TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Ulf Markdwardt, Martin Schroschk

Site Update

NEC-Systems-Bouquet at TU Dresden

June 13th, 2024 / NUG

Agenda

- Environment: TU Dresden, ZIH, and data center
  - NHR and procurement strategie and cycle
- HPC systems
- Trends and demands

TU Dresden and ZIH

Center for Information Services and High Performance Computing

TU Dresden

- Founded in 1828
- Largest university in Saxony
- One of Germany's eleven *Universities of Excellence*
- Numbers
  - 28.000 students
  - 8.700 employees
  - Over 600 professorships
  - 120 courses of studies

TU Dresden Campus

LZR - The Lehmann Center Data Center

LZR - The Lehmann Center Data Center II

- 1200 qm of floor space for IT systems
- HPC systems and central IT systems of the TU Dresden
- Highest efficiency is standard here; reuse of waste heat
- 4 MW electrical power for IT services (upgrade to 7 MW is planned)
- Sophisticated, redundant power supply for critical infrastructure
- Annual black building tests are conducted to verify the correct functioning of all infrastructure

ZIH – Center for Information Services and HPC

- Central scientific unit of TU Dresden

Basic IT Services for TU Dresden

- Campus network, email, groupware, data exchange, backup, …

Computational Science Services

- Virtual machines, hosting and housing
- High Performance Computing (HPC) since 1997

Research and Development

- Parallel programming and algorithms
- Data Analytics applications, machine learning, …
- Tools for performance and energy efficiency optimization

Research Project EECliPs

- Increase energy-efficiency of the ICON weather model through the use of heterogeneous clusters
- Target architectures incl. NVIDIA GraceHopper, **NEC SX-Aurora Tsubasa**, Intel Alder Lake, etc.
- ZIH's part: enhance software analysis tools for the target architectures

Project

- Partners: DKRZ (German Climate Computing Center), Atos, ParTec, TU Dresden
- Period: 09.2022 - 08.2025
- Funding: Federal Ministry of Education and Research (BMBF)
- Contact and further information: robert.schoene@tu-dresden.de

Tools

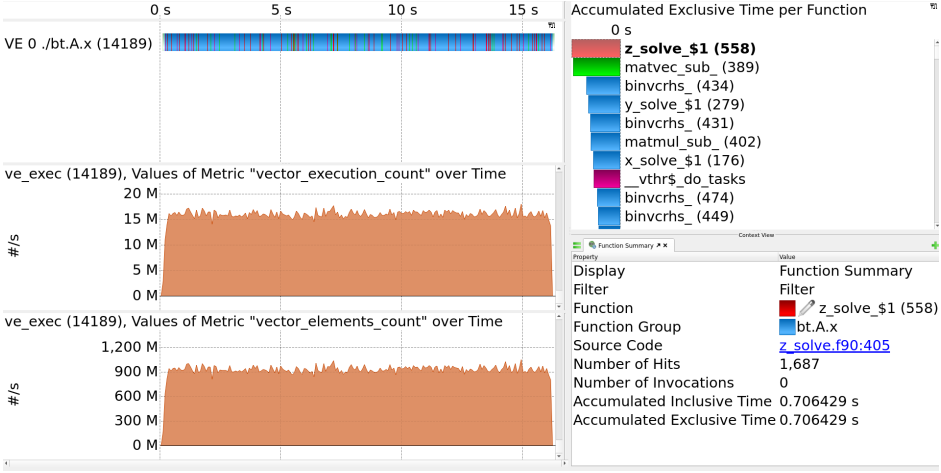
Score-P

- Source code instrumentation of HPC applications for profiling and event tracing
- Difficulties with NEC toolchain, C++ standard compliance
- Preliminary support for compiler instrumentation/MPI implemented

`lo2s`

- Node-level analysis tool utilizing non-invasive low-overhead sources
- Uses **libveos** to sample Vector Engine program counter, record performance monitoring counters
- Vector Engine support upstreamed: https://github.com/tud-zih-energy/lo2s

Example Sampling Trace from *lo2s*



NHR - National High Performance Computing

- Bundle the resources and competencies of university high-performance computing and make them available to scientists at German universities
- User support by providing advice and training in the use of high performance computing in their fields of application
- NHR@ZIH: AI + Live Sciences + Earth System Sciences

Procurement Strategy

W/o NHR

- Every five years
- One large system for about 15 Mio EUR
- Uncertainties due to competition for the same funding pot with other large research projects

W/ NHR

- Planning period is 10 years
- Hardware, maintenance, software licenses, staff, etc.
- New system every two years for about 8-9 Mio EUR
- Island concept: multiple clusters with different capabilities and burst buffer connected together via network

HPC Systems at ZIH

HPC Systems

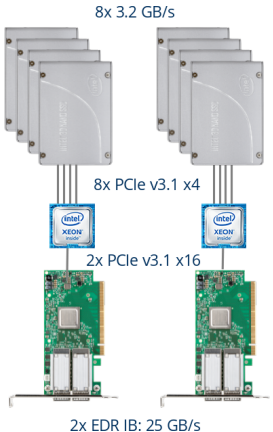
- (two housing systems from other research facilities)
- DLR system; isolated in DLR network
- Five clusters from different vendors all connect via very same InfiniBand fabric
- Good and solution-focused collaboration of all vendors
- NEC provided three clusters and two storage systems
  - NEC is a competent and reliable partner for ZIH

Quobyte (2019)

- Warm archive with 15 PB capacity (incl. S3 option)
- Easy to operate and use
- Issues
  - Small community at the time; covered by agile support from NEC
  - Misconfiguration at initial setup lead to loss of data
- Will be re-used for non-HPC-services

NVMe over Fabrics (2019)

- Fast scratch filesystem for e.g. AI workflows
- 2 PB storage in NVMe; 4 TB/s bandwidth
- 90 NVMe storage nodes
- Plan for future
  - Further operation for three more years w/o maintenance support
  - Challenge: spare parts
  - Fast storage for *Barnard* cluster using WEKA (NEC)
  - Configuration and setup started



Romeo CPU-Cluster (2019)

- General purpose cluster based on AMD Rome CPUs
  - Basically same nodes as CARA
- 192 nodes, each with
  - 2 x AMD EPYC CPU 7702 (64 cores)
  - 512 GB RAM
  - 200 GB local memory on SSD at `/tmp`
- Slurm, Rocky Linux 8.7
- High availability, high reliability

Alpha Centauri (2021)

- AMD Rome CPUs and NVIDIA A100 GPUs
- Designed for AI and ML tasks
- 37 nodes, each with
  - 8 x NVIDIA A100-SXM4 Tensor Core-GPUs
  - 2 x AMD EPYC CPU 7352 (24 cores)
  - 1 TB RAM
  - 3.5 TB local memory on NVMe device
- Slurm, Rocky Linux 8.7
- Rank 32 in Green500 (June 2022) with 15.8 Gflops/watts

Alpha Centauri (2021) II

Great system, but

- High failure rate and long repair cycle (~ 6 w)
- Multiple causes
  - Supply bootleneck and high demand
  - Quality of GPUs
  - Not sufficient testing phase
  - Too few manpower in L2/L3 service
- No follow-up maintaince contract (years 4-5)

CARA

- General purpose cluster based on AMD Naples and Rome CPUs
- Specifically designed for CFD simulations and workflows
- 2168 nodes, each with
  - 2 x AMD EPYC CPU 7601 (32 cores)
  - 128 GB up to 256 GB RAM
- 664 nodes, each with
  - 2 x AMD EPYC CPU 7702 (64 cores)
  - 256 GB up to 1 TB RAM
  - 40 nodes with local memory on NVMe SSD
  - 10 nodes with 4 x NVIDIA A100 GPUs

CARA II

Storage

- Two Lustre systems
  - Bulk
    - 16.518 PB in total
    - Flash-based metadata system
  - Fast
    - Scratch filesystem for high IOPs and also `/home`
    - 600 TB in total
    - Flash-based metadata system with high metadata performance

Users

- $>$ 200 active users per month
- $>$ 700 unique active users since start of operation
- 29 DLR institutes can access and make use of CARA

Trends and Demands

- Continuing increase in demand for HPC resources
  - New research groups and areas test and apply AI methods for their fields
- Disaster recovery
  - Time limit 24 h
  - Ansible and git
- Network
  - InfiniBand not set; open for alternatives
- DLR: Increasing diversity in use cases and workflows

Thank you for your attention.