Karlsruhe Institute of Technology

First experiences with an Arm-based HPC-testbed at NHR@KIT

14.11.2022, AHUG@SC22, Dallas

René Caspart, Steinbuch Centre for Computing, Karlsruhe Institute of Technology

Outline

- HPC at NHR@KIT
- Operational experiences with an ARM testbed
- Experiences and first results with benchmarks and mini-apps

HPC at KIT

National High-Performance Computing Center

- KIT is a center in the German national high performance computing alliance (NHR)
  - NHR: Alliance of national academic Tier 2 HPC provider
  - NHR@KIT focuses on expert level support and software sustainability among other aspects
  - Novel and disruptive technologies as a key aspect

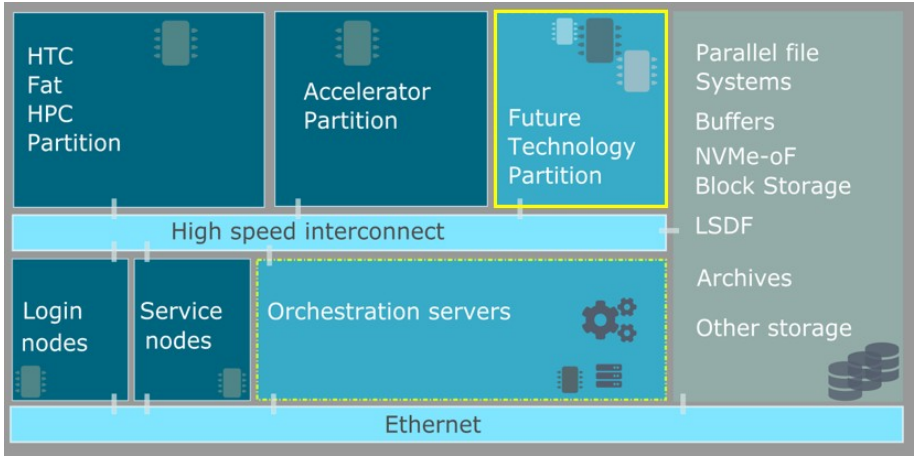
- KIT operates two HPC systems
  - General purpose Tier-3 system for the state of Baden Württemberg
  - Tier-2 system with two subsystems
    - HAICORE → AI specific setup
    - Future Technologies Partition

Hochleistungsrechner Karlsruhe - HoreKa

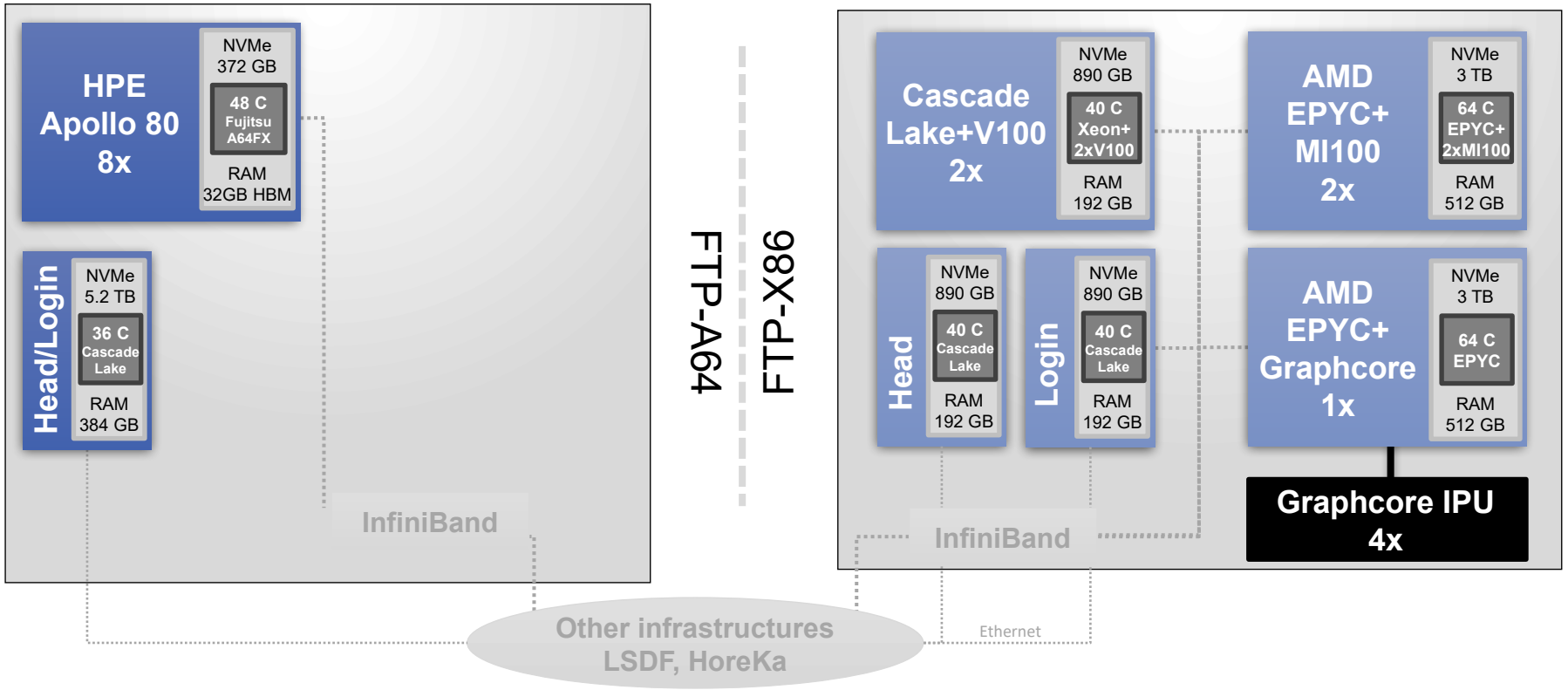
- Tier 2 System in the National High Performance Computing alliance
- Modern hybrid system
  - ~60k CPU cores
  - 668 Nvidia A100
  - 200 Gbit/s interconnect
  - 16 PB storage
- Peak #13 in the Green500 and #52 in the TOP500 (June 2021)
- In operation for the scientific community since 1st June 2021

The Future Technologies Partition

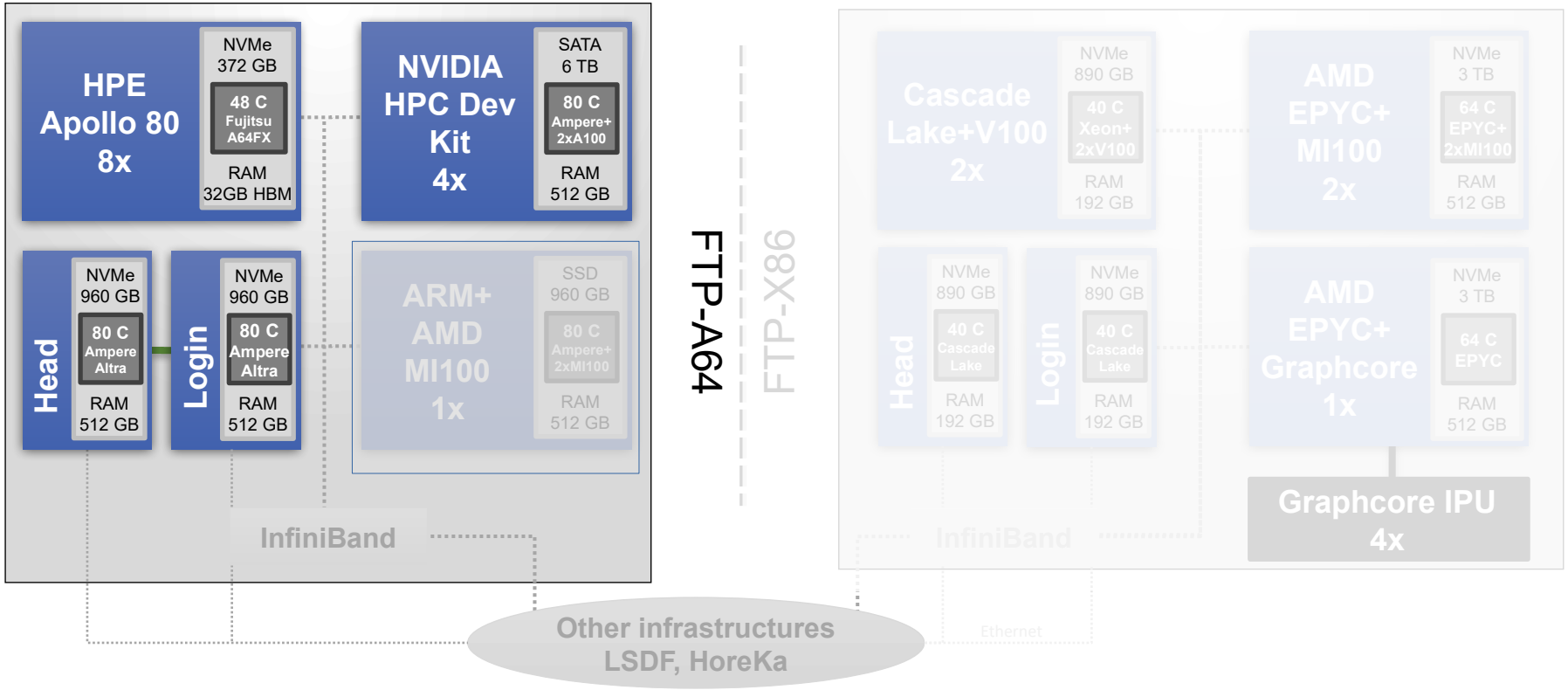
- Our goal:
  - Enabling Scientists and Research Software Engineers to explore new hard- and software options
  - Preparing for possible future clusters
- We provide a testbed for this type of HPC hardware and software
  - Future Technologies Partition
  - Accessible to all users of HoreKa upon request



The Future Technologies Partition before 2022



The Future Technologies Partition in 2022



Architectures and Accelerators at NHR@KIT

	x86	ARM
Only CPU	HoreKa	FTP
Nvidia	HoreKa	FTP
AMD	FTP	FTP
Graphcore	FTP	

Source: Simon Raffeiner

Operational Aspects – Cluster Management

- KITE as tooling for cluster deployment and management
  - Based on Open Source Software components
  - Required adaptions for ARM platform straight forward
  - Little changes needed
- Base OS Rocky Linux 8

Operational Aspects – Userspace Software

- Installing userspace software, mostly straight forward
  - Compiler: GNU, LLVM, NVIDIA HPC SDK, ARM
  - MPI: OpenMPI, MVAPICH
  - OpenMP
  - BLIS, openBLAS, cuBLAS
  - ...

Cluster-Job-Performance Monitoring

- Cluster-Job-Performance Monitoring
  - Cluster Cockpit Monitoring stack largely developed by NHR@FAU and NHR@KIT
  - cc-metric-collector node-agent to collect metrics for each node
  - Many different types of metrics (compute, network, data, ...)

- Largely works out-of-the box
  - Some dependencies not (yet) available on ARM
  - Some information from the CPU differs to x86 processors

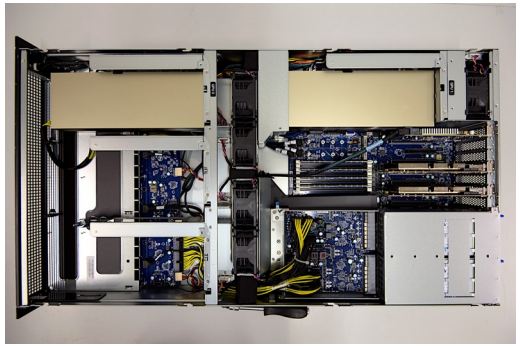
    → Small adaptions needed

Benchmarking

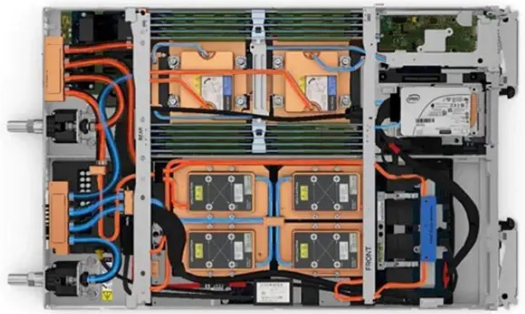
- Running selection of synthetic benchmarks
  - Babelstream
  - DGEMM
  - Graph500

Benchmarking

- For comparison results from two platforms
  - Nvidia ARM HPC Dev Kits
    - Single socket, Ampere Altra Q80-30
    - 2 Nvidia A100-40 PCIe
    - 512 GB DDR4-3200
  - HoreKa compute nodes
    - Dual socket, Intel Xeon Platinum 8368
    - 4 Nvidia A100-40 SXM4
    - 512 GB DDR4-3200



Source: Simon Raffeiner



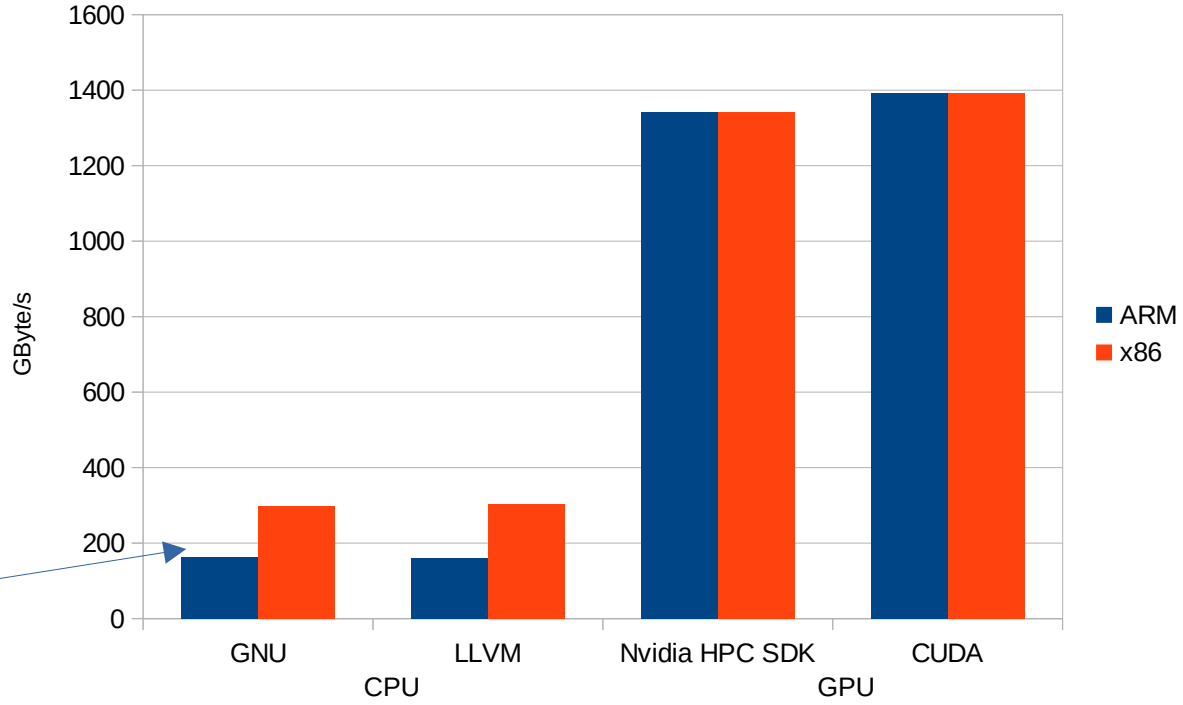
Source: Lenovo

Benchmarking

- Babelstream

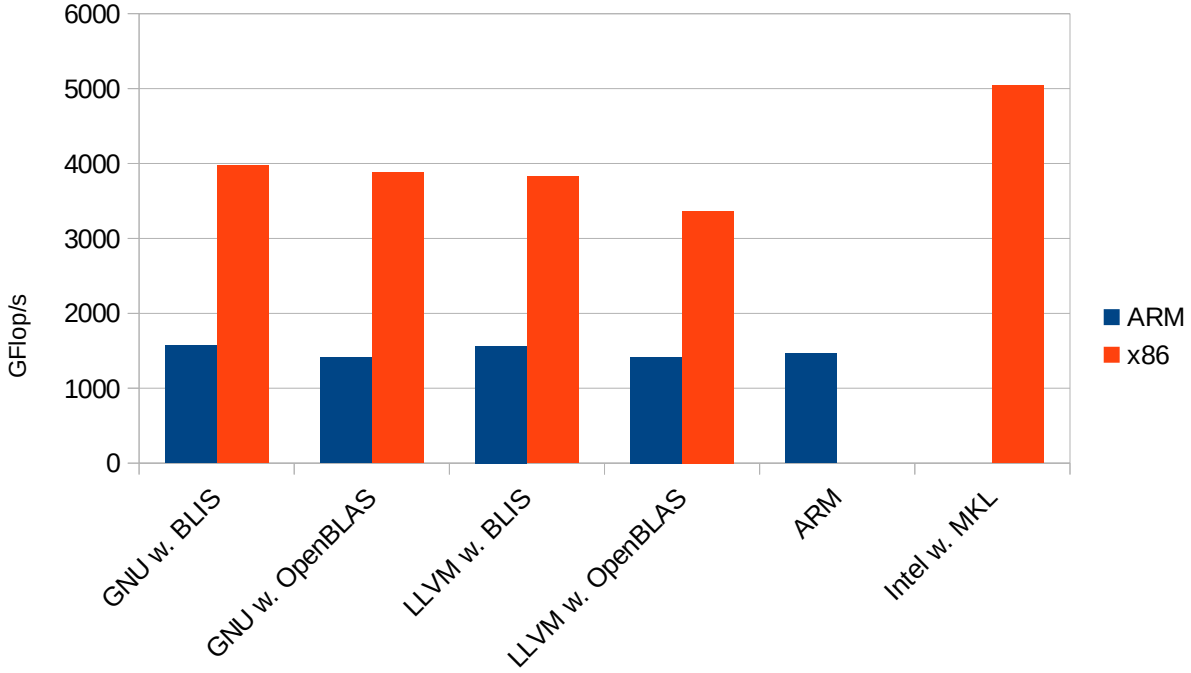
https://github.com/UoB-HPC/BabelStream

Note:
x86 system is dual socket with twice as many memory channels



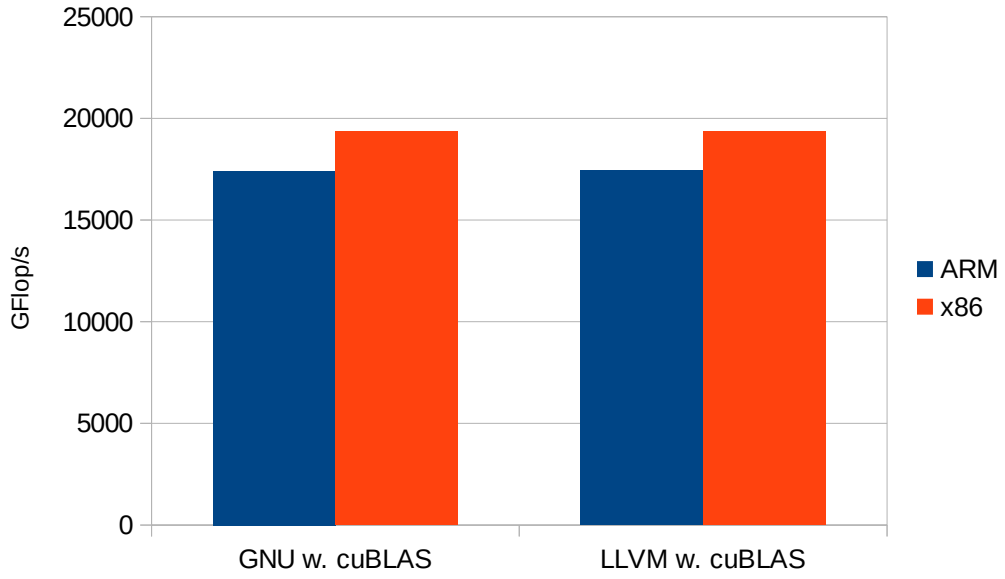
Benchmarking

- DGEMM



Benchmarking

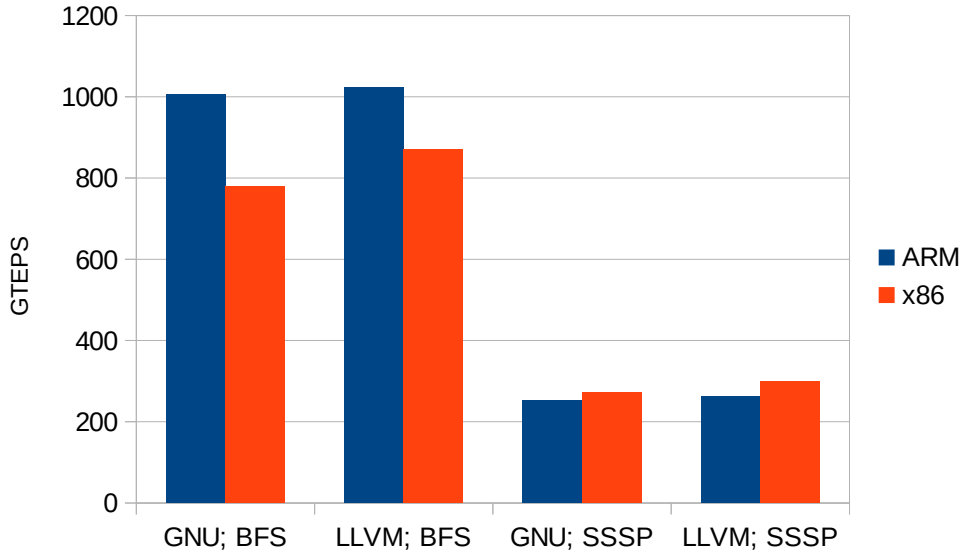
- DGEMM
  - Using the A100 GPUs with cuBLAS
  - Similar performance on both platforms
    - Known different powerbudget for A100 with PCIe vs SXM4



Benchmarking

- Graph500

https://graph500.org



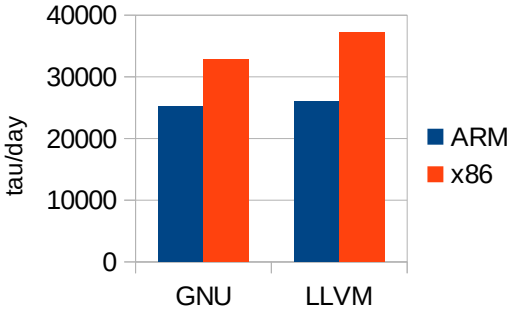
Mini-Applications

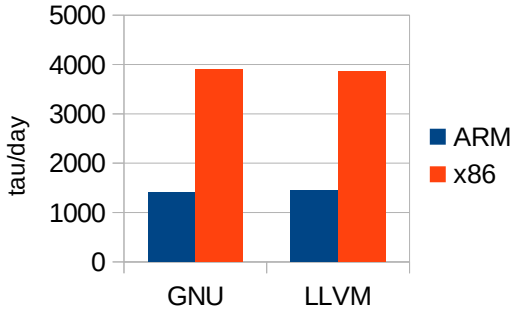
- Run exemplary workload applications from users on HoreKa
  - LAMMPS
  - OpenLB
  - OpenFOAM
  - Machine Learning (CIFAR10 with ResNet18)
  - High Energy Physics (HEPScore)

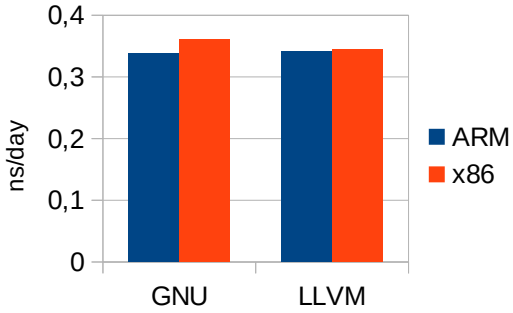
LAMMPS

- Run three different workloads
  - FENE beadspring
  - Chute flow with frozen base at 26 degrees
  - Rhodopsin model

https://docs.lammps.org/Speed_bench.html





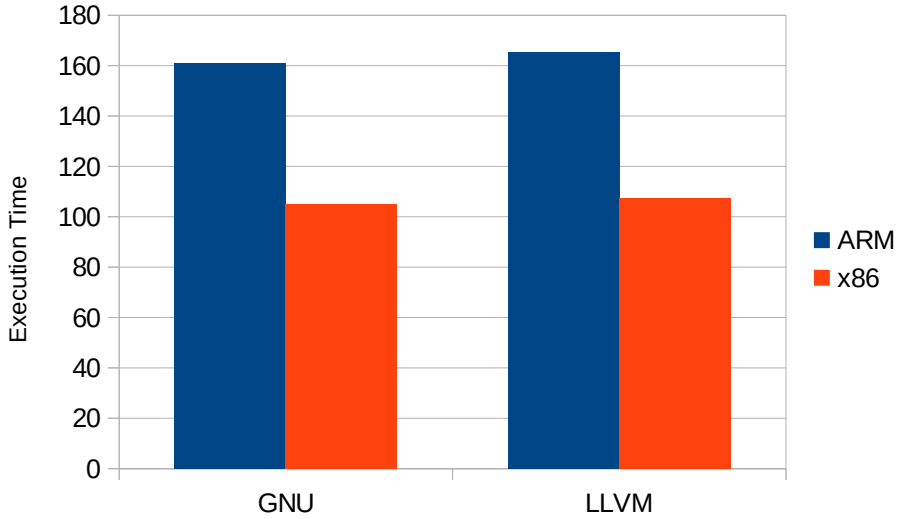


OpenFOAM

- 3D Lid driven cavity
- Modelsize: S
- 500 iterations

OpenFOAM®

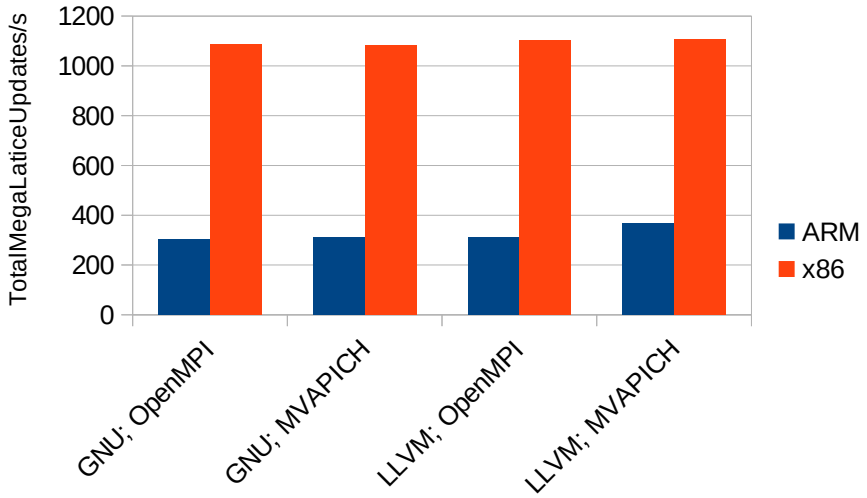
https://develop.openfoam.com/committees/hpc



OpenLB

- Standard OpenLB example
  - Laminar 3D cavity

https://www.openlb.net/lid-driven-cavity/

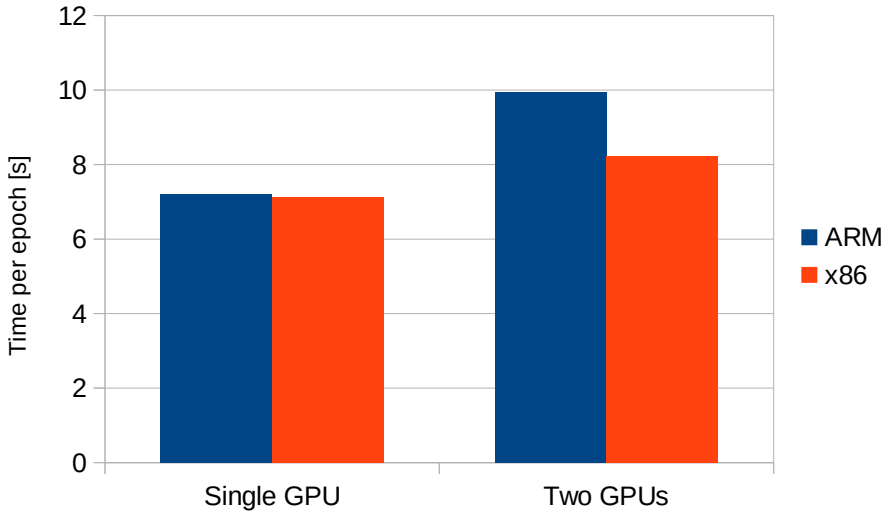


Machine Learning

- Train a ResNet18 on the CIFAR10 dataset
  - Compare time per epoch for two setups with Nvidia A100-40 PCIe GPUs
    - Nvidia ARM HPC Developer Kits
    - x86 server; Intel Ice Lake
- Setting up
  - Native setup not available
    - No pytorch modules available with ARM and CUDA
  - Containers for pytorch available from Nvidia
    - Our two supported containerization tools enroot and singularity work out of the box

Machine Learning

- Train ResNet18 on the CIFAR10 dataset
- Evaluate time per epoch using 1 and 2 GPUs



High-Energy-Physics Use case

- KIT has a strong particle physics community
  - No traditional HPC workflows (mainly HTC)
  - Hosting one of the WLCG Tier 1 computing centers
  - So far mainly x86 resources
- Interest in seeing performance and portability to ARM
  - Two benchmark sets
    - Old: HEPSpec06, based on Spec06
    - New: HEPScore, successor of HEPSpec06

High-Energy-Physics Use case

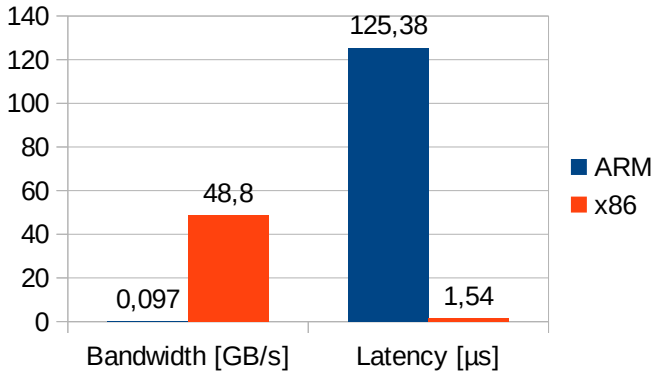
- HEPScore
  - Based on mini-apps by the user communities
  - Partially available for ARM (ongoing efforts for others)
  - Only utilizing the CPU

    https://gitlab.cern.ch/hep-benchmarks/hep-score

- Results
  - Nvidia ARM HPC Dev Kit: 5.3
  - Dual socket AMD EPYC 7742: 6.9

One General Observed Problem

- Our HPC Dev Kits experience issues with internode and GPU-to-GPU communication
  - Unfortunately, so far no solution found
  - Active support from e.g. Nvidia (thanks a lot!)
- Clearly visible in benchmarks
  - E.g. OSU Micro-Benchmarks



Summary

- The Nvidia ARM HPC Dev Kits provide an interesting testbed for HPC administrators and users
  - Most setups need only minor adaptions
  - Benchmarks provide expected performance results
  - However: some issues still pending
- The Future Technologies Partition at NHR@KIT provides a valuable testbed for HPC users and operators
  - Looking forward to see more HPC specialized ARM CPUs and nodes in the (near) future!