National High-Performance Computing Center

Introduction to NHR@KIT and HoreKa

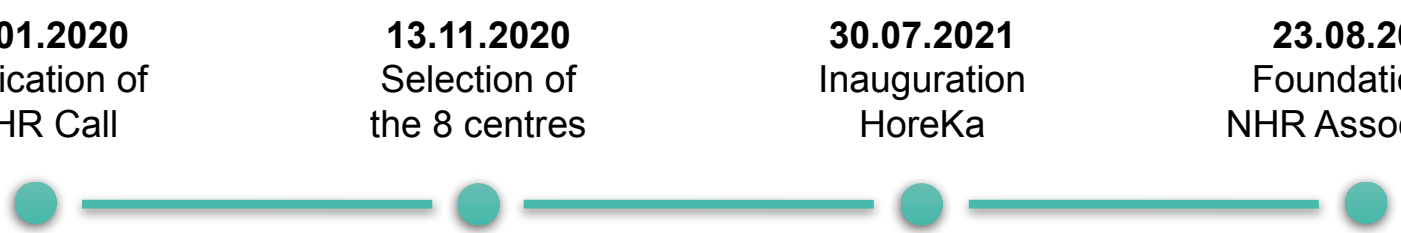
Robert Barthel – KIT, SCC

NHR Overview (1)

„**Nationales Hochleistungsrechnen**“ (NHR) replaces the federal 91b DFG funding for academic Tier 2 HPC systems

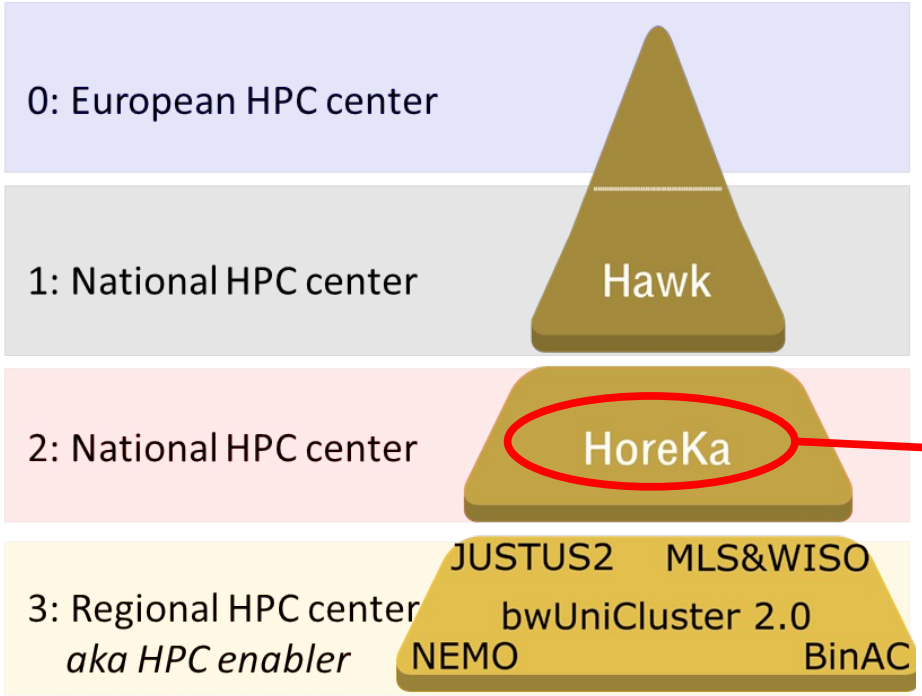
625 million Euros over 10 Years for 9 NHR centers

Coordination of Applications, Method, Hardware and Operational competencies



NHR Overview (2)

HPC in Baden-Württemberg



National HPC at Tier 2

NHR4CES@RWTH, NHR4CES@TuDa	RWTH Aachen TU Darmstadt
NHR@Göttingen, NHR@ZIB	GWDG und Universität Göttingen Zuse-Institut Berlin
NHR@FAU	FAU Erlangen-Nürnberg
NHR@KIT	Karlsruher Institut für Technologie
NHR@TUD	ZIH TU Dresden
PC2	Universität Paderborn

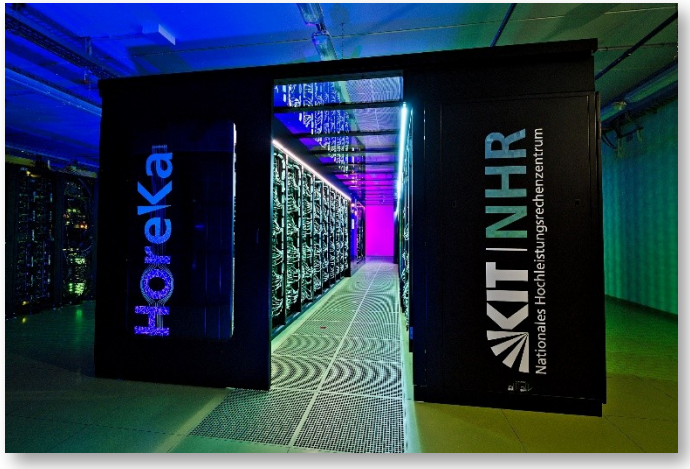
Overview NHR@KIT

„HoreKa“ operated by NHR@KIT since 01.06.2021 is one of the **most powerful supercomputers in Europe**.

HoreKa was the first new system dedicated to NHR.

The **successor systems** (2025/2030) are already funded.

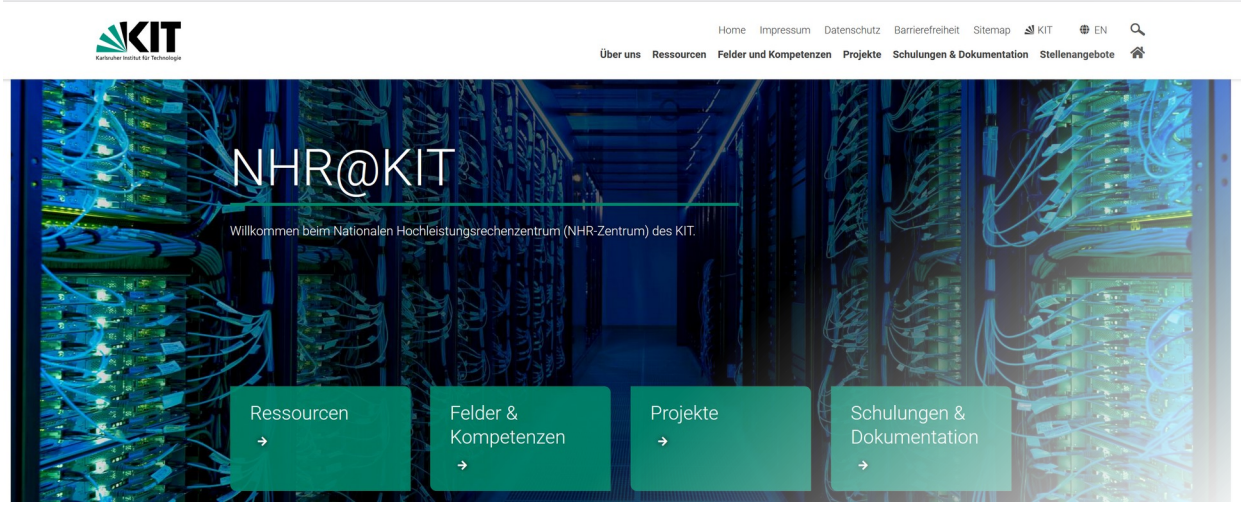
NHR also enables an increase of **expert support**.



NHR website

https://nhr.kit.edu/

Resources, Documentation, Consulting, Training, Support…



NHR@KIT: Additional Support for users

Experimental/Bleeding edge resources
- Future Technology Partition

Software Sustainability and Performance Engineering Team (SSPE)
- Porting codes to new Programming Languages, Environments, Libraries
- Porting codes to new hardware (Accelerators, ARM CPUs etc.)
- Support with Continuous Integration/Testing/Benchmarking/Deployment (Cx)

Simulation and Data Life Cycle Labs (SDLs)
- Support with scientific data mgnt., data intensive computing & analysis

Training
- More trainings, workshops, Hackathons etc. at KIT
- NHR@KIT Training Overview: https://www.nhr.kit.edu/english/66.php
- NHR Alliance Training Overview: https://www.nhr-verein.de/en/courses-and-workshops

NHR@KIT trainings

https://www.nhr.kit.edu/english/66.php

Title	Date	Time	Speaker	Venue
2nd AI-Hero Hackathon on Energy Efficient AI	12.-14.12.2022	8:30-17:30	Helmholtz AI, HMC, Helmholtz Imaging, NHR@KIT	Online
Introduction to OpenFOAM	11.11.2022	09:30-15:00	Jordan Denev	Online
Advanced High Performance Computing (HPC) and Data Management topics	19.10.2022	9:30 - 15:45	bwHPC@KIT team	Online
Introductory Course NHR@KIT and HoreKa	12.10.2022	9:30-15:45	NHR@KIT Team	online

Karlsruher Institut für Technologie

HoreKa

Budget: 15 Mio. €
Procurement: Q3/19 – Q2/20, Installation: Q3/20 – Q2/21

769 nodes, **~60.000 cores total**
668 NVIDIA A100 GPUs
17 PetaFLOPS Peak

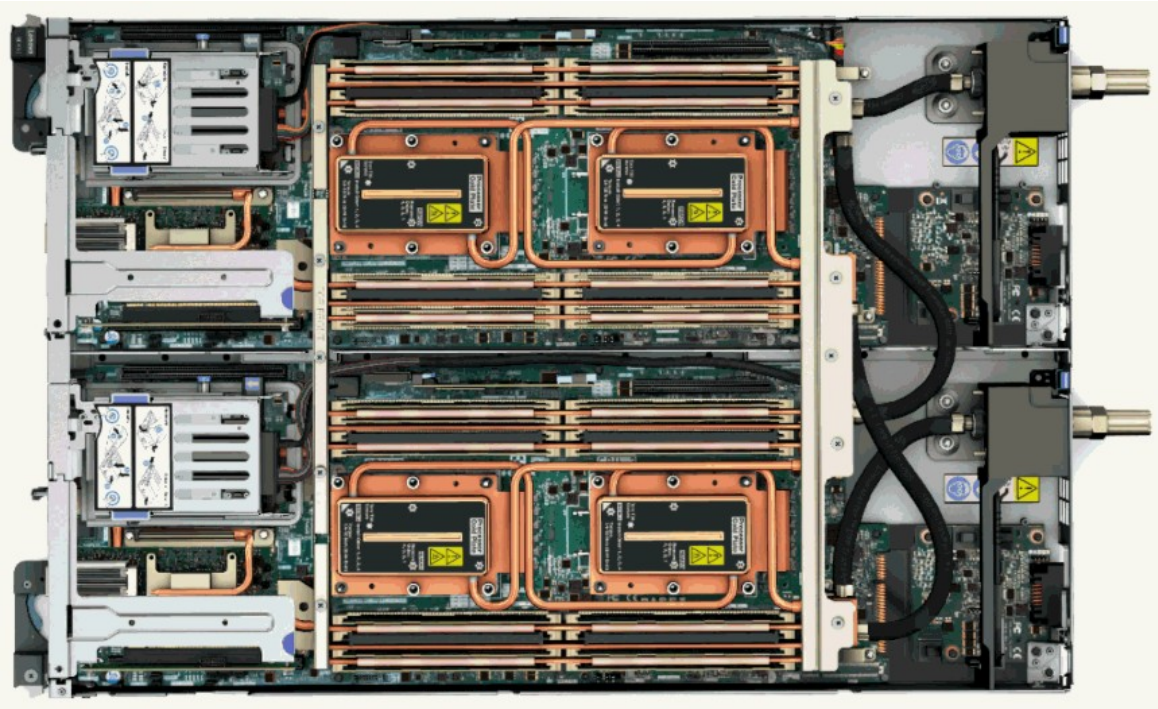
830 kW Warm Water Cooling (+Re-Use)
16 PB Spectrum Scale, InfiniBand HDR

Green500 #13
Top500 #53 (#220), EU Top 15

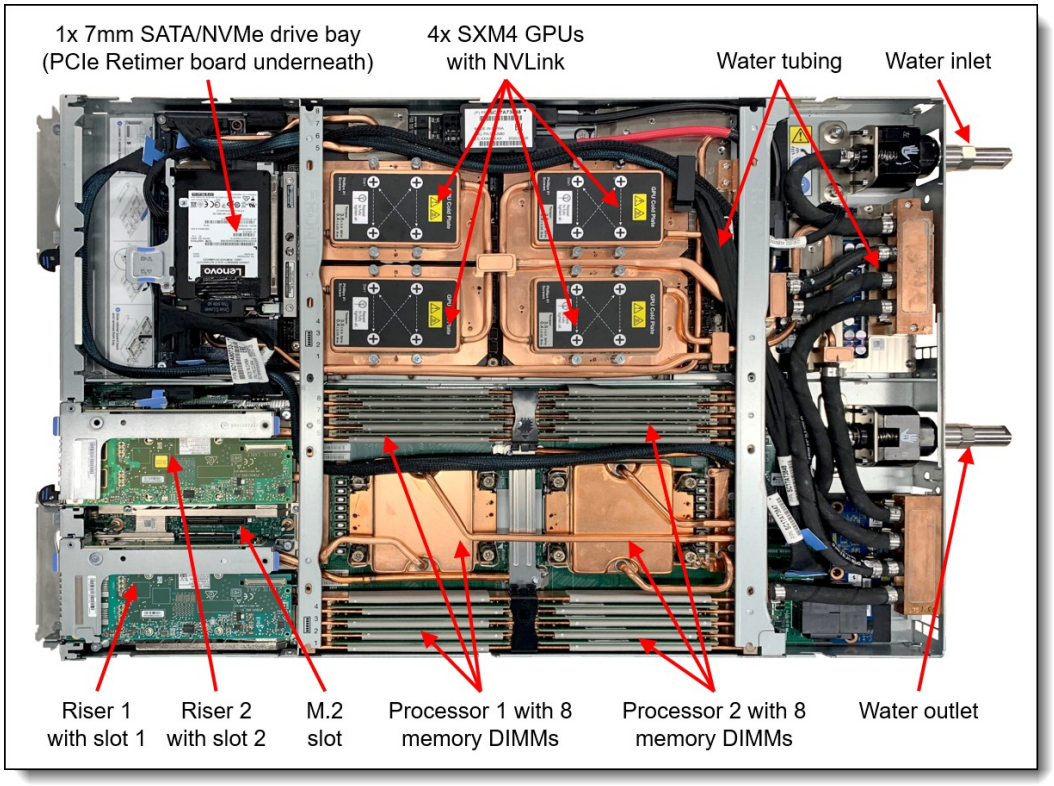
HoreKa node types

	Lenovo SD650 v2	Lenovo SD650-N v2
# Nodes	570+32	167
CPUs	2x Intel „Ice Lake“ 76 Core/152 Thread	2x Intel „Ice Lake“ 76 Core/152 Thread
Memory	256/512 GB	512 GB
GPUs		4x NVIDIA A100
Interconnect	InfiniBand HDR200 Fat Tree	
Storage	16 PB, ~150 GB/s GPFS	

HoreKa CPU Node



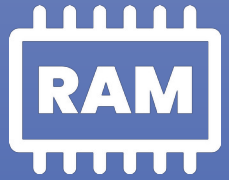
HoreKa GPU Node



HoreKa Hardware “Look-and-Feel”

Good balance
25% CPU / 75% GPU

CPU-part 4x ForHLR II



Large memory
3,3 to 6,7 GB per Core

1-2x ForHLR II

High Throughput
Main memory 400 GB/s
GPU mem up to 1 TB/s
Local SSDs up to 3 GB/s

2-6x ForHLR II

HoreKa Software “Look-and-Feel”

Well-known environment

Red Hat Linux 8.x, SLURM, Modules, Parallel file systems…

Standard libraries and tools

CPUs: Intel MKL, OpenBLAS, OpenMPI, MPI …

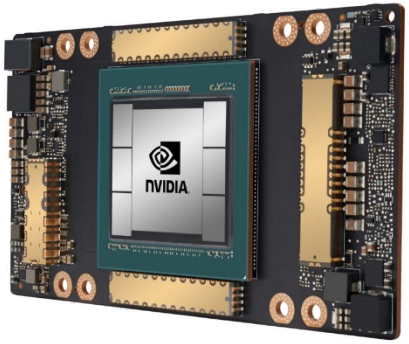
GPUs: CUDA, cuBLAS, cuFFT, cuDNN, NPP, OpenACC, MPI, …

Standard compilers

CPUs: Intel, GCC, LLVM, NVIDIA

GPUs: NVIDIA, GCC, LLVM

HoreKa: CPU vs. GPU



Intel Ice Lake Xeon
10 nm, 38 Cores, 8x DDR4

Optimised for „General Purpose“
Average Floating Point Performance
Large memory with medium throughput
External Interconnect (InfiniBand)

NVIDIA Ampere A100
7nm, 6912 Cores, 6x HBM2

Optimised for highest performance
High floating point performance
Small, fast memory
Internal NVLink mesh for Multi-GPU

HoreKa file systems

Differences to ForHLR II

- IBM Spectrum Scale instead of Lustre
- $WORK replaced by Workspaces
- Quota limits for Workspaces

Quota limits

- For capacity and inodes
- Per project on project file system / Per user on workspace file system

New Features / Characteristics

- Snapshots allow fast restore of files by users
- Metadata performance is significantly higher due to SSDs

HoreKa and LSDF

Increased Stability

through redundant gateways

Mounted permanently

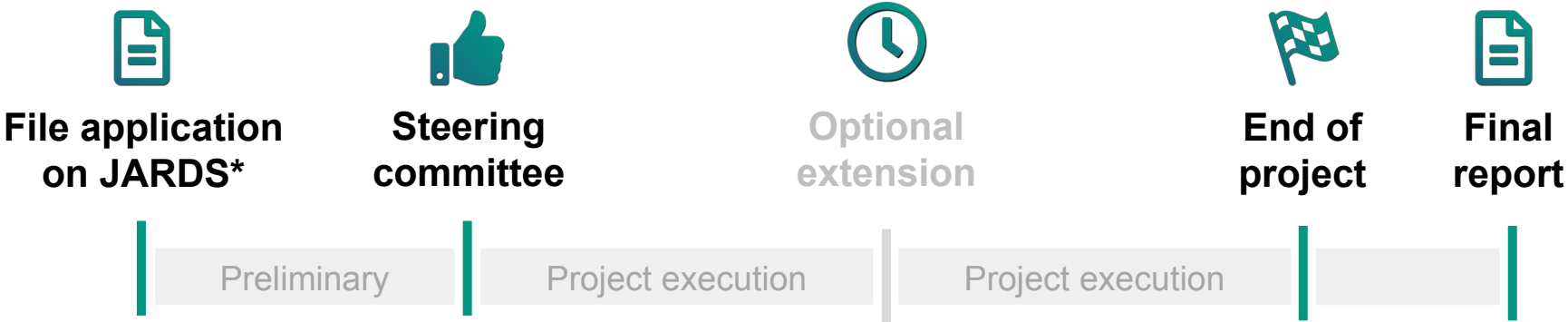
On all nodes

Higher speed

Up to 45 GiB/s

(~7x *ForHLR II*)

HoreKa: Full projects



PI gets access to user management

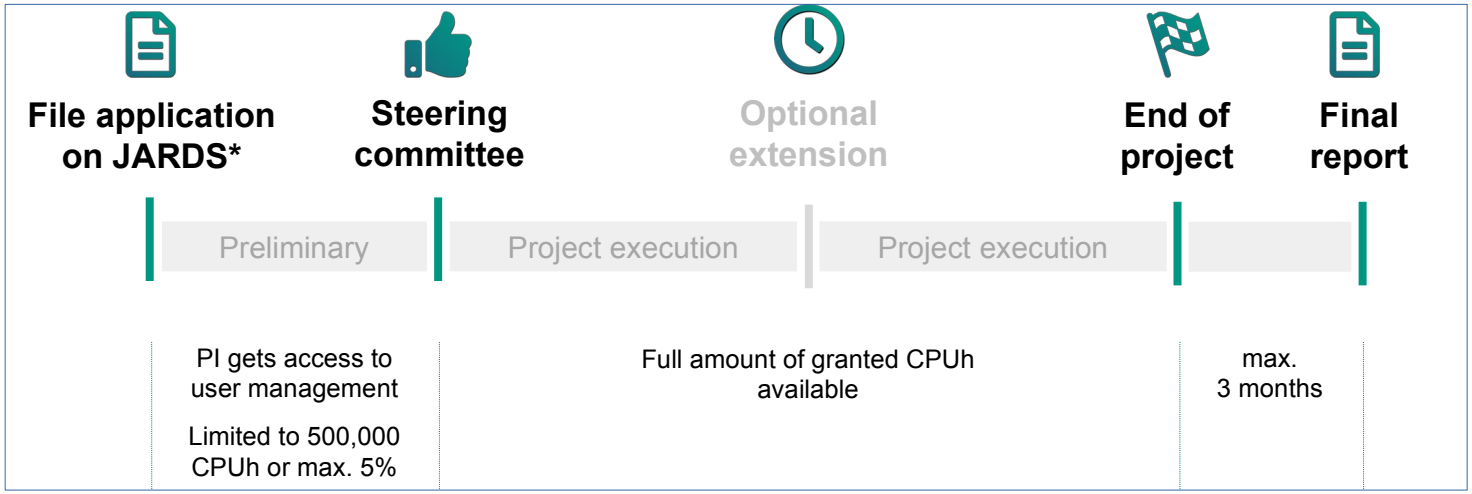
Limited to 500,000 CPUh or max. 5%

Full amount of granted CPUh available

max. 3 months

*https://jards.nhr.kit.edu/

HoreKa: Large projects



CPU: 15 – 70 Mill CPUh, GPU: 0.2 – 1.0 GPUh
Quarterly calls, e.g. submit deadline 01.01. for project start at 01.04.
Additionally reviewed by Nation Steering Committee

*https://jards.nhr.kit.edu/

HoreKa: Test projects

File application on JARDS*

Optional extension

End of project

Project execution

Project execution

Limited to max. 500,000 compute hours

Limited to 6 months**

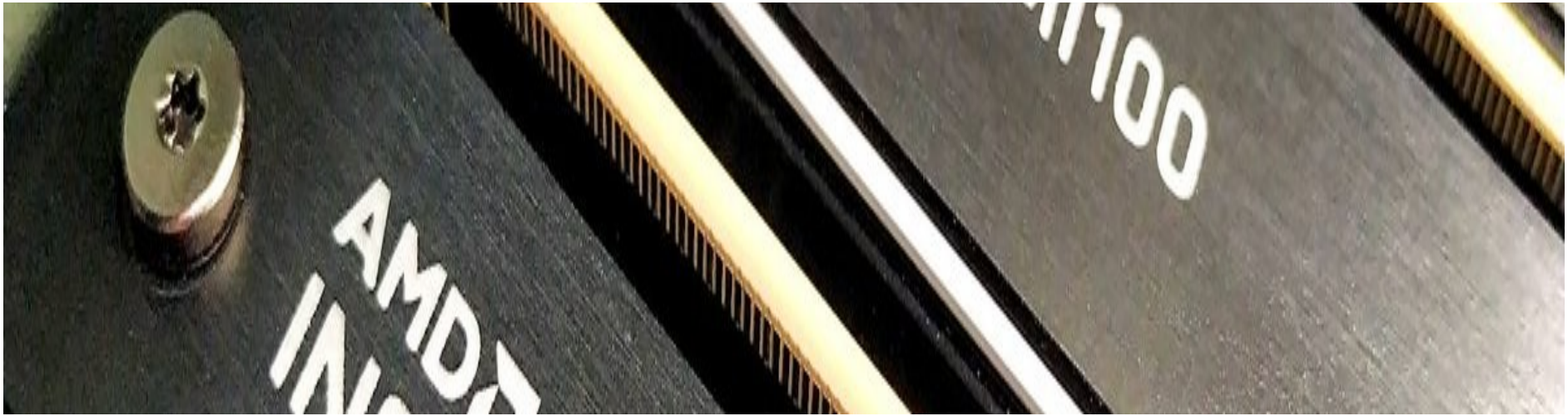
Optional another max. 500,000 compute hours

Optional another max. 6 months**

*https://jards.nhr.kit.edu/ **from 2022 on, test projects will be limited to 1 month

Karlsruher Institut für Technologie

Future Technology Partition



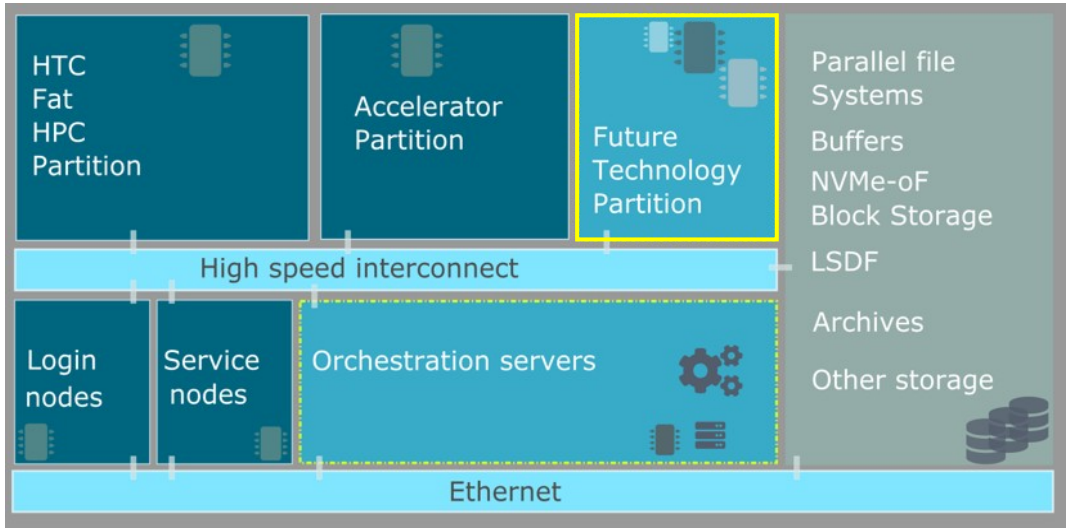
Future Technology Partition (FTP)

Effective Support of scientific software development

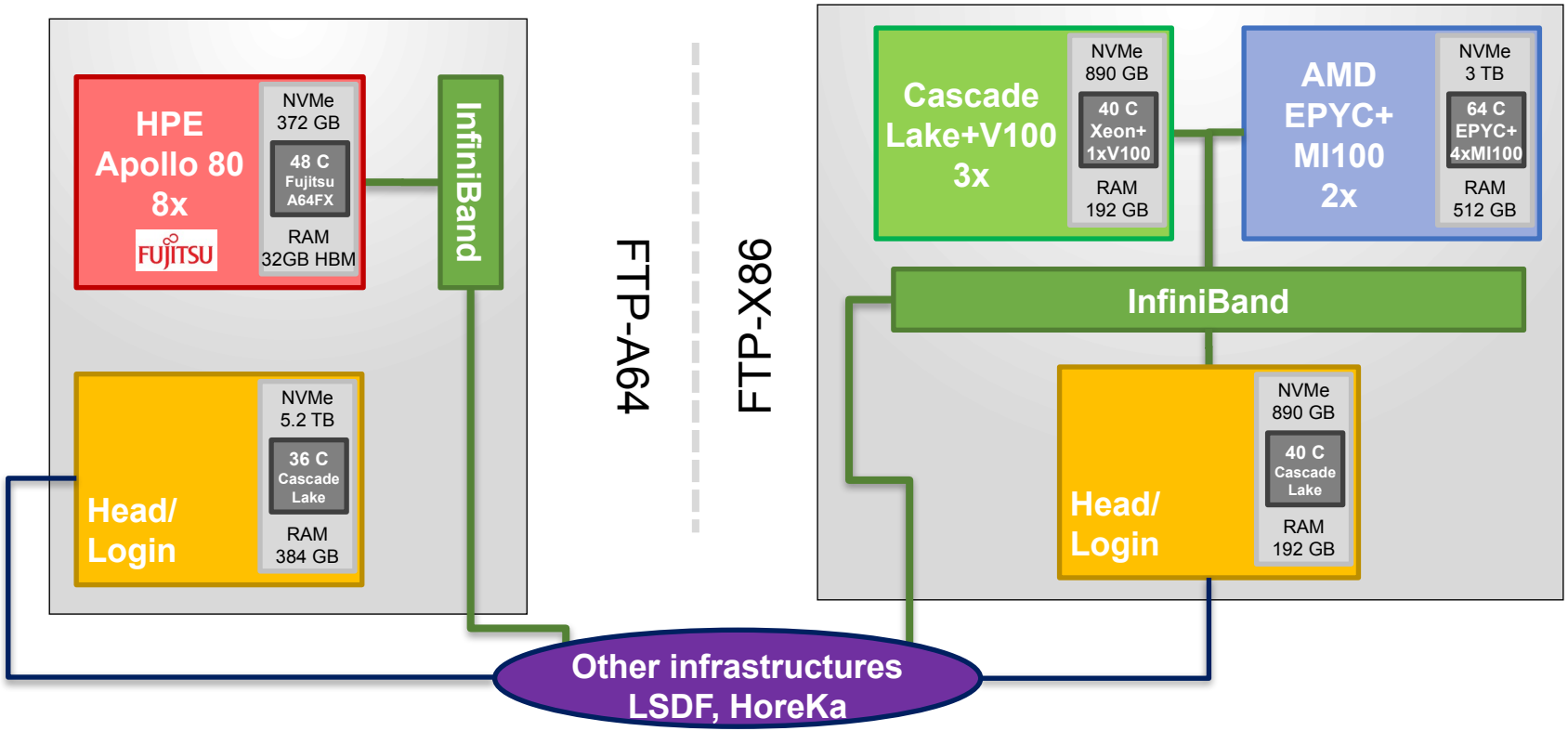
Porting to new hardware

Acceleration of development cycles

Development of efficient algorithms, libraries and applications



FTP: Initial State in 2021



FTP: Planned extensions in 2021

GRAPHCORE IPU-POD16

Dedicated accelerator systems for AI/ML

NVIDIA HPC Dev Kits

Prototype for Grace ARM CPUs (2023)
Ampere Altra Q80-30 CPU (80 Cores),
2x NVIDIA A100, 2x BlueField-2 DPU

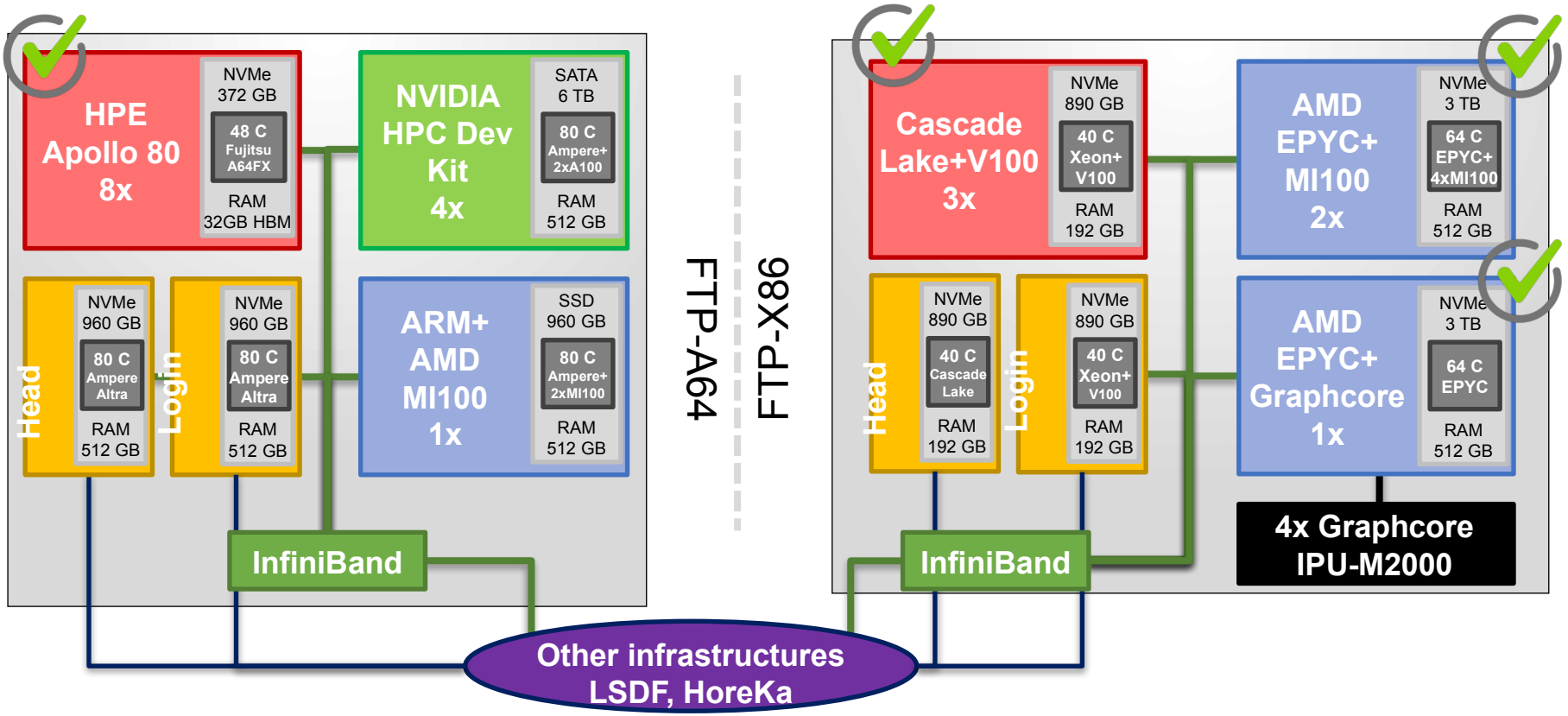
ARM+AMD GPU

Evaluation of ROCm software stack on ARM
Ampere Altra Q80-30 CPU (80 Cores),
2x AMD MI100, 2x BlueField-2 DPU

ARM Service Nodes

Ampere Altra Q80-30, BlueField-2 DPU
Replacing Intel service nodes
Porting KITE cluster stack

FTP: Next Evolution



Future Technology Partition

Software

As usual
(Red Hat, Slurm,
Compiler, Modules etc.)

Access

Granted HoreKa project
(Test or Full) necessary

Expansions

In regular Intervals
(~2x/year)