

# Object Storage

Erfahrung mit Anwendungsfällen

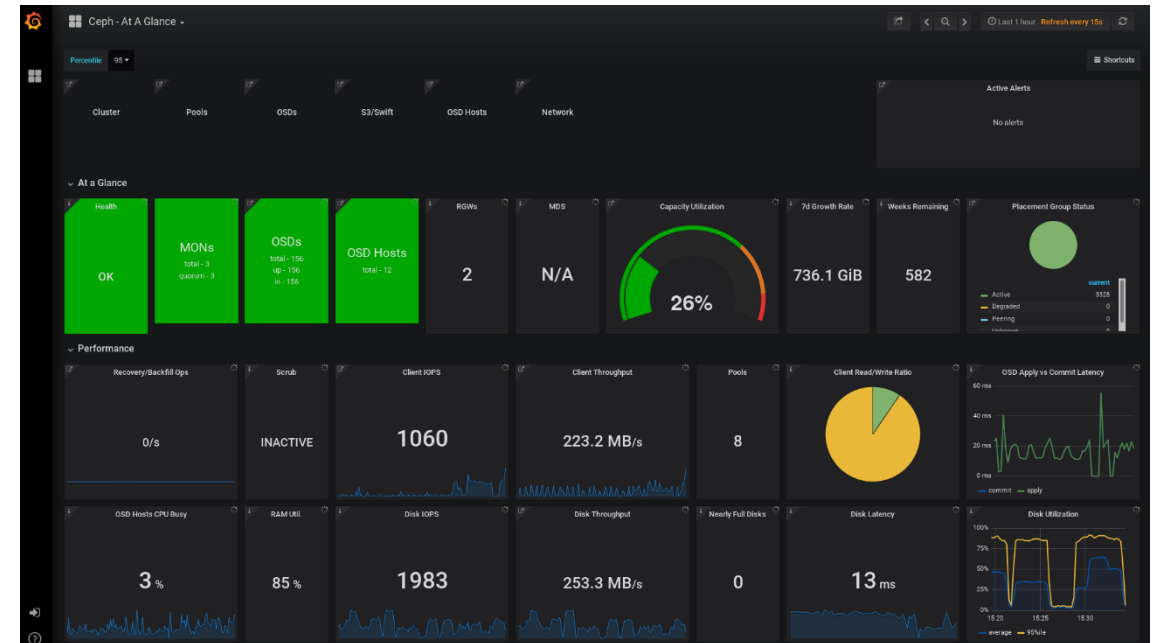
# Object Storage - Systemübersicht



Quelle: mapstyle.withgoogle.com

# Object-Storage - Systemeigenschaften

- Systemkomponenten
  - je 3 Stageserver an 4 Standorten
  - 2 S3-Gateways
  - 3 virtualisierte MONs
- Software
  - Ceph (Version 12.2.8 luminous) von RedHat
  - Inkl. Support
- Eigenschaften
  - Inbetriebnahme Anfang 2018
  - Nutzbare Kapazität ca. 250TB
  - Erasure Coding über alle Standorte
  - Ausfall eines Standortes tolerierbar
  - Maximale Objektgröße aktuell 5TB
  - Maximaler gemessener S3-Datendurchsatz 400MB/s



# Object Storage – Nutzersicht

---

- Etablierte Storage-Merkmale
  - Kapazität (Gesamtkapazität, maximale Größe einzelner Elemente, Anzahl Elemente)
  - Performance (Bandbreite, Latenz, OP/s)
  - Zugriffsprotokoll
  - Berechtigungsstrukturen (POSIX)
  - Backup-Strategie
- Merkmale Object Storage
  - Objekte
    - Operationen beschränkt auf Schreiben, Lesen, Löschen
    - Access-Listen (Zugriffsrechte pro Objekt)
  - Strukturelement Bucket
    - Bucket-Policy (Definition erlaubter Operationen)
    - Steuerung von Zugriffsrechten auf Bucketebene
    - Credentials pro Bucket
  - Backupless-Konzept (Bucket-Replication)
  - Emulation von Zugriffsprotokollen

# Object Storage - Anwendungsfälle

---

- Einsatz als temporäres Backup
  - Speicherung kompletter virtualisierte Server während der SAP-Umstellung im Wochenrhythmus
  - Ablage von .vmdk-Dateien mittels S3-Client
- Einbettung in Anwendung
  - Nutzung im Projekt Internet Observatory (<https://netray.io>), Speicherung umfanglicher Webscans
  - Nutzung eines Python-Moduls zur Speicherung der Daten mittels S3 aus der Anwendung heraus
- Einbettung in Webanwendung
  - Nutzung im Cloudservice Gigamove 2.0 zur Verteilung einzelner, großer Dateien
  - Upload erfolgt per https und S3-Gateway (CORS) transparent für die Nutzenden
- Speicherung der Daten von Großgeräten
  - Ablage von Elektronenmikroskopie-Aufnahmen
  - Ablage von Daten mittels S3-Client, Skripte
  - Herausforderungen
    - Ablage vieler Millionen Dateien (Bucket-Obergrenze)
    - Objektgröße << 128KB
    - Auseinandersetzung mit dem Datenmanagement (FDM)

# Object Storage - Anwendungsfälle

---

- GitLab
  - „Sorglos“-Packet für Nutzende bezüglich Repo-Quota und Größe einzelner Dateien
  - Nutzung von LargeFileStorage (Git LFS) und Container-Registry
  - Herausforderungen
    - Umstellung auf Object Storage im laufenden Betrieb
    - Awareness auf Nutzendenseite
- Nutzung für den Service Backup/Restore
  - Nutzung mittels IBM-Spectrum Protect als Storage-Backend im Pilotbetrieb
  - Backup-Server nutzen den Object Storage als Cloud Container Storage-Pool
  - Herausforderung
    - Umstellung der TSM-Serverhardware entsprechend gestiegener Performance-Anforderungen
- Cloudservice Object Storage
  - Bereitstellung von Buckets mittels einer Weboberfläche
  - Eigenentwicklung eines entsprechenden Selfservice-Portals

# Object Storage - Anwendungsfälle

---

- Einsatz im HPC-Umfeld
  - (Weltweiter) Datentransfer
  - Life Science-Anwendungen
  - Jupyter Notebook
  - Herausforderungen
    - Erwartungshaltung Anwender
    - Identifikation Anwendungsszenarien
- Anwendungsfälle in Vorbereitung
  - Gensequenzierung
  - Bildbearbeitung
- Weitere?

**Vielen Dank  
für Ihre Aufmerksamkeit**