Storage systems at GWDG

and the user’s problems to use them

Sebastian Krey, Hendrik Nolte

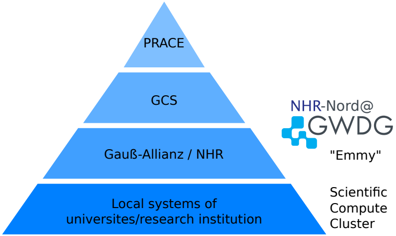
16.05.2024

Outline

About GWDG

- IT service center and data center operation for **University Göttingen** and **Max Planck Society** (MPG) since 1970
- Since 2021 working group “Computing” for HPC operation.
- Operating site of “North German Supercomputing Alliance” (HLRN) since 2018, since 2021 part of NHR
- HPC operating site for the “German Aerospace Center” (DLR) since 2022

HPC systems at GWDG



- Tier 2: **HLRN/NHR "Emmy"** Top500 #47 Nov. 2020, now #133
- Tier 2: **NHR "Grete"** Top500 #142 Nov. 2023, Green500 #22
- Tier 3: **Scientific Compute Cluster** (SCC)
- **"CARO"** for **DLR** Top500 #135 Nov. 2021, now #228
- AI Service Center KISSKI for critical infrastructure
- Several smaller systems for Max-Planck- and Uni Göttingen Inst.

HLRN-IV “Emmy”

Uni Göttingen/GWDG

- TOP 500: #47 in 2020-11 (5.95 PFlop/s), now #133,
  approx (inofficial) 4.56 GFlops/Watt (would have been #55 in 2020-11)
- phase 1 compute nodes (air cooled), EoL Q2/2024
  - 2x Intel Xeon Gold 6148 (SKL), 40 cores per node, 480 GB SSD
  - 432x 192 GB, 16x 768 GB
- phase 2 compute nodes (warm water DLC)
  - Intel Walker Pass System
  - 2x Intel Xeon Platinum 9242 (CLX-AP), 96 cores per node
  - 1100x 384 GB, 16x 768 GB, 2x 1536 GB
  - CoolIT DLC

NHR “Grete+”

- GPU cluster consisting of three procurement modules
- Performance optimized: 5.46 PFlop/s, TOP 500: #142 in 2023-11
- Energy optimized: 34.647 GFlop/Watt
  (Green500 #16, second in Germany, best at inauguration)
- 103 nodes
- 2 AMD Epyc Milan 7513
- 4 A100 GPUs per node (36 nodes with 40 GB)
- Dual rail Infiniband HDR interconnect
- Cluster local GPU Direct enabled storage
- CoolIT DLC
- Upcoming additional module with 25 4xH100 nodes

NHR “Emmy Phase 3”

- 411 nodes to replace Emmy Phase 1
- 2 Sapphire Rapids 48 core CPUs (Xeon Platinum 8468)
- Memory: 164x256GB, 16x1TB, 3x2TB, remaining 512GB
- Cornelis Omnipath 100G interconnect
- Connection to storage of other islands via routing
- CoolIT DLC with direct free cooling with outside air for residual heat
- Recently handed over to the users for production

DLR "CARO"

- Operated for the German Aerospace Center
- 1370 nodes with 2 AMD Epyc Rome 7702
- 3.46 PFlop/s, TOP 500 #135 in 2021-11, now #228
- 364 TB memory
- 24 Quadro RTX 5000 for visualization
- Infiniband HDR100 interconnect
- 8.4 PiB DDN Lustre (200 TiB SSDs)
- CoolIT DLC

Storage Systems: Current

- HOME/SW: 350 TiB DDN Gridscaler, EoL 07/24
- WORK MDC: DDN ExaScaler 5 EoL 07/24
  - Metadata SFA7700X
  - 8 PiB HDD 2x ES14KX
  - 113 TiB NVME 2x SFA200NV
- WORK RZGÖ: DDN ExaScaler 6 113 TiB NVME 2x ES400NVX
- HOME/SW/WORK KISSKI: VAST Data 500TiB NVME (1x dBox, 2x cBox)
- WORK SCC: 2.2 PiB BeeGFS based on DDN SFA7990 block storage
- HSM/Tape: Quantum StorNext HSM 3 PiB (EoL 01/25)

Storage concept for NHR/SCC/KISSKI

- Replace HDD based WORK storage with central Ceph instance
- Compute island specific high performance storage, all flash (Lustre, VAST or BeeGFS, DAOS maybe a candidate in the future)
- Unify HOME/SW to central VAST or GPFS (exported via NFS)
- HPC S3 object storage for “Cloud” workloads and easy data ingest/export with central S3 storage of infrastructure group and external parties

Ceph for HPC?

Common opinion:

- Are you insane?
- Ceph is slow, complex, unreliable,. . .
- Only TCP connections

On closer look:

- Ceph is reliable standard in cloud environments
- Some institutes use it successfully in HPC (e.g. CERN, IZUM)
- Ceph allows complete hardware vendor independence
- Hardware migrations in live operation, without user interaction
- Recent performance improvements show respectable performance (work from Clyso and Croit)

Storage Systems: New Coldstorage

Hardware:

- 53 Servers, 23 PB HDD, 3.5 PB NVME
- HDD Cluster with 45 Servers:
  - 24x 22TB HDD, 4x 7.68 NVME
  - 2x24 Core Sapphire Rapids CPUs, 512 GB memory
  - 2x25G Ethernet
- NVME Cluster with 8 Servers
  - 20x 15.36TB NVME
  - 2x32 Core Milan CPUs, 512GB memory
  - 100G Ethernet
- HDD cluster capacity optimized → Erasure Coding
- NVME cluster performance optimized → Replication
- Installation support from “Clyso”

A few performance numbers

Ceph test environment with 7 nodes, each with 2x8 Core Skylake, 192GB memory and 2x 2TB NVME as the full system not yet available.

- Benchmarking RBD, not yet CephFS and S3
- With replication up to 6 GiB/s writing and 15 GiB/s reading (4M IO size)
- About 50% of reached perf. with BeeGFS or MinIO S3 on this hardware
- Around 40k random IO/s with 4K IO size.
- EC much more difficult, only single node tests
- Small IO sizes unexplainable slow $\rightarrow$ debugging necessary
- 4MB IO sizes look good, up the 4.5 GiB/s reading, 2.5 GiB/s writing
- Performance highly dependendant of EC configuration, highest performance with ISA library, standard jerasure library 10-20% lower
- Comparison: Single node replication reading 4.5 GiB/s, writing 1.8 GiB/s

Why Storage Tiering?

- Users have different requirements depending on the type of data
  - Think of Software as compared to hot data
- The different storage systems differ in many attributes, e.g.
  - Size
  - Speed
  - Data Durability
  - Backups
  - Lifetime, e.g. only available during job runtime, certain TTL, etc.

An Example at GWDG

Project Origin	Name	Storage Kind	Storage Type	Clusters	Path	Disk Kind	Filesystem	Backed Up	Description
all	Project Directory	MAP	Filesystem	Emmy, Grete	`/projects/PROJECTPATH`	SSD	VAST NFS	yes+snapshot	Symlink farm pointing to all the data stores
NHR	NHR Archive	ARCHIVE	Filesystem	Emmy, Grete	`/perm/projects/PROJECT`	Tape	Stornext	yes	Archival storage (very robust, very slow)
NHR (legacy)	Legacy Project HOME	HOME	Filesystem	Emmy, Grete	`/home/projects/PROJECT`	HDD	GPFS	yes+snapshot	HOME storage for the project (robust, but slow and small)
NHR	Project HOME	HOME	Filesystem	Emmy, Grete	/mnt/ddn-gpfs/projects/PROJECT	HDD	GPFS	yes+snapshot	HOME storage for the project (robust, but slow and small)
NHR	Lustre Emmy HDD	SCRATCH	Filesystem	Emmy, Grete	`/mnt/lustre-emmy-hdd/projects/PROJECT`	HDD	Lustre	no	Large and reasonably fast storage optimized for Emmy
NHR	Lustre Emmy SSD	SCRATCH	Filesystem	Emmy, Grete	`/mnt/lustre-emmy-ssd/projects/PROJECT`	SSD	Lustre	no	Small and fast storage optimized for Emmy
NHR	Lustre Grete	SCRATCH	Filesystem	Grete	`/mnt/lustre-grete/projects/PROJECT`	SSD	Lustre	no	Small and fast storage optimized for Grete
NHR (legacy)	scratch-emmy	SCRATCH	Filesystem	Emmy, Grete	`/scratch-emmy/projects/PROJECT`	HDD	Lustre	no	Large and reasonably fast storage optimized for Emmy
NHR (legacy)	scratch-grete	SCRATCH	Filesystem	Grete	`/scratch-grete/projects/PROJECT`	SSD	Lustre	no	Small and fast storage optimized for Grete

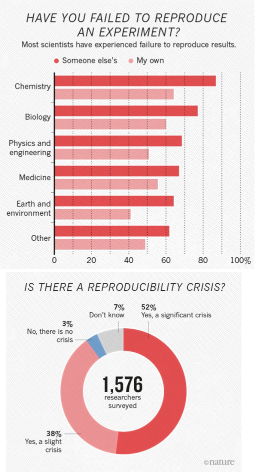
- 8 storage spaces in 4 tiers, and the local `tmpfs` and SSD’s are even neglected

Resulting Problem

- Users are overwhelmed and do wrong data placement
  - Hot data sits on cold storage
  - Standard datasets sit on expensive backed up storages
  - Important results are on fragile storage
  - The wrong storage system for the wrong cluster island is used
    - GWDG might be an edge case here, but also think of a Dragonfly Topology
- Many storage tiers quickly lead to a loss of oversight
  - Data is not cleaned up
  - Data is not reproducible, unclear where it belongs to
  - Data loss, hard to find

Reproducibility Crisis

- There is a general reproducibility crisis
- For HPC one needs to distinguish
  - Deterministic execution of a job
  - Proper provenance auditing
- Deterministic execution is hard
- Proper lineage recording shouldn't be
  - Due to insufficient data management
- Specific HPC tools are often not used
  - e.g. *PASS*, *LPS*, or *ReproZIP*
  - Domains developed own standards
  - Integrated into remote DMS



Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a

Folder Mapping

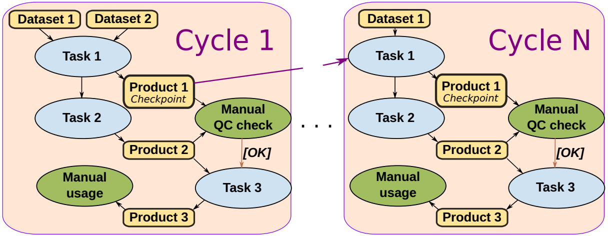
- As an immediate solution we introduced a *Map* data store
  - `/projects/PARENTPROJECTS/PROJECT`
- This gets also symlinked within each user home
  - `$HOME/.projects`
- These folders point to the same or different storage tiers
  - Which do not necessarily replicate the structure
- There are compute and data projects
  - Data projects can be used for archiving and to organize access rights

Challenges and Characteristics Data-Intense Projects

- Data-intense Projects have a few characteristics:
  - A workflow is being executed rather then one large, expensive single task
    - Data-Parallelism vs. Task-Parallelism
  - Data-Sharing with different people, cp. previous roles
    - And doing so with different permissions
  - Data has a lifecycle, which requires a data flow between different tiers
  - Provenance is key, since single workflow is executed multiple times

Processing of Data-Intense Projects

- We plan that users with large data, storage and IO requirements can be filtered out by their applications
- We contact these users to work with them
  - They provide is with an overarching workflow description
  - We would help them to map these on our
    - Storage tiers
    - User roles
    - Compute infrastructure



Summary

- Different storage systems are needed to handle requirements regarding
  - Performance
  - Cost
  - Data durability
- Users need guidance
- Some users have to be forced into doing proper data management
- Admins should try to simplify the storage landscape as much as possible