Carl von Ossietzky
Universität
Oldenburg

BI Betriebseinheit für technisch-wissenschaftliche Infrastruktur

Introduction to High-Performance Computing

Session 01
Course Organization and Introduction

Carl von Ossietzky Universität Oldenburg

but before we begin…

Course Organisation

- Contact Information
- Outline of the Course

Contact Information

Stefan Harfst (Scientific Computing)

email: stefan.harfst@uni-oldenburg.de

office: W03 1-139

web: http://www.uni-oldenburg.de/fk5/wr/

Wilke Trei (ForWind)

email: wilke.trei@forwind.de

office: W33 3-323

Course Organisation

Monday	Tuesday	Wednesday
10:15 – 11:15 (SH) Introduction to HPC	10:15 – 11:15 (WT) Introduction to Parallel Programming	10:15 – 11:15 (SH) Introduction to Matlab Distributed Computing (MDCS)
11:30 – 13:00 (WT) Basic Cluster Usage	11:30 – 13:00 (WT) Introduction to OpenMP	11:30 – 13:00 (SH) Parallel Programming in Matlab
13:45 – 15:15 (SH) HPC Environment	13:45 – 15:15 (SH) Introduction to OpenMP	13:45 – 15:15 (SH) Parallel Programming in Matlab
15:30 – 17:00 (SH) Advanced Cluster Usage	15:30 – 17:00 (SH) Performance Optimization	

- times will probably change a little bit

- lectures and practical parts (on your own)
- slides and files for practical parts will be available in the Wiki, see https://wiki.hpcuser.uni-oldenburg.de/index.php?title=HPC_Introduction

Carl von Ossietzky Universität Oldenburg

Hands-on Sessions

- we will be using the local HPC clusters CARL and EDDY
- if you don‘t have an account yet go to http://www.uni-oldenburg.de/fk5/wr/
  - there are useful links to request an account and to the HPC wiki, which is the central resource for information
  - in the form to request your account select the research group or “student” if you are not yet in a group (you can change your group later if needed by the same process)

Carl von Ossietzky Universität Oldenburg

Introduction to High-Performance Computing

- Supercomputers and Parallel Architectures
- HPC Resources in Germany and Europe
- HPC at University Oldenburg

What is High-Performance Computing?

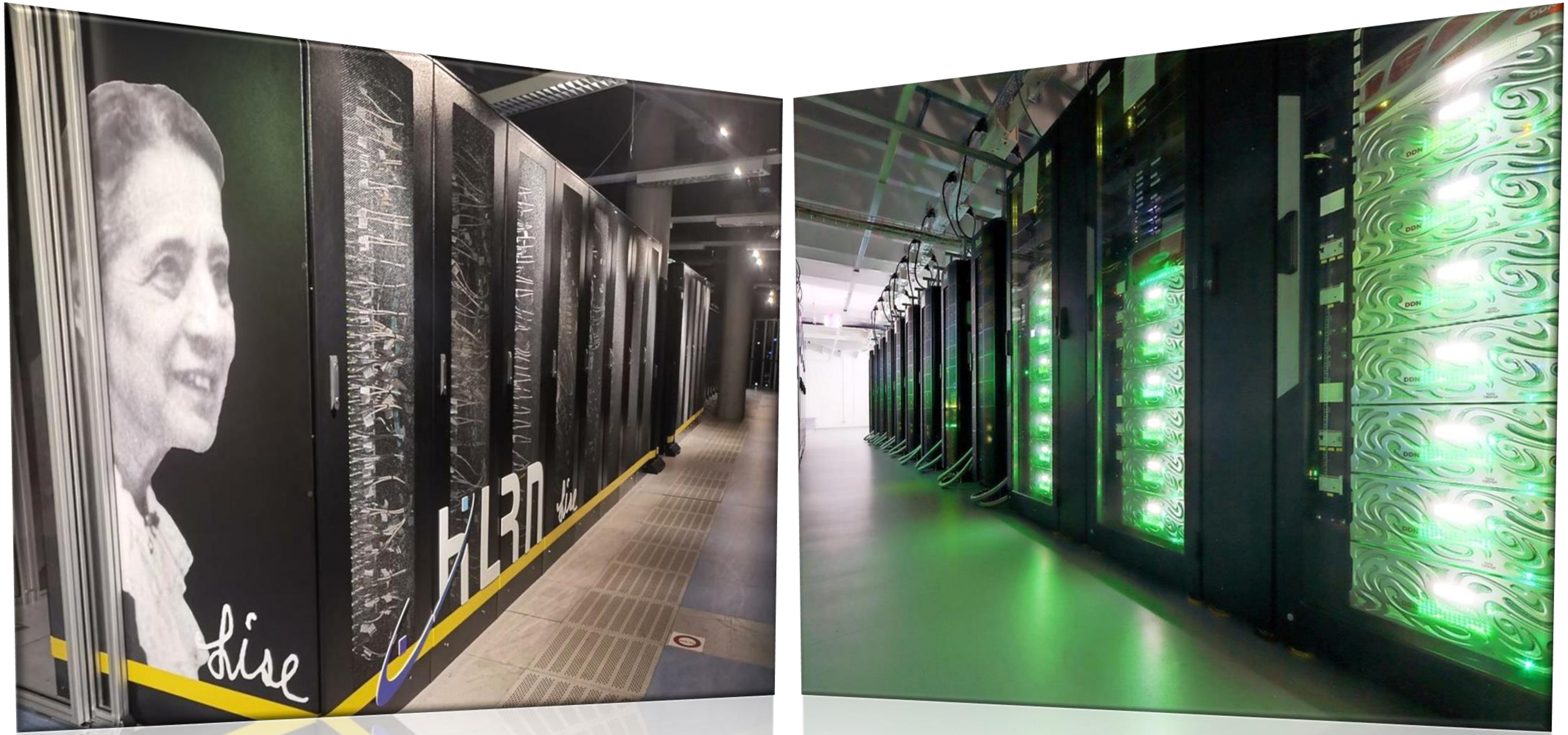
- possible answer (e.g. from asking Google):

> High Performance Computing most generally refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business.
>
> (http://insidehpc.com/hpc-basic-training/what-is-hpc/)

  – in short: use a supercomputer (HPC cluster) to solve a big problem

- another answer:
  – „computing at the bottleneck“ (from G. Hager @ RRZE, see this book)
  – what is the optimal performance I can achieve on a single core or a multi-core processor?
  – what is the limiting factor? e.g. CPU-speed vs. memory access

What is High-Performance Computing?

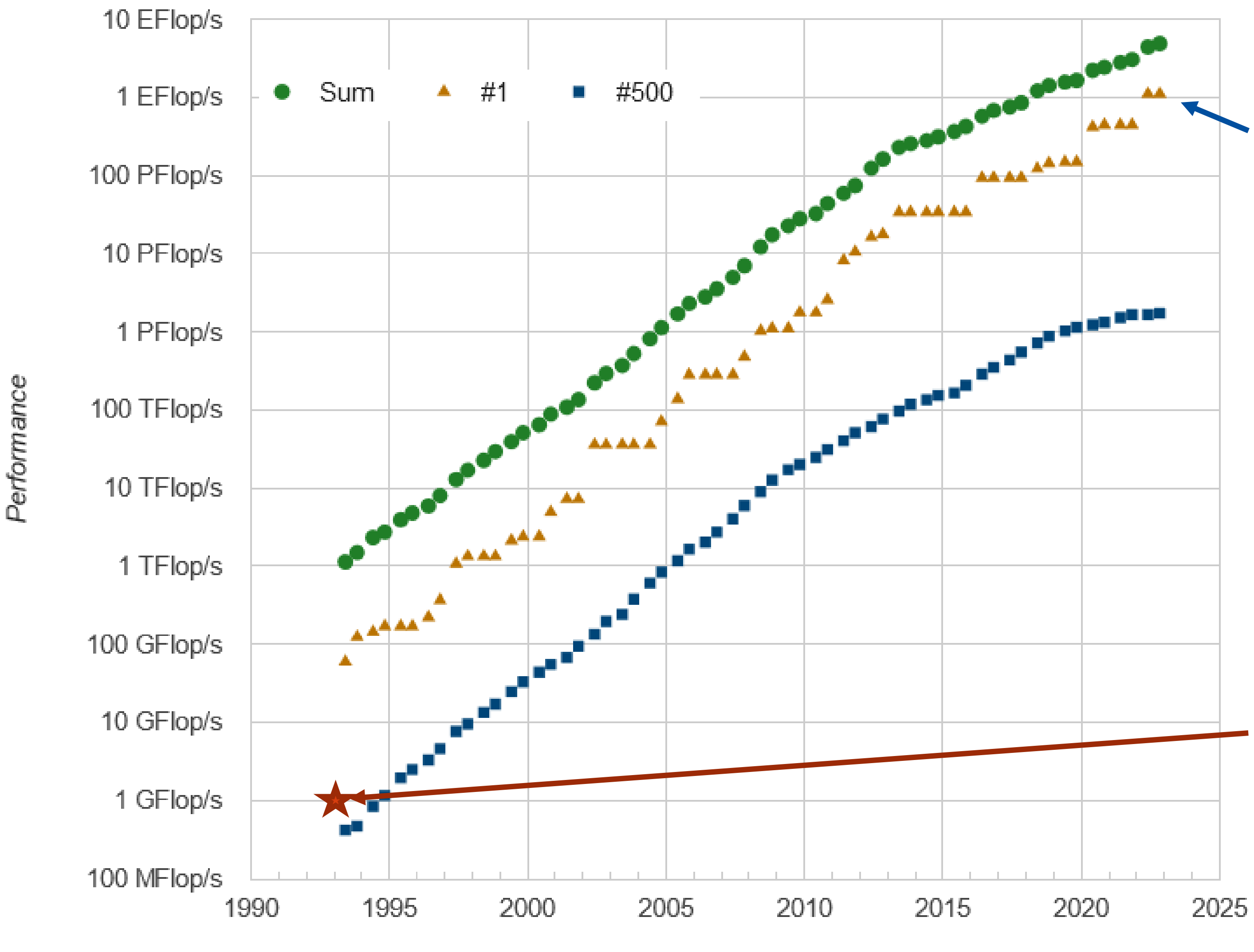


Supercomputers LISE (Lise Meitner) and EMMY (Emmy Noether) @ HLRN
(ranked in places 55 and 47 in the Top500 in November 2020)

High-Performance Computing (HPC)

- typical HPC resources
  - HPC Cluster (e.g. CARL and EDDY)
  - Grid-Computing (distributed resources, e.g. Folding@home)
  - GPUs
- requirements for using HPC resources
  - parallelization and optimization of programs/algorithms
  - understanding of the used computational architecture
  - identification of computing bottlenecks

The Fastest Computers on Earth



(taken from https://top500.org)

#1 November 2022: Frontier
Oak Ridge National Laboratory
(HPE, 1,102 PFlop/s, 8.73M cores, AMD "Milan/Trento" 64C 2.7GHz + AMD MI250x GPUs, Slingshot interconnect)

Expected measured Peak Performance

Current Tablet ≈ 1.5 Gflop/s

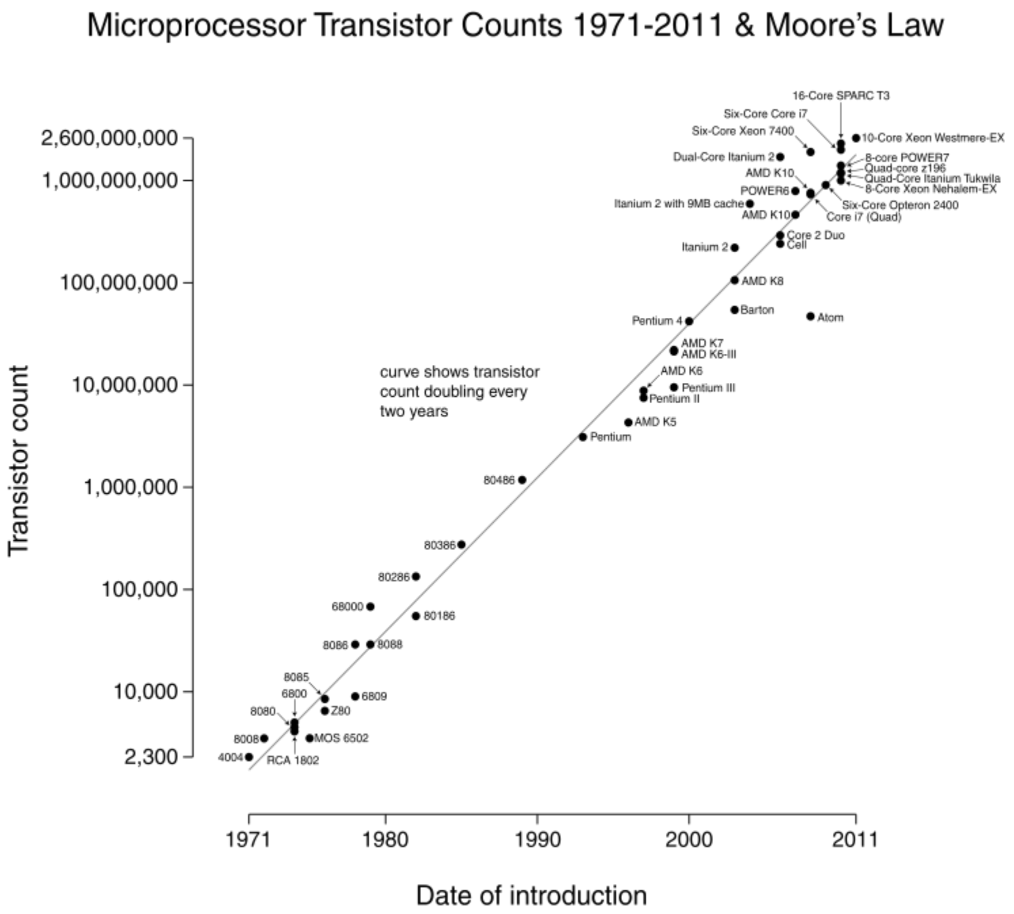
would have made it into the list in 1994

http://bits.blogs.nytimes.com/2011/05/09/the-ipad-in-your-hand-as-fast-as-a-supercomputer-of-yore/?_r=1

Moore’s Law

- the number of transistors in a dense integrated circuit doubles every 12 to 24 months
- often interpreted as computing speed doubles every ~18 months
- empirical law and self-fulfilling prophecy
- Moore’s Law must eventually break down

there are indications that we are already seeing the end of Moore’s Law



(taken from Wikipedia)

The End of Moore‘s Law?

- see e.g. http://www.nature.com/news/the-chips-are-down-for-moore-s-law-1.19338
- it is probably not economic to shrink below a few nm
- what comes next?
  - not known yet but it will be interesting
  - one possible way is to go 3d, already seen in memory
  - quantum computer? see https://singularityhub.com/2019/02/26/quantum-computing-now-and-in-the-not-too-distant-future or https://www.research.ibm.com/ibm-q/learn/what-is-quantum-computing/

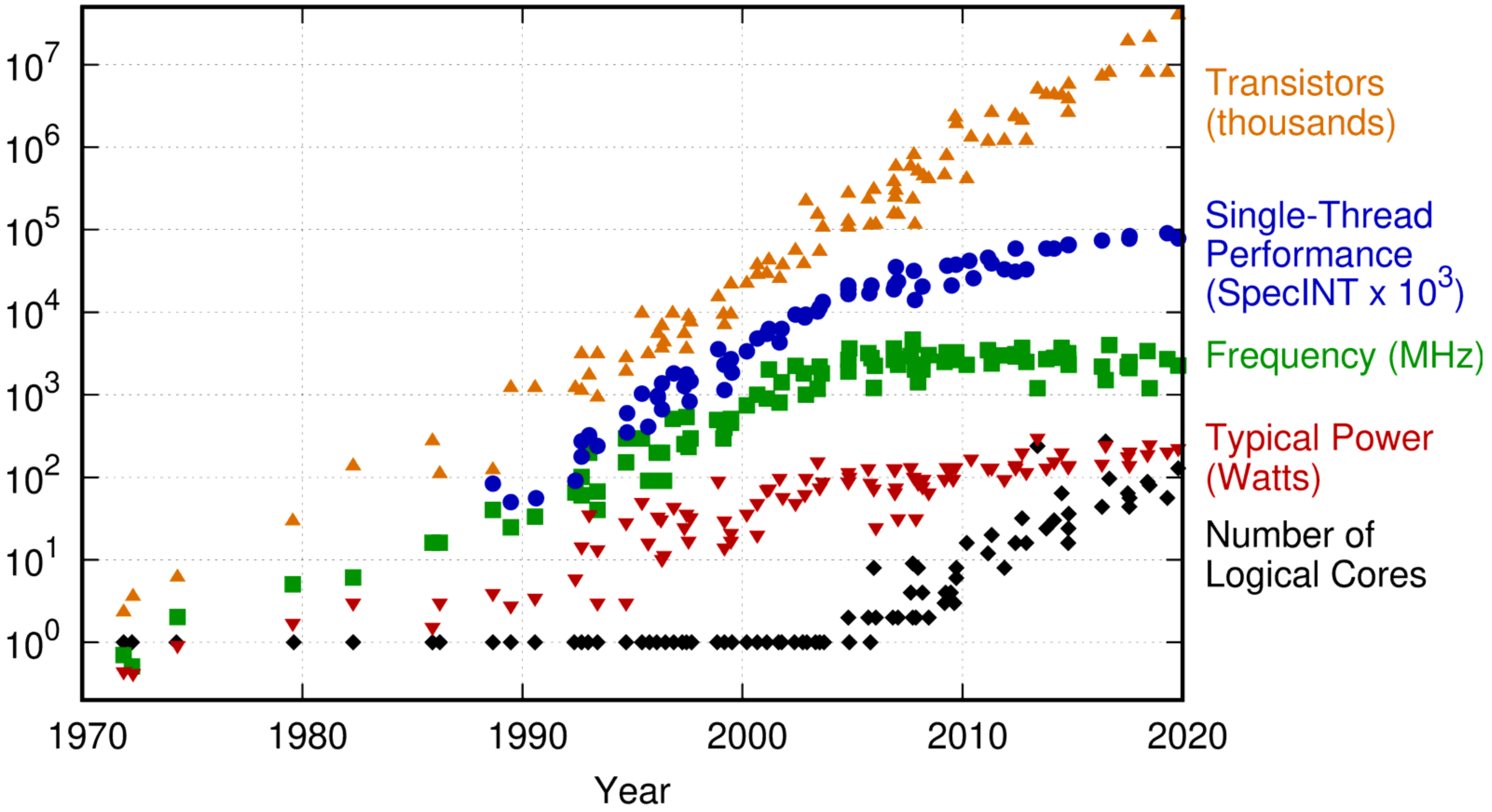
Carl von Ossietzky Universität Oldenburg

Parallelization

- parallelization is needed if the computational power of a single computer is not enough
  - distribution of data and/or work among several computers
  - different strategies depending on the problem at hand and also the available parallel architecture
  - parallel computing is very important today and will be more important in the future
  - modern PCs have multiple core CPUs → can be utilized by parallel applications

48 Years of Microprocessor Trend Data

https://github.com/karlrupp/microprocessor-trend-data



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2019 by K. Rupp

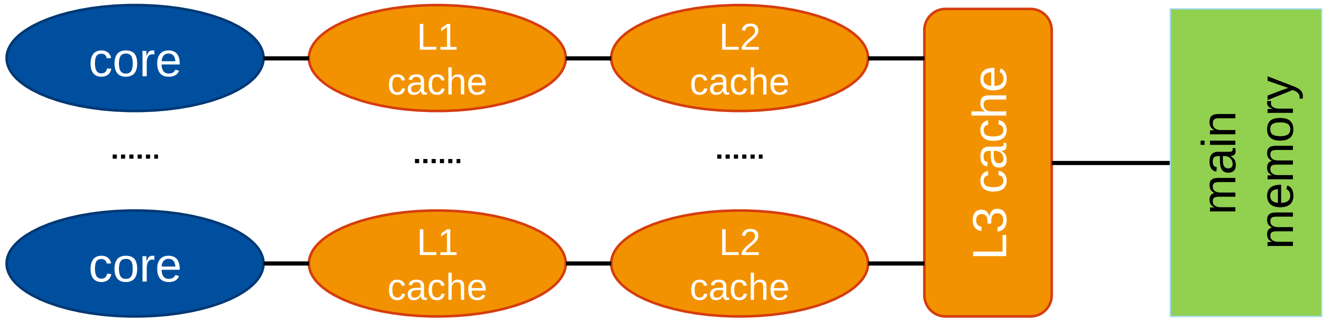
Parallel Architectures

- the knowledge of the underlying parallel architecture is important
  - allows to write more efficient code
  - avoid bottlenecks
  - understand the limiting factors for performance

- typical HPC cluster (e.g. CARL and EDDY)
  - multi-core CPU
  - multi-socket node
  - many nodes with node interconnect

Parallel Architectures

- multi-core CPU
  - contains several cores (currently up to 64 or so)
  - cores connected to caches for fast memory access, low latency (order of ten times faster than direct memory access)
  - cache coherence



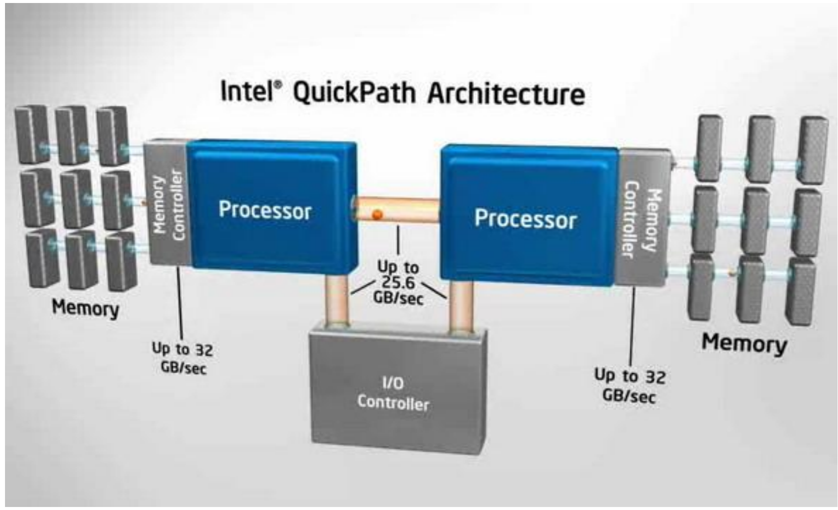
Broadwell E5-2650 | 12 cores, 2.2GHz | 64kB | 256kB | 30MB

Parallel Architectures

- multi-core CPU
  - number of cores increases, drives the performance increase
  - uniform memory access (UMA)
  - each core can execute single instructions on multiple data (SIMD), e.g. SSE, AVX, AVX2
  - clock frequency stagnates or even decreases (thermal effects)
  - computations can be limited by memory access (bandwidth) or by the number of instructions per cycle

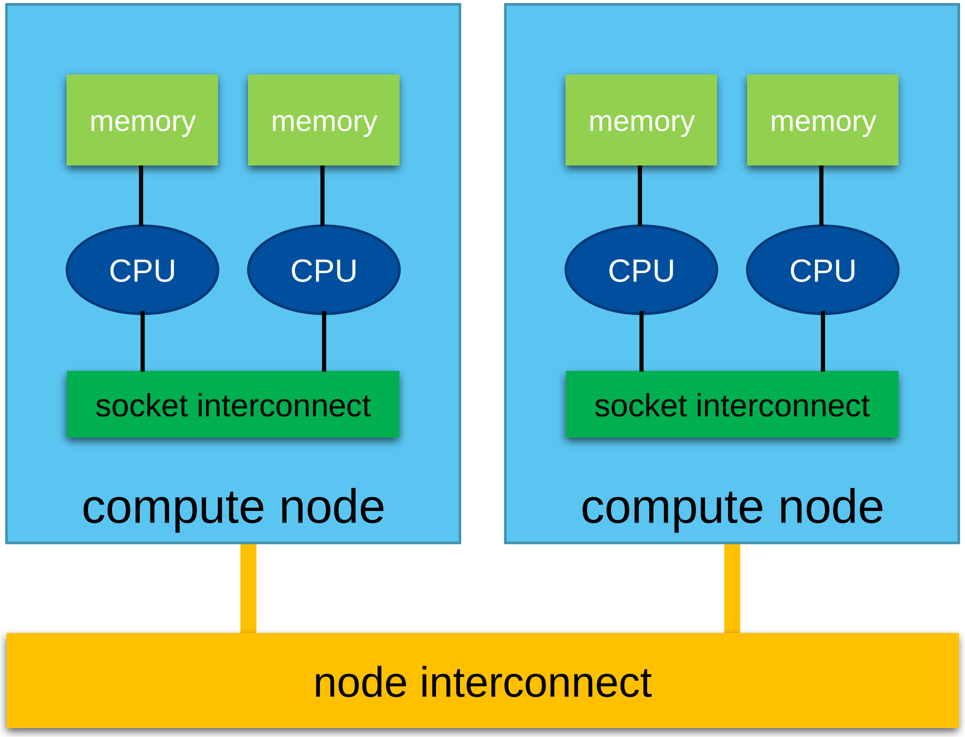
Parallel Architectures

- multi-socket node
  - several CPUs on a single main board
  - each CPU has access to its own main memory and cache levels
  - different, non-uniform memory access (NUMA)
  - with cache coherence → ccNUMA

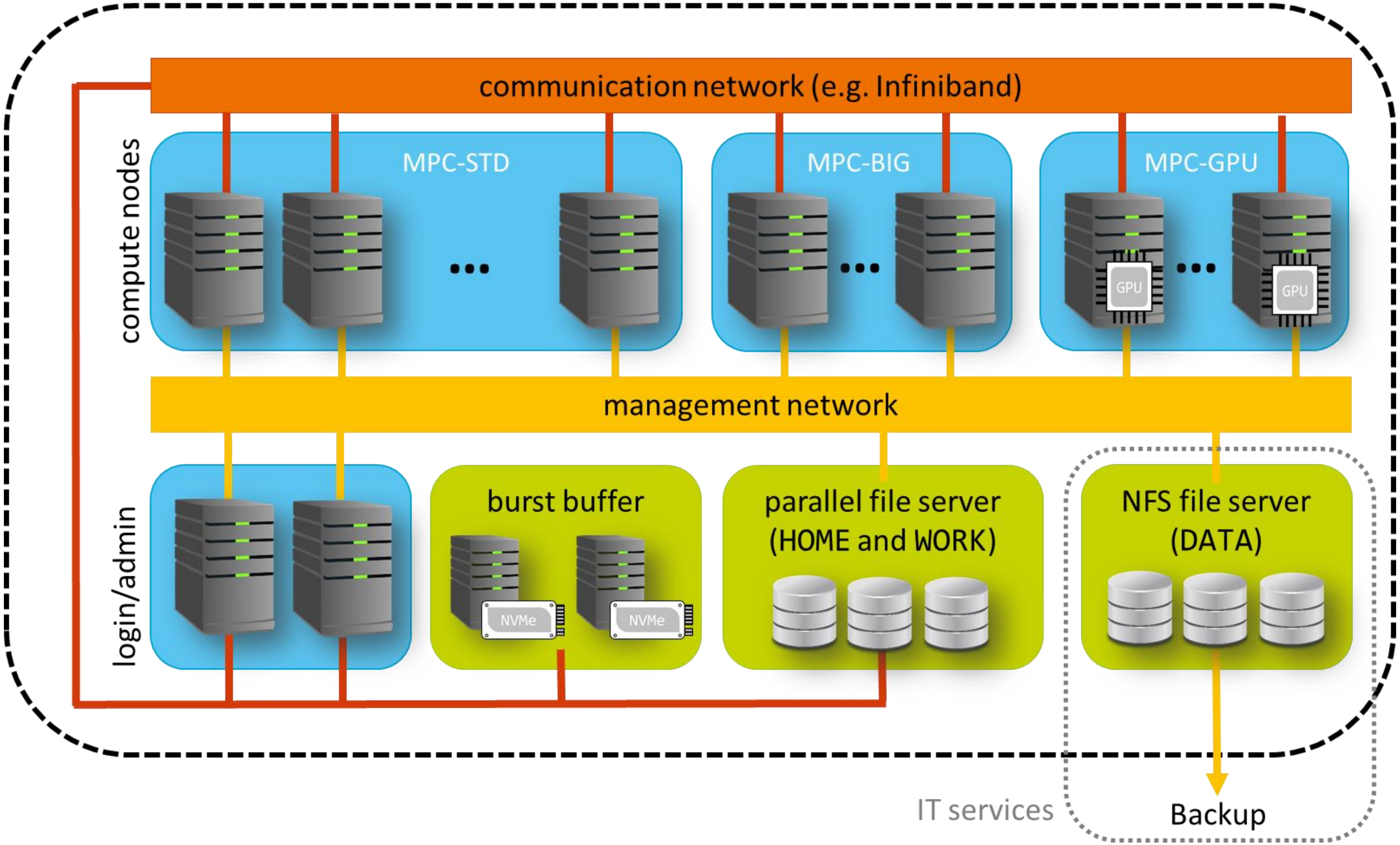


Parallel Architectures

- multi-computer
  - many (ccNUMA) nodes
  - distributed memory
  - node interconnect



HPC Cluster


communication network (e.g. Infiniband)
compute nodes
MPC-STD
MPC-BIG
MPC-GPU
GPU
GPU
management network
login/admin
burst buffer
NVMe
NVMe
parallel file server
(HOME and WORK)
NFS file server
(DATA)
IT services
Backup

HPC Resources in Germany and Europe

Tier 0

- ➢ Partnership for Advanced Computing in Europe
  - see PRACE website

Tier 1

95 PFlop/s
3 locations

- ➢ national Supercomputing Centers
  - locations Jülich, Stuttgart und Garching
  - organized in the Gauss Centre for Supercomputing

Tier 2

28 PFlop/s
about 18 locations

- ➢ (trans)regional Supercomputing Centers
  - e.g. Norddeutscher Verbund für Hoch- und Höchstleistungsrechnen (HLRN)
  - associated in the Gauß Alliance (GA)

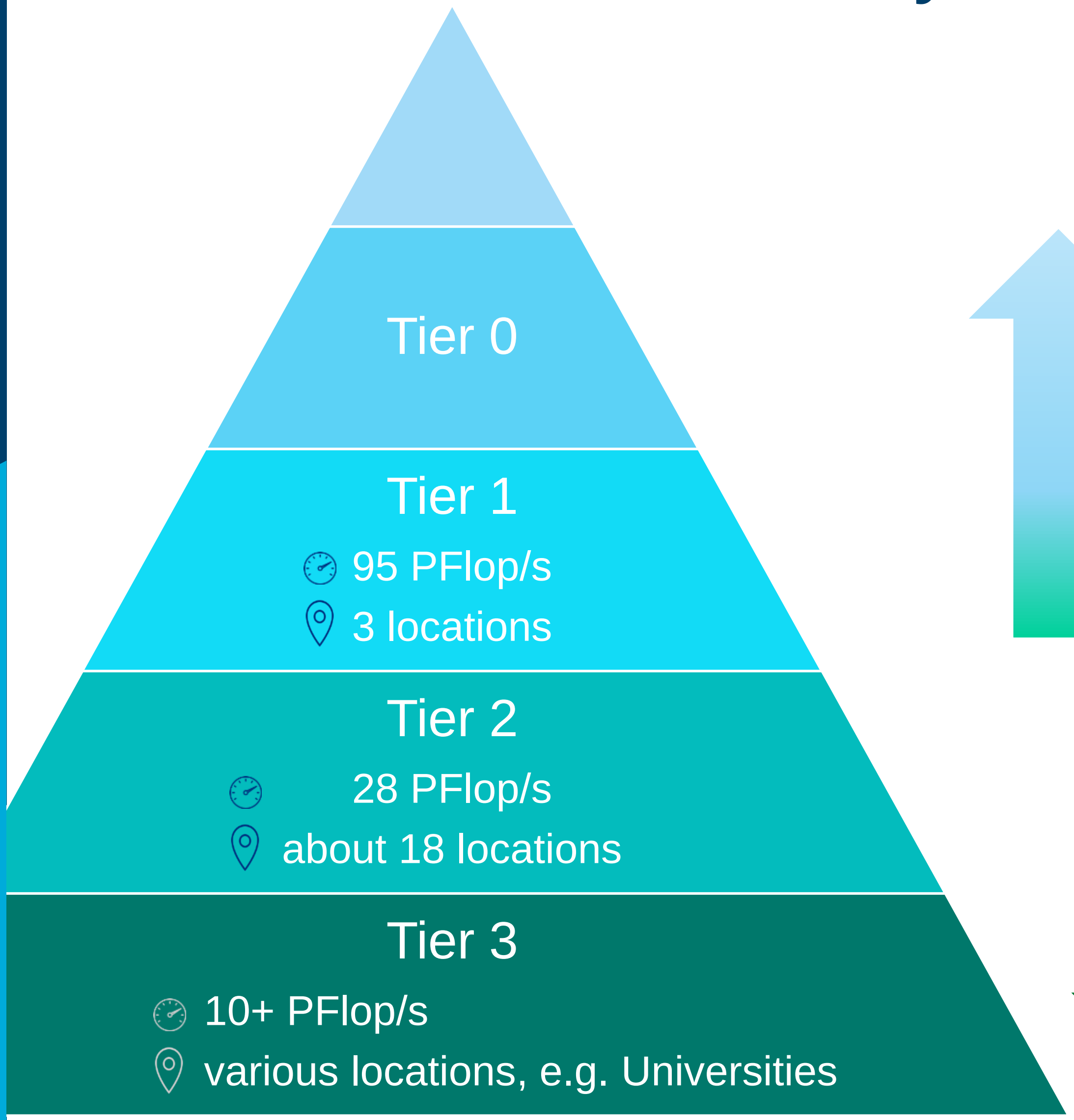
Tier 3

10+ PFlop/s
various locations, e.g. Universities

- ➢ local HPC clusters
  - e.g. CARL und EDDY with R_{max} of 457 TFlop/s

Carl von Ossietzky Universität Oldenburg

HPC Performance Pyramid



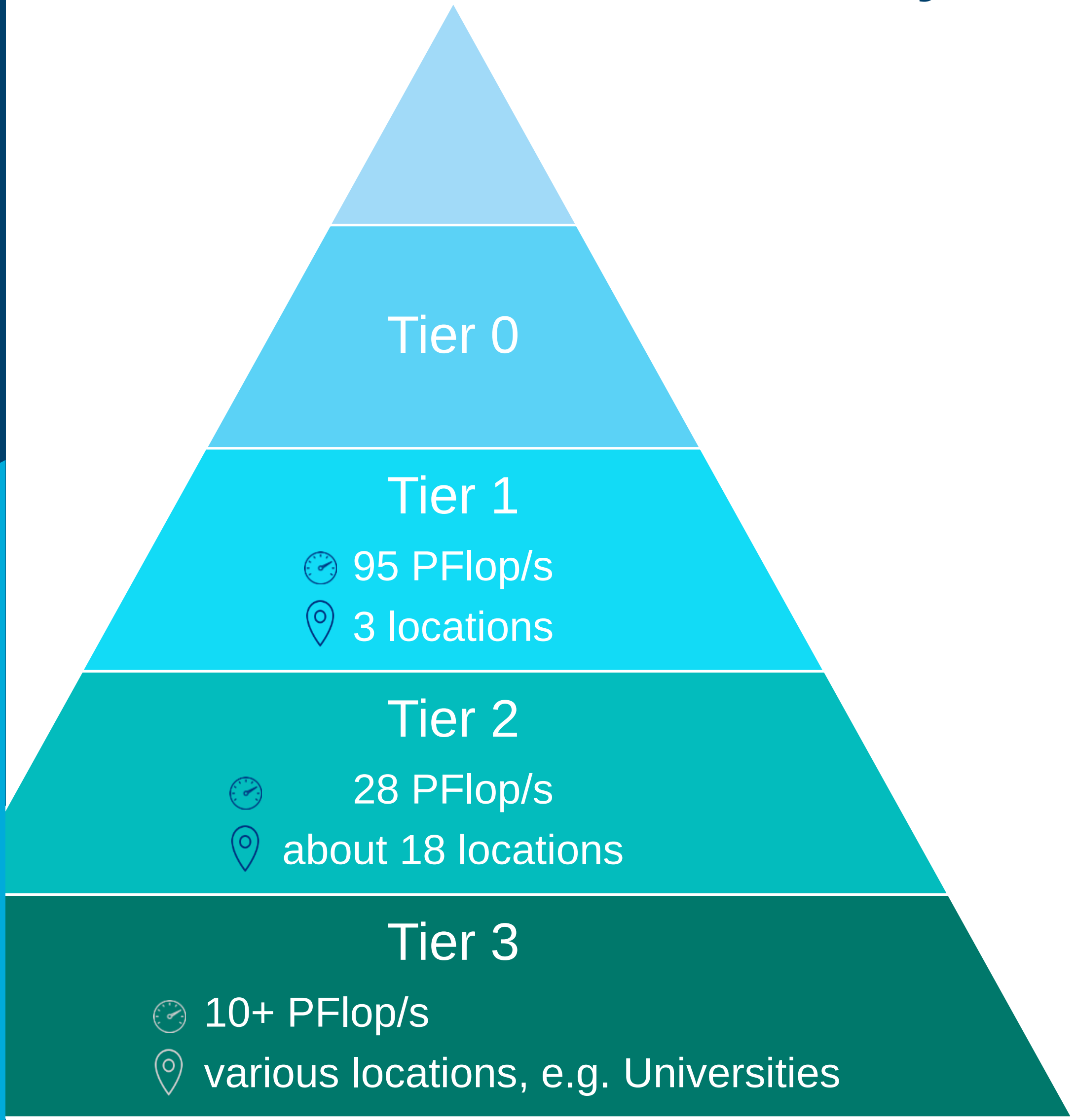
„capability computing“

- using massively parallel computing power to solve a problem in minimal time

„capacity computing“

- using cost-efficient computing power to solve a large number of smaller problems

HPC Performance Pyramid



- permeability between the levels
  - preparation of projects on the lower levels
  - development of algorithms on the lower levels („capabilty test computing“)
  - transfer of knowledge in both directions
  - special requirements can be better fulfilled in the lowest tier
  - education and training from the bottom to the top

Carl von Ossietzky Universität Oldenburg

Accessing HPC Resources

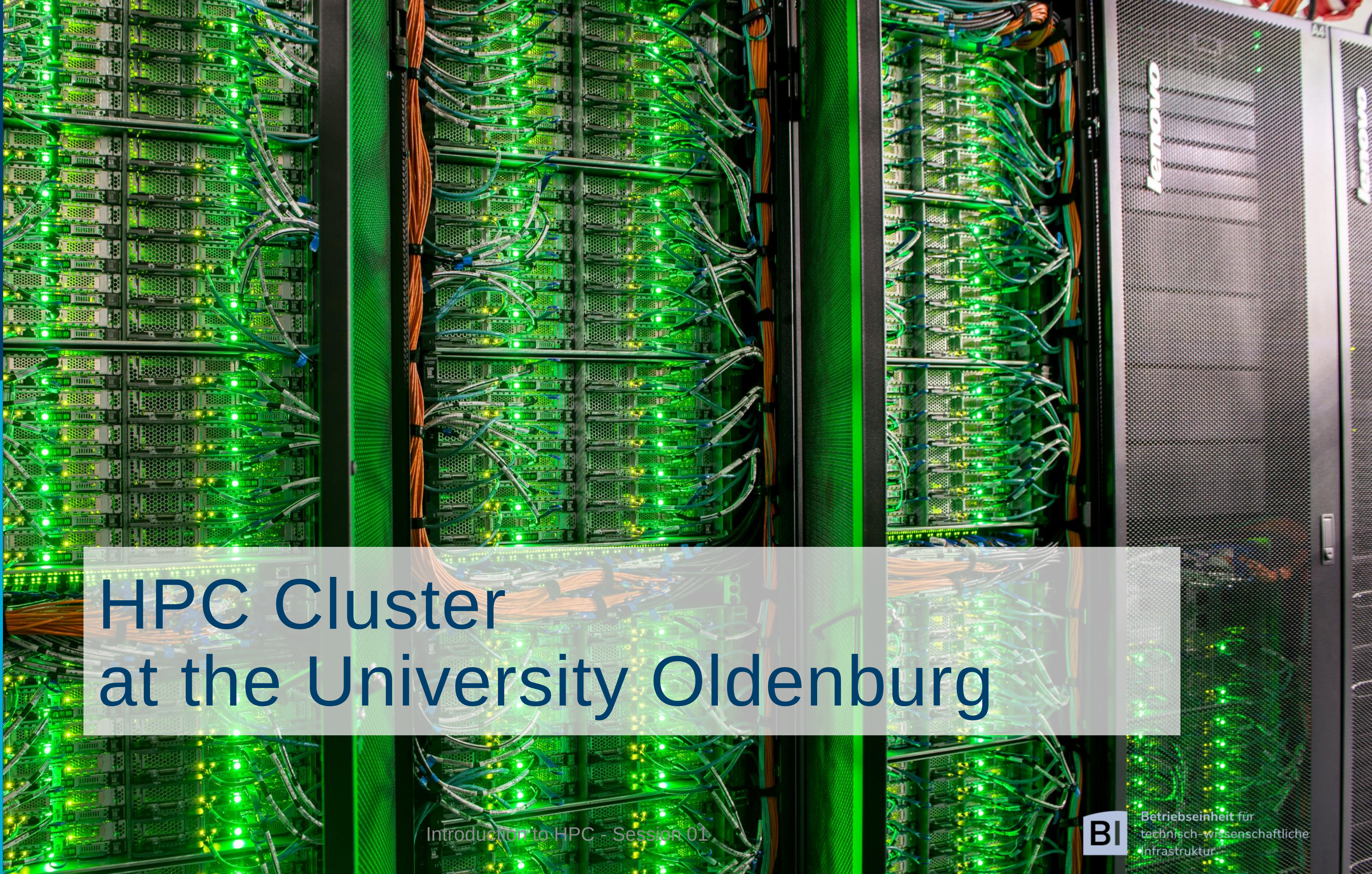
- researchers at UOL can always use CARL and EDDY
- for larger projects you can apply for compute time at
  - HLRN
  - one of the Tier 1 centers
  - PRACE
- in 2020 NHR-grands have been awarded to eight supercompting centers
  - a new HPC infrastructure will be created to improve access to HPC resources
  - eventually you will be allowed to apply for compute time at all NHR centers

Carl von Ossietzky Universität Oldenburg

HLRN NORDDEUTSCHER VERBUND FÜR HOCH- UND HÖCHSTLEISTUNGSRECHNEN

North German Supercomputing Alliance (HLRN)

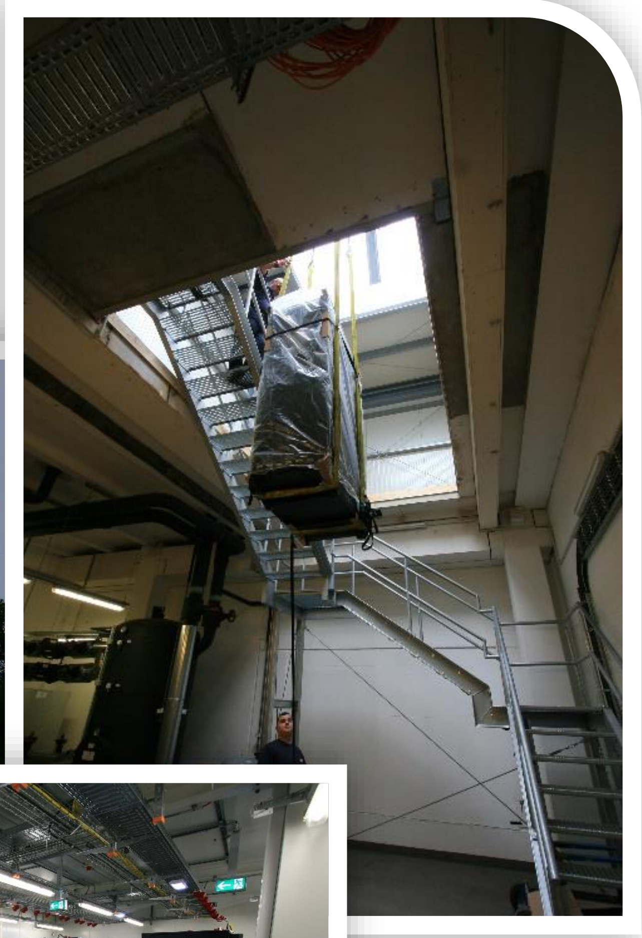
- supercomputing centers at two locations (Berlin and Göttingen)
  - two of the eight NHRs
  - large (25+) network of consultants for scientific computing and HPC
- two HPC clusters – EMMY and LISE
  - fourth generation (HLRN-IV) became operational in 2019/20, total peak performance of more than 16 PFlop/s
  - EMMY: 974 compute nodes (2x Intel Cascade Lake, 48C @ 2.3 GHZ, 384+ GB RAM, OmniPath)
    448 compute nodes (2x Intel Skylake, 20C @ 2.4 GHZ, 384+ GB RAM, OmniPath)
  - LISE: 1270 compute nodes (2x Intel Cascade Lake, 48C @ 2.3 GHZ, 384+ GB RAM, OmniPath)
- accessing HLRN (see https://zulassung.hlrn.de/)
  - request a user account (to prepare project proposal, valid for 9 months)
  - submit project proposal (dead-lines 28th of January, April, July, October)
  - compute time is allocated for one year (in form of NPLs per quarter)

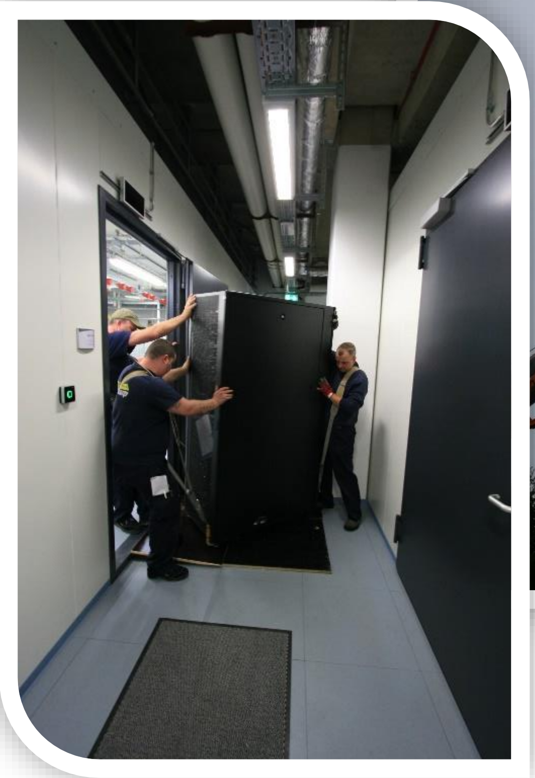


HPC Cluster
at the University Oldenburg

Delivery Day

- the current HPC cluster was delivered in August, 2016





http://www.uni-oldenburg.de/fk5/wr/aktuelles/artikel/art/neue-hochleistungsrechner-fuer-die-universitaet-oldenburg-2380/

Overview HPC Clusters

- CARL
  - multi-purpose cluster as a basic computing resource
  - funded by the University/MWK and the DFG under grant number INST 184/157-1 FUGG (Forschungsgroßgerät nach Art. 91b GG)
- EDDY
  - CFD cluster for wind energy research
  - funded by the BMWi under grant number 0324005
- used as a shared HPC cluster
  - common infrastructure is shared (e.g. file systems, network)
  - shared administration

do not forget to acknowledge the use of the HPC cluster in your publications

http://wiki.hpcuser.uni-oldenburg.de/index.php?title=Acknowledging_the_HPC_facilities_2016

Carl von Ossietzky Universität Oldenburg

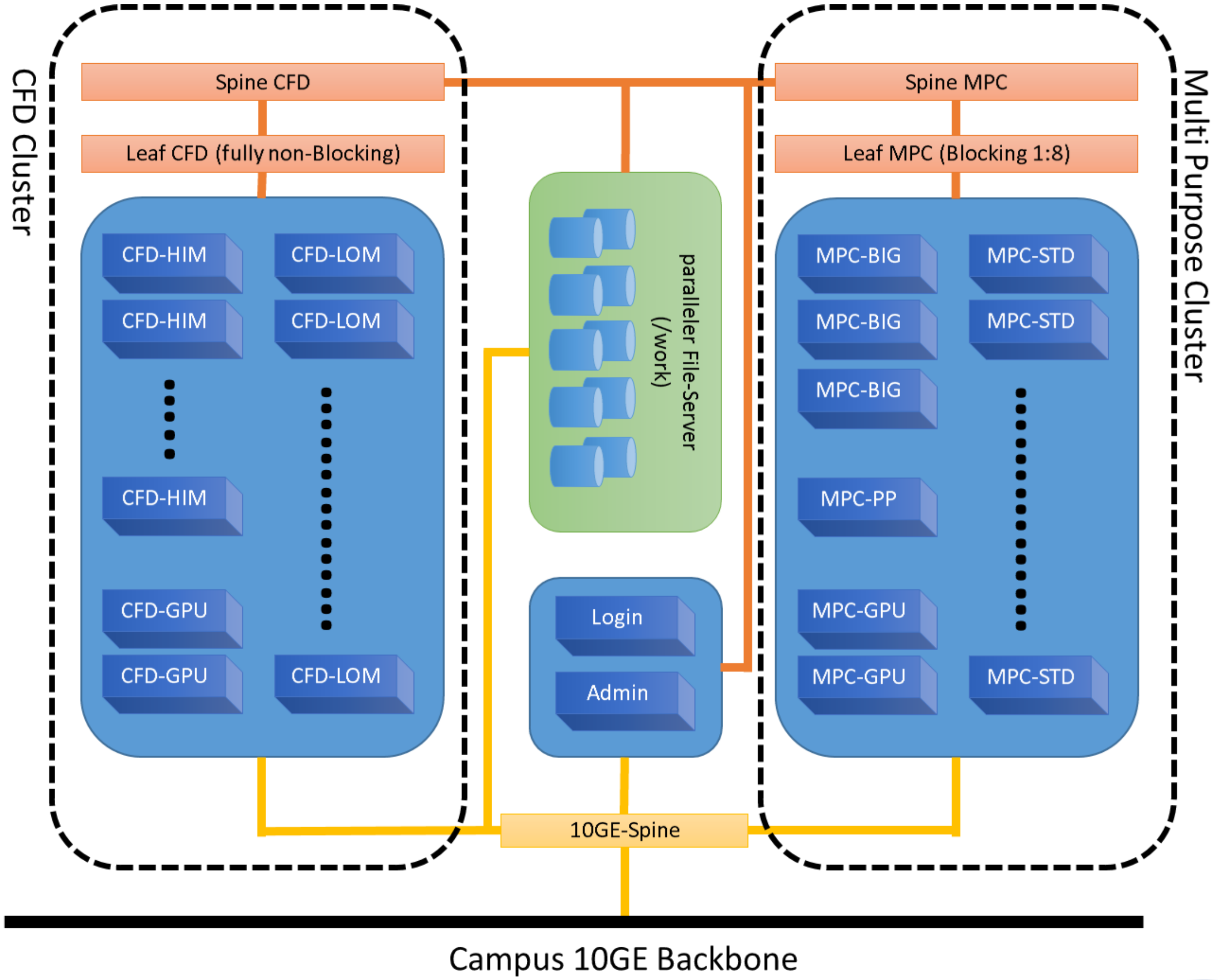
HPC Facilities @ University Oldenburg

- shared HPC cluster CARL/EDDY
  - close to 600 compute nodes
  - 4 login and 2 administration nodes
  - Infiniband FDR interconnect for parallel computing
  - 10/1GE network
  - parallel file system (GPFS) with 900TB capacity
  - NFS mounted central storage

  - Linux (RHEL) as OS
  - many scientific applications and libraries available
  - Job Scheduler (SLURM)

https://wiki.hpcuser.uni-oldenburg.de/index.php?title=HPC_Facilities_of_the_University_of_Oldenburg_2016

Schematic View of HPC Cluster



Compute Node Configurations

	Node Type	Count	Description	CPU	RAM
CARL	MPC-STD	158	standard	2x Intel Xeon E5-2650v4 12C with 2.2GHz	256 GB
	MPC-LOM	128	low memory		128 GB
	MPC-BIG	30	big memory	2x Intel Xeon E5-2667v4 8C with 3.2GHz	512 GB
	MPC-PP	2	post-processing	4x Intel Xeon E7-8891v4 10c with 2.8GHz	2048 GB
	MPC-GPU	9	GPU	2x Intel Xeon E5-2650v4 12C with 2.2GHz	256 GB
EDDY	CFD-LOM	160	low memory		64 GB
	CFD-HIM	81	high memory		128 GB
	CFD-GPU	3	GPU		256 GB

https://wiki.hpcuser.uni-oldenburg.de/index.php?title=HPC_Facilities_of_the_University_of_Oldenburg_2016

Carl von Ossietzky Universität Oldenburg

Overview new HPC-Clusters

- CARL-2
  - multi-purpose compute (MPC) cluster as a basic computing resource
  - funded by the University/MWK and the DFG under grant number INST 184/225-1 FUGG (Forschungsgroßgerät nach Art. 91b GG)
  - total funding: 2.400.000€ (+460.000€ HIFMB)
  - responsible PIs: Thorsten Klüner, Jörg Lücke, Stefan Harfst
- RE-HPC and MOUSE (ForWind)
  - CFD cluster for wind energy research
  - funded by REACT-EU under grant number ZW7-95186744 and BMWK under grant number 03EE3067A
  - total funding: 1.667.000€ (REACT-EU, including 167.000€ Uni) and 478.956,21€ (BMWK, HPC-Cluster only, project total is 1.628.058,80€)
  - responsible PIs: Laura Lukassen, Joachim Peinke, Martin Kühn, Hendrik Heißelmann
- used as a shared HPC cluster
  - common infrastructure is shared (e.g. file systems, network)
  - shared administration

Carl von Ossietzky Universität Oldenburg

HPC-Cluster @ University Oldenburg

- shared HPC-cluster
  - close to 150 compute nodes
  - 5 login and 2 administration nodes
  - Infiniband HDR interconnect for parallel computing
  - 25/1GE network for management
  - parallel file system (GPFS) with >4 PiB capacity
  - NFS mounted central storage

  - Linux (RHEL) as OS
  - many scientific applications and libraries available
  - Job Scheduler (SLURM)

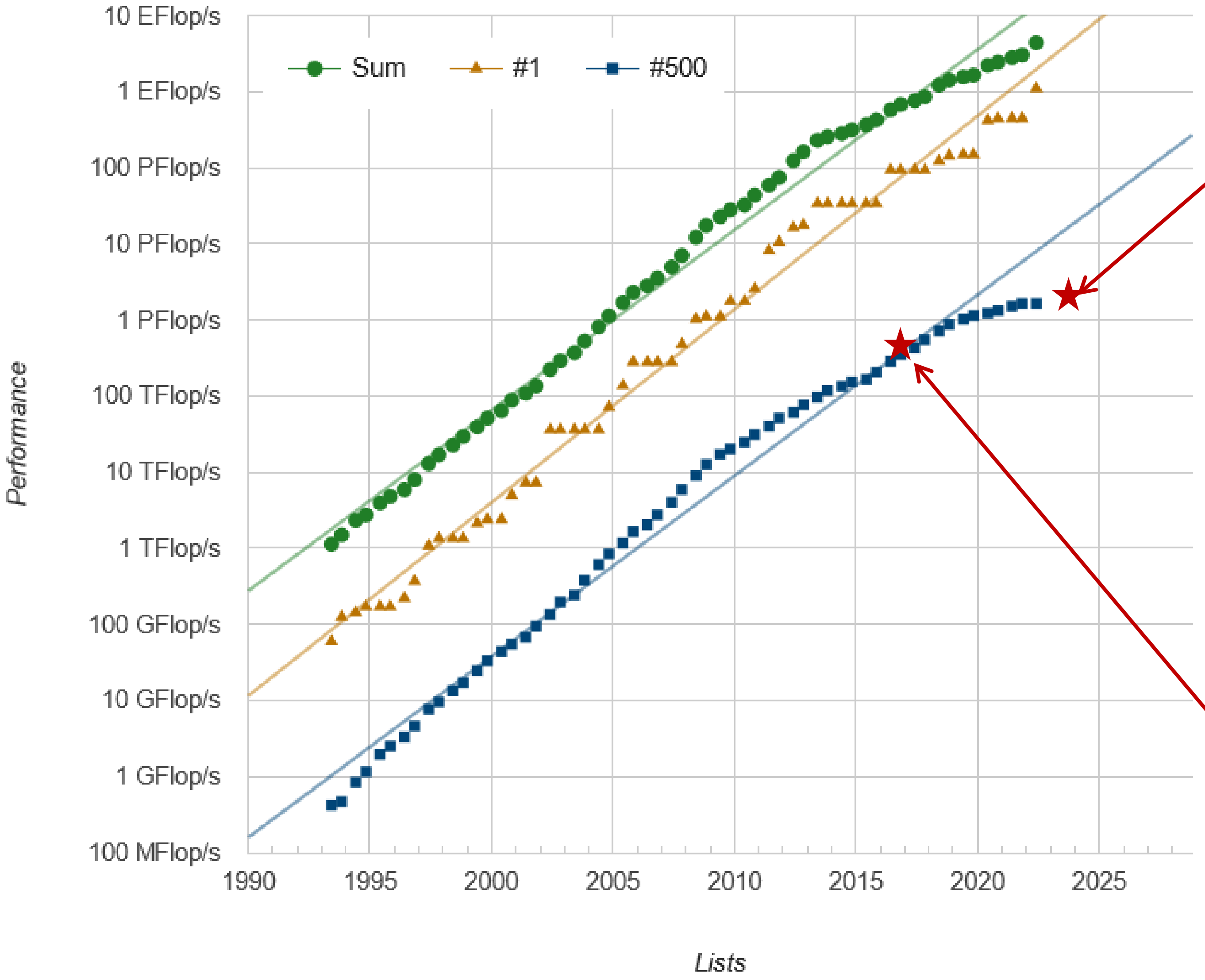
Carl von Ossietzky Universität Oldenburg

Summary MPC & CFD

Feature	MPC	CFD	Total
Nodes	91	70	161
CPU-Cores	11.648	8.960	20.608
RAM	91 TiB	54 TiB	145 TiB
GPUs	24x H100 94GB	8x A100 80GB 4x H100 94GB	8x A100 80GB 28x H100 94GB
GPFS	>2 PiB	>2 PiB	>4 PiB
Burst Buffer	92 TB	-	92 TB
Rmax (CPU)	~546 TFlop/s	~414 TFlop/s	960 TFlop/s
Rpeak (GPU)	~1.440 TFlop/s	~320 TFlop/s	1.760 TFlop/s

Betriebseinheit für technisch-wissenschaftliche Infrastruktur

Top500 Performance Development



MPC/CFD
- not ranked
- Rpeak 2,720 TFlop/s (CPU+GPU)
- 161 nodes, 20,608 cores
- 145 TiB RAM

4.7x Rpeak
0.28x nodes
1.53x cores
1.48x RAM

CARL/EDDY
- Rank 363
- Rmax (CPU) 457.2 TFlop/s
- Rpeak 577 Tflop/s (CPU+GPU)
- 571 nodes, 13,500 cores
- 98 TiB RAM

Carl von Ossietzky Universität Oldenburg

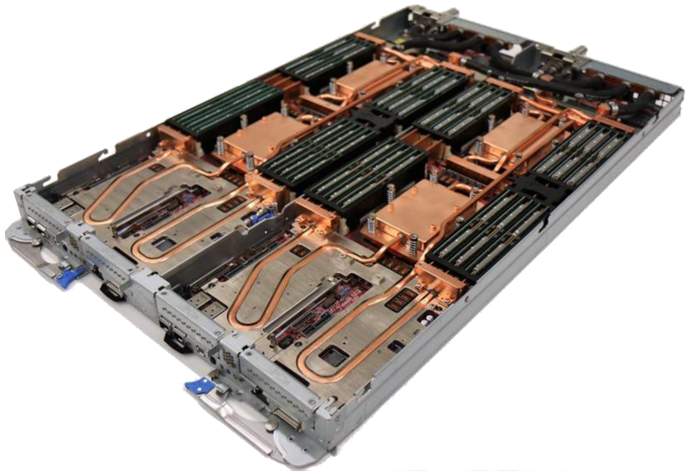
Compute Node Configurations

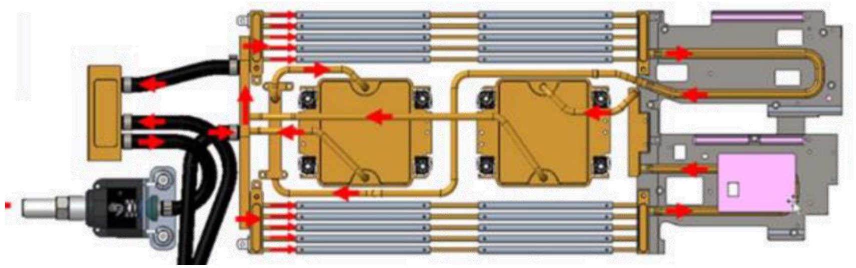
	Node Type	NodeCnt	Description	CPU / GPU	RAM
MPC	MPC-STD	56	standard	2x AMD Genoa EPYC 9554, 64C, 3.1 GHz, 360W	768 GB
	MPC-BIG	26	big memory		1124 GB
	MPC-PP	3	post-processing		4096 GB
	MPC-GPU	6	GPU	CPU identical to MPC-STD, 4x Nvidia H100 94GB SXM	1124 GB
CFD	CFD-STD	68	standard	2x AMD Genoa EPYC 9554, 64C, 3.1 GHz, 360W	768 GB
	CFD-GPU	1	GPU	Identical to MPC-GPU	1124 GB
		1		2x AMD Milan EPYC 7713, 64C, 2.0 GHz, 225W, 8x Nvidia A100 80GB SXM	1024 GB

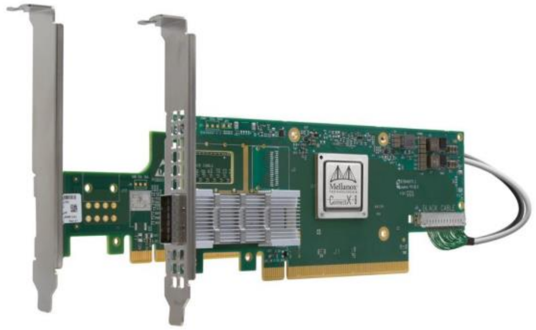
Betriebseinheit für technisch-wissenschaftliche Infrastruktur

Details on MPC/CFD-STD Compute Nodes

- compute nodes equipped with
  - CPU: 2x AMD „Genoa“ EPYC 9554, 64C @ 3.1GHz, 360W TDP
  - RAM: 768 GB (24x 32GB DDR5 @ 4800MHz)
  - Infiniband: ConnectX-6 HDR/200GbE (SharedIO, see below)
  - direct water-cooling (DWC) for CPUs, RDIMMs, and more
- two nodes on a single tray
  - shared DWC
  - shared IB-connector







Details on MPC/CFD-GPU Compute Nodes

- Lenovo Neptune DWC Tray
  - CPU: 2x AMD „Genoa“ EPYC 9554, 64C @ 3.1GHz, 360W TDP
  - RAM: 1124 GB (24x 48GB DDR5 @ 4800MHz)
  - Infiniband: ConnectX-7 NDR/200GbE
  - direct water-cooling (DWC) for CPUs, GPUs, RDIMMs, and more

- Supermicro 4U Server
  - CPU: 2x AMD EPYC™ “MILAN” 7713, 64C @ 2.0GHz, 225W TDP
  - RAM: 1024GB (16x 64GB DDR4 @ 3200 MHz)
  - Nvidia Delta GPU-Board with 8x A100-80GB SXM4 GPUs @ 400 Watt
  - Infiniband: ConnectX-6 HDR100
  - 960GB NVMe SSD
  - Air-cooled

Racks in Server Room (CARL and EDDY)

Carl von Ossietzky Universität Oldenburg

Login to the HPC Cluster

http://wiki.hpcuser.uni-oldenburg.de/index.php?title=Login

- Linux
  - use the **`ssh`**-command to login to one of the login nodes of **`carl`** or **`eddy`**

    `ssh -X abcd1234@carl.hpc.uni-oldenburg.de`

- Windows
  - use MobaXterm or PuTTY
  - with Windows 10 you can also use Windows Subsystem for Linux (WSL, see https://docs.microsoft.com/en-us/windows/wsl/install-win10)
  - MobaXterm or WSL are recommended
- login host names

  `hpcl00[1-4].hpc.uni-oldenburg.de`

  - can be used instead of **`carl`** or **`eddy`** (for login to specific node)
  - no difference between **`carl`** and **`eddy`** as login
- from outside of the campus network use VPN connection
  - see instructions at http://www.itdienste.uni-oldenburg.de/21240.html