Programming Techniques for Supercomputers

Prof. Dr. G. Wellein[a,b] , Dr. G. Hager[a]
A. Afzal[a], D. Lacey[a], J. Laukemann[a]

[a] Erlangen National Center for High Performance Computing (NHR@FAU)
[b] Department für Informatik

FAU Erlangen-Nürnberg
Sommersemester 2025

Audience & Contact

- Audience
  - **Computational Engineering**
  - **Computer Science**
  - **Computational & Applied Mathematics (CAM)**
  - Physics, Engineering, Materials Science, Chemistry,…

- Contact:
  - Gerhard Wellein: gerhard.wellein@fau.de (Lecture&Tutorial)
  - Georg Hager: georg.hager@fau.de (Lecture&Tutorial)

- For questions first use the moodle forum (see next slide)

How to **register** for the **lecture** & **tutorials**

- Enroll into the course at our moodle LMS
  (login via "DFN AAI" and use your IdM credentials for SSO):

  https://go-nhr.de/PTfS

  After enrolment, fill out the poll:
  https://moodle.nhr.fau.de/mod/choice/view.php?id=3161

- Registration for exam (starts later): campo.fau.de

Format of lecture / exams

- PTfS: Lecture & Exercises: 7.5 ECTS
  - Written exam (90 Minutes)
    - Material covered in lecture AND tutorial
    - Register in campo for: **278169**

- PTFS-CAM: Lecture & Excercises & Additional programming project: 10 ECTS
  - Oral exam (30 minutes)
    - Material covered in lecture AND tutorial AND Programming Project
    - Register in campo for **58751**

- Both exams:
  - No additional materials allowed
  - Bonus system based on total credits of returned exercise sheets:
    - 60% - 79% an upgrade in the final mark of 1 stage, e.g. 2.7 → 2.3
    - 80% and higher upgrade in the final mark of 2 stage, e.g. 2.7 → 2.0
    - You must pass the exam before “boost” is applied! 1.0 is still best marks

Organization & Format

- 4 hrs of lecture (2 x 2 hrs) / week : !!! In person !!!
  - **Wednesday (8:30 – 10:00): H10**
  - **Thursday (8:30 – 10:00): H5**
  - Monday (14:15 – 15:45): H10 (**backup slot** – see moodle for lecture dates)

GW needs to travel several times during the semester → **backup slot**

- 2 hrs of tutorial per week – **choose one**
  - **Thursday (10:15 – 11:45) 02.133-113**

OR

  - **Friday (8:15 – 9:45) 00.151-113**

Tutorials kick off this week (April 16/17)
Attend one of them!

- CAM office hours: Friday 13:00 – 14:00 in the office 04.139-113
  starting **after** first project intro (will be announced on Moodle)

During lectures / tutorial: DON'T BE SHY AND ASK QUESTIONS!

Organization & Format: Lecture

- Lecture/Tutorial is completely documented in moodle LMS:

  https://go-nhr.de/PTfS (see also StudOn entry)

- Please enroll into the lecture and specify your matriculation number!
  - Homework assignments, announcements etc. all handled via moodle
  - Use forum for questions or discussion with other students

- Recordings of 2020 and 2021 PTfS lectures are available at https://www.fau.tv
  - https://www.fau.tv/course/id/2351 (2021 - Zoom lecture)
  - https://www.fau.tv/course/id/1233 (2020 – CAMTASIA recording)

Material has changed since then!!!!!

Organization & Format: Tutorials and Homework

- New homework assignments are released every Wednesday around 12:00 p.m. (noon)
- Report submission
  - Deadline: the week after on **Thursday at 10:00 a.m.** **No extensions!**
  - Deadline for Assignment 0: **Thursday, April 30**
  - Upload a single file in Moodle
    - PDF report or
    - compressed archive including a PDF report and supporting material
  - Grading will be done based on PDF report
- Problems? Questions?
  - Ask during tutorial Q&A
  - Use Moodle forum

Thursday / Friday Tutorial Sessions: Content

- Presentation of solution to previous assignment
- Presentation of current (new) assignment
- !!! Opportunity to ask questions !!!
- Strongly recommended to attend the tutorial weekly

Organization & Format: **PTfS-CAM**

- CAM students have to do a programming and performance modeling project:
- Shared memory parallel Preconditioned Conjugate Gradient (PCG) solver

- Extra events for PTfS-CAM students:
  - Basic shell introduction
  - Introduction to C programming
  - Introduction to project work (OpenMP parallelization)
  - Introduction to project work (performance modelling)
  - PTfS-CAM-Consultation hour every week (Jan / Dane)
  - Dates & times will be announced in moodle → SEE MOODLE

Can also be attended by other students

Format of the course

- Prerequisite (for exercises):
  - Basic programming knowledge in C/C++ or FORTRAN
  - Using LINUX / UNIX OS environments (including ssh)

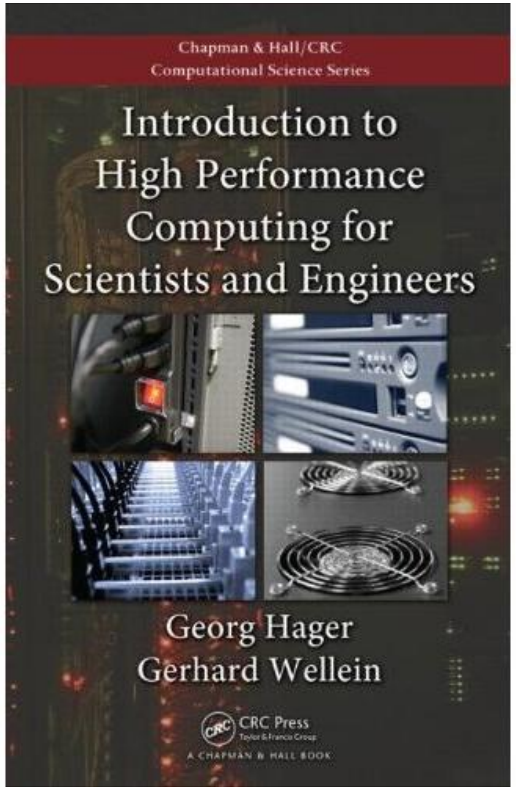
- Recommended
  - First experiences with parallel programming – though we will introduce necessary basics

- Tutorials:
  - You **must** submit your homework assignments for inspection to qualify for bonus!
  - Do the exams and programming exercises even if you do not submit!
  - Some topics will be covered (in more detail) in the tutorials
  - Topics additionally covered in the tutorials are part of the 7.5 ECTS exam
  - Practical parallelization skills will be tested in the 7.5 ECTS exam!

Supporting material

- Books:
  - G. Hager and G. Wellein:
    Introduction to High Performance Computing for Scientists and Engineers.
    CRC Computational Science Series, 2010.
    ISBN 978-1439811924

    - 10 copies are available in the library
    - discounted copies – ask us



  - J. Hennessy and D. Patterson: *Computer Architecture. A Quantitative Approach.* Morgan Kaufmann Publishers, 2017. ISBN 978-0128119051
  - W. Schönauer: *Scientific Supercomputing*.
    (cf. http://www.rz.uni-karlsruhe.de/~rx03/book/)

Supporting material

- Documentation:
  - http://www.openmp.org
  - http://www.mpi-forum.org
  - http://developer.intel.com/products/processor/manuals
  - http://developer.amd.com/documentation/guides

- The big ones and more useful HPC related information:
  - http://www.top500.org

Related teaching activities

- Regular seminar on

  "Efficient numerical simulation on multicore processors" (**MuCoSim**)
  - 5 ECTS
  - 2 hrs per week
  - 2 talks + written summary

  - Topics from code optimization, code parallelization and code benchmarking on latest multicore / manycore CPUs and GPUs

  - This semester: Wednesday 16:00 – 17:30

  - See moodle: http://go-nhr.de/MuCoSim

  - Also offered during winter term

Related teaching activities

- Lecture on

Practical parallel algorithms with MPI (**PAMPI**)

- Winter term
- 5 ECTS
- 2 hrs of lecture
- 2 hrs of tutorials

- Lecturer: Dr. Jan Eitzinger

Scope of the lecture

Scope of the lecture

Understand relevant hardware features/concepts of modern HPC systems and derive efficient parallelization & implementation strategies

→Identify **basic hardware concepts** (CPU & GPU)
single core, chip/device, node-level

→Learn strategies how to **efficiently use (program) the hardware**
(code transformations/optimizations, OpenMP/CUDA)

→Establish appropriate **performance expectations/models** to

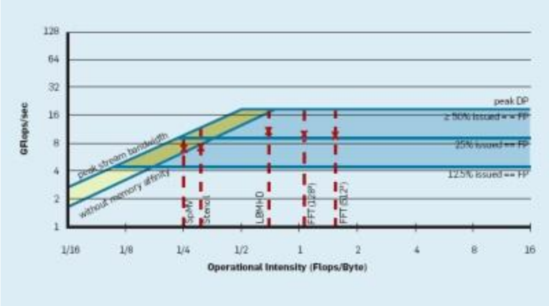
- →assess the attainable performance and
- →identify the “hardware bottleneck”

Scope of the lecture

- Hardware coverage:
  - Single-core + Multi-Core — Intel Xeon & AMD EPYC & ARM-based
  - Many-core / GPU: — NVIDIA A100
  - Shared memory nodes — (multiple multi-/many-core & GPUs)
  - Distributed memory computers — (multiple/many nodes)

- Programming models (mostly basic introduction; performance issues):
  - OpenMP — (shared memory nodes)
  - CUDA — (GPUs)
  - MPI — (distributed memory) → **PAMPI lecture**

- Performance Modelling
  - Expectations
  - Roofline Model [Williams&Patterson, 2009]
  - (ECM Model)



Structure of the lecture

- Introduction
  - Performance: Basics, Measuring & Reporting, Benchmarks
- Modern (multicore) processors
  - Single core: Basics, Pipelining, Superscalarity, SIMD
  - Memory Hierarchy: Caches & Main Memory
  - Multicore: Technology & Basics
  - GPU
- Parallel computers: Shared Memory
  - Shared-memory system architectures: UMA, ccNUMA
  - OpenMP basics
- Performance Modelling / Engineering:
  - Roofline Model
  - Case Studies: Dense&Sparse Matrix-Vector-Multiplication / Stencils / …
- Shared Memory in depth
  - Advanced OpenMP, Pitfalls, Data Placement
- Hardware performance monitoring and model validation → LIKWID

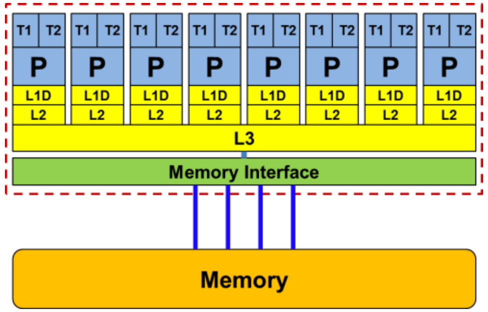
Performance Analysis and Modeling

Scope of the lecture – a typical example

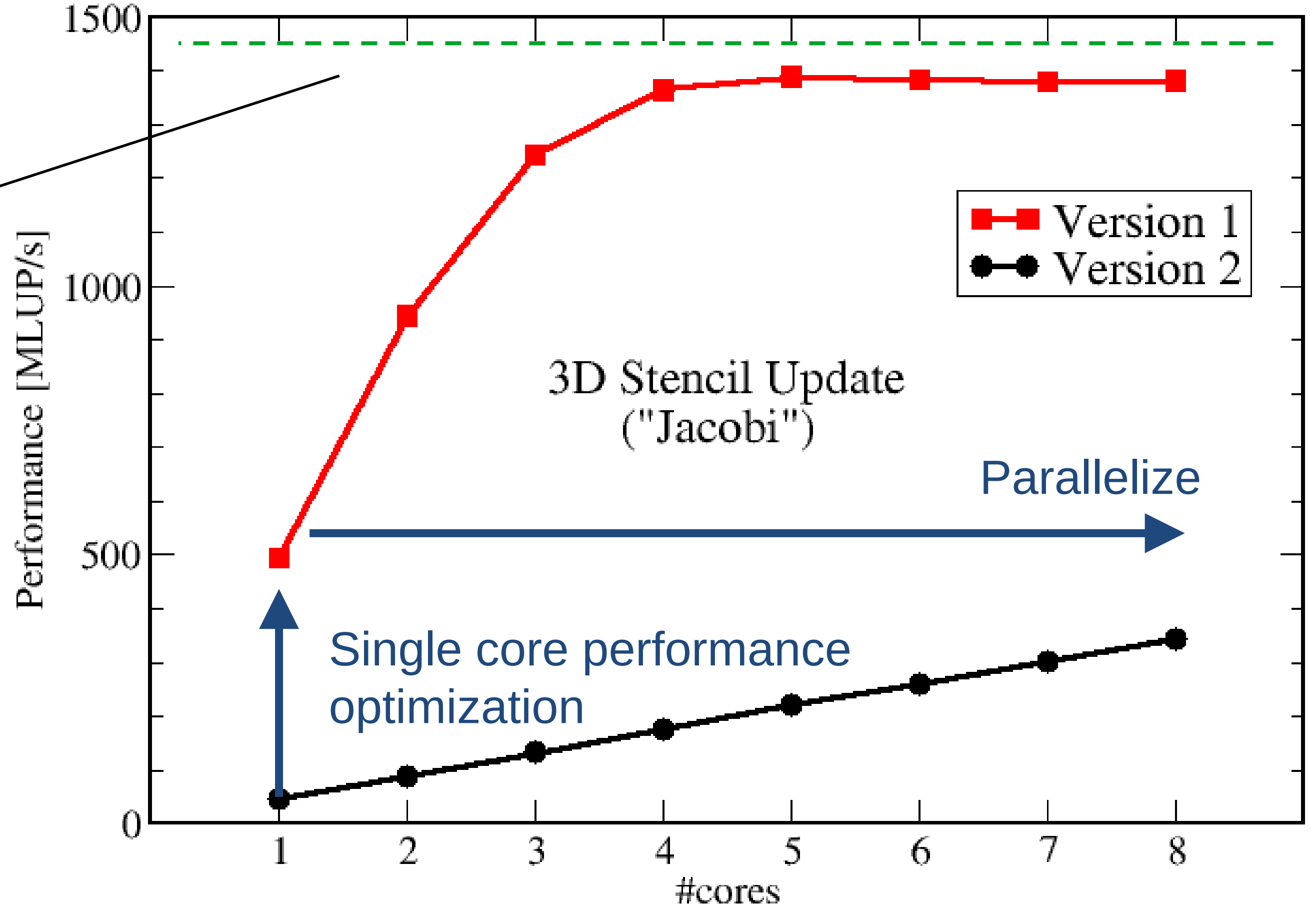
```
!$OMP PARALLEL DO
do k = 1 , 400
  do j = 1 , 400; do i = 1 , 400
     y(i,j,k)= b*(  x(i-1,j,k)+ x(i+1,j,k)+ x(i,j-1,k)+
                    x(i,j+1,k)+ x(i,j,k-1)+ x(i,j,k+1))
    enddo; enddo
enddo
!$OMP END PARALLEL DO
```

Parallelize

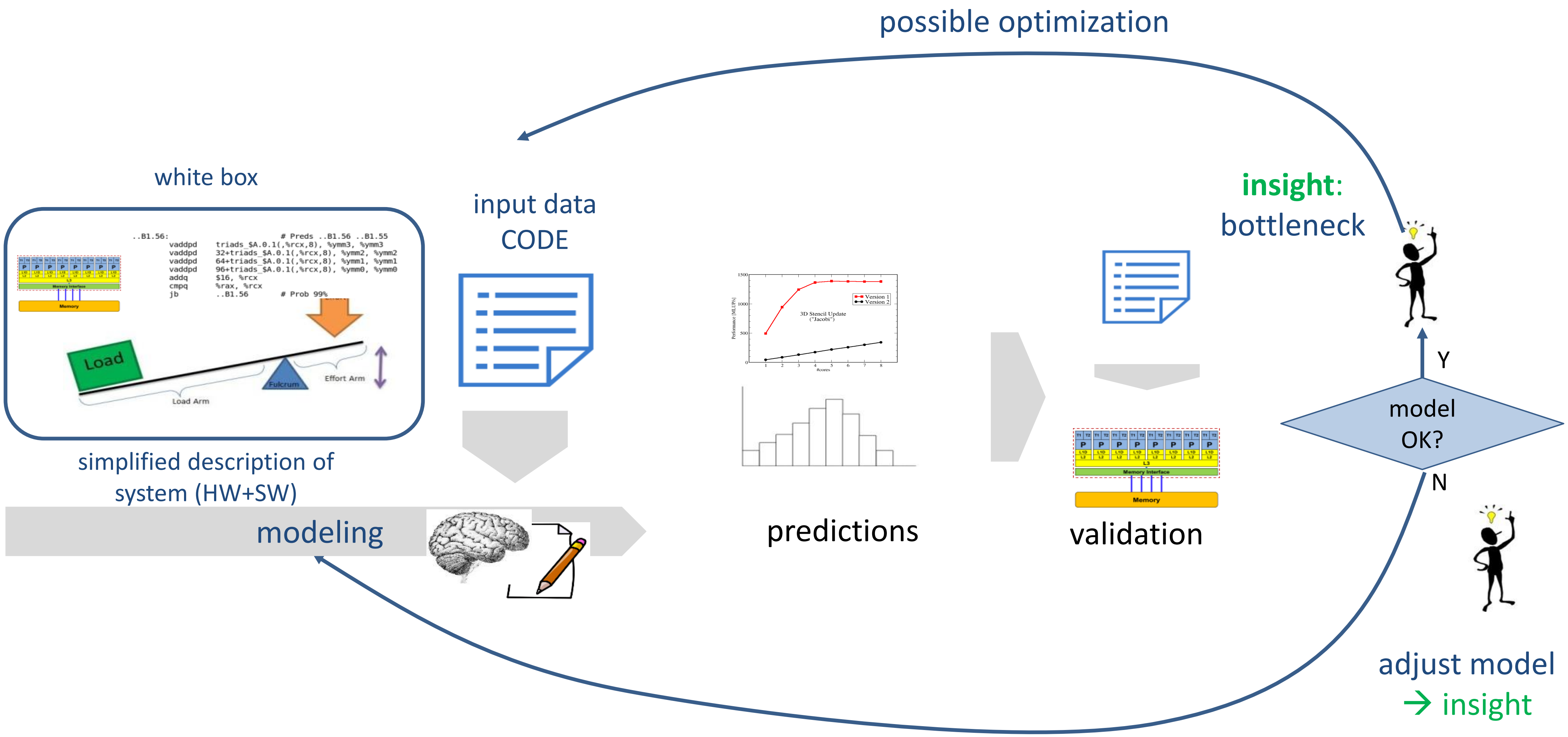
Establish limit simple performance model (Here: Roofline)



Intel® Xeon® Prozessor E5-2670



Code optimization/parallelization – no black boxes!



„Performance Engineering“

Introduction

Supercomputers:
The Big Ones and the working horses

Most powerful computers in the world: TOP500

- Top 500: Survey of the 500 most powerful supercomputers
  - http://www.top500.org
  - How to rate the performance?

  → Solve large dense system of equations: `A x = b` („LINPACK“)

  - Max. performance achieved with 64-Bit Floating Point Numbers: R_{max}
  - Published twice a year (ISC in Germany, SC in USA)
  - First: **1993** (CM5/1,024 procs.): 60 GFlop/s (TOP1)
  - **Nov. 2024** (HPE – AMD MI300A GPUs): 1,742,000,000 GFlop/s (TOP1)
  - Performance increase: 74% p.a. from 1993 – 2024

Most powerful computers in the world: TOP500

- Performance measures: MFlop/s, GFlop/s, TFlop/s, PFlop/s, EFlop/s
  - Number of FLOATING POINT operations per second
  - FLOATING POINT operations: Typically Add & Multiply operations (→Chapter 3 of the lecture)
  - Performance may depend on accuracy (of input operands):
    - double precision, double: 64 Bit
    - single precision, float: 32 Bit
    - half precision: 16 Bit
    - **default: 64 Bit**
    - See Chapter 3 for details

10^6: MFlop/s	10^{12} :**T**Flop/s	← Single node or GPU
10^9: GFlop/s	10^{15}: **P**Flop/s	← TOP500 systems
	10^{18}: **E**Flop/s	← TOP1 in 2023

TOP3 as of November 2024

Number of processors

LINPACK Performance

Peak Performance

Power@ LINPACK

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**El Capitan** - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	11,039,616	1,742.00	2,746.38	29,581
2	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS, HPE DOE/SC/Oak Ridge National Laboratory United States	9,066,176	1,353.00	2,055.72	24,607
3	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	9,264,128	1,012.00	1,980.01	38,698

Investment for such systems: > 200 M€

Power bill @ 30ct/kWhr: 1 MW ←→ € 2,500,000 p.a.

Source: www.top500.org

TOP10 as of November 2024

Leading systems: US (5 entries) –
National Labs and Industry

Four systems from Europe in Top 10: 5, 7, 8, 9

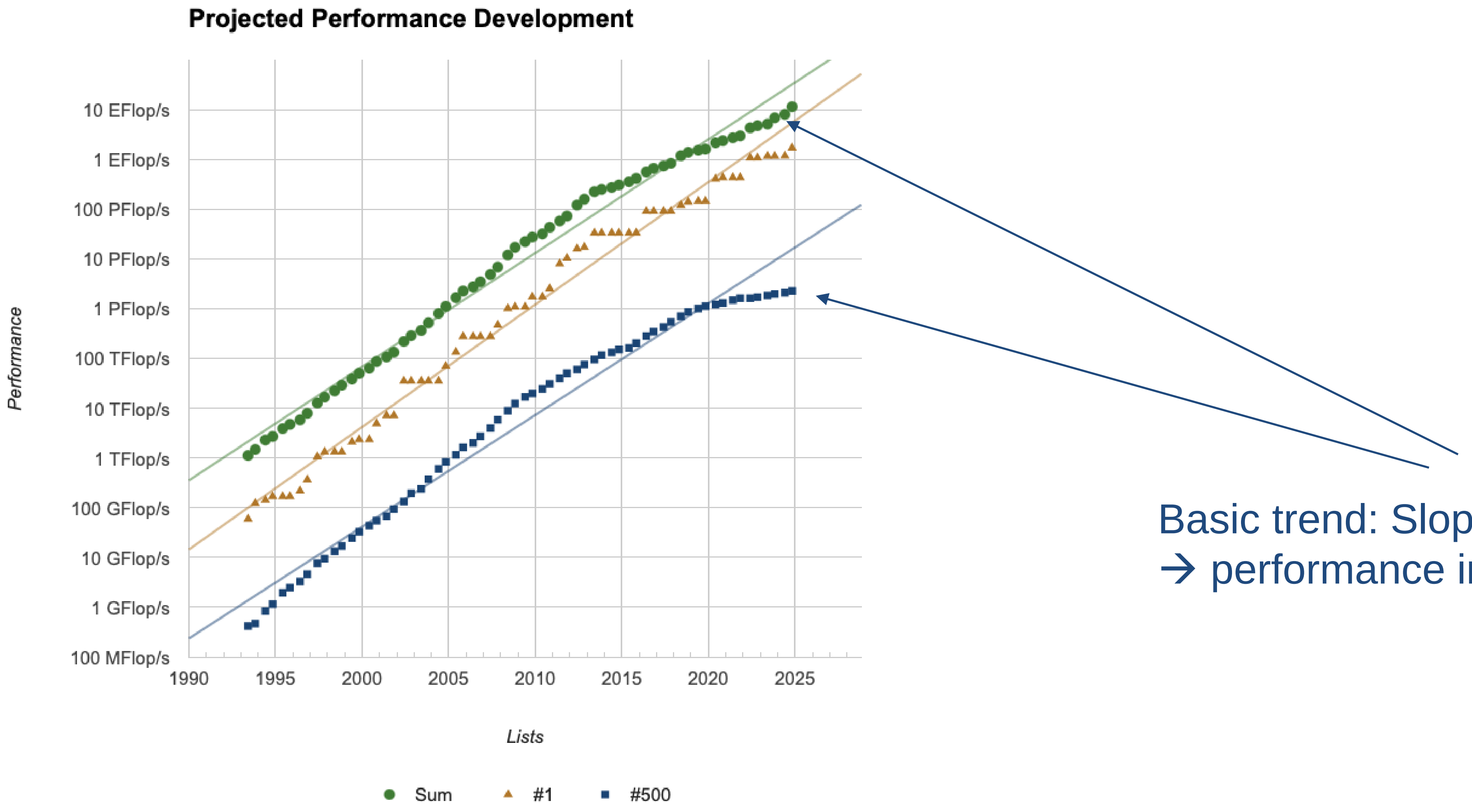
Germany:
JETI - JUPITER Exascale Transition Instrument @FZJ Jülich:
#18 → Transition system to first EFlop/s system in Europe

Only a single CPU based system left

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**El Capitan** - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	11,039,616	1,742.00	2,746.38	29,581
2	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS, HPE DOE/SC/Oak Ridge National Laboratory United States	9,066,176	1,353.00	2,055.72	24,607
3	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	9,264,128	1,012.00	1,980.01	38,698
4	**Eagle** - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Azure Microsoft Azure United States	2,073,600	561.20	846.84	
5	**HPC6** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, RHEL 8.9, HPE Eni S.p.A. Italy	3,143,520	477.90	606.97	8,461
6	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
7	**Alps** - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cray OS, HPE Swiss National Supercomputing Centre (CSCS) Switzerland	2,121,600	434.90	574.84	7,124
8	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,752,704	379.70	531.51	7,107
9	**Leonardo** - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, EVIDEN EuroHPC/CINECA Italy	1,824,768	241.20	306.31	7,494
10	**Tuolumne** - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	1,161,216	208.10	288.88	3,387

Source: www.top500.org

Performance Trend & Projection (2024)



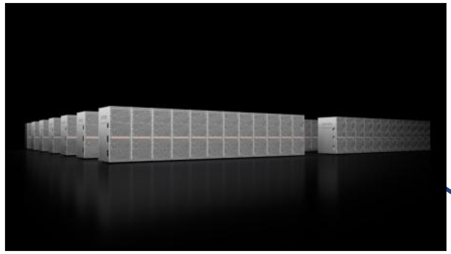
Source: www.top500.org

Basic trend: Slope changes
→ performance increase slows down

HPC Centers in Germany: A view from Erlangen

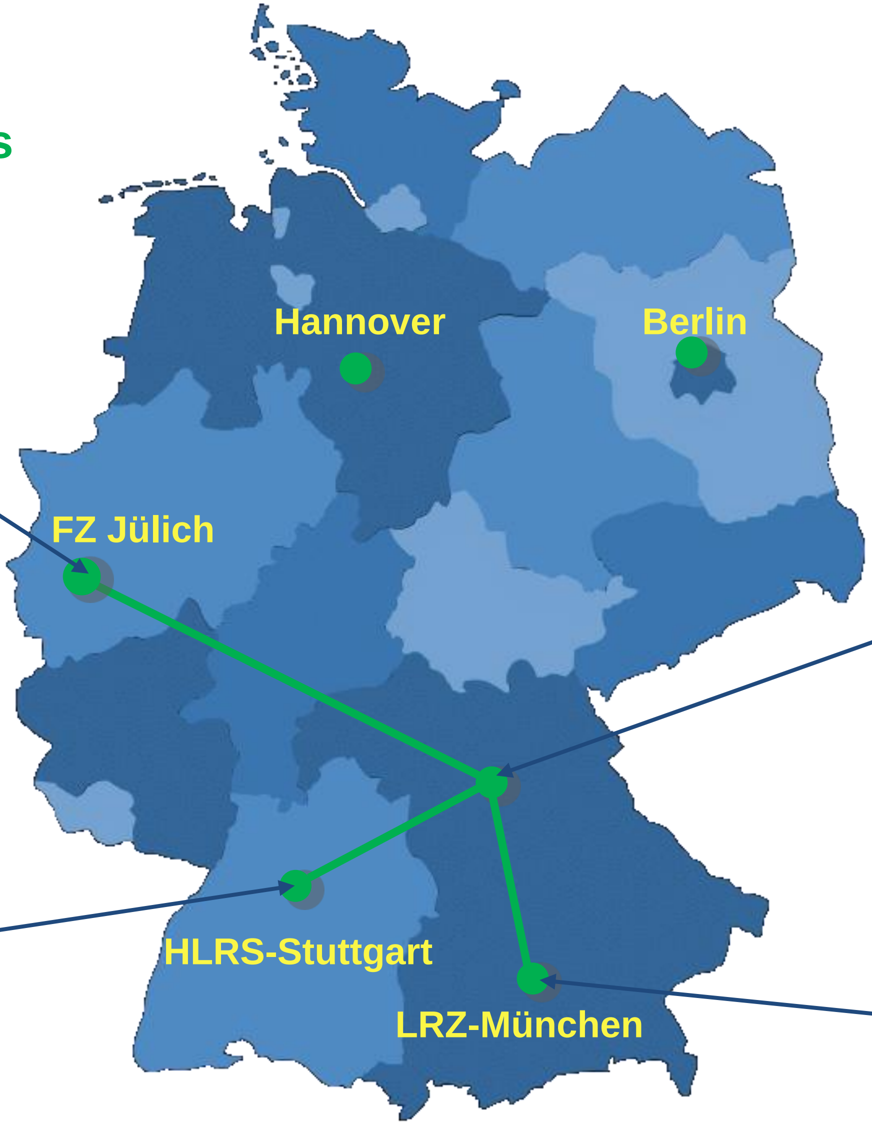
Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

Expecting JUPITER – 1,000 PF/s



HLRS

HLR Stuttgart
Hawk: 26 PF/s
New system 25/26



NHR FAU

FAU Erlangen/-Nürnberg
Helma (50 PF/s)
Fritz (5 PF/s) & Alex (5 PF/s)

lrz

SuperMUC-NG: 26.8 PF/s
New System 2027

Erlangen National Center for High Performance Computing

www.nhr.fau.de

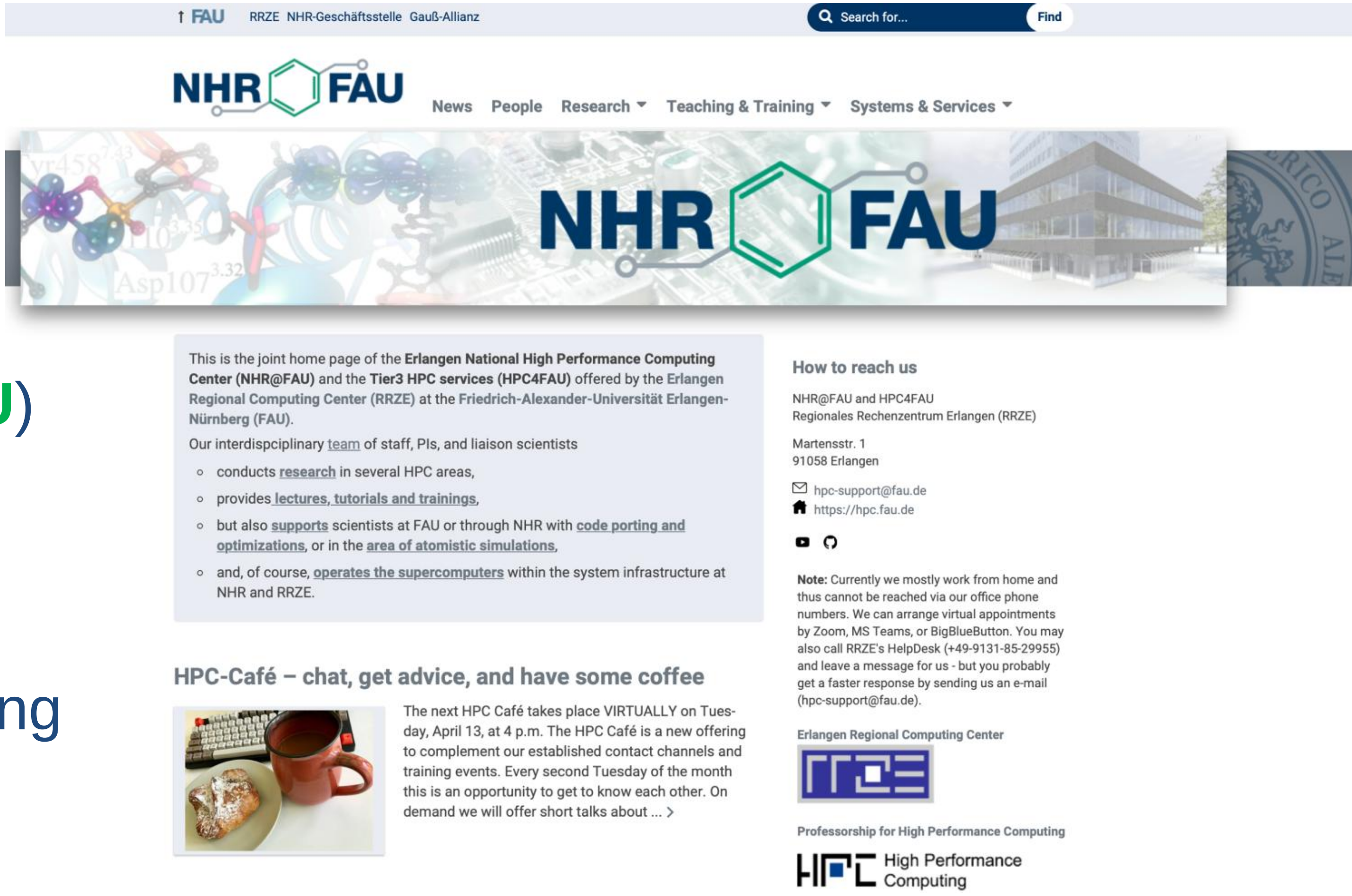
One out of nine national HPC centers at German universities

Ongoing:

- New system installed (**Helma-GPU**)
- New HPC center - 2030

Focus topics:

- Node-level Performance engineering
- Atomistics simulations
- AI/ML



Student positions, master theses, doctoral positions available

Fritz & Alex: Fact Sheet

	#nodes	Node conf.	Storage	Typical job sizes	Peak (FP64)
Fritz	992 Intel ICL (71,424 cores)	2 * 36 c (8360Y) 256 GB 1 x HDR100	Shared PFS • 3 PB • >20 GB/s	1 – 64 nodes	5.9 PF/s (4.1 PF/s)
	64 Intel SPR (6,656 cores)	2 * 52 c (8470) **1 TB / 2 TB** 1 x HDR100		1 – 4 nodes	n.y.a.

178	**Fritz** - Megware D50TNP, Xeon Platinum 8360Y 36C 2.4GHz, InfiniBand HDR100, MEGWARE Universitaet Erlangen - Regionales Rechenzentrum Erlangen Germany	71,424	3.58	5.45	672

Power consumption (kW) for LINPACK

Fritz: https://doc.nhr.fau.de/clusters/fritz/

Fritz & Alex: Fact Sheet

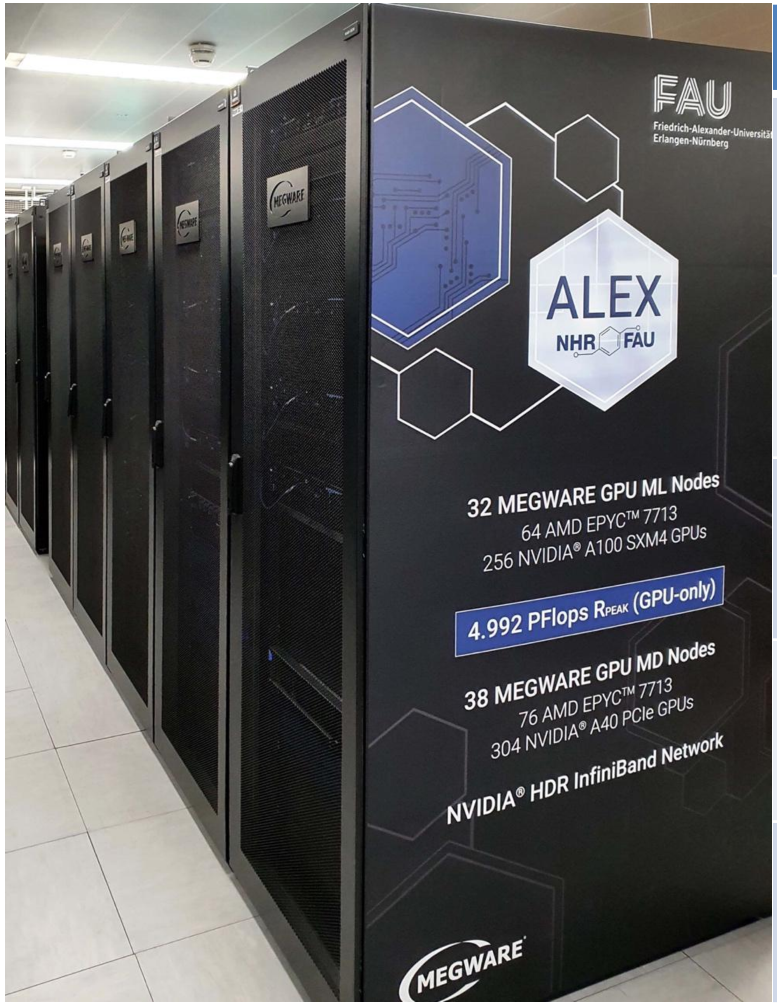
	#nodes	Node conf.	Storage	Typical job sizes	Peak (FP64)
Alex	38 AI/ML (**304 NVIDIA A100**) **(+320 H100 in 2024)**	8 * NVIDIA A100 2 * 64 c (AMD) 1 TB 2 x HDR200	Node local 14 TB NVMe	1 – 8 GPUs	6.1 PF/s
	44 MD (**352 NVIDIA A40**)	8 * NVIDIA A40 2 * 64 c (AMD) 0.5 TB	Node local 7 TB NVMe	1 – 8 GPUs	---

157	**Alex** - MEGWARE NF5488A5, AMD EPYC 7713 64C 2GHz, NVIDIA A100 SXM4 80 GB, Infiniband HDR, MEGWARE Universitaet Erlangen - Regionales Rechenzentrum Erlangen Germany	37,696	4.03	6.08	179

Power consumption (KW) for LINPACK (A100)

Alex: https://doc.nhr.fau.de/clusters/alex/

HPC compute infrastructure – TOP500



Alex	List	Fritz
#184 2.9 PF/s	06/22	**#323** 2.2 PF/s
#174 3.2 PF/s	11/22	**#151** 3.6 PF/s
#157 4.0 PF/s	06/23	**#178** 3.6 PF/s
#187 4.0 PF/s	11/23	**#213** 3.6 PF/s
#250 4.0 PF/s	11/24	**#281** 3.6 PF/s

RRZE-Serverraum

NatFak Kältezentrale

Helma+: Fact Sheet

Helma: **Wilhelmine von Brandenburg-Bayreuth**
(* 3. Juli 1709 in Berlin; † 14. Oktober 1758 in Bayreuth)

	#nodes	node configuration			job sizes	peak (FP64)
Helma — Q3/24	96 AI/ML (384 NVIDIA H100)	4 * H100 94 GB 2 * 64 c (AMD) 4 x NDR200	0.7 TB DDR 15 TB NVMe	NDR Infiniband 5PB NVME Storage	1–96 GPUs	26 PF/s
Q1/25	96 AI/ML (384 NVIDIA H200)	4 * H200 141 GB 2 * 64 c (AMD) 4 x NDR200	0.7 TB DDR 15 TB NVMe		1–96 GPUs	26 PF/s
Q3/25	*312 CPU* *(119,808 cores)*	*2 * 192 c (AMD)* *0.7 – 2.1 TB DDR* *1 x NDR200*	---		*1–96 nodes*	

79 **Helma** - Lenovo ThinkSystem SD665-N V3, AMD EPYC 9554 64C 3.1GHz, Nvidia H100 SXM5 94Gb, Infiniband NDR200, AlmaLinux 9.4, MEGWARE 62,976 16.94 25.96

TOP500 11/24

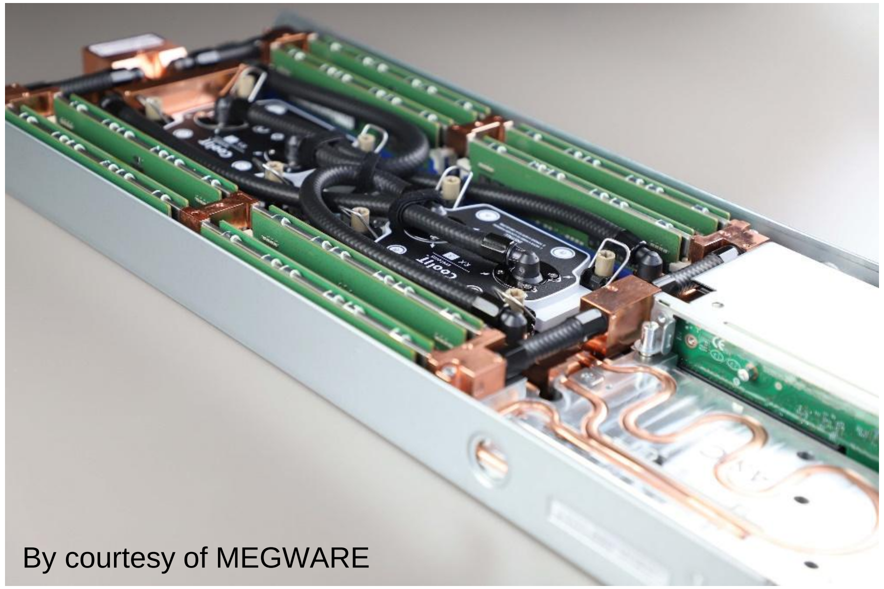
NHR@FAU: Infrastructure requirements

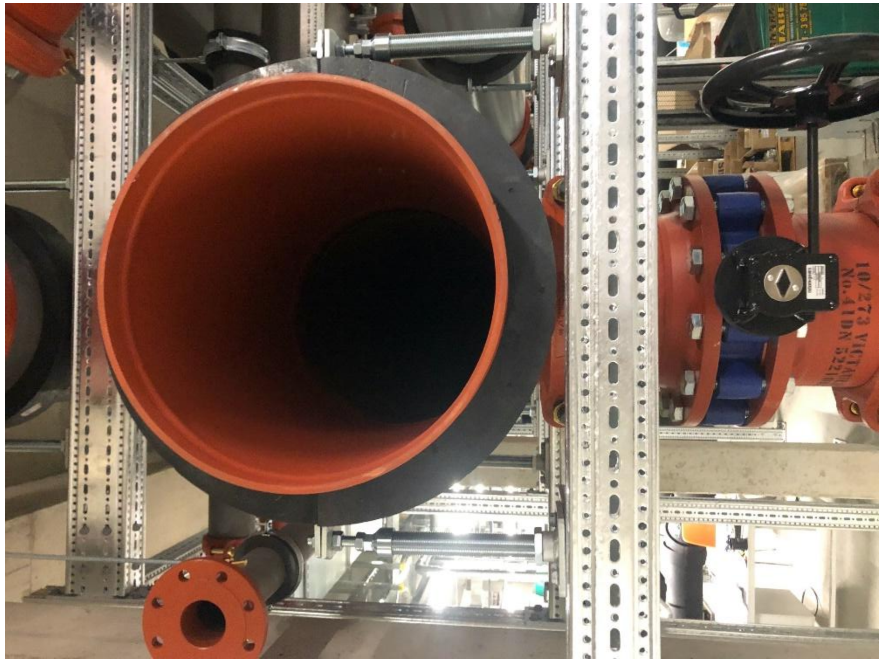
Helma-GPU: 384 H100 + 384 H200 – 700 KW at work

NF-Technikzentrale: “data center” for Fritz & Helma

NHR FAU

By courtesy of MEGWARE


By courtesy of MEGWARE



Lecture plan until end of May (1)

- **24.04.2025: Lecture**
- **25.04.2025: Lecture + tour to Helma (backup slot)**

- **29.04.2025: JE: Introduction to C**
- **01.05.2025: PUBLIC HOLIDAY**
- 02.05.2025: 8:15-9:45 **and** 10:15-11:45 Tutorial + intro to Linux Shell

- **06.05.2025: Lecture**
- **08.05.2025: Lecture**
- 09.05.2025: backup slot

Lecture plan until end of May (2)

- **13.05.2025: Georg Hager – GW on travel**
- **15.05.2025: Lecture**
- **16.05.2025: Lecture**

- **20.05.2025: Lecture**
- **22.05.2025: Lecture**
- 23.05.2025: No Lecture (backup slot)

- **27.05.2025: Lecture** (DAREXA-Workshop at FAU)
- **29.05.2025: Public Holiday**
- 30.05.2025: (backup slot)

Lecture plan until end of June

- **03.06.2025: Lecture**
- **05.06.2025: Lecture**
- **06.06.2025: Lecture**

- **10.06.2025: Public Holiday**
- **12.06.2025: Sebastian Kuckuck (GPU)**
- **13.06.2025: Sebastian Kuckuck (GPU)**

- **17.06.2025: Lecture (17)**
- **19.06.2025: Public Holiday**
- 20.06.2025: No Lecture (backup slot)

Lecture plan June / July

- **24.06.2025: Lecture**
- **26.06.2025: Lecture**
- 27.06.2025: No Lecture (backup slot)

- **01.07.2025: Lecture**
- **03.07.2025: Lecture**
- 04.07.2025: No Lecture (backup slot)

- **08.07.2025: Lecture – Vertretung: Berlin Strategieworkshop**
- **10.07.2025: Lecture (23)**
- 11.07.2025: No Lecture (backup slot) – Montag nachholen?

Lecture plan until end of June

- **15.07.2025: Lecture**
- **17.07.2025: Lecture (24)**
- 18.07.2025: No Lecture (backup slot)

- **22.07.2025: Lecture**
- **24.07.2025: Lecture (FAQ)**
- 25.07.2025: No Lecture (backup slot)