Apache Ozone Recon

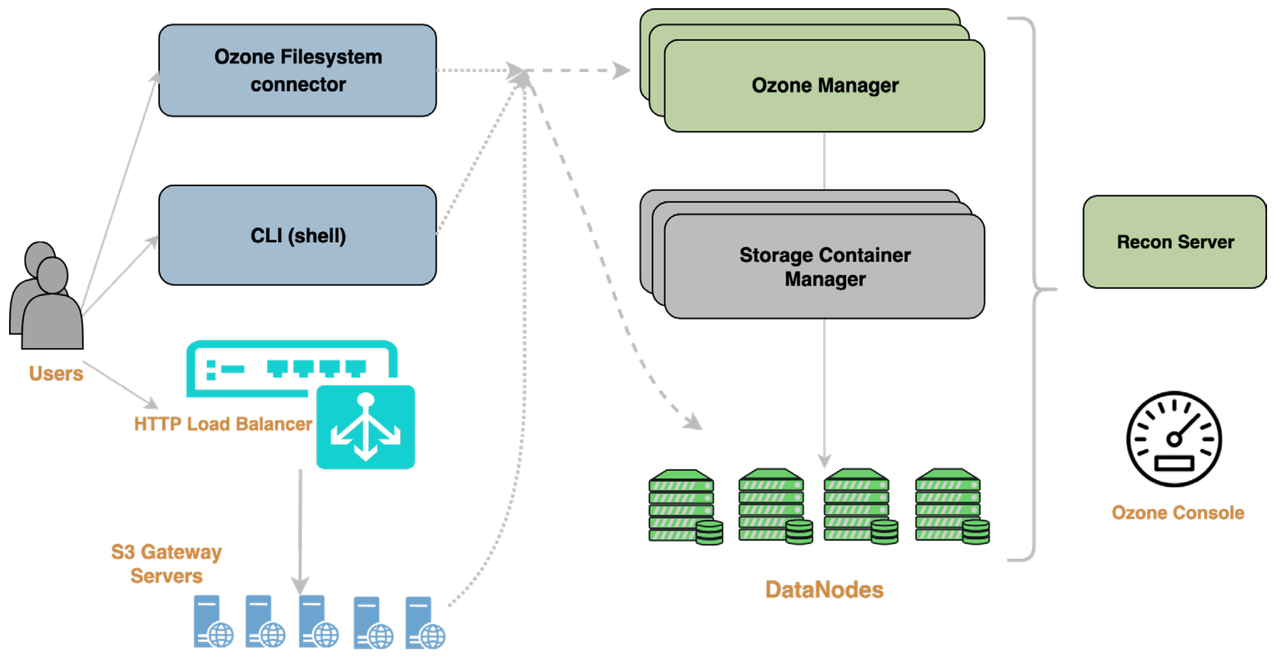
CLOUDERA

Apache Ozone

Bring cloud-native storage into your datacenter

CLOUDERA

- S3 protocol
- Hadoop FS
- CSI



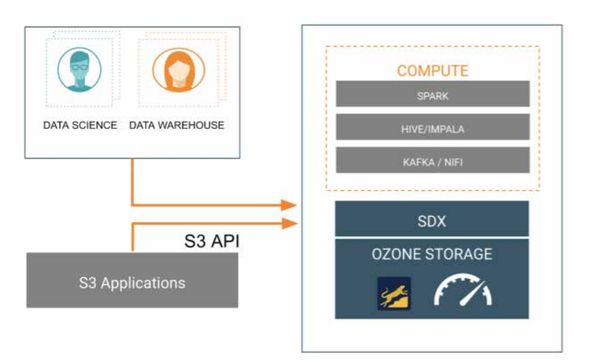
Salient Features Of Ozone

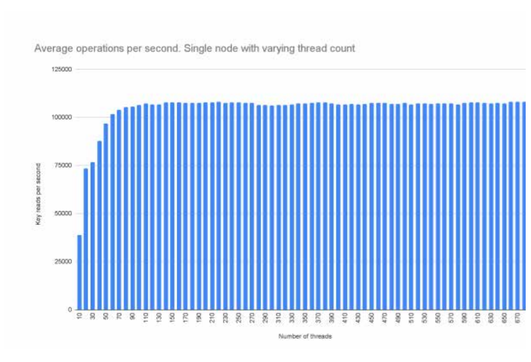
- Ozone is a scalable, redundant, and distributed object store for Hadoop
- Ozone can scale up to billions of objects and work effectively in containerized environments like Yarn or Kubernetes.
- Provides strong consistency and provides the benefits of traditional HDFS and S3 Object Store
- Built-in security: Kerberos authentication, pluggable authorizer, encryption
- Scale to 1000's of nodes with dense storage configurations
- Reduce cost per TB using commodity hardware
- Erasure Coding with 50% Storage Space Savings
- Quick Snapshot capability and avoid extra storage for snapshots
- Ozone can now support around 105k read operations per second

Recon (Monitoring and Observability for Ozone)

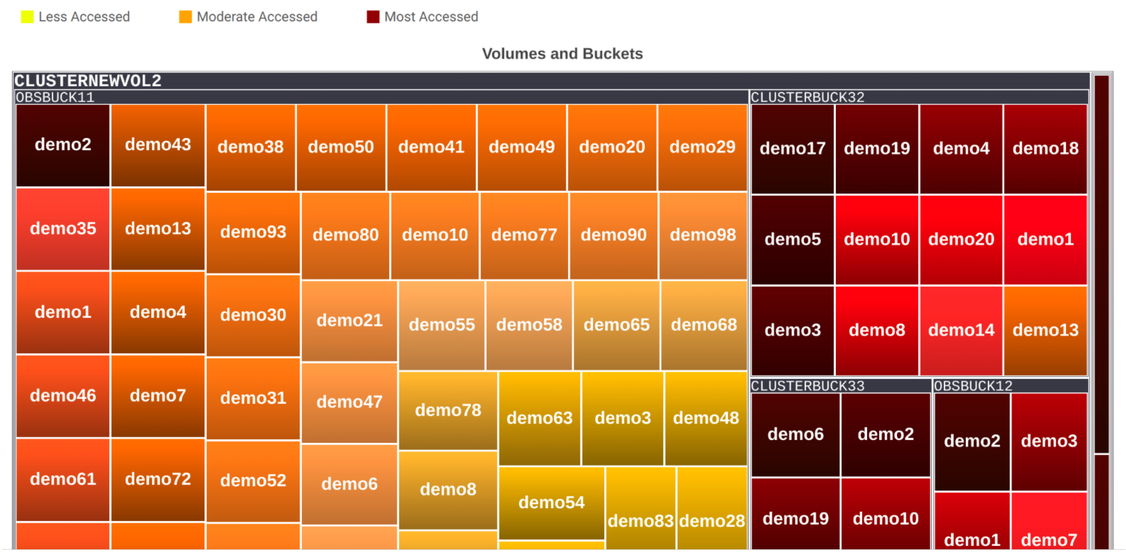
Recon acts as a passive node for Ozone and offloads the responsibility to monitor and analyze the health of the Ozone system. This provides lightweight synchronization from ozone services and datanodes and efficiently provides the capability.

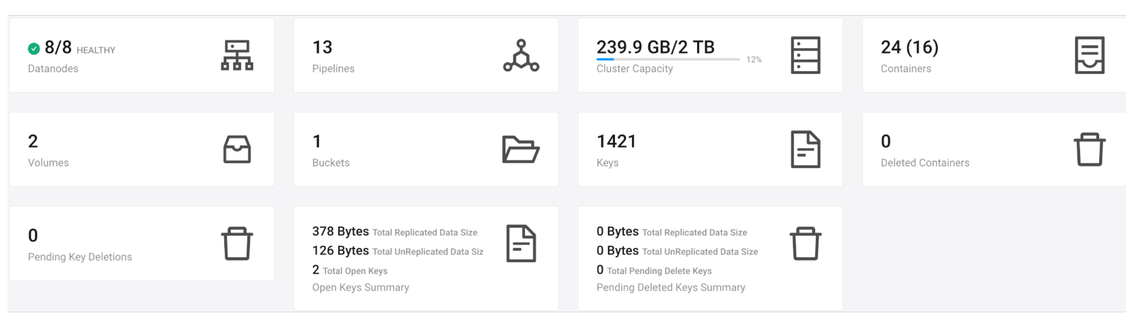
- Cluster Overview - Number of nodes, storage capacity, utilization, pipelines information.
- Node & Container Metrics - Monitors individual datanodes, reports their health and operational status, distribution of data across nodes and containers.
- Report data distribution: over-replicated containers, under-replicated containers, mis-replication of containers (not following placement and rack policy), orphan and missing containers
- Container size and file size distribution chart for all volumes and buckets.
- OM Insights such as open keys (client writing data), data pending to be reclaimed and free up space.
- Disk Usage calculations of all types of volumes (including S3 volumes) and all bucket types.
- HeatMap shows most heated volumes, buckets, keys/files.

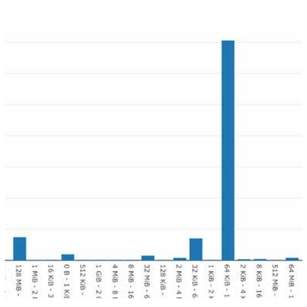




Recon And Prometheus/Grafana







Recon can integrate with any Prometheus instance configured to collect metrics and can display useful information in Recon UI in Datanodes and Pipelines pages.