Parallelstore

Fully-managed DAOS Storage in Google Cloud

Dean Hildebrand, Technical Director, Office of the CTO
Barak Epstein, Production Manager, Storage
corwin Coburn, SWE, Storage
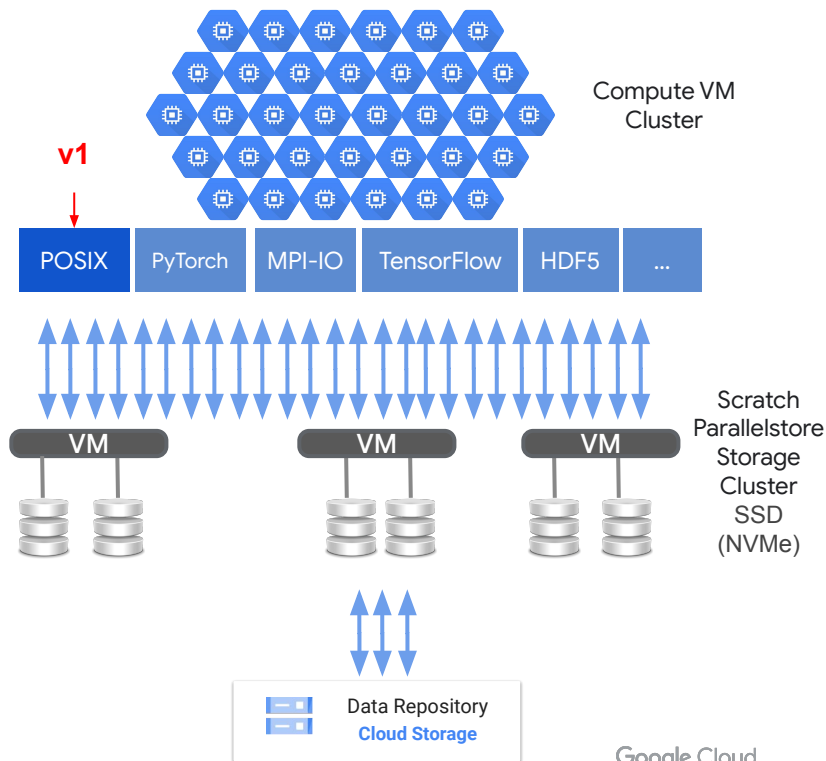
Google Cloud

Parallelstore
Now in Private Preview!

Accelerate HPC and AI workloads that require extreme scale and/or low latency I/O operations

🔑 Key Advantages

- **Scratch-space, GCS-integrated 1st party parallel filesystem**
- Unique open-source DAOS storage architecture improves performance over existing POSIX storage options
- Well-aligned to emerging patterns in AI workloads with **distributed metadata, extreme IOPS, and K/V architecture**
- Using RAM for MD/Small, NVMe for data
- EC2+1 with 6TB per storage server
- Demonstrated >1.2GB/s per TiB, >1.5M create/s with open-source on GCP (only limited by available HW)

Compute VM Cluster

v1

| POSIX | PyTorch | MPI-IO | TensorFlow | HDF5 | ... |

Scratch Parallelstore Storage Cluster SSD (NVMe)

VM   VM   VM

Data Repository
Cloud Storage

Google Cloud

Caching of Object Store

Combine DAOS and Object Storage (GCS)

Caching model with GCS has many advantages

- Hybrid deployments
- Lower overall costs
- Dynamically burst to compute located in any zone, region, or cloud
- Improve overall job TCO vs GCS directly

Proposal: Initially target ML Training Use Case

- High Value
- Avoid eviction
- Whole file caching of small files
- Focus on read-caching

daos
On-Premises

Google Cloud Platform

Parallelstore   GCS

Compute
Cluster

Google Cloud
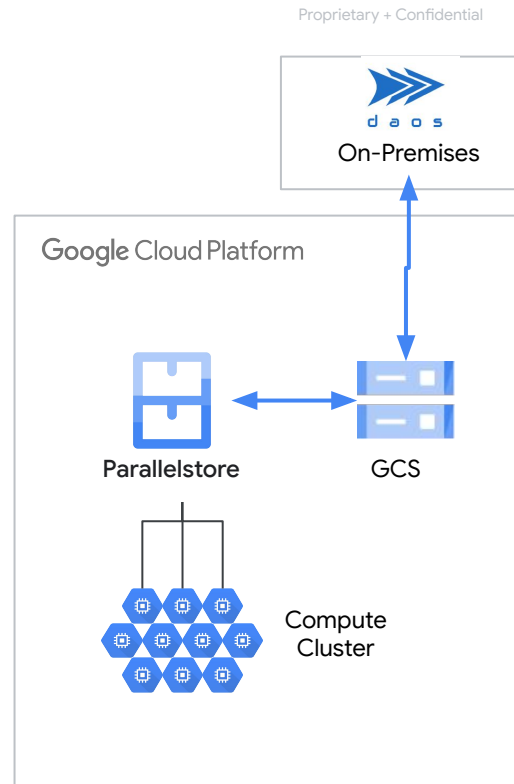
DAOS in Preview - Experience So Far

DAOS in Preview

Customer feedback is always a gift…

- Lots of demand for POSIX interface
  - POSIX is still the easy gateway prior to app modernization
  - dfuse config/tuning and multiple interception libraries confusing to navigate
  - Interception library not always applicable (e.g., Golang, static compile) and concerns some teams
  - Are FUSE and Interception Library capable of being the long term solution?

- HPC is not just Centos/Rocky
  - E.g., Ubuntu dominant in AI and Kubernetes

- First impressions are critical
  - FIO is the first DAOS experience for most users…

Google Cloud

DAOS in Production
Specific Focus Areas for Collaboration

- POSIX - Need more API support and predictable performance
  - Too many tradeoffs
  - Per-workload configurations nailed down
  - Potentially map configurations to ~3 target workloads?
- Memory is actually quite durable (given there is capacity for rebuild)
  - All-NVMe support will still improve resiliency in certain cases (e.g., zonal failure)
  - Caching all metadata in RAM is also expensive…
- Separation of user and admin responsibilities
  - DAOS gives lots of control to users…is it too much?
  - For example, is per-file EC/striping necessary?

DAOS In Production in the Cloud
Specific Focus Areas for Collaboration

- Users control the compute VMs (not administrators)
  - Need clients to keep operating through many versions of server upgrades
  - Client platform support
- While 'maintenance windows' are ok, they limit the cloud dream
  - Non-disruptive upgrades to deliver cloud expectations of high availability
- Config recommendations
  - Out-of-the-box experience is key
  - For example, we have customers with 10s of 1000s of clients, and assuming the application is manually sharing handles may be unrealistic in many scenarios
- Security
  - Cloud inherently provides security boundaries
  - But IAM integration will be critical sooner rather than later… (no Kerberos please)

Parallelstore

Features

How It Works

Common Uses

    AI/ML training

    Quantitative trading
    analysis

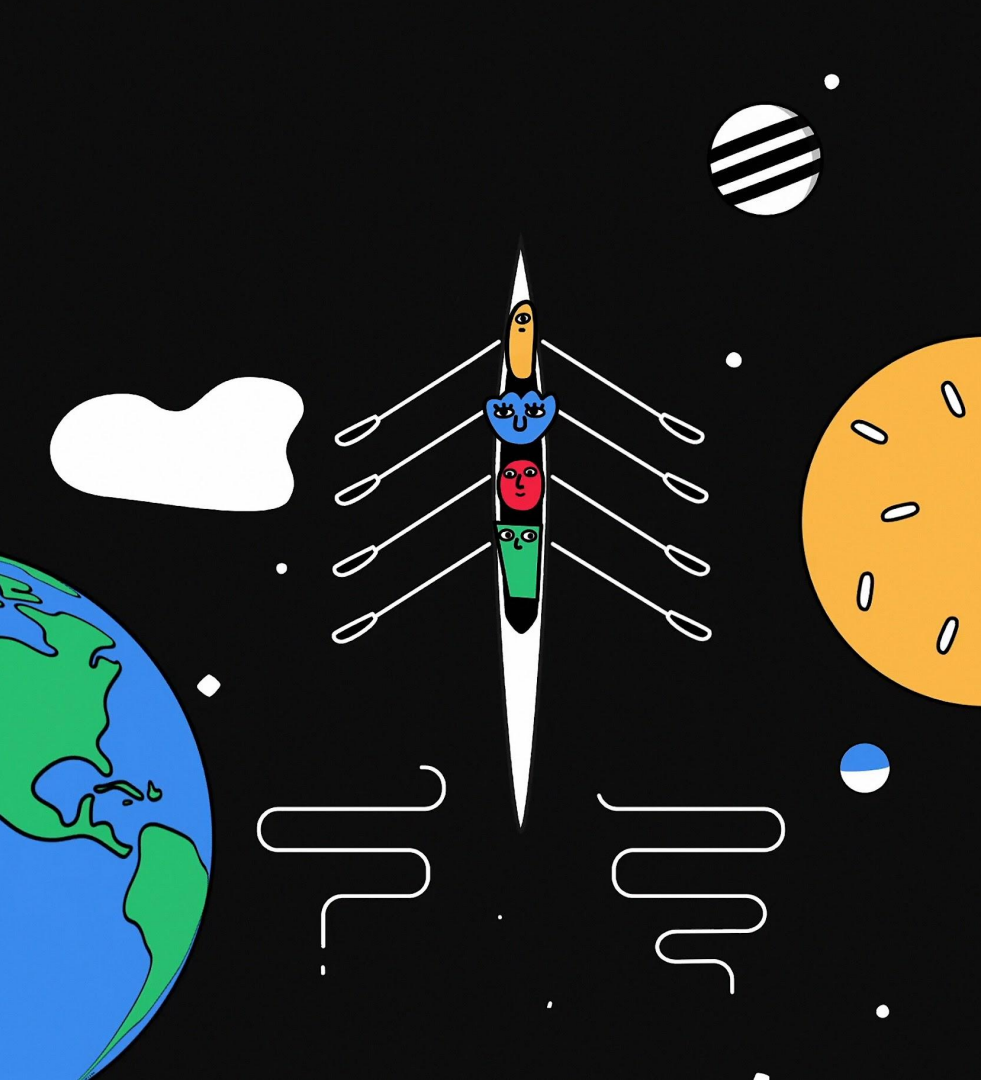   Computer-Aided
    Engineering (CAE)

Pricing

Join us at Next '23 from August 29-31 for the latest product launches, demos, trainings, ar

Parallelstore

High performance, managed parallel file service

Parallelstore is based on **Intel DAOS** and delivers up to 6.3x greater read throughput performance compared to competitive Lustre scratch offerings.

Get started

Thank you.

cloud.google.com/parallelstore

Google Cloud