GWDG
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Scalable HPC node deployment with Warewulf and S3 storage

Sebastian Krey, Timon Vogt

24.04.2024

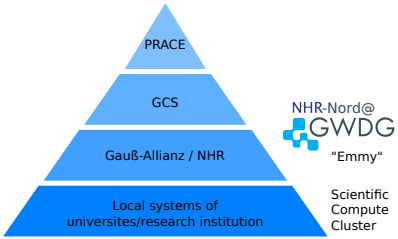
ZKI Supercomputing Ilmenau

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 Institutes

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 16-25 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
- Acceptance completed, production usage from May 2nd.

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

Hardware Summary

- In total around 10 HPC systems/compute islands in operation
- Hardware from currently 6 different vendors
- Vendors of larger systems provide their own management stack
- Compute resources have to be provided to several independent user groups
- Every user group and funding agency has different accounting requirements
- Increasing requirements regarding operational efficiency (especially energy)
- Increasing reporting requirements
- Human resources limited

Vendor independence and operational unification

- Replace system specific deployment system with own modified Warewulf instance (mostly completed)
- Unified central monitoring of all systems (in progress)
- Geo redundancy for deployment and management systems (in progress)
- Unify and simplify user and compute project application and management in self developed project management tool (completion in Q2 2024)
  - Unified batch system
  - Unified HPC Home storage
  - Compute and storage ressource accounting

Deployment system

- Learning a vendor dependent deployment system takes time
- Long term usage of Bright Cluster Manager $\rightarrow$ increasing costs
- Evaluation of alternatives based on several feature requirements
- Migration effort from Bright node images should be minimized
- Customizability to own requirements
- Candidates: xCat, Qlustar, Foreman, BlueBanquise, Warewulf

Warewulf

- Simple cluster management and deployment system
- Open source and Go based
- Provides DHCP, PXE
- Text based configuration files
- Templating engine for indivudalization of node cnofiguration
- Shell script based installer environment $\rightarrow$ easy to customize
- Default diskless deployment based on OCI container

Warewulf boot phase

- DHCP/PXE request
- Load iPXE via `tftp`
- Download kernel, modules and container via HTTP
- Execute kernel and container
- Run scripts from system/wwinit overlay to install and configure node

$\rightarrow$ Single HTTP server bottleneck for large deployments

Deployment scaleout

- Warewulf documentation indicates establishment of Warewulf slaves for horizontal scaling, but description very vague
- DHCP/PXE servers unique in a VLAN $\rightarrow$ no parallelization possible
- iPXE download via tftp very small $\rightarrow$ little benefit of scaleout
- HTTP downloads of large containers are major bottleneck

$\rightarrow$ Additional fast HTTP servers provides most benefit

Problem:
How to replicate and update the directory structure of the Warewulf HTTP server

S3 storage for node deployment

- S3 storage is HTTP(S) based
- Anonymous S3 access via ACL configurable → GET is a simple HTTP donwload
- VAST storage provides unified namespace between S3 and NFS
- Mount a unified share on Warewulf
- Create hardlink based Warewulf drectory infrastructure on NFS share
- Modify download URL in iPXE script to reference S3 bucket

→ Performance critical HTTP download relocated to fast and scalable storage

First results

Currently two compute islands deployed with this method

Small island with 12 GPU nodes

Node reboot time independent of number of parallel reboots
$\rightarrow$ perfect linear scaling up to 12G

Larger island with limited per rack uplink bandwidth

Linear scaling up to uplink limit (28 nodes $\rightarrow$ 10G)

Next week: Full parallel bootup of Emmy P3 after maintenance to test limits

Outlook: Future plans

- VAST-link-farm: Creating/Updating the required hard-links directly from within a `wwctl configure` by patching warewulf (instead of a script)
- HTTPS: iPXE does support securing its downloads via HTTPS by supplying a trusted certificate right at build-time of the iPXE binary.
  - Build the certificate of the VAST storage system directly into the iPXE binary
  - Secure traffic via HTTPS.
- Kubernetes-based deployment:
  - Currently Warewulf as `systemd-nspawn` container on two admin nodes installed in the same rack as active/passive pair.
  - Future: georedundant Kubernetes cluster across 3 data centres of the GWDG.
  - Storage of the Warewulf container would be replicated using Project Longhorn
  - Ansible based deployment of the entire setup allowing easy disaster recovery.