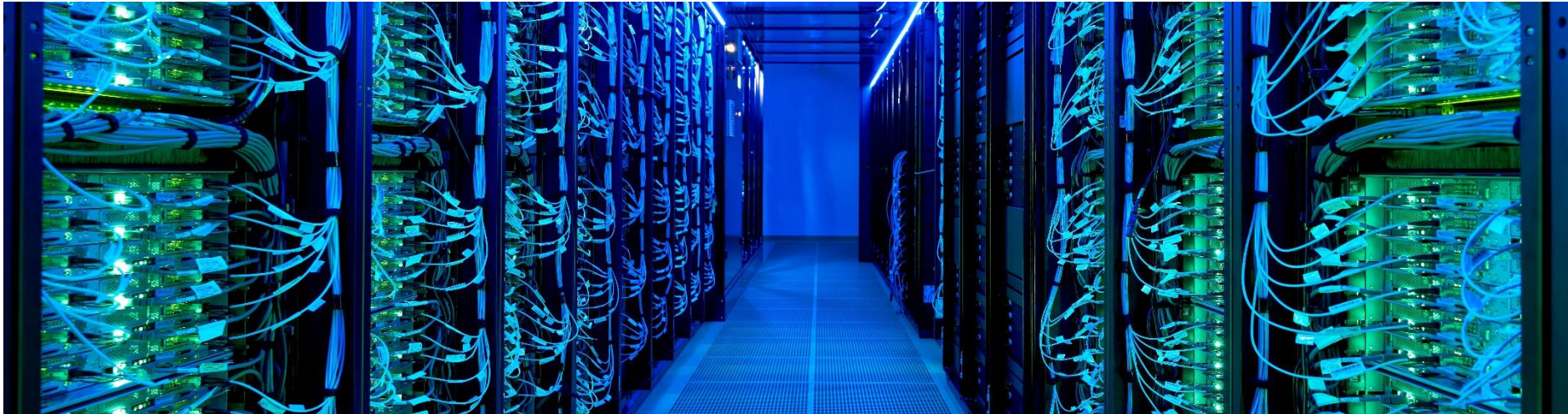


NHR@KIT und Future Technologies

Simon Raffener



Überblick NHR

- „**Nationales Hochleistungsrechnen**“ (NHR) ersetzt die 91b DFG-Förderung für universitäre Tier-2-Systeme
- **625 Millionen Euro über 10 Jahre** für 8 NHR-Zentren
- Abstimmung von Anwendungs-, Methoden-, Hardware- und Betriebskompetenzen



07.01.2020

Veröffentlichung
NHR-Ausschreibung

13.11.2020

Auswahl der
acht Zentren

30.07.2021

Einweihung
HoreKa

23.08.2021

Gründung
NHR-Verein



Überblick NHR@KIT

- Mit „HoreKa“ betreibt NHR@KIT seit 01.06.2021 einen der **leistungsstärksten Supercomputer Europas** am KIT.
- HoreKa ist das erste neue System, das in den NHR-Verbund eingebracht wurde.
- Die **Nachfolgesysteme** (2025/2030) sind durch NHR bereits finanziert.
- NHR ermöglicht auch eine Aufstockung der **fachlichen Unterstützung**.



NHR@KIT: Unterstützung für Nutzende

■ Computational Science and Data Labs

■ Neue Ressourcen und Dienste

- Future Technology Partition, Cx as a Service, Föderierte Authentifizierung, JupyterHub, ...

■ Software Sustainability and Performance Engineering Team (SSPE)

- Portierung von Codes auf neuere Programmiersprachen, Hardware, Libraries
- Unterstützung bei Continuous Integration/Testing/Benchmarking/Deployment (Cx)

■ Breiteres Schulungsangebot

- Kurse, Workshops, Hackathons, Code-Summers

HoreKa

Budget: 15 Mio. €

Ausschreibung: Q3/19 – Q2/20, Aufbau: Q3/20 – Q2/21

769 Knoten, ~**60.000 Cores** gesamt
668 NVIDIA A100 GPUs

17 PetaFLOPS Peak

Green500 Platz 13

Top500 Platz 53 (+220), EU Top 15

830 kW Warmwasserkühlung (+Nachnutzung)

16 PB Spectrum Scale, InfiniBand HDR



HoreKa Knotentypen



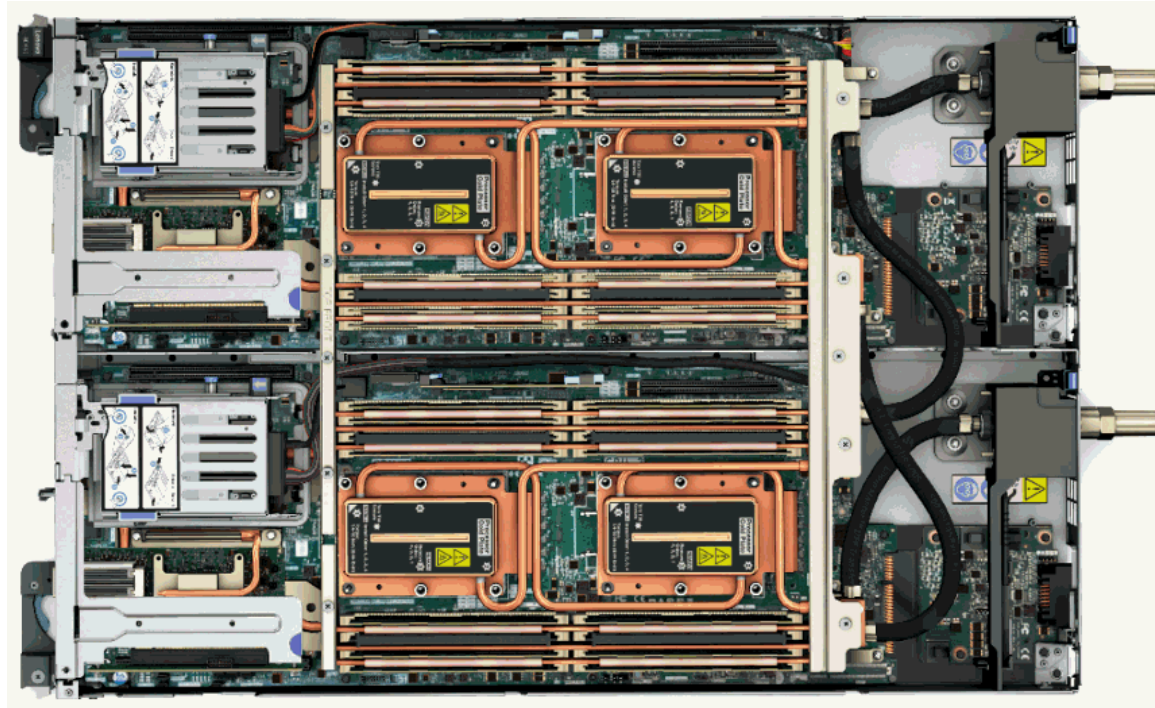
Lenovo SD650 v2



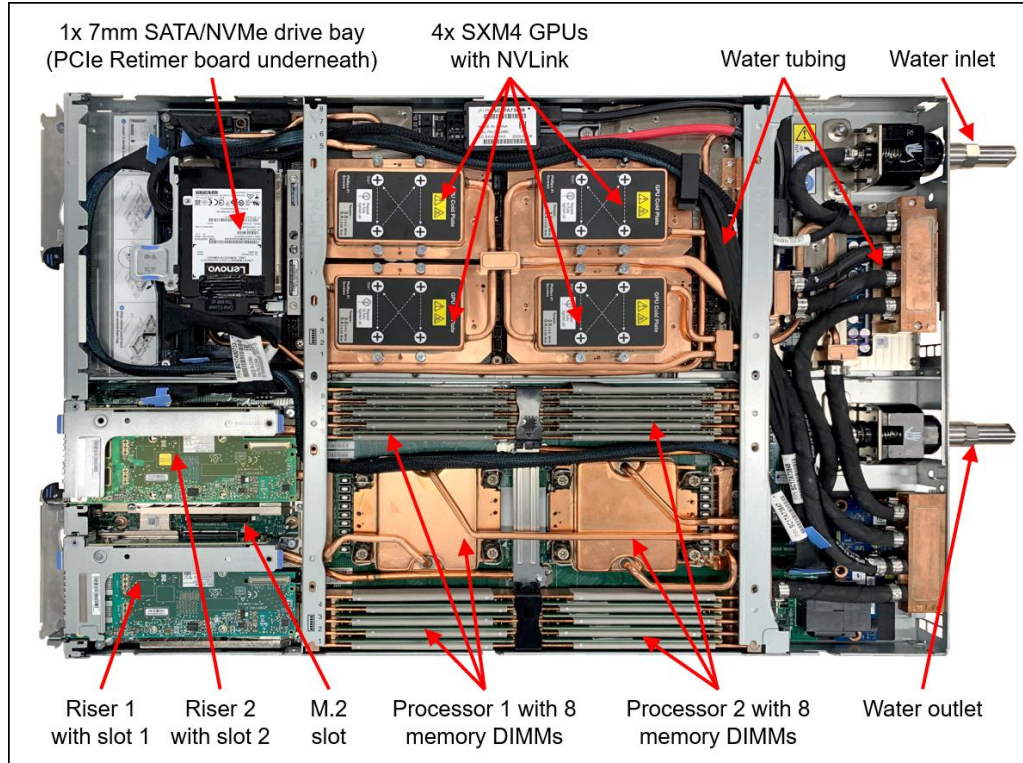
Lenovo SD650-N v2

Knoten	602	167
CPUs	2x Intel „Ice Lake“ 76 Core/152 Thread	2x Intel „Ice Lake“ 76 Core/152 Thread
Speicher	256/512 GB	512 GB
GPUs		4x NVIDIA A100
Interconnect	InfiniBand HDR200 Fat Tree	
Dateisystem	16 PB, ~150 GB/s GPFS	

HoreKa CPU Node



HoreKa GPU Node

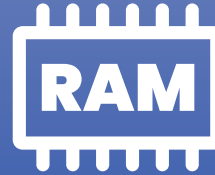


HoreKa - Hardware “Look-and-Feel”



Gute Balance
25% CPU / 75% GPU

CPU-Teil 4x ForHLR II



Großer Speicher
3,3 bis 6,7 GB per Core

1-2x ForHLR II



Hoher Durchsatz
Hauptspeicher 400 GB/s
GPU-Speicher bis 1 TB/s
Lokale SSDs bis 3 GB/s

2-6x ForHLR II

HoreKa - Software “Look-and-Feel”

Bekannte Umgebung

Red Hat Linux 8.x, SLURM, Modules, Parallele Dateisysteme...



Standard-Bibliotheken und -Tools

CPUs: Intel MKL, OpenBLAS, OpenMPI, MPI ...

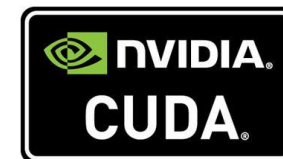
GPUs: CUDA, cuBLAS, cuFFT, cuDNN, NPP, OpenACC, MPI, ...



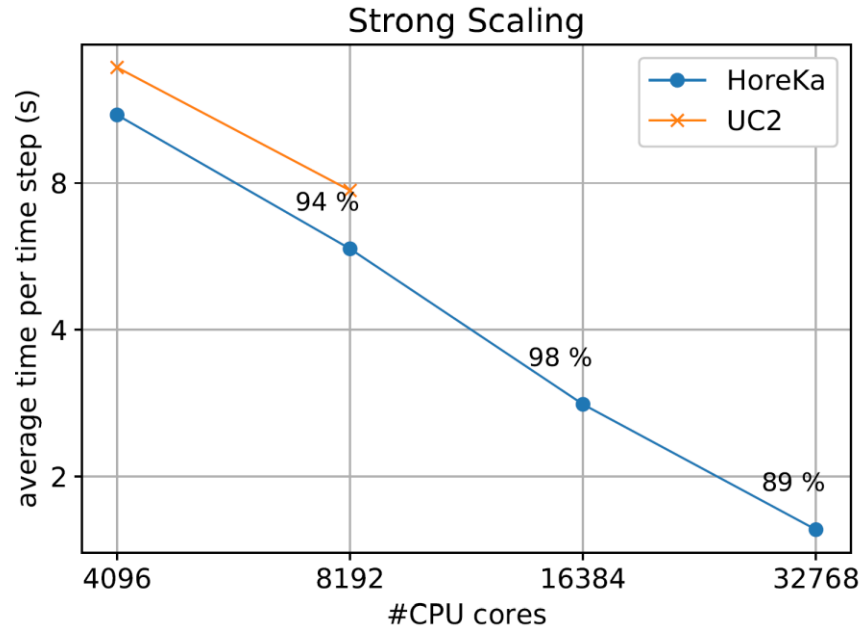
Standard-Compiler

CPUs: Intel, GCC, LLVM, NVIDIA

GPUs: NVIDIA, GCC, LLVM

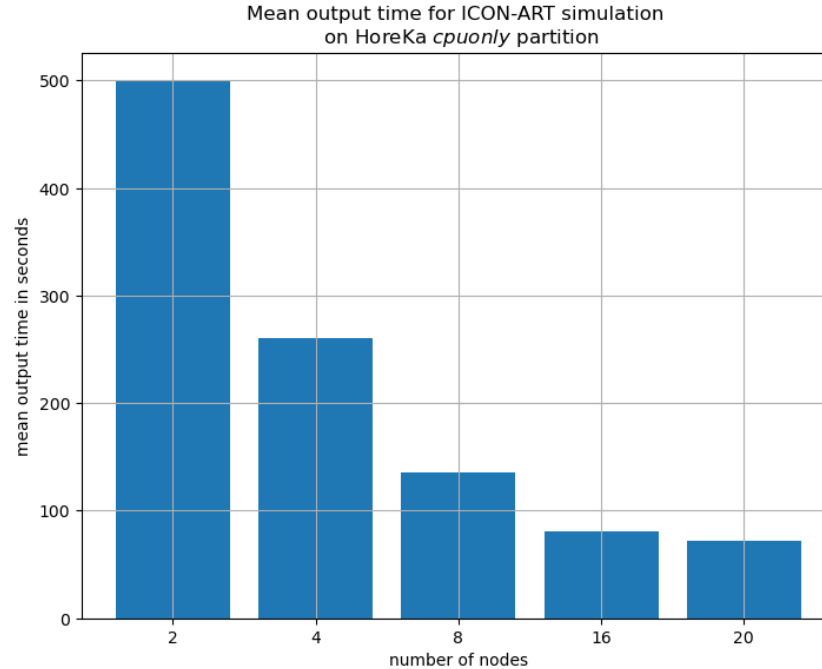


HoreKa – Skalierung EBI dnsFoam



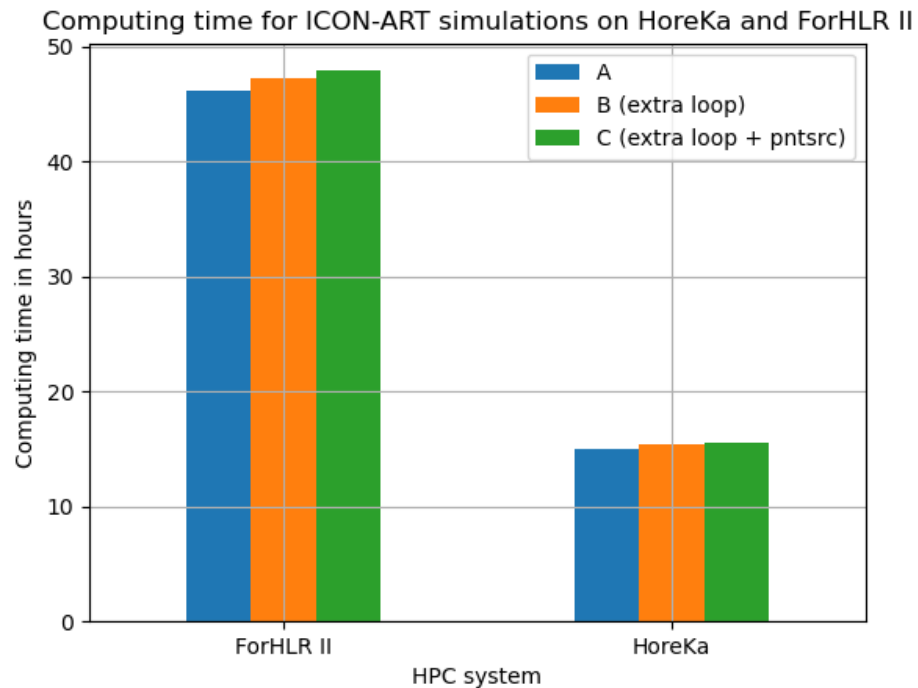
T. Zirwes, F. Zhang, P. Habisreuther, M. Hansinger, H. Bockhorn, M. Pfitzner, and D. Trimis, "Quasi-DNS dataset of a piloted flame with inhomogeneous inlet conditions," Flow, Turbulence and Combustion, vol. 104, pp. 997–1027, 2020 (<https://doi.org/1007/s10494-019-00081-5>)

HoreKa – Skalierung ICON-ART



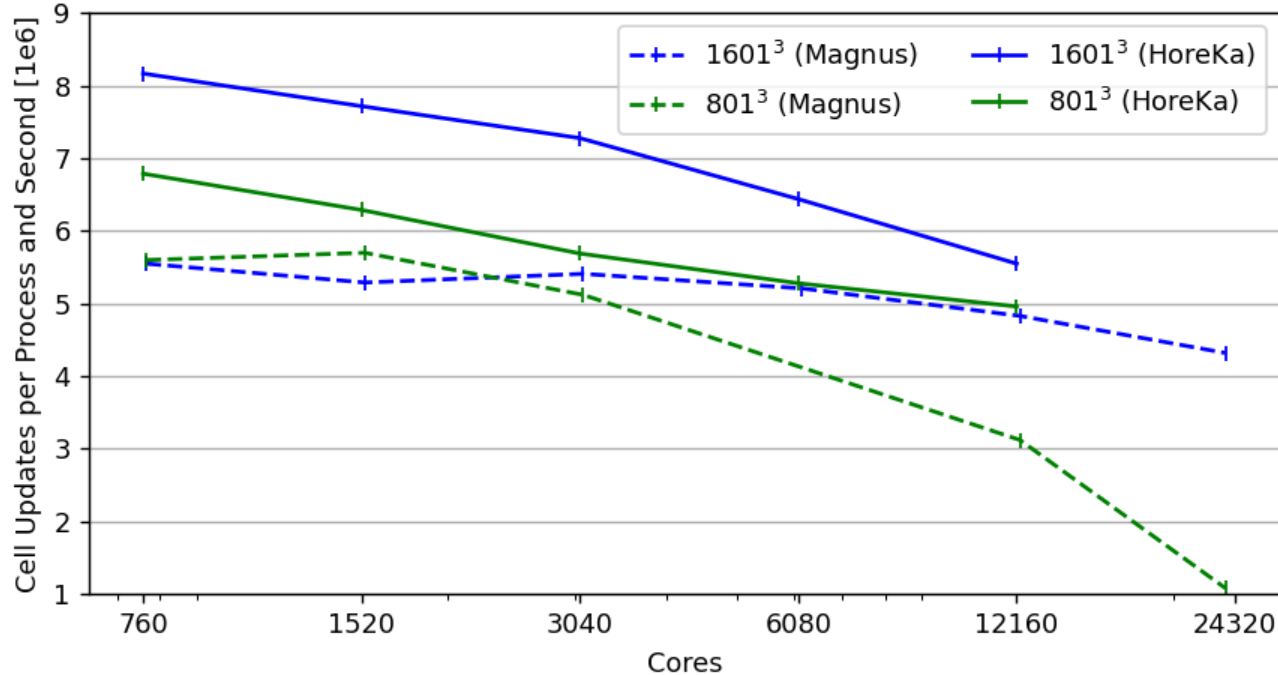
Graph: Christian Scharun, Institut für Meteorologie und Klimaforschung – Atmosphärische Spurengase und Fernerkundung (IMK-ASF), KIT

Vergleich HoreKa / ForHLR II: ICON-ART



Jeweils 20 Knoten pro System - Graph: Christian Scharun, Institut für Meteorologie und Klimaforschung – Atmosphärische Spurengase und Fernerkundung (IMK-ASF), KIT

HoreKa – Skalierung Kommunikation



Vergleich von HoreKa mit der Cray XC40 „Magnus“ am Pawsey Supercomputing Center – Graph: Adrian Kummerländer, Institut für Angewandte Numerik und Mathematik, KIT

Future Technology Partition



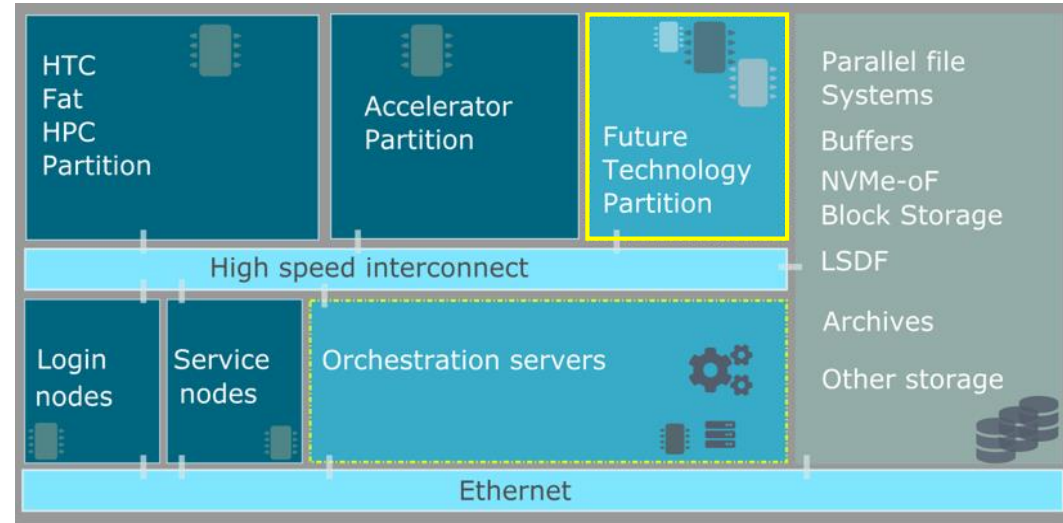
Future Technology Partition (FTP)

Effektive Unterstützung wissenschaftlicher Softwareentwicklung

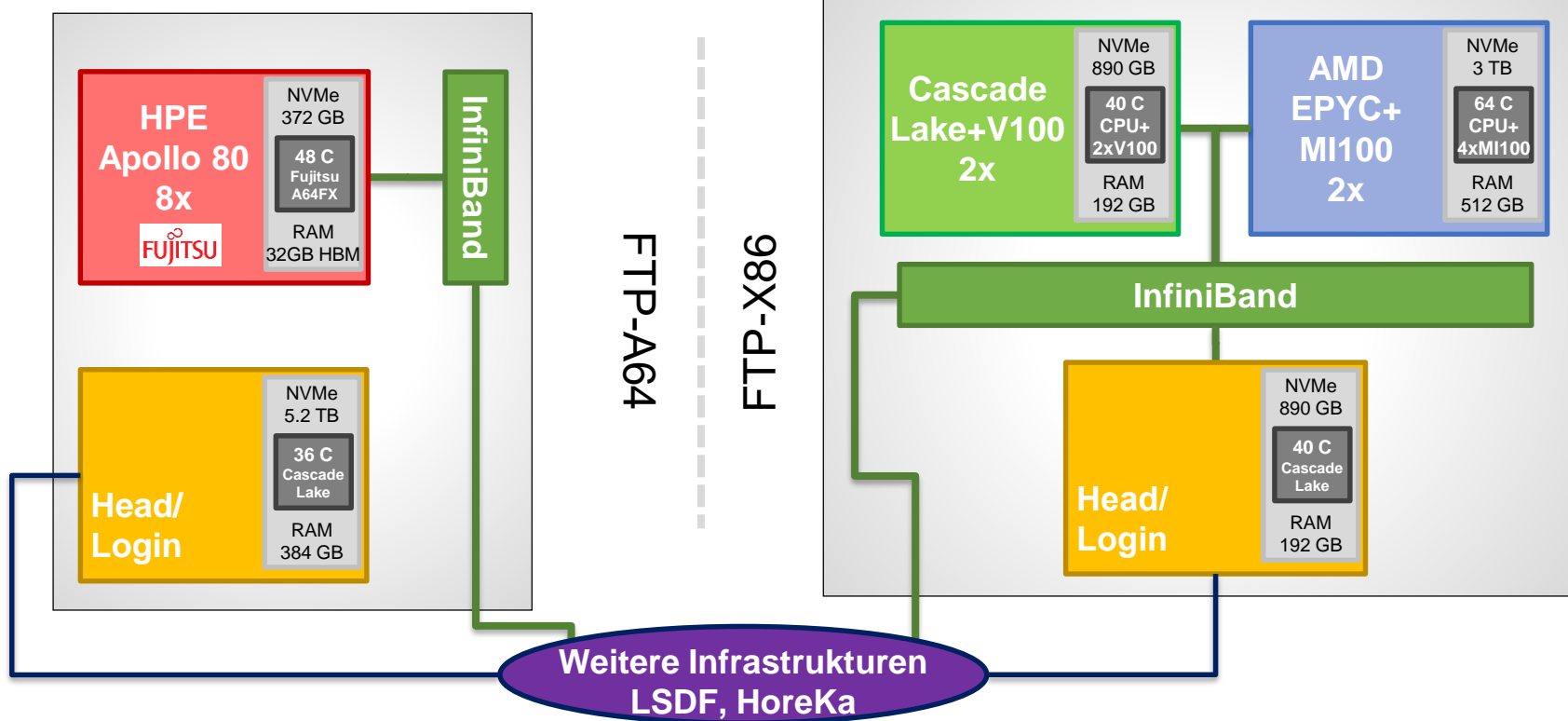
Zielgerichtete Hardware-Portierung

Beschleunigung der Entwicklungszyklen

Entwicklung effizienter Algorithmen, Bibliotheken und Anwendungen



FTP: Ist-Zustand



FTP: Geplanter Ausbau bis Ende 2021

GRAPHCORE IPU-POD16

Spezielle Beschleunigersysteme für KI/ML

NVIDIA HPC Dev Kits



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

ARM+AMD GPU



Evaluation ROCm-Softwarestack auf ARM
Ampere Altra Q80-30 CPU (80 Cores),
2x AMD MI100, 2x BlueField-2 DPU

ARM-Serviceknoten

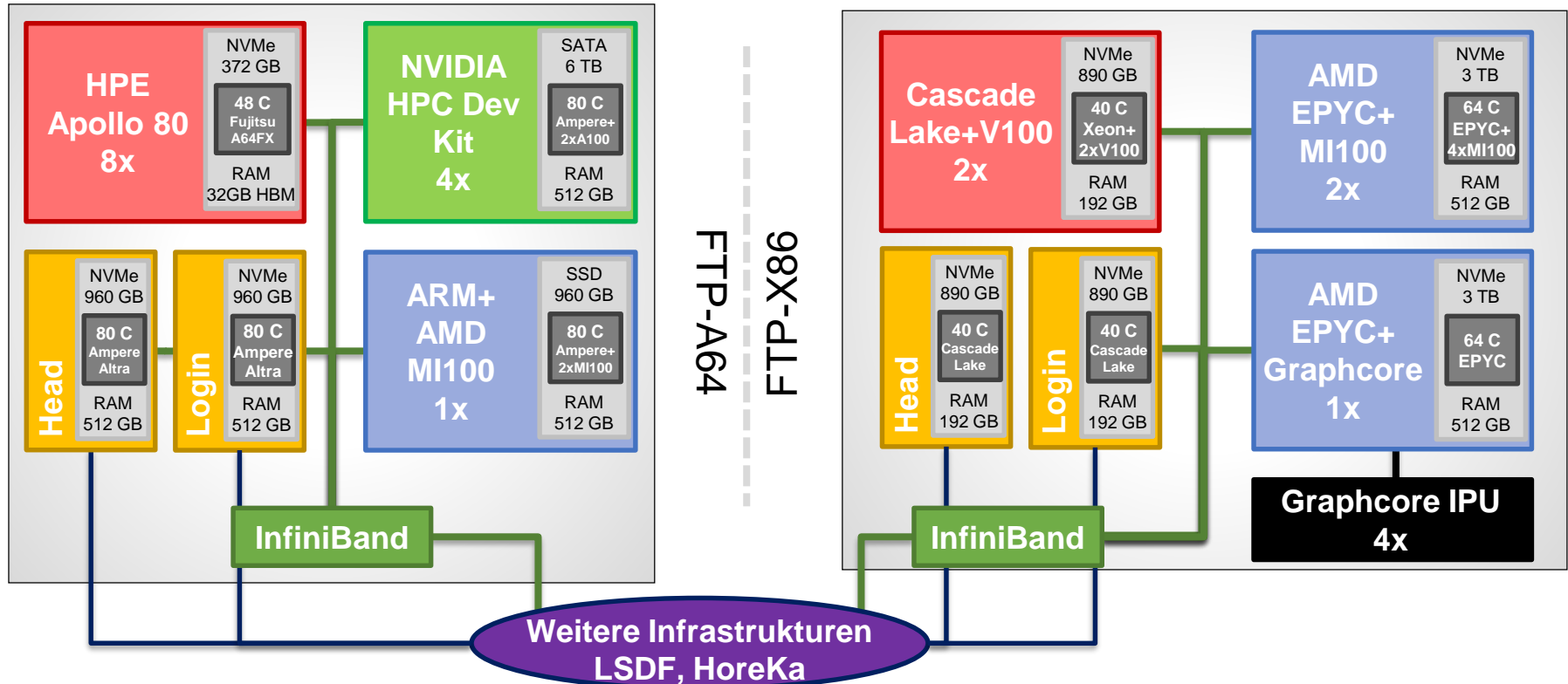


Ampere Altra Q80-30, BlueField-2 DPU
Ablöse Intel-Serviceknoten
Portierung KITE-Clusterstack

Graphcore IPU-M2000 POD16



FTP: Soll-Zustand Ende 2021



Pläne für 2022

**Intel
Sapphire
Rapids**



**Intel Xe/
Ponte
Vecchio**



AMD MI200



**Ampere
Altra Max/
Next-Gen**

