General Purpose High Performance Computing as Competitive Advantage for Scientists

Dieter Kranzlmüller

Munich Network Management Team
Ludwig-Maximilians-Universität München (LMU) &
Leibniz Supercomputing Centre (LRZ)
of the Bavarian Academy of Sciences and Humanities

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities

With approx. 250 employees
for more than 100.000 students and
for more than 30.000 employees
including 8.500 scientists

- European Supercomputing Centre
- National Supercomputing Centre (GCS)
- Regional Computer Centre for all Bavarian Universities
- Computer Centre for all Munich Universities

Photo: Ernst Graf

Gauss Centre for Supercomputing (GCS)

High(est) Performance Computing in Germany

- Combination of the 3 German national supercomputing centers:
  - John von Neumann Institute for Computing (NIC), Jülich
  - High Performance Computing Center Stuttgart (HLRS)
  - Leibniz Supercomputing Centre (LRZ), Garching n. Munich
- Founded on 13. April 2007
- Hosting member of PRACE
  (Partnership for Advanced Computing in Europe)

SuperMUC Phase 1 + 2

- 3,2 + 3,6 PetaFlop/s
- 147.456 (Thin) + 8.200 (Fat) + 3.840 (Phi) + 86.016 (Phase 2) Kerne
- 288 (Thin) + 52 (Fat) + 2.56 (Phi) + 194 (Phase 2) TByte Speicher
- 2,3 + 1,1 MWatt

Top 500 Supercomputer List (June 2012)

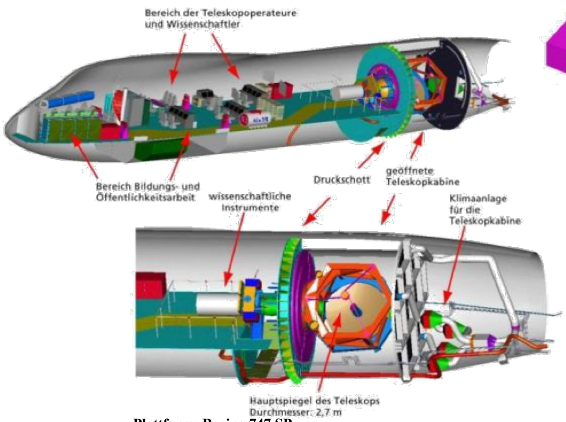
Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom / 2011 IBM	1572864	16324.75	20132.66	7890.0
2	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect / 2011 Fujitsu	705024	10510.00	11280.38	12659.9
3	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom / 2012 IBM	786432	8162.38	10066.33	3945.0
4	Leibniz Rechenzentrum Germany	**SuperMUC** - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR / 2012 IBM	147456	2897.00	3185.05	3422.7
5	National Supercomputing Center in Tianjin China	**Tianhe-1A** - NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050 / 2010 NUDT	186368	2566.00	4701.00	4040.0
6	DOE/SC/Oak Ridge National Laboratory United States	**Jaguar** - Cray XK6, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA 2090 / 2009 Cray Inc.	298592	1941.00	2627.61	5142.0
7	CINECA Italy	**Fermi** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom / 2012 IBM	163840	1725.49	2097.15	821.9
8	Forschungszentrum Juelich (FZJ) Germany	**JuQUEEN** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom / 2012 IBM	131072	1380.39	1677.72	657.5
9	CEA/TGCC-GENCI France	**Curie thin nodes** - Bullx B510, Xeon E5-2680 8C 2.700GHz, Infiniband QDR / 2012 Bull	77184	1359.00	1667.17	2251.0
10	National Supercomputing Centre in Shenzhen (NSCS) China	**Nebulae** - Dawning TC3600 Blade System, Xeon X5650 6C 2.66GHz, Infiniband QDR, NVIDIA 2050 / 2010 Dawning	120640	1271.00	2984.30	2580.0

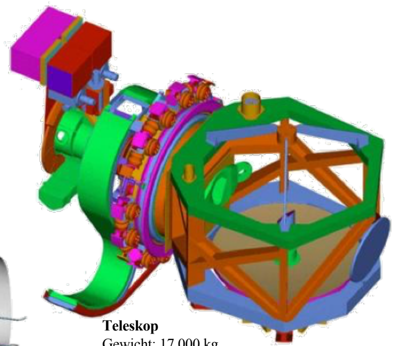
www.top500.org

LRZ Application Mix

- **Computational Fluid Dynamics**: Optimisation of turbines and wings, noise reduction, air conditioning in trains
- **Fusion**: Plasma in a future fusion reactor (ITER)
- **Astrophysics**: Origin and evolution of stars and galaxies
- **Solid State Physics**: Superconductivity, surface properties
- **Geophysics**: Earth quake scenarios
- **Material Science**: Semiconductors
- **Chemistry**: Catalytic reactions
- **Medicine and Medical Engineering**: Blood flow, aneurysms, air conditioning of operating theatres
- **Biophysics**: Properties of viruses, genome analysis
- **Climate research**: Currents in oceans
- ...

SOFIA - Stratospheric Observatory For Infrared Astronomy





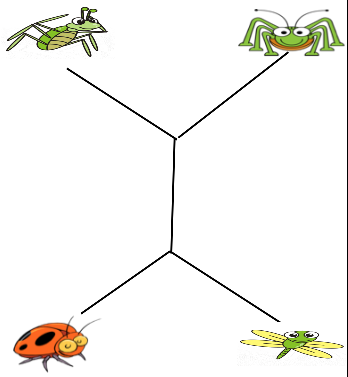
Teleskop
Gewicht: 17.000 kg
Durchmesser: 2.7m
Ansichtswinkel: 20° - 60°
Entwickelt und gebaut in Deutschland

Plattform: Boeing 747 SP
Max. Bruttogewicht: 300.000 kg
Spannweite: 60m
Max. Entfernung: 15.000km

Phylogenetic Tree Computation

ACGT
ACC
ACGG
AAGC

ACGT
ACC -
ACGG
AAGC



Sequencing → Alignment → Phylogenetic Tree

Alexandros Stamatakis, H-ITS

of possible trees for 150 species

- 4226517247809112252219618802377042809718932383449
  8822942857479880831434032178759024536798491951168
  3076494692867414802738570221298292428457687814873
  4552121861861600804474608426626044448936698500560
  2468116186441264227425440726676614927906540649360
  2976397461917469326750931190889241406694054603576
  66015625

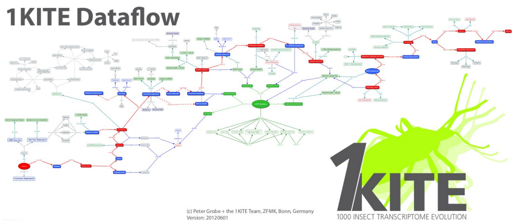
- ≈ 4.22×10^{301}

Alexandros Stamatakis, H-ITS

1KITE Project
(1,000 Insect Transcriptome Evolution)

- Alexandros Stamatakis
  Scientific Computing Group,
  Heidelberg Institute for Theoretical Studies (HITS) /
  Exelixis Lab

- „Big Data“ and High Performance Computing
- Novel software and applications needed
- Reading the data: only 1 minute (instead of 15 minutes)
- 1000 Processors: 17 hours (instead of 10 days)
- Load balancing



Scientific Results - Publications

Alexandros Stamatakis, H-ITS

Biodiversity

- Neotropical Rainforests are hyperdiverse ecosystems
- Since Humboldt and Bonpland, we know about the high animal and plant richness
- New study now finds that unicellular eukaryotes are even more diverse
- Particularly the parasitic Apicomplexa dominate these forests
- Their presence might drive the
- diversity of macro-organisms

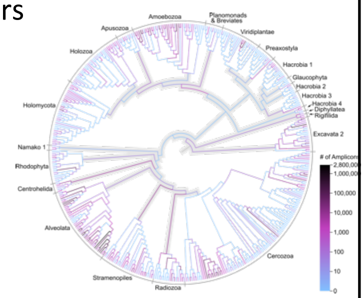
Micah Dunthorn/TU Kaiserslautern

https://natureecoevocommunity.nature.com/channels/521-behind-the-paper/posts/15402-a-larger-microbial-perspective-of-tropical-rainforests

Data Analysis (TU Kaiserslautern, HITS, KIT)

- More than 130 million DNA sequences were analysed
- Most of them belong to yet unknown microbial species
- Thus, a thorough method was necessary for classifying those sequences
- The method takes the evolutionary history of known species into account
- But this comes at the cost of increased computational needs
- Approximately 1 million computation hours on **SuperMUC** were necessary

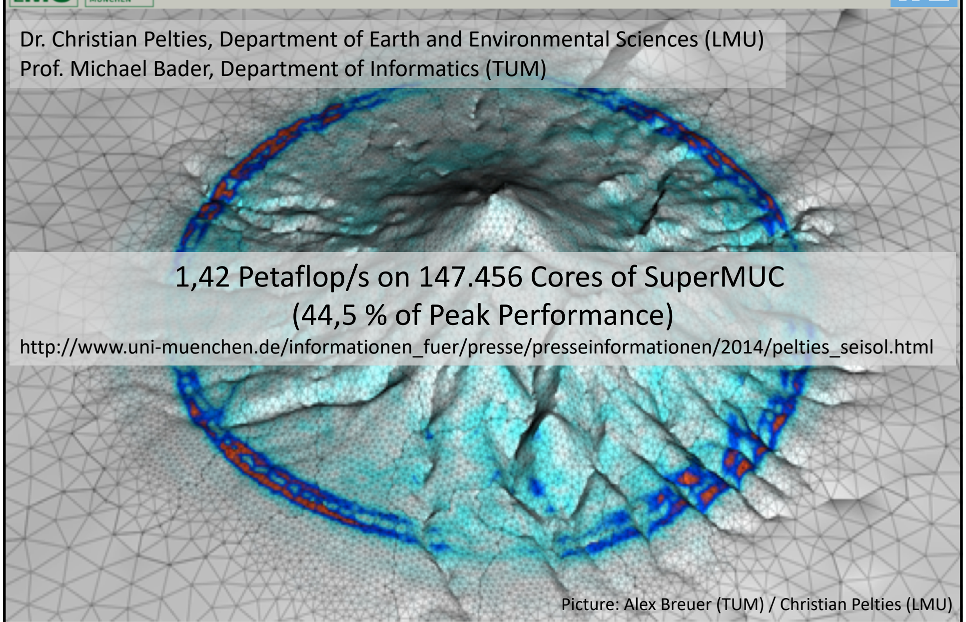
Mahé et al. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests.
Nature Ecology and Evolution 1:09.
DOI: 10.1038/s41559-017-0091



https://www.uni-kl.de/aktuelles/news/news/detail/News/neue-arten-entdeckt-mikroparasiten-tragen-zur-stabilitaet-des-oekosystems-im-regenwald-bei/

SeisSol - Numerical Simulation of Seismic Wave Phenomena

Dr. Christian Pelties, Department of Earth and Environmental Sciences (LMU)
Prof. Michael Bader, Department of Informatics (TUM)



1,42 Petaflop/s on 147.456 Cores of SuperMUC
(44,5 % of Peak Performance)

http://www.uni-muenchen.de/informationen_fuer/presse/presseinformationen/2014/pelties_seisol.html

Picture: Alex Breuer (TUM) / Christian Pelties (LMU)

Increasing numbers

Date	System	Flop/s	Cores
2000	HLRB-I	2 Tflop/s	1512
2006	HLRB-II	62 Tflop/s	9728
2012	SuperMUC	3200 Tflop/s	155656
2015	SuperMUC Phase II	3.2 + 3.6 Pflop/s	229960

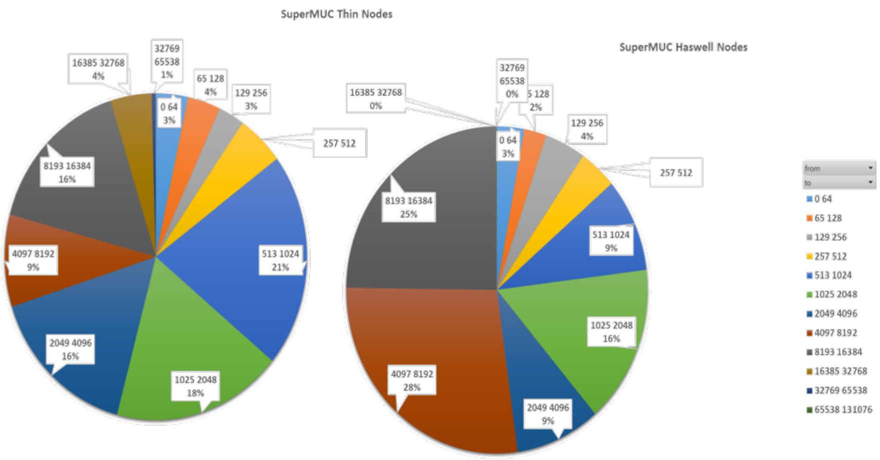
LRZ Extreme Scaling Workshop Series

- Results:

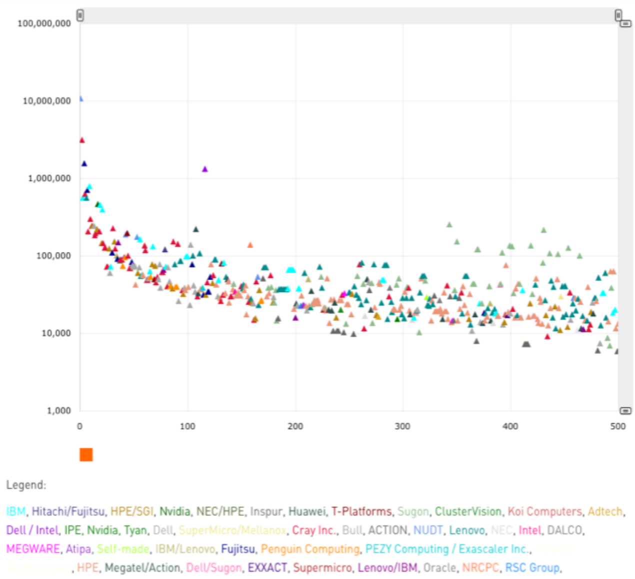
Name	MPI	# cores	Description	TFlop/s/island	TFlop/s max
Linpack	IBM	128000	TOP500	161	2560
Vertex	IBM	128000	Plasma Physics	15	245
GROMACS	IBM, Intel	64000	Molecular Modelling	40	110
Seissol	IBM	64000	Geophysics	31	95
waLBerla	IBM	128000	Lattice Boltzmann	5.6	90
LAMMPS	IBM	128000	Molecular Modelling	5.6	90
APES	IBM	64000	CFD	6	47
BQCD	Intel	128000	Quantum Physics	10	27

- Sustained TFlop/s on 64000/128000 cores

SuperMUC Jobsize 2015 (in Cores)

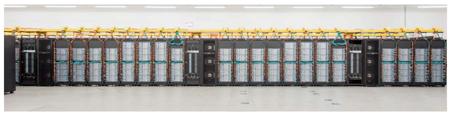


Statistics from www.top500.org



2016

SuperMUC System @ LRZ



Phase 1 (IBM System x iDataPlex):

- 3.2 PFlops peak performance
- 9216 IBM iDataPlex dx360M4 nodes in 18 compute node islands
- 2 Intel Xeon E5-2680 processors and 32 GB of memory per compute node
- 147,456 compute cores
- Network Infiniband FDR10 (fat tree)

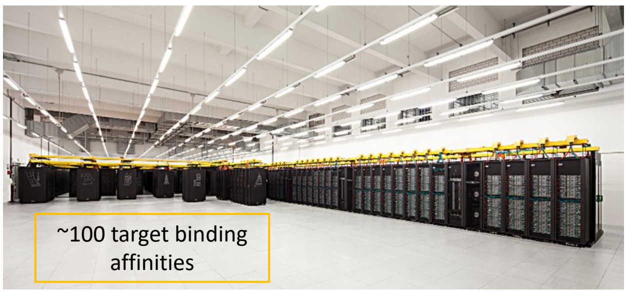
Phase 2 (Lenovo NeXtScale WCT):

- 3.6 PFlops peak performance
- 3072 Lenovo NeXtScale nx360M5 WCT nodes in 6 compute node islands
- 2 Intel Xeon E5-2697v3 processors and 64 GB of memory per compute node
- 86,016 compute cores
- Network Infiniband FDR14 (fat tree)

Common GPFS file systems with 10 PB and 5 PB usable storage size respectively
Common programming environment
Direct warm-water cooled system technology

Molecular Simulation for Personalized Medicinem (Prof. Peter Coveney, UCL)

- Running on all cores of SuperMUC Phase1+2

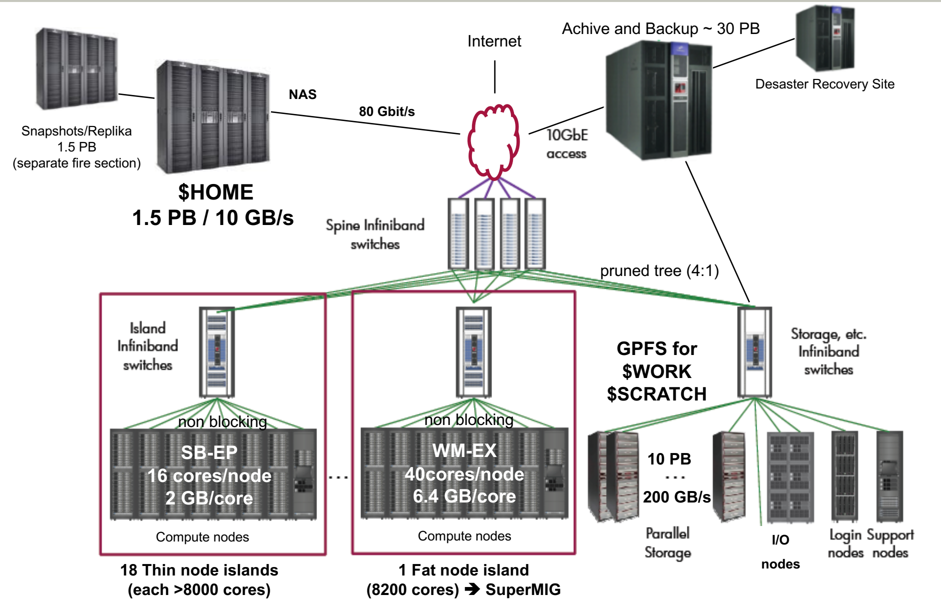


- Docking simulation of potentials drugs for breast cancer
- 37 hours total run time
- 241,672 cores
- 8.900.000 CPU hours
- 5 Terabytes of data produced

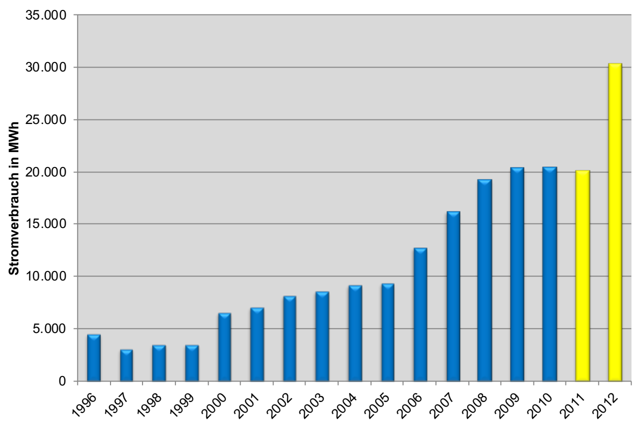
http://www.compat-project.eu

http://www.compbiomed.eu

SuperMUC Architecture



Power Consumption at LRZ



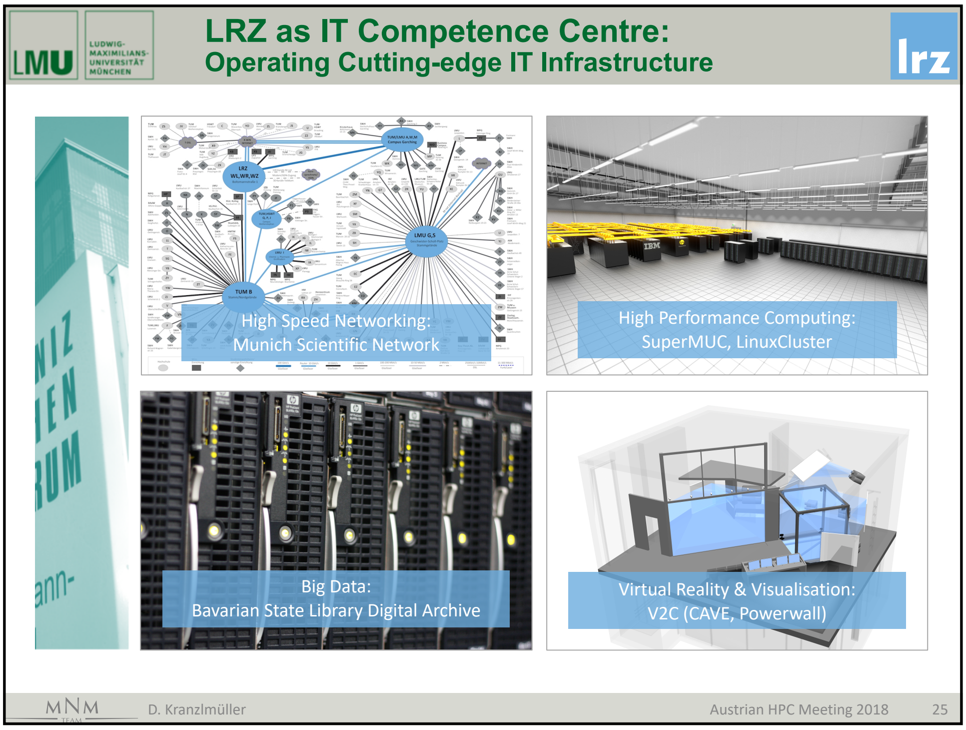
Cooling SuperMUC

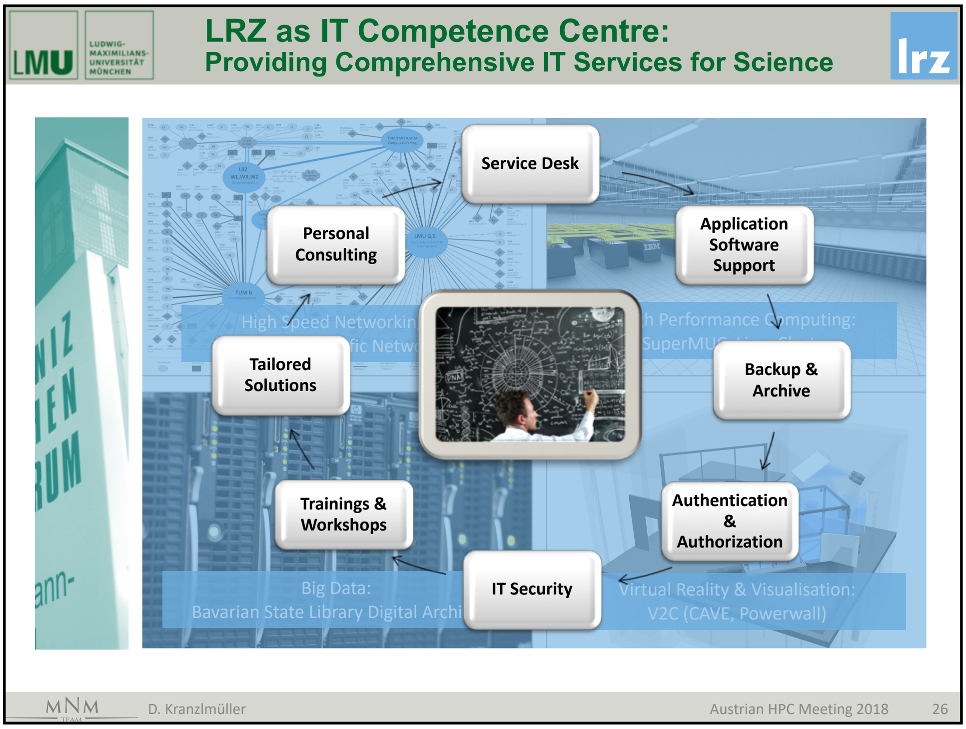
SuperMUC Phase 2 @ LRZ

- Intel Xeon E5 2697v3 CPU
- Direct liquid cooling
- Energy-aware scheduling

- Total annual savings: ~2 Mio. € for SuperMUC Phase 1 and 2

Photos: Torsten Bloth, Lenovo


LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
LRZ as IT Competence Centre:
Operating Cutting-edge IT Infrastructure
lrz
High Speed Networking:
Munich Scientific Network
High Performance Computing:
SuperMUC, LinuxCluster
Big Data:
Bavarian State Library Digital Archive
Virtual Reality & Visualisation:
V2C (CAVE, Powerwall)
MNM
D. Kranzlmüller
Austrian HPC Meeting 2018
25


LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
LRZ as IT Competence Centre:
Providing Comprehensive IT Services for Science
lrz
Service Desk
Personal Consulting
Application Software Support
Tailored Solutions
Backup & Archive
Trainings & Workshops
Authentication & Authorization
IT Security
MNM
D. Kranzlmüller
Austrian HPC Meeting 2018
26

Partnership Initiative Computational Sciences πCS

- **Individualized services** for selected scientific groups – **flagship role**
  - Dedicated point-of-contact
  - Individual support and guidance and targeted training &…
  - Planning dependability for use case specific optimi… …uctures
  - Early access to latest IT infrastructure (hard- a… …developments and specification of future requirements
  - Access to IT competence network a… …t CS and Math departments
- **Partner contribution**
  - Embedding IT expe… …oups
  - Joint research… …uding funding)
  - Scientifi… …– equal footing – joint publications
- **LRZ b…**
  - U… …ding the (current and future) needs and requirements of the res…ective scientific domain
  - Developing future services for all user groups
  - Thematic focusing: **Environmental Computing**

The LRZ Partnership Initiative Computational Science (piCS)
http://www.sciencedirect.com/science/article/pii/S1877050914003433

piCS Cookbook

1. Choose focus topics to serve as lighthouse
   - National agreement within GCS: LRZ focuses on Environment (& Energy)
2. Choose user communities
   - Already active at LRZ?
   - Not active at LRZ?
3. Invite them for introductory piCS Workshops
   - Show faces & tour
   - Discussion on joint topics, requirements, interests, ...
4. Establish links between communities and specific points-of-contact
   - Whom to talk to, if there are questions?
   - When to talk to them? In general, as early as possible
   - Maybe, place people into the research groups (weekly, for a certain period)
5. Run joint lectures (e.g. hydrometeorology and computer science)
6. Apply for joint projects
7. Use **HPC Machines** efficiently

What‘s next? A View into the Future

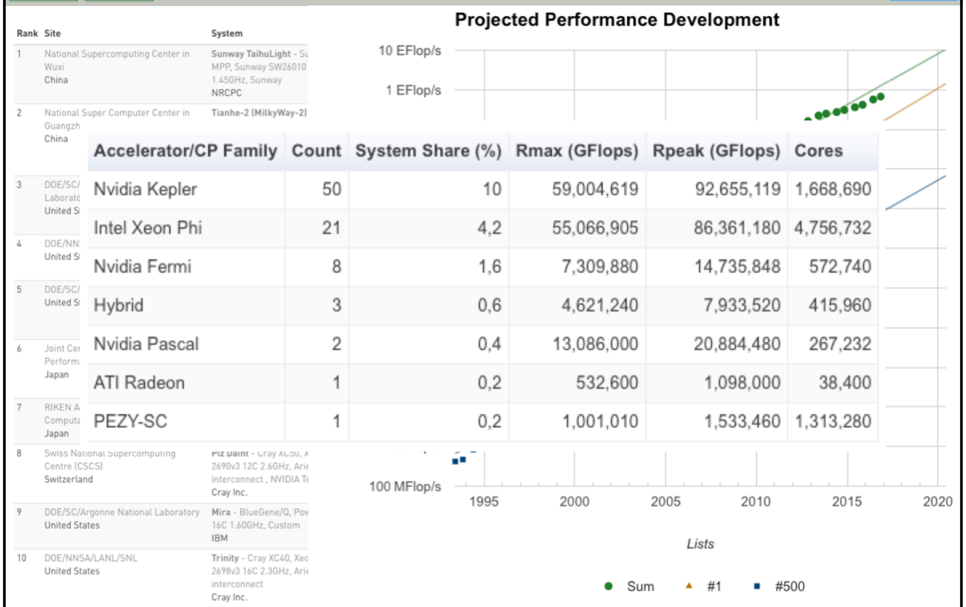
Until today:

- HLRB-II (pre-SuperMUC): Top 500 06/2007: 56,5 Tflop/s
- SuperMUC Phase 1: Top 500 06/2012: 2897 Tflop/s

Coming up:

- **SuperMUC NG** (Next Generation)

Consulting the Top 500 List - www.top500.org



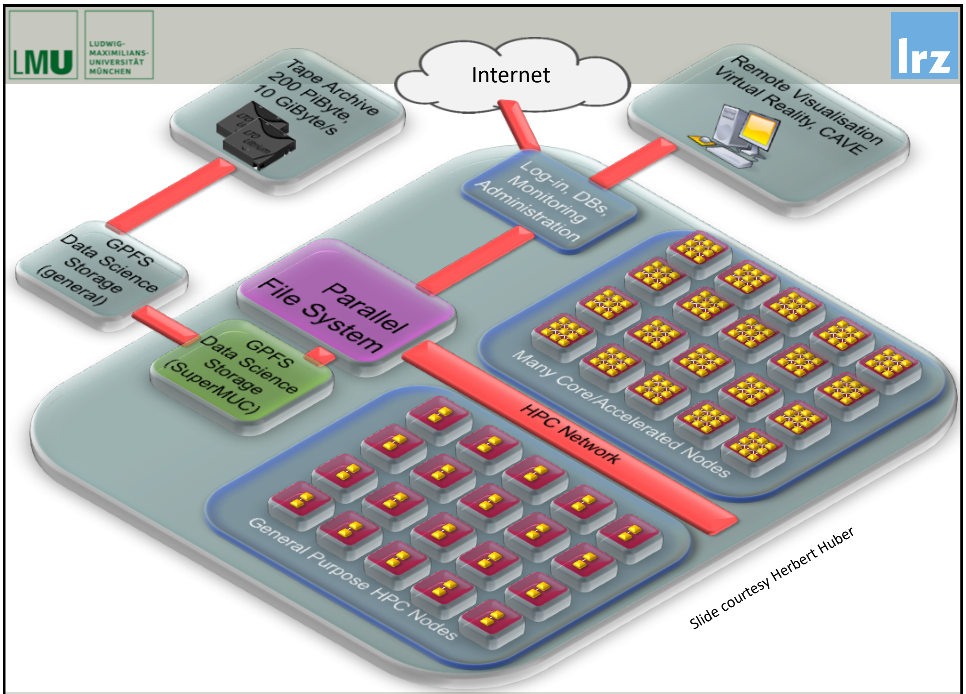
Accelerator/CP Family	Count	System Share (%)	Rmax (GFlops)	Rpeak (GFlops)	Cores
Nvidia Kepler	50	10	59,004,619	92,655,119	1,668,690
Intel Xeon Phi	21	4,2	55,066,905	86,361,180	4,756,732
Nvidia Fermi	8	1,6	7,309,880	14,735,848	572,740
Hybrid	3	0,6	4,621,240	7,933,520	415,960
Nvidia Pascal	2	0,4	13,086,000	20,884,480	267,232
ATI Radeon	1	0,2	532,600	1,098,000	38,400
PEZY-SC	1	0,2	1,001,010	1,533,460	1,313,280

What's next? A View into the Future

Until today:
- HLRB-II (pre-SuperMUC): Top 500 06/2007: 56,5 Tflop/s
- SuperMUC Phase 1: Top 500 06/2012: 2897 Tflop/s

Coming up:
- **SuperMUC NG** (Next Generation)

- **Selection criteria:**
  - LRZ application mix (compute,memory,bandwidth characteristics)
    - Number of cores
    - Memory per core
    - Interconnect
  - Accelerators (Manycore, GPGPU, ...)
  - Virtualization (Docker, Cloud, ...)
  - Workflow engines, HTC applications, ...
  - Power consumption (in total, over time, ...)



What‘s next? A View into the Future

Until today:

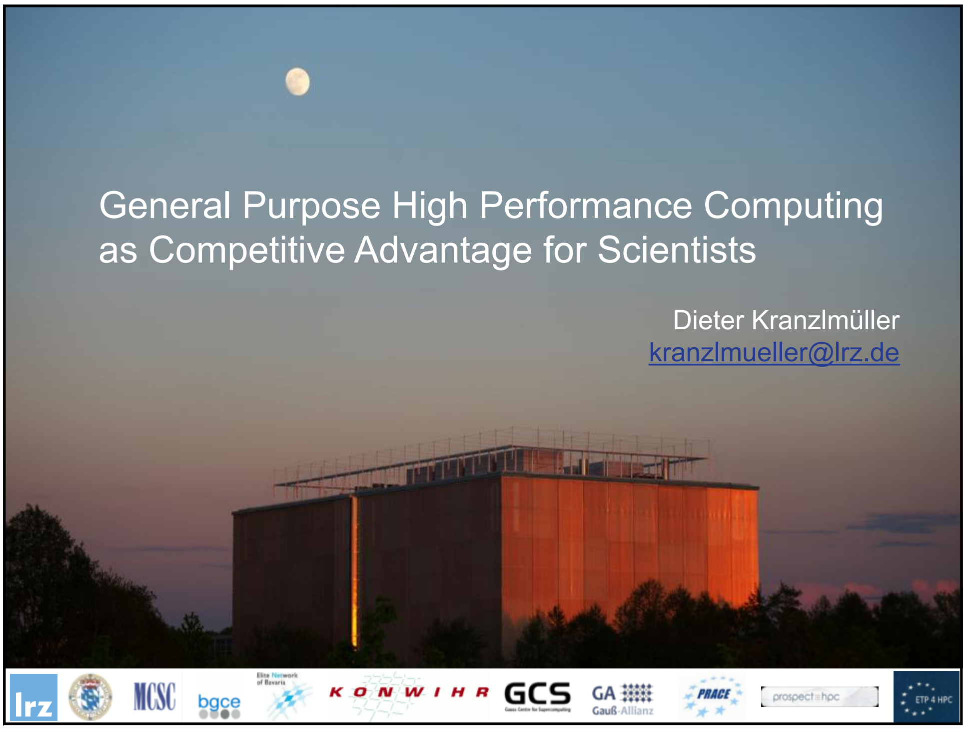
- HLRB-II (pre-SuperMUC): Top 500 06/2007: 56,5 Tflop/s
- SuperMUC Phase 1: Top 500 06/2012: 2897 Tflop/s

Coming up:

- **SuperMUC NG** (Next Generation)
  - Contract signed on 14 December 2017
  - Intel/Lenovo System
  - Peak performance: 26,7 Pflop/s
  - > 300.000 Compute Cores (Intel Xeon Scalable Processor)
  - Intel Omni-Path Architecture
  - > 700 TByte main memory
  - > 70 PByte disk storage
  - User operation: End 2018

Conclusions

- High Performance Computing provides **competitive advantages** for scientists
- **General purposes architectures** can deliver high performance to a broad range of applications
- **User support is essential** and requires competent approaches and **partnership collaborations**
- HPC strategy needs to include
  - Sufficient user support man power and collaboration models
  - Scalability across different layers of the HPC performance pyramid
  - Cyclic renewal of systems to keep up to developments


General Purpose High Performance Computing
as Competitive Advantage for Scientists
Dieter Kranzlmüller
kranzlmueller@lrz.de