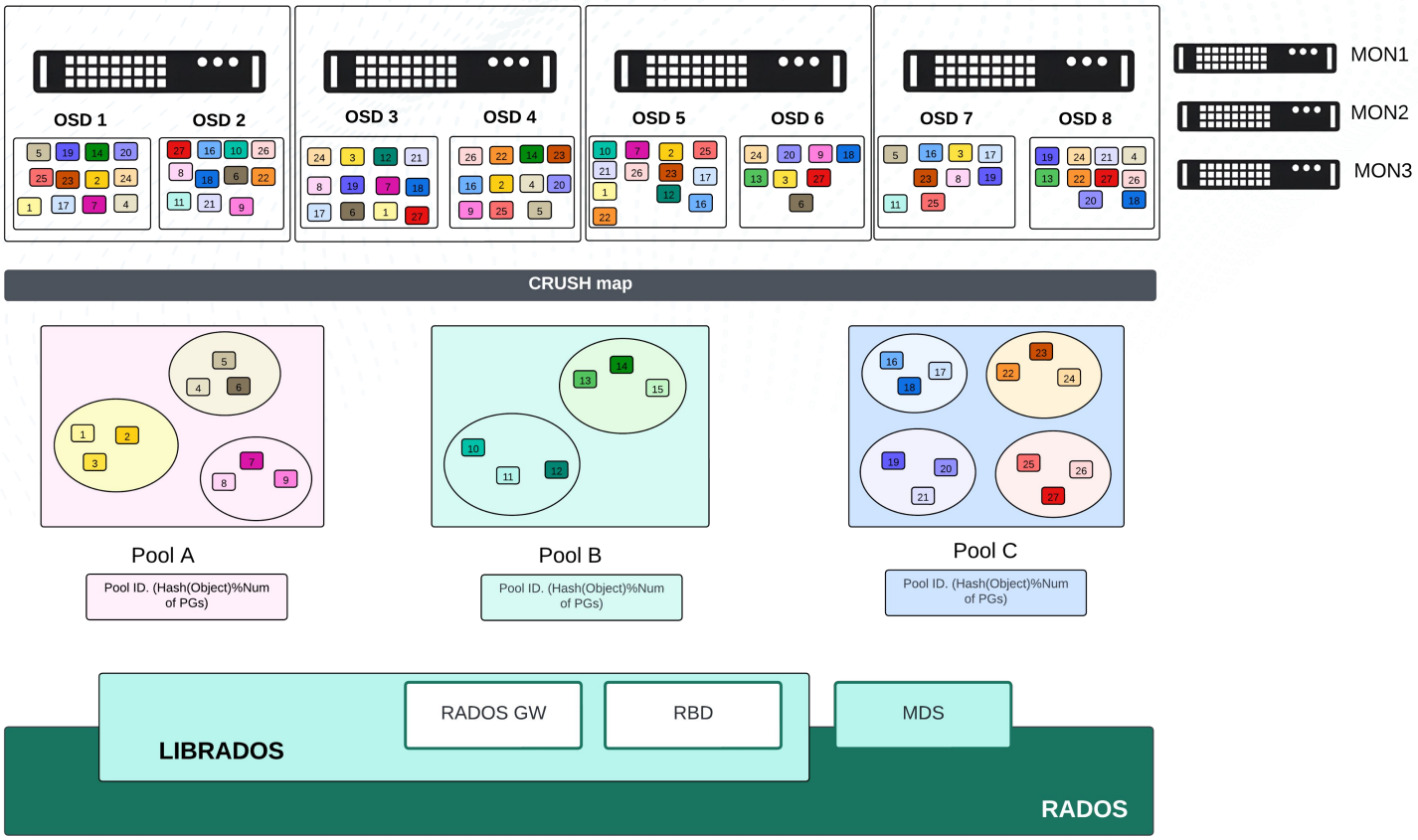


Components of Ceph

- **Objects:** Objects are the smallest unit of data storage in Ceph cluster. Each object has a unique identifier and can range from a few bytes to several terabytes in size. Objects are stored in pools and are mapped to PG, and these Objects and their copies are spread on different OSDs.
- **Pools:** Pools are logical groupings of objects in Ceph. These pools are made up of PG (Placement Groups). They serve as containers for storing data (objects) with similar characteristics or requirements. Each pool has its own set of rules and policies for data replication, placement, and management.
- **Placement Groups (PGs):** PGs are subdivisions of pools and are responsible for data distribution and redundancy. Each PG contains a subset of objects from its parent pool. These PG containing objects are spread across multiple OSD to improve reliability. The crush algorithm determines how Ceph distributes data across OSDs (Object Storage Daemons).
- **Cluster Setup (OSD, MONs):**
  1. Object Storage Daemons (OSDs): OSDs are responsible for storing and managing data objects in Ceph. They handle tasks such as data replication, recovery, and rebalancing across the cluster. OSDs run on storage nodes and communicate with each other to ensure data integrity and availability.
  2. Monitors (MONs): MONs maintain cluster state and configuration information. They keep track of OSD status, cluster membership, and monitor health. MONs also facilitate communication between OSDs and clients, helping clients locate data within the cluster.
- **CRUSH Map:** The CRUSH (**Controlled Replication Under Scalable Hashing**) algorithm is used by Ceph to determine data placement and distribution across OSDs. The CRUSH Map defines the hierarchy of the cluster and dictates how data is mapped to OSDs based on their weight and capabilities. By using CRUSH, Ceph ensures that data is evenly distributed and fault-tolerant even as the cluster scales.

Key Points

- Ceph storage cluster can host more than one Pools.
- Each pool has multiple Placement Groups. More the PG, better your cluster performance, more reliable your setup.
- A PG contains multiple Objects.
- A PG is spread on multiple OSD, i.e. Objects are spread across OSD. The first OSD mapped to PG will be its primary OSD and the other ODS's of same PG will be its secondary OSD.



Mirroring and Data Backup in Ceph

Mirroring stands as a pivotal feature for ensuring data redundancy and high availability across distributed environments. This functionality allows for the continuous replication of data between two separate Ceph clusters, typically situated in different geographical locations. By synchronizing data between primary and secondary clusters, mirroring ensures that critical data remains accessible and redundant, even in the event of primary cluster failures. This asynchronous replication process minimizes performance impact on the primary cluster while providing a reliable failover mechanism for seamless business continuity.

Data backup mechanisms provide additional layers of data redundancy and disaster recovery capabilities ensuring data integrity and availability by implementing backups of the entire Ceph cluster. This involves creating copies of the cluster's data, configurations, and metadata to safeguard against data loss and corruption. These backups serve as crucial safeguards, particularly in scenarios where mirroring may not be sufficient, such as catastrophic failures or widespread data corruption. By regularly backing up the entire Ceph node, organizations can mitigate risks and ensure comprehensive data protection and disaster recovery preparedness.