

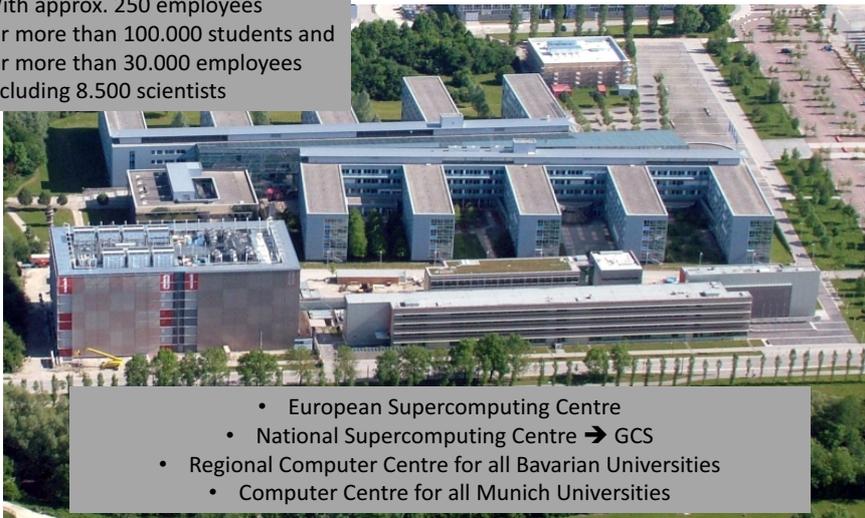
Is Heterogeneity A Solution To The HPC Dilemma?

Dieter Kranzlmüller

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



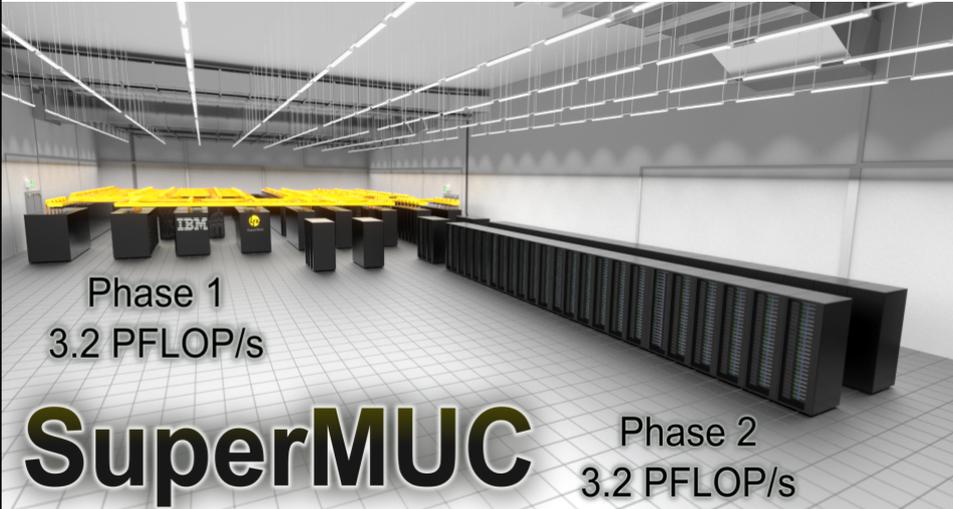
With approx. 250 employees
for more than 100.000 students and
for more than 30.000 employees
including 8.500 scientists



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

Photo: Ernst Graf

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN **SuperMUC Phase 1 + 2** lrz



Phase 1
3.2 PFLOP/s

SuperMUC Phase 2
3.2 PFLOP/s

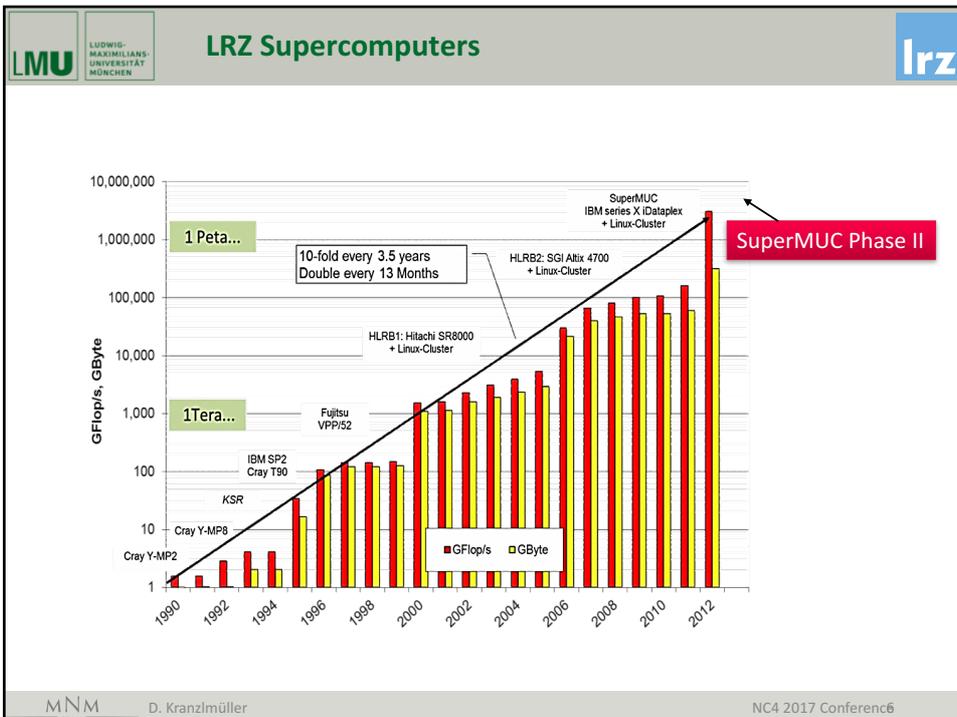
MNM D. Kranzmüller NC4 2017 Conference 3

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN **LRZ Application Mix** lrz

- **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

MNM D. Kranzmüller NC4 2017 Conference 4

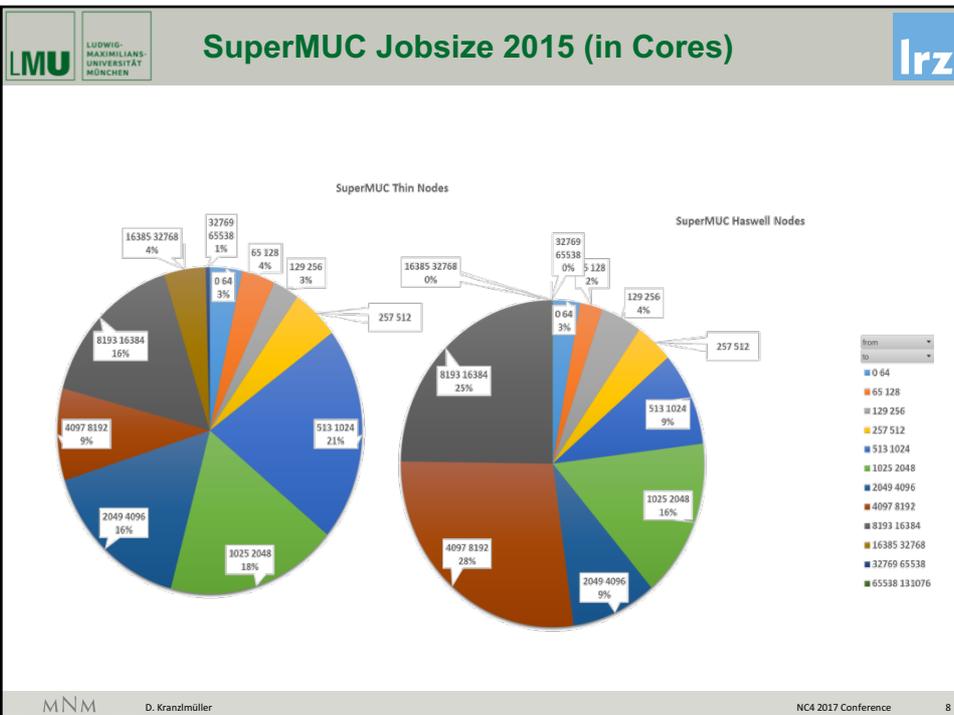
Rank	Site	Computer/Year Vendor	Cores	R _{max}	R _{peak}	Power
1	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom / 2011 IBM	1572864	16324.75	20132.66	7890.0
2	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIII/fx 2.0GHz, Tofu interconnect / 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

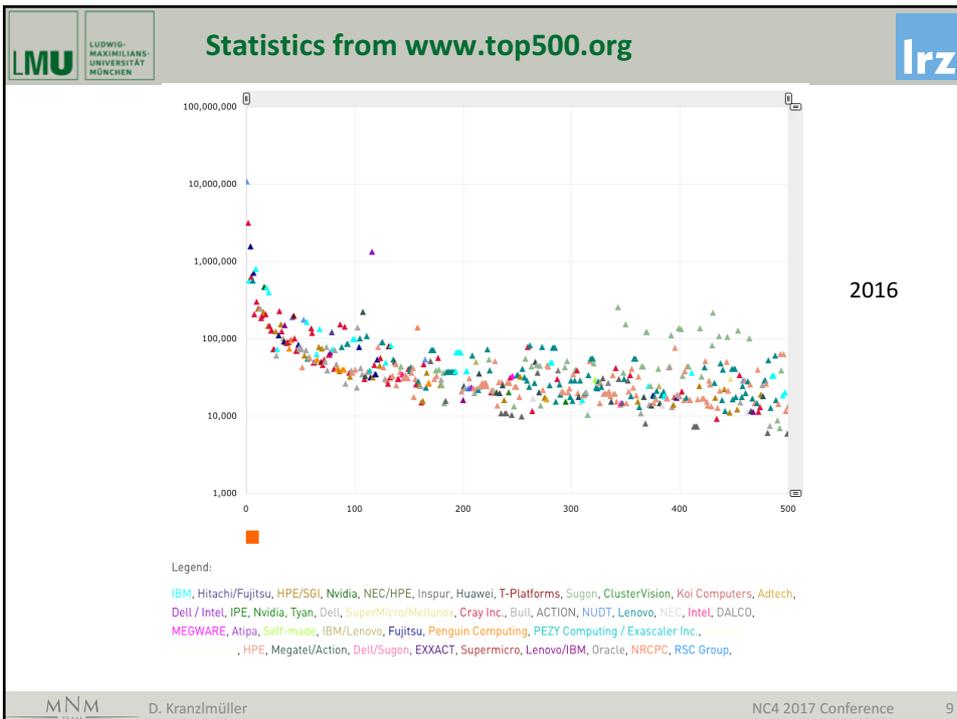


Increasing numbers



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 + 3.2 Pflop/s	229960





LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

SeisSol - Numerical Simulation of Seismic Wave Phenomena

lrz

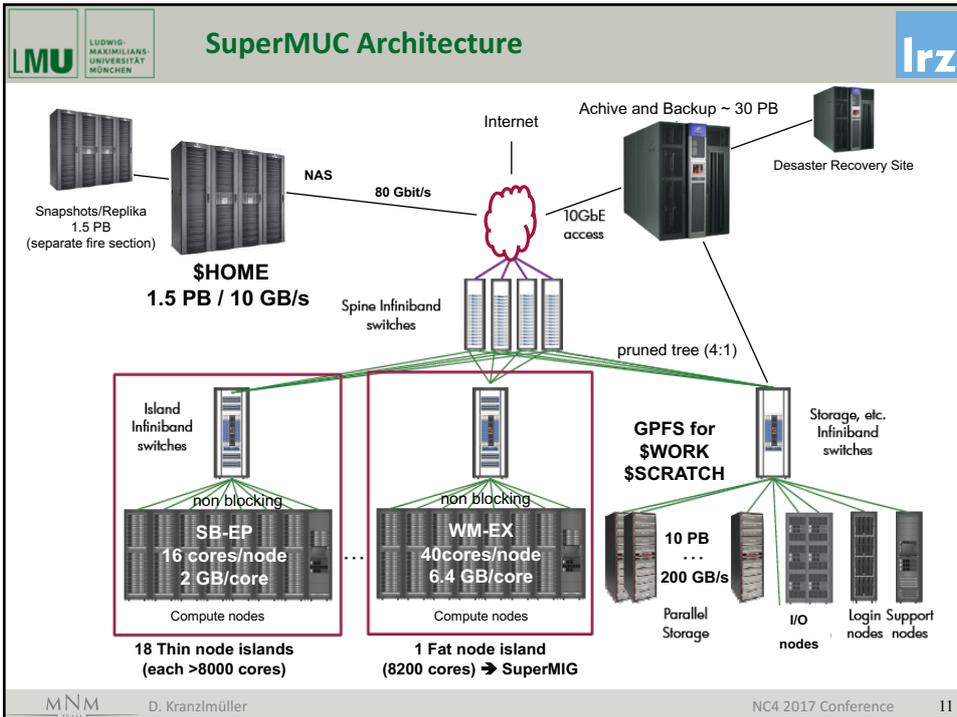
Dr. Christian Pelties, Department of Earth and Environmental Sciences (LMU)
 Prof. Michael Bader, Department of Informatics (TUM)

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)

MNM D. Kranzmüller NC4 2017 Conference 10



SuperMUC System @ LRZ

The image shows two phases of the SuperMUC system at LRZ:

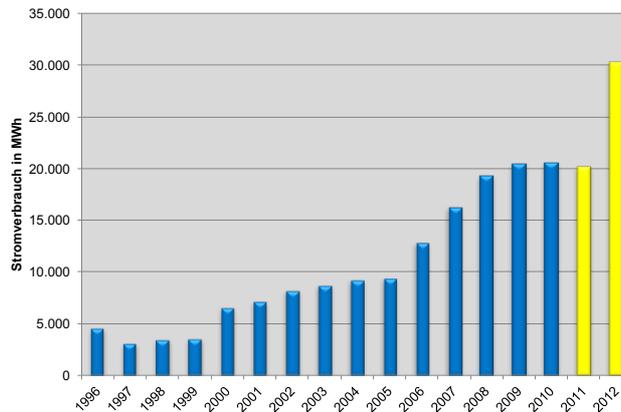
- Phase 1 (IBM System x iDataPlex):**
 - 3.2 PFlops peak performance
 - 9216 IBM iDataPlex dx360M4 nodes in 18 compute node islands
 - 2 Intel Xeon E5-2680 processors and 32 GB of memory per compute node
 - 147,456 compute cores
 - Network Infiniband FDR10 (fat tree)
- Phase 2 (Lenovo NeXtScale WCT):**
 - 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

Logos: LMU (LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN), lrz, MNM, D. Kranzmüller, NC4 2017 Conference, 12

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)



LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN **Cooling SuperMUC** lrz

MNM D. Kranzmüller NC4 2017 Conference 15

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN **SuperMUC Phase 2 @ LRZ** 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

MNM D. Kranzmüller Slide: Herbert Huber NC4 2017 Conference 16

- 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>



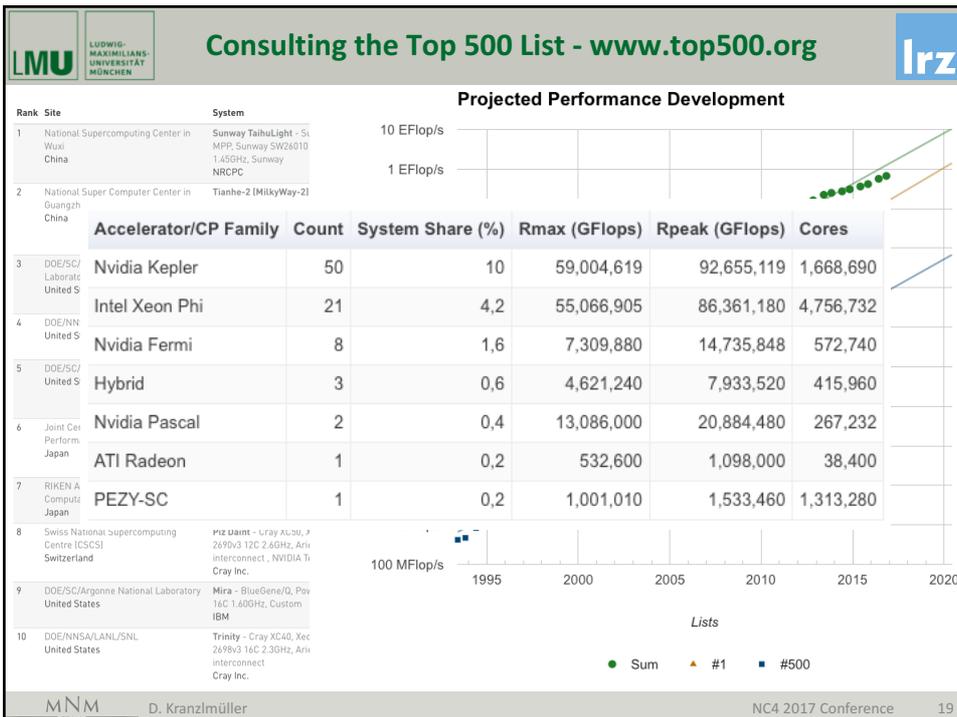
<http://www.compbioimed.eu>

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) – Procurement on-going



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) – Procurement on-going
- **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, ...)

MNM D. Kranzlmüller NC4 2017 Conference 20

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN Partnership Initiative Computational Sciences πCS lrz

- **Individualized services** for selected scientific groups – **flagship role**
 - Dedicated point-of-contact
 - Individual support and guidance and targeted training & education
 - Planning dependability for use case specific optimization and development
 - Early access to latest IT infrastructure (hard- and software) and specification of future requirements
 - Access to IT competence network and expertise at CS and Math departments
- **Partner contribution**
 - Embedding IT expertise into scientific groups
 - Joint research projects (including funding)
 - Scientific publications – equal footing – joint publications
- **LRZ best practice**
 - Understanding the (current and future) needs and requirements of the respective scientific domain
 - Developing future services for all user groups
 - Thematic focusing: **Environmental Computing**

<http://www.sciencedirect.com/science/article/pii/S1877050914003433>

MNM D. Kranzmüller NC4 2017 Conference 21

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN piCS Cookbook lrz

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 **Heterogeneity on HPC Machines** efficiently

MNM D. Kranzmüller NC4 2017 Conference 22

Is Heterogeneity A Solution To The HPC Dilemma?

Dieter Kranzmüller
kranzmueller@lrz.de

