GöHPCoffee

HPC/GPU resources and hosting/housing options

Christian Köhler
christian.koehler@gwdg.de

14. September 2022

hpc@gwdg.de

Systems

Applications

Current Procurements

Sharing/Hosting/Housing

Systems

HPC systems operated by GWDG

LGWDG

- Scientific Compute Cluster (SCC)
  - Replacement Procurement for the Delta cluster (d∗ nodes, Faßberg, from 2016) in preparation
  - Installation in the new data center
  - Continued operation of the Atos cluster (a∗ nodes, MDC, from 2020)
- DLR system ''CARO''
  - Cooperation between GWDG, University of Göttingen and German Aerospace Center (DLR)
  - System delivered and installed by NEC
  - in operation since CW21
  - GWDG: Installation, Operation and Consulting
- Göttingen Campus Institute for Dynamics of Biological Networks (CIDBN)
  - vendor SVA/DELL, currently test operation
- HLRN-IV system ''Emmy''

Mobile Data Center (MDC)
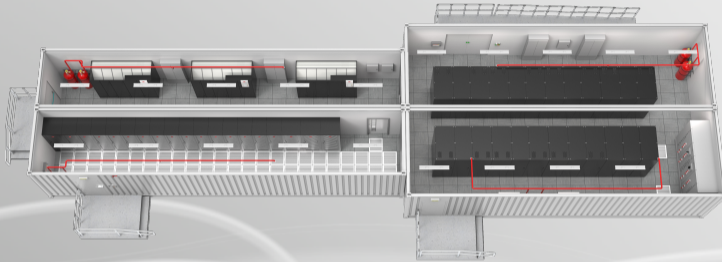
- left area: air-cooled systems (max 19 Racks)
  HLRN-IV phase 1, GWDG GPU nodes, storage, burst buffer

- right area: hot water-cooled systems (max 14 racks)
  HLRN-IV phase 2, GWDG CPU compute nodes

- already operational: DLR system CARO
- starting 10/2022: „Emmy" GPU expansion
- starting 2023: SCC expansion

Scientific Compute Cluster (SCC)

- several compute islands at these sites:
  - Faßberg (FAS) (until 2022)
    - → returning facilities to MPG
      - was hosting GWDG since 1970
  - Fernmeldezentrale (FMZ)
  - Modular Data Center (MDC)
  - Göttinger Rechenzentrum (RZGö) (starting 2023)
- 3 login nodes (reachable from GöNET or via VPN)
- integration of systems from various institutes (housing/hosting)
- central resource management (SLURM)

SCC Partitions
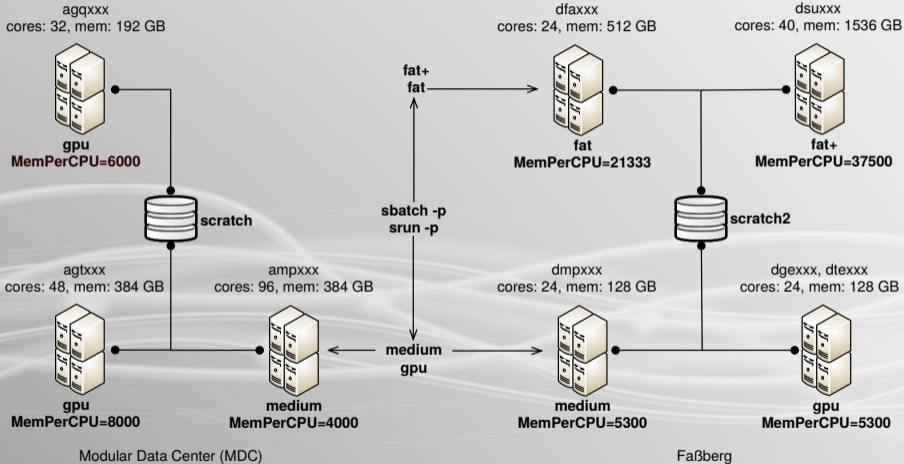
Scientific Compute Cluster
Replacement procurement 2020 - Overview

GWDG

- 96 standard nodes (2 Xeon Platinum 9242, 2x48 CPU cores, 384 GB memory)
    - → cf. ''Emmy'' Phase 2 standard nodes
- 2 GPU-Knoten with 8 Tesla V100/32
    - →2 Xeon Gold 6252, 2x24 CPU cores, 384 GB memory
- 12 GPU-Knoten with 4 Quadro RTX5000
    - →2 Xeon Gold 6242, 2x16 CPU cores, 192 GB memory
- 100 GBit/s Intel Omni-Path interconnect
- Storage: 2,15 PB /scratch
    - → BeeGFS, with a 100 TB SSD part
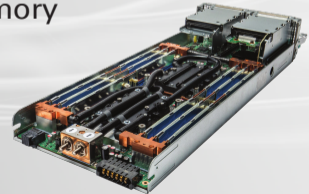
Scientific Compute Cluster
Replacement procurement 2020 - GPUs

GWDG

- 48x Quadro RTX5000 (Turing architecture)
  - TU104 GPU: 48 SMX → 3072 CUDA cores, 384 Tensor cores
  - 11,2/89,2 TFLOPs (SP/Tensor)
  - 16 GB VRAM (GDDR6)
  - + 48 RT cores for raytracing acceleration=handing off lots of intersection tests from CUDA cores/shaders)
- 16x Tesla V100/32 (PCIe version) (Volta architecture)
  - GV100 GPU: 80 SMX → 5120 CUDA cores, 640 Tensor cores
  - 14/7/112 TFLOPs (SP/DP/Tensor)
  - 32 GB VRAM (HBM2)
- from 2016: 32x GTX980, 14x GTX 1080, 20x Tesla K40m
- GPUS can be requested individually in the batch system (Slurm)

Applications

Typical Usage Scenarios

AI Research

- Labs with many students doing small-scale experiments
- Mid-scale development/experiments by researchers or students
- Large-scale/scalable execution of developed pipelines

HPC Research

- Small-scale / single-node testing of an application
- Labs with many students learning to use HPC
- Scalable/Multi-node execution of existing/developed application

GPU Usage Modes

- **Asynchronous batch jobs**
  - submitted jobs are executed later
  - depending on load, individual priority (cf. FairShare)
  - once running resources are granted exclusively
  - most resources, main form of GPU usage

- **Interactive batch jobs**
  - resources are requested in batch mode
  - live usage once jobs starts
  - exclusive usage, potentially long waiting time
  We partition GPUs for maximum sharing (into up to 7 "devices")

- **Shared resources**
  - partition with no exlusive usage (overprovisioning)
  - no waiting time, but no guaranteed performance
  - for testing and small-scale experiments
  - usage via Jupyter HPC →

Data Analytics Applications
Jupyter on SCC

GWDG

Spawner Options

Select a job profile:

GWDG HPC with own Container

Set your own Singularity container location (allowed characters: [a-zA-Z-Z.~-])

$HOME/jupyterhub-gwdg/jupyter.sif

Set the duration (in hours):

8

Set the number of cores:

10

Set the amount of memory (in GB):

32

Jupyter Notebook's Home directory

$HOME/jupyterhub-gwdg

Documentation

Spawn

- `jupyter-hpc.gwdg.de` spawns Jupyter on HPC
- supports IPython Parallel
- Users can choose resources, start individual Singularity Container
- Support for Jupyter on (shared) GPU nodes

jupyter

Data Analytics Applications
Apache Spark / Tensorflow

- Apache Spark
  - Spark clusters can be set up automatically
  - interaktive sessions (e.g. Scala)
  - Monitor cluster status via web interface
  - in preparation: Horovod on Spark
- TensorFlow
  - Integration in users' Python environment
  - Using GPU nodes with CUDA possbile

Overview & Documentation of HPC applications: https://hpc.gwdg.de

Current Procurements

SCC replacement procurement
Hardware specification

- Standard nodes:
    - 48 -- 64 CPU Cores, 256 GB memory (''2x dmp nodes'')
    - 2x 1 TB NVMe SSD
- High Mem nodes: 4x 1 TB, 1x 2 TB
- GPU nodes (like HLRN-IV expansion), i.e. each equipped with
    - 4x NVIDIA A100 GPU (SXM4, 80 GB HBM2 memory), 512 GB system memory
- Interconnect:
  InfiniBand (possibly RoCE (100 GBit/s), Cornelis OPA)
- SSD storage
- **D**irect **L**iquid **C**ooling

(expected configuration according to current bidding process)

NHR Procurements

- **GPU expansion for the NHR system "Emmy"**
  - 07/2021: Tender with expansion options
  - 09/2021: Acceptance of the bid by MEGWARE
  - 10/2022: Delivery and installation

- **Storage**
  - 12/2021: Delivery
  - 06/2022: Operation of SSD pool, burst buffer
  - 10/2022: Operation of GPU expansion flash storage

- **NVIDIA ARM DevKits → Future Technology Platform**
  - 10/2021: Requesting quotes from 3 vendors
  - 11/2021: Acceptance/order from MEGWARE
  - 05/2022: Delivery
  - 07/2022: Test operation

NHR Procurements
GPU expansion for the NHR system "Emmy"

GWDG

- Technical specification: 36 nodes (3 racks), each equipped with
  - 2x AMD Epyc 7513 CPU (32 "Milan" cores, Zen 3 microarch.)
  - 512 GB memory (DDR4, 3200 MHz)
  - 2x 1 TB NVMe SSD
  - 4x NVIDIA A100 GPU (SXM4, 40 GB HBM2 memory)
  - 2x Mellanox InfiniBand HCA (HDR)
- Installation at RZGö
- Connection to "Emmy" storage
  - Local flash storage (Atos/DDN) for the GPU cluster at RZGö
  - Link to existing WORK (at MDC)
  - LNet routers for connecting to Lustre (IB↔eth↔OPA)

NHR Procurements
Future Technology Platform (FTP)

GWDG

- Start of Future Technology Platform for testing cutting-edge HW
- Experimental applications on an alternative microarchitecture (ARM) and Security/Storage offloading (DPUs)
- Procurement of two NVIDIA ARM DevKits, each equipped with
  - 1x Ampere Altra Q80-30 (80 cores, ARM)
  - 512 GB memory (DDR4, 3200 MHz)
  - 2x NVIDIA A100 GPU (SXM4, 40 GB HBM2 memory)
  - 2x NVIDIA BlueField-2 E-Series DPU (200 GbE/HDR, 16 GB memory)
- Starting operation in late 2022

Sharing/Hosting/Housing

Getting access to the HPC systems

- **Scientific Compute Cluster (SCC)**
  - Requirements
    - GWDG Full Account (Nutzer UGOE/MPG)
    - Thesis projects: Apply for an account via supervisor's institute
    - Courses: temporarily usable guest accounts
  - Account activation: request via hpc@gwdg.de
  - Job prioritization via FairShare
- **NHR system "Emmy"**
  - Requirements: Research at German universities
  - Apply for an account at https://zulassung.hlrn.de
  - Starting from 300.000 Coreh per quarter
    - Project application for further compute time
    - review by scientific board (+techn. review by HPC sites)
    - sci. review can be omitted , if project has been granted by BMBF/DFG/EU/... (Whitelisting)
  - Compute time usable at NHR@ZIB ("Lise")and NHR@Göttingen ("Emmy")

Integration of own systems

- **Housing** (administration by users) includes:
  - Rack space
  - Power
  - Cooling
  - Networking connectivity
- In addition, with **Hosting** GWDG takes care of:
  - Provisioning Node images
  - optional integration with the batch system (FairShare compensation or exclusive usage)
- How to include **own procurements**
  - Stating your requirements, initial consulting
  - Deciding on housing/hosting, batch system integration
  - GWDG: Procurement, order, preparing infrastructure, installation
  - further coordination: granting user access, project directories

Q&A

All presentations in the GöHPCoffee series:

https://docs.gwdg.de/doku.php?id=en:services:application_services:
high_performance_computing:hpc_coffee