Carl von Ossietzky Universität Oldenburg

Betriebseinheit für technisch-wissenschaftliche Infrastruktur

Introduction to High-Performance Computing

Session 01
Course Organization and Introduction

but before we begin…

Course Organisation

- Contact Information
- Outline of the Course

Carl von Ossietzky Universität Oldenburg

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 Dis-tributed 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

- 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

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

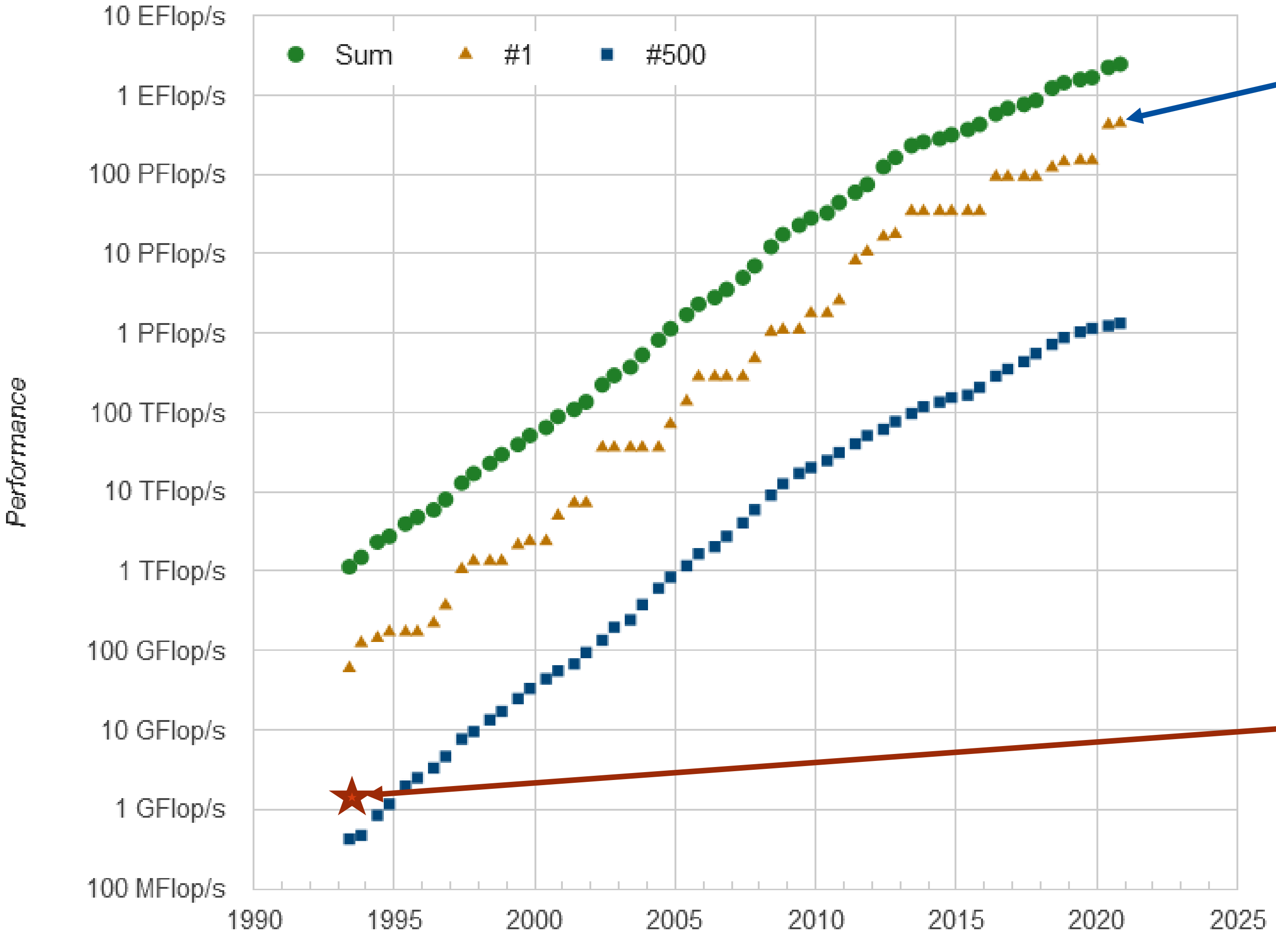
What is High-Performance Computing?

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. SETI@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



#1 November 2020: Fugaku
RIKEN Center for Computational Science
(442 PFlop/s, 7.3M cores, A64FX 48C 2.2GHz, Tofu interconnect D)

Expected measured Peak Performance

Current Tablet ≈ 1.5 Gflop/s

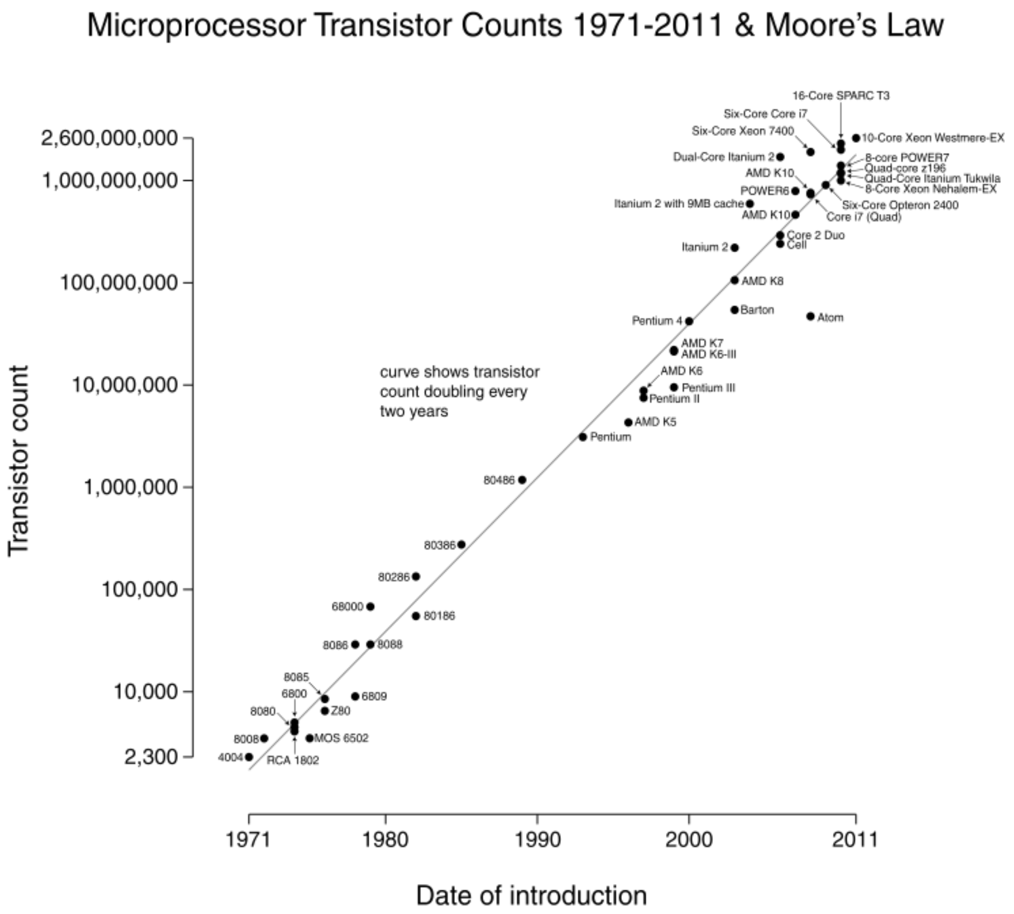
would have made it into the list in 1994

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

(taken from https://top500.org)

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/

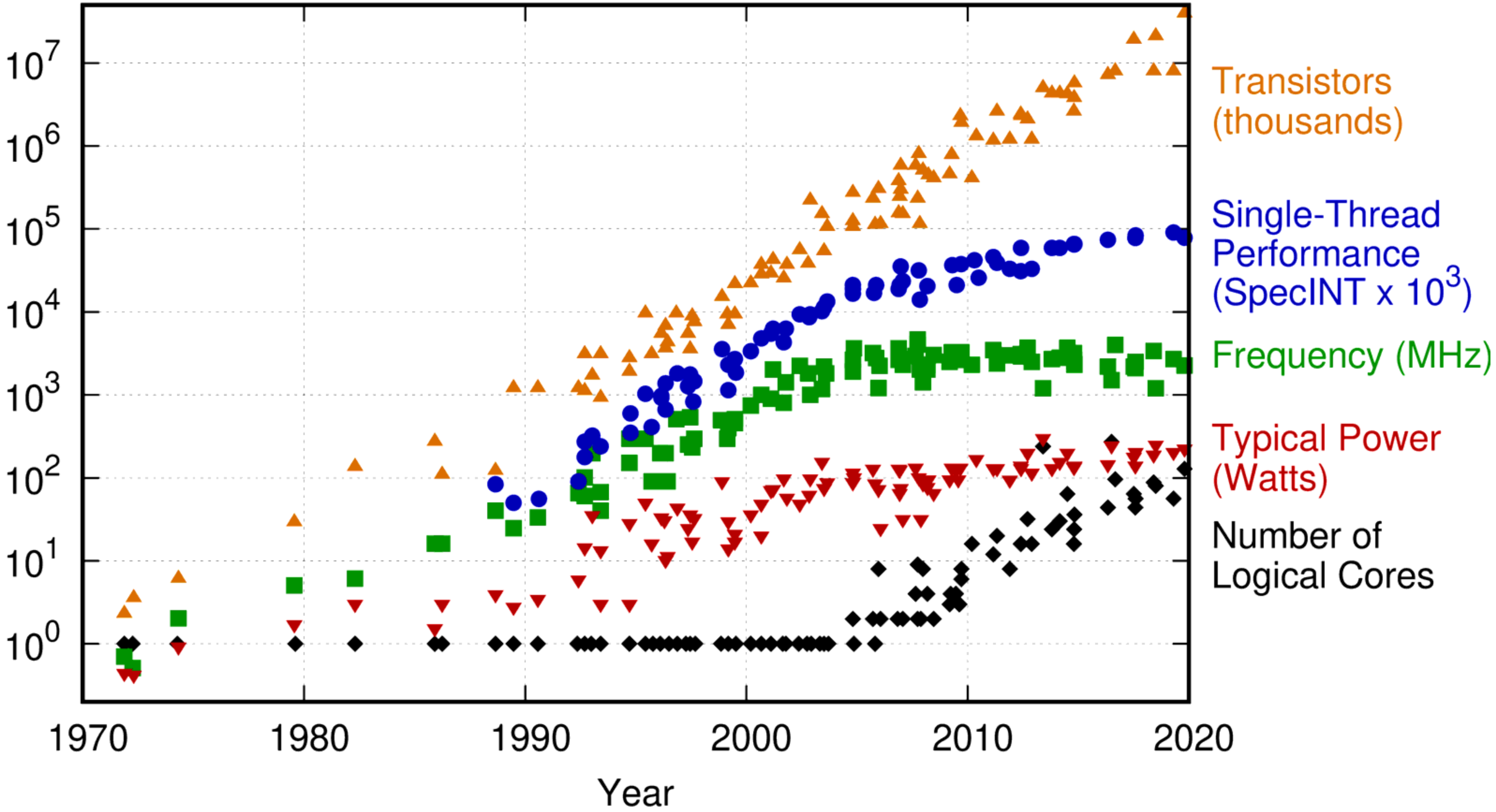
Parallelization

- 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

Carl von Ossietzky Universität Oldenburg

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

Betriebseinheit für technisch-wissenschaftliche Infrastruktur

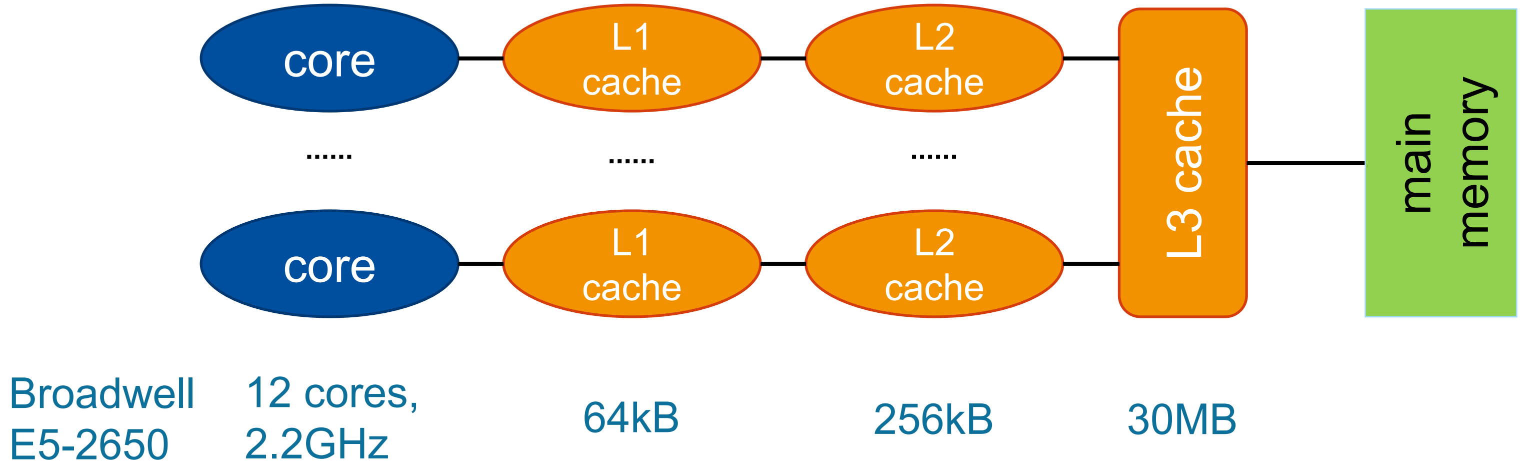
Parallel Architectures

- 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 48 or so)
  - cores connected to caches for fast memory access, low latency (order of ten times faster than direct memory access)
  - cache coherence

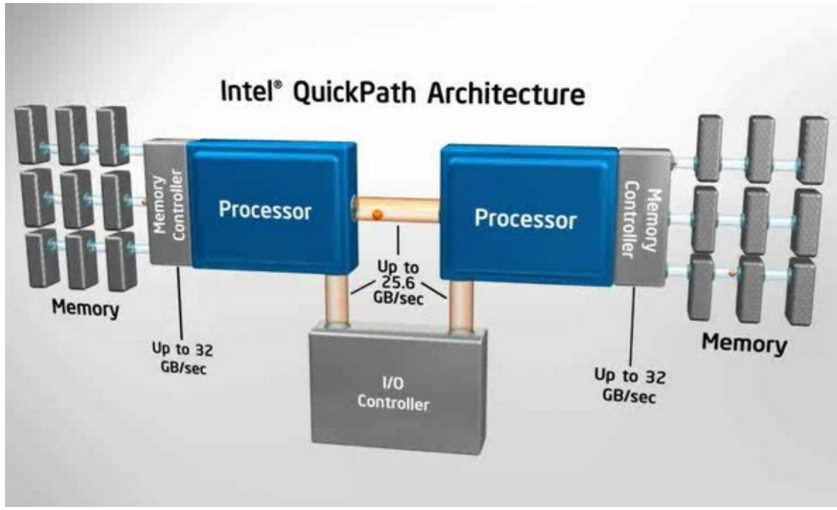


Parallel Architectures

- 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

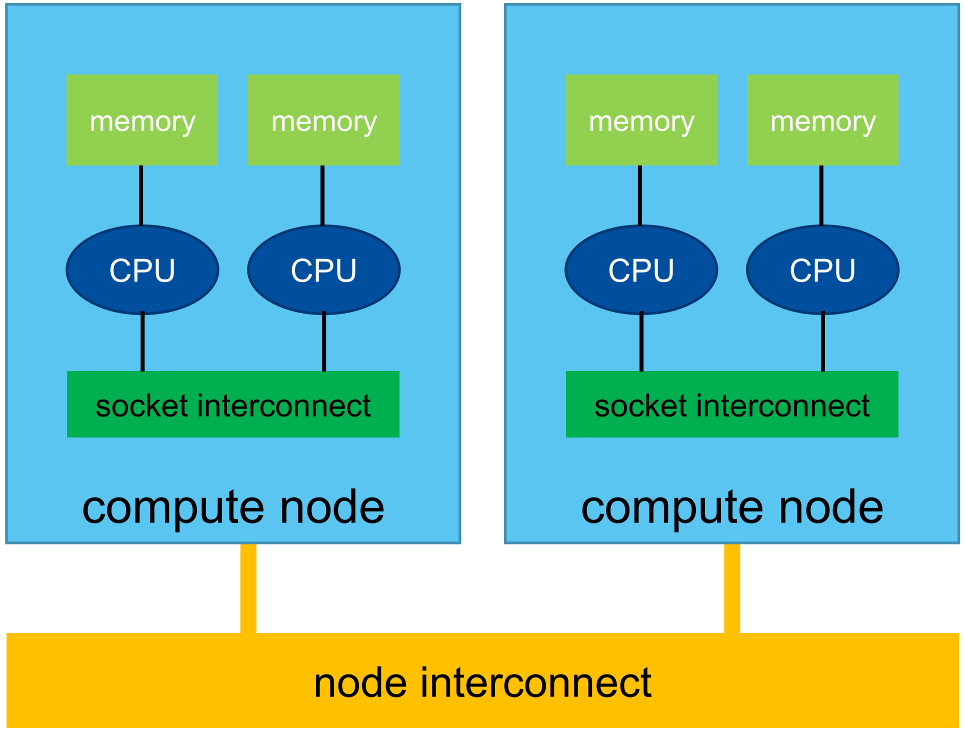
Parallel Architectures

- 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

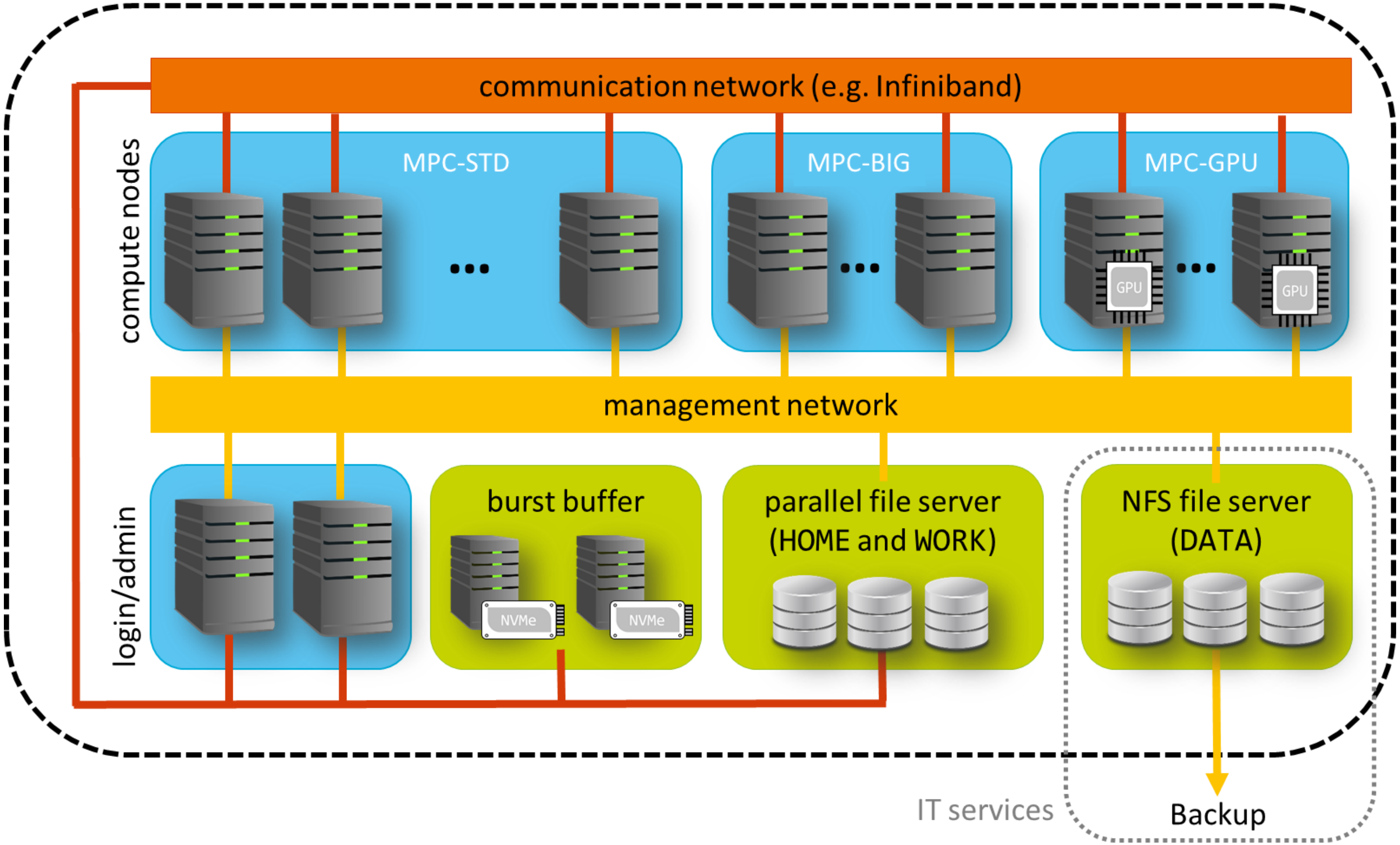


Parallel Architectures

- 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)

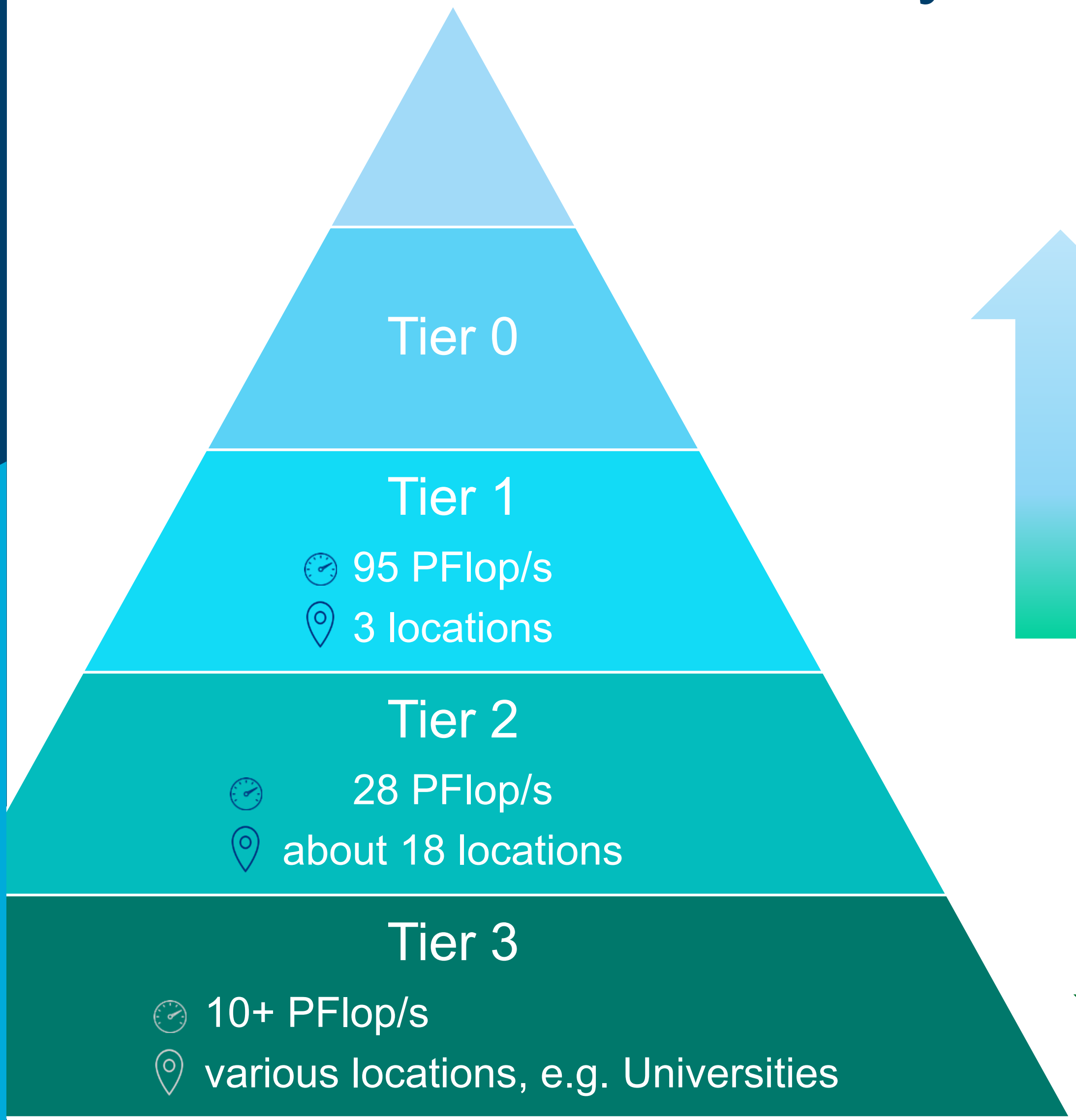
Tier 3

10+ PFlop/s

various locations, e.g. Universities

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

HPC Performance Pyramid



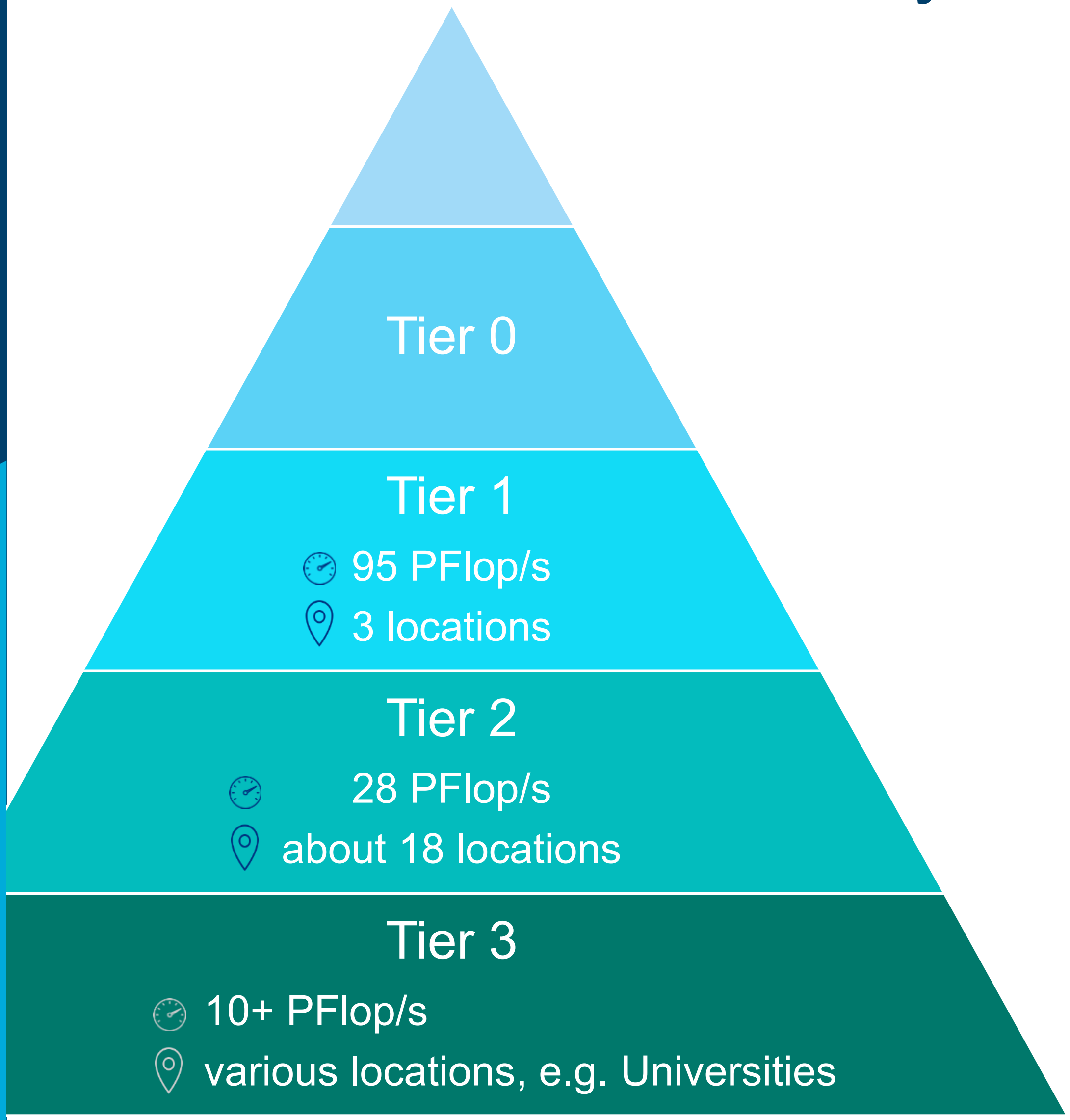
„capability computing“

- 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

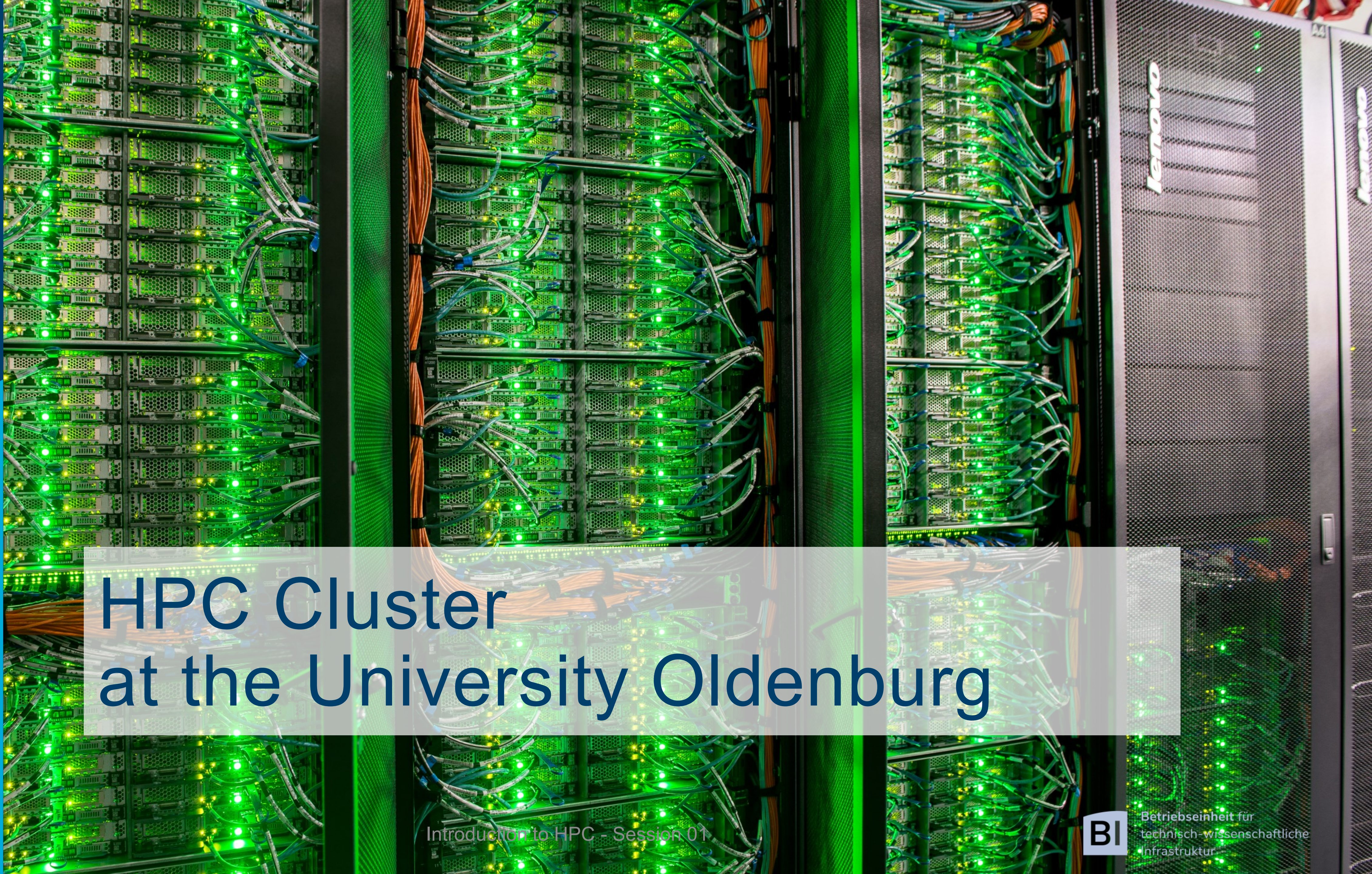
Accessing HPC Resources

- 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

North German Supercomputing Alliance (HLRN)

HLRN – NORDDEUTSCHER VERBUND FÜR HOCH- UND HÖCHSTLEISTUNGSRECHNEN

- 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/

Carl von Ossietzky Universität Oldenburg

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

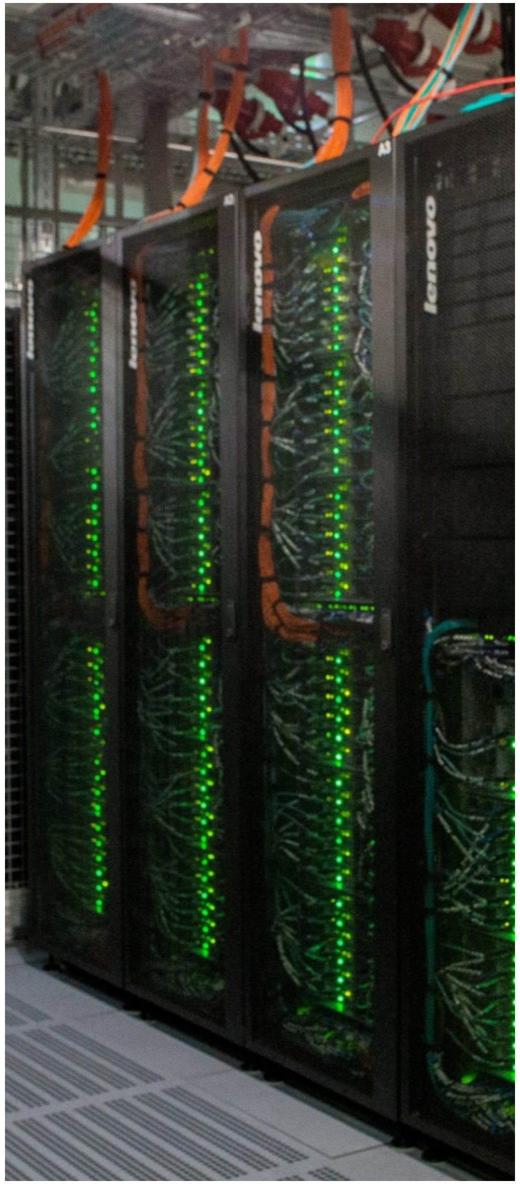
http://wiki.hpcuser.uni-oldenburg.de/index.php?title=Acknowledging_the_HPC_facilities_2016

Carl von Ossietzky Universität Oldenburg

HPC Facilities @ University Oldenburg

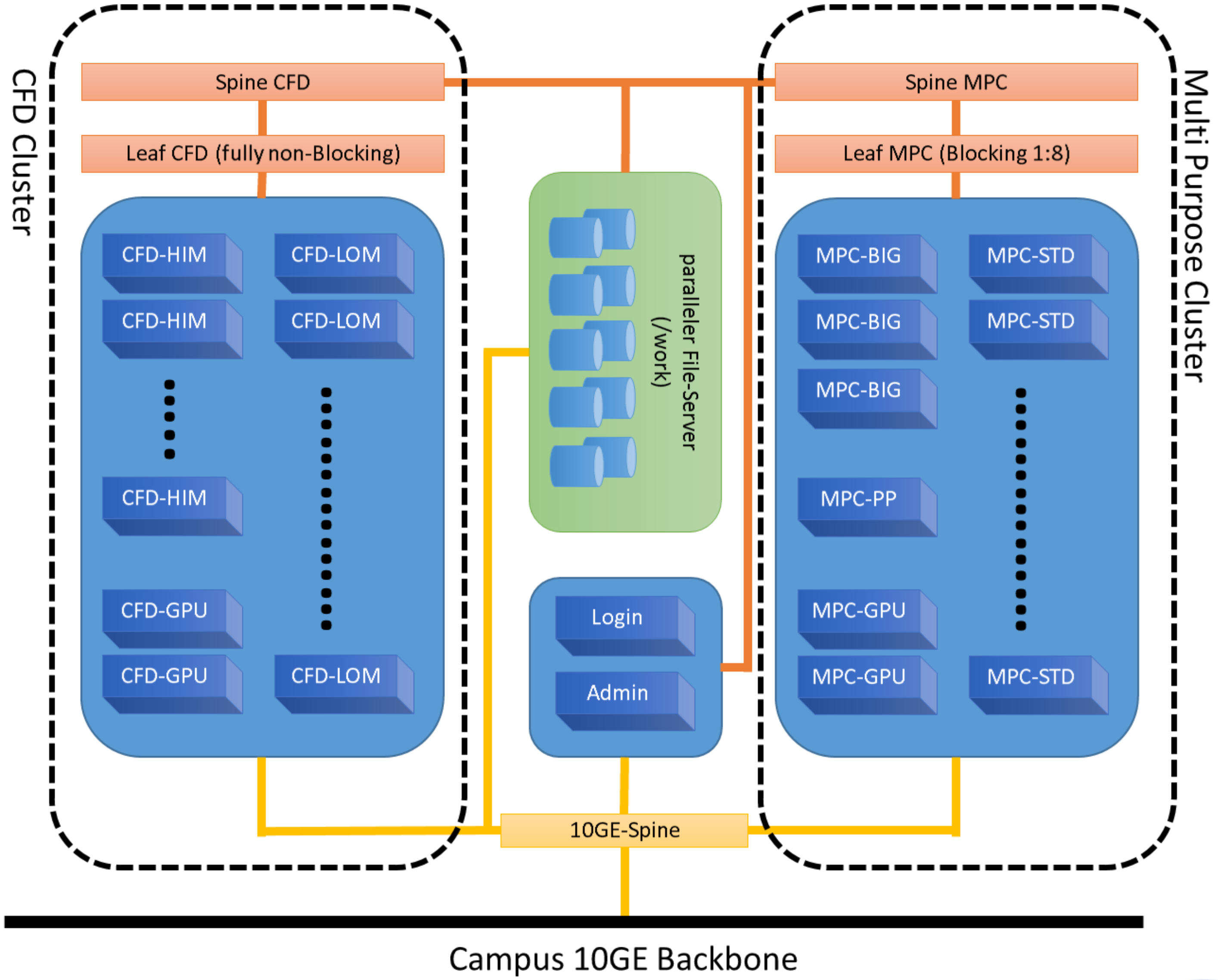
- shared HPC cluster CARL/EDDY
  - close to 600 compute nodes
  - 4 login and 2 administration nodes
  - Infiniband FDR interconnect for parallel computing
  - 10/1GE network
  - parallel file system (GPFS) with 900TB capacity
  - NFS mounted central storage

  - 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



Summary CARL & EDDY

Feature	CARL	EDDY	Total
Nodes	327	244	571
Cores	7.640	5.856	13.496
RAM	77 TB	21 TB	98 TB
GPFS	450 TB	450 TB	900 TB
local disks	360 TB	-	360 TB
Rpeak	277 Tflop/s	206 Tflop/s	483 Tflop/s
Rmax			457.2 Tflop/s

Rank 363 in Top500 https://www.top500.org/system/178942

Carl von Ossietzky Universität Oldenburg

TOP500 CERTIFICATE

The List.

CARL - Lenovo NeXtScale nx360M5, Xeon E5-2650v4 12C 2.2GHz, Infiniband FDR

Carl von Ossietzky University of Oldenburg, Germany

is ranked

No. 363

among the World's TOP500 Supercomputers

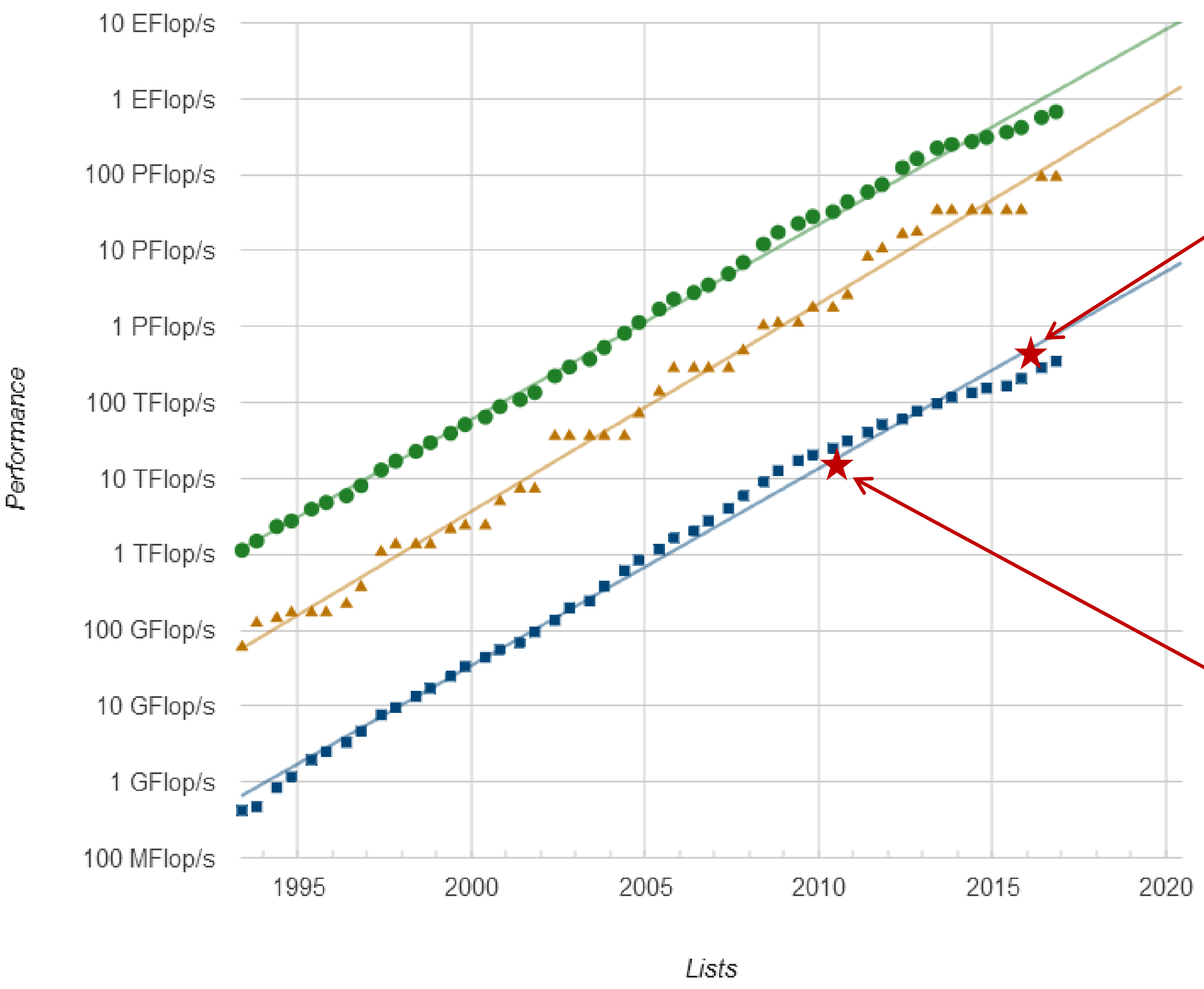
with 457.23 Tflop/s Linpack Performance

in the 48th TOP500 List published at SC16, Salt Lake City, UT on November 14th, 2016.

Congratulations from the TOP500 Editors

Erich Strohmaier	Jack Dongarra	Horst Simon	Martin Meuer
NERSC/Berkeley Lab	University of Tennessee	NERSC/Berkeley Lab	Prometeus

Top500 Performance Development



CARL/EDDY

- Rmax 457.2 Tflop/s
- Rank 363
- 571 nodes
- 13,500 cores
- 100 TB RAM

17x Rmax
1.6x nodes
3.3x cores
10x RAM

HERO/FLOW

- Rmax 27.2 Tflop/s
- not ranked
- 350 nodes
- 4,000 cores
- 10 TB RAM

Carl von Ossietzky Universität Oldenburg

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

Racks in Server Room

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