

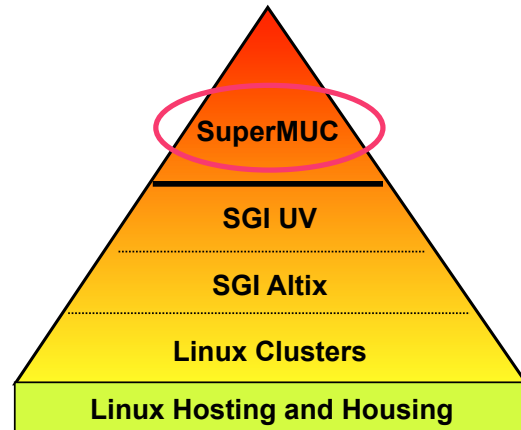
Trends in Computation, Communication and Storage: Consequences for Data-Intensive Science

Dieter Kranzlmüller

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



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



Video: SuperMUC rendered on SuperMUC by LRZ

<http://youtu.be/OIAS6iiqWrQ>

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

- **Curie @ GENCI:**
Bull Cluster, 1.7 PFlop/s
- **FERMI @ CINECA:**
IBM BG/Q, 2.1 PFlop/s
- **Hermit @ HLRS:**
Cray XE6, 1 Pflop/s
- **JUQUEEN @ FZJ:**
IBM Blue Gene/Q, 5.9 PFlop/s
- **MareNostrum @ BSC:**
IBM System X iDataPlex, 1 PFlop/s
- **SuperMUC @ LRZ:**
IBM System X iDataPlex, 3.2 PFlop/s



- Single pan-European Peer Review
- <http://www.prace-project.eu/Call-Announcements?lang=en>
- Early Access Call in May 2010
 - 68 proposals asked for 1870 Million Core hours
 - 10 projects granted with 328 Million Core hours
 - Principal Investigators from D (5), UK (2) NL (1), I (1), PT (1)
 - Involves researchers from 31 institutions in 12 countries
- Further calls being scheduled (every 6 months)
 - Call open February > Access September same year
 - Call open September > Access March next year
- Example from 8th Regular Call closed on 15 October 2013
 - Spatially adaptive radiation-hydrodynamical simulations of reionization
 - Project leader: Dr Andreas Pawlik, Max Planck Society, GERMANY
 - Research field: Universe Sciences
 - Resource Awarded: 33,800,000 core hours on SuperMUC, Germany;



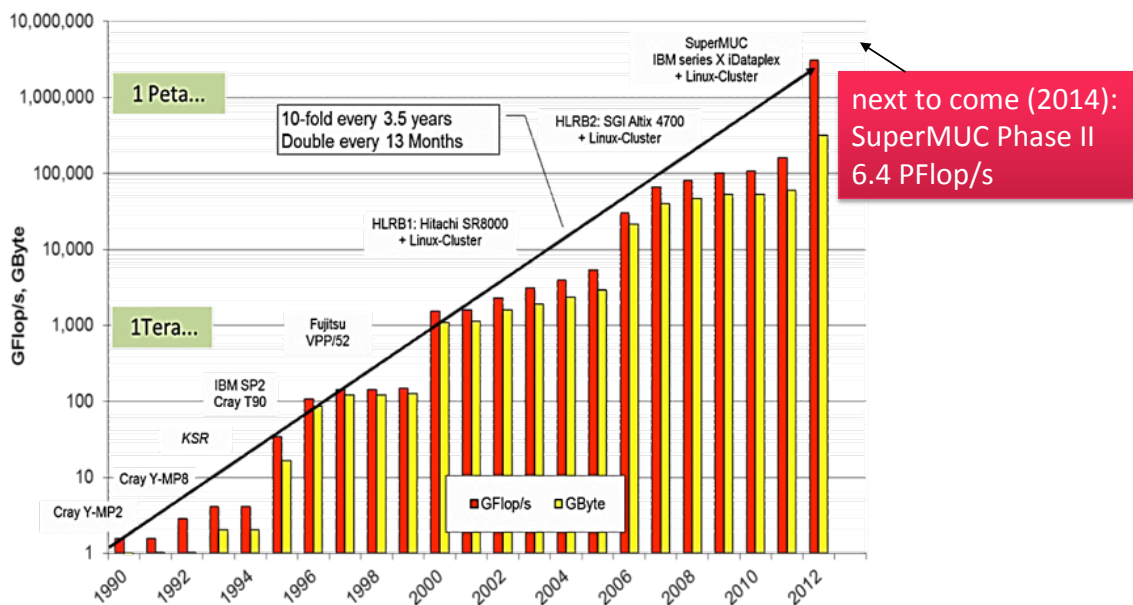
Trends in Computation, Communication and Storage

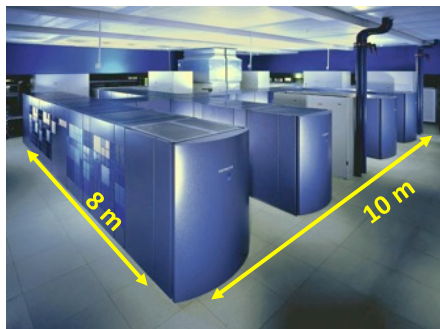
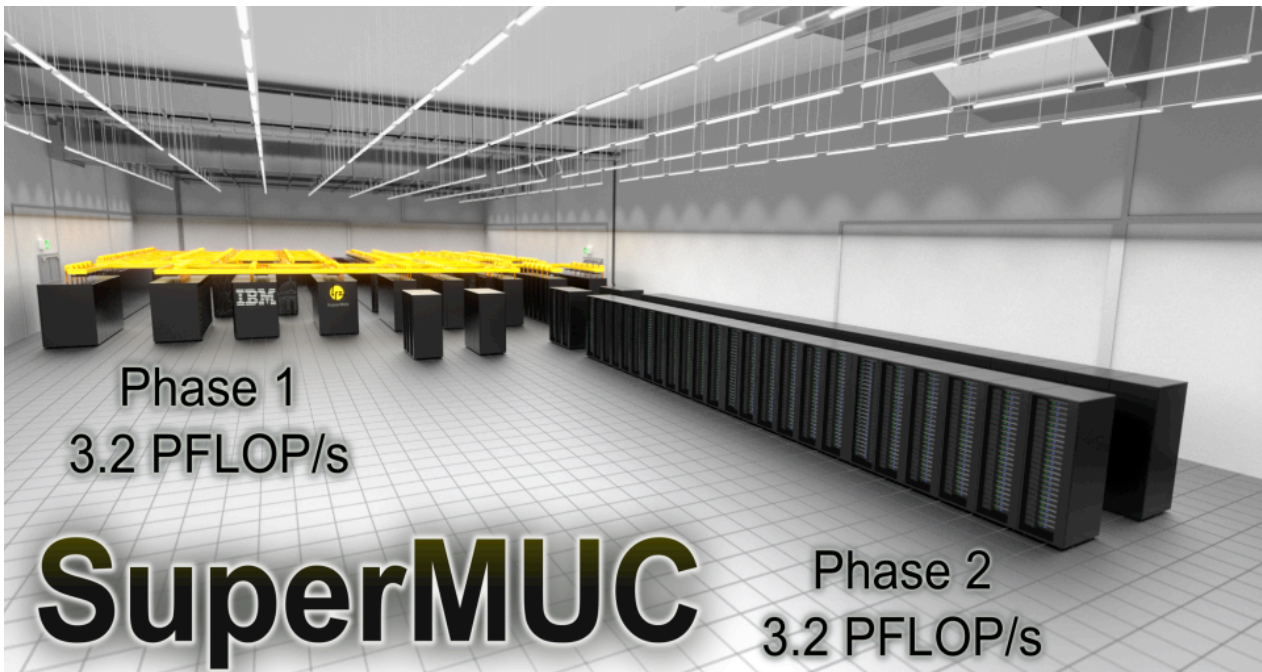
RANK	SITE	SYSTEM	CORES	RMAX (TFLOP/S)	RPEAK (TFLOP/S)	POWER (KW)
1	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945

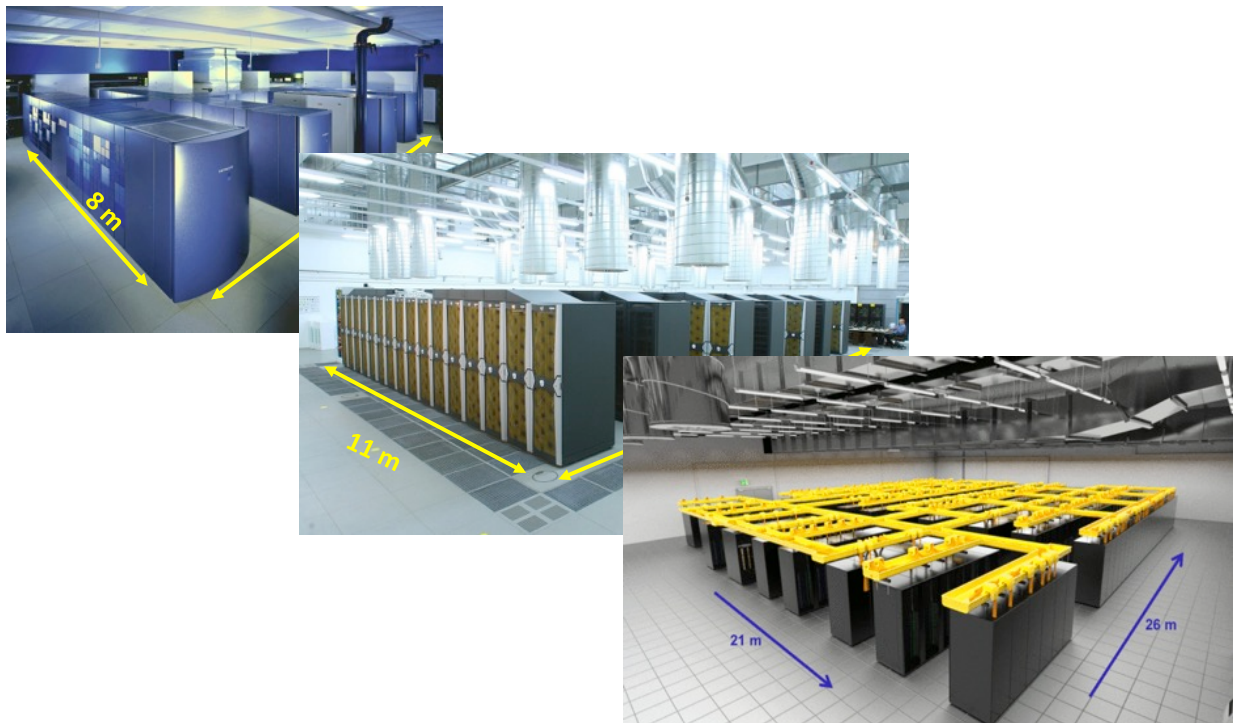


RANK	SITE	SYSTEM	CORES	RMAX (TFLOP/S)	RPEAK (TFLOP/S)	POWER (KW)
1	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808

RANK	SITE	SYSTEM	CORES	RMAX (TFLOP/S)	RPEAK (TFLOP/S)	POWER (KW)
1	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
14	Leibniz Rechenzentrum Germany	SuperMUC - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM	147,456	2,897.0	3,185.1	3,423







Picture: Horst-Dieter Steinhöfer

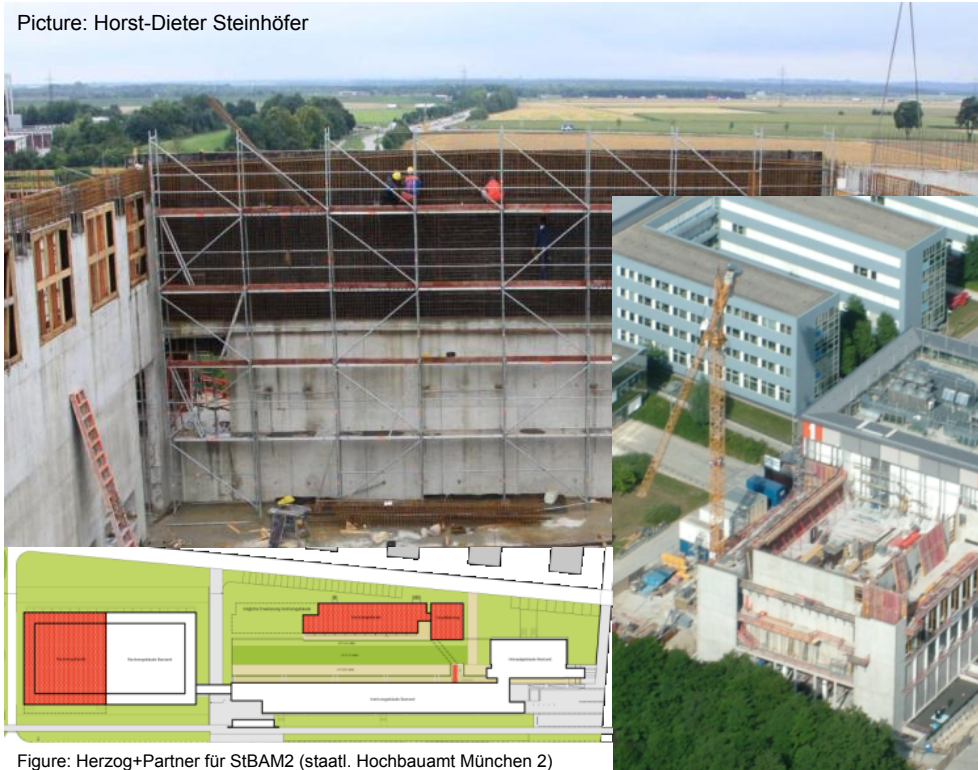
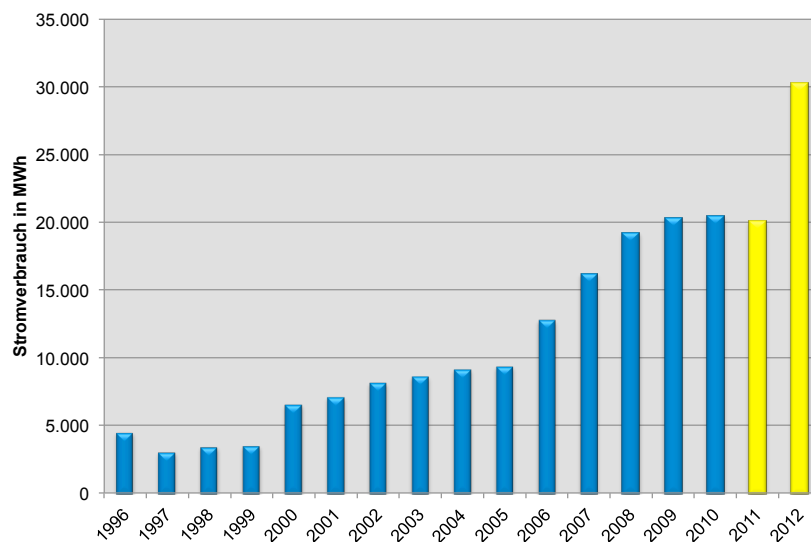
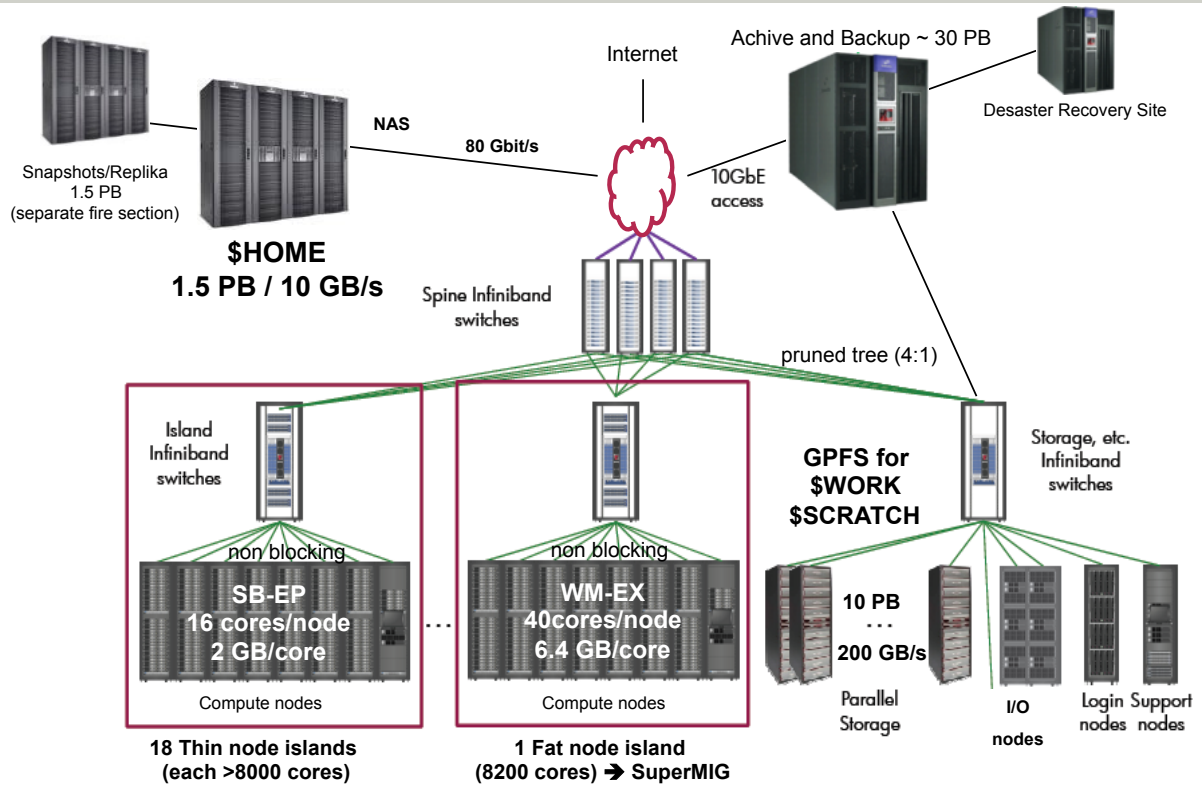
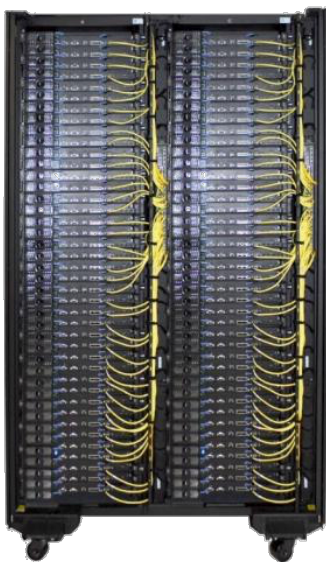


Figure: Herzog+Partner für StBAM2 (staatl. Hochbauamt München 2)

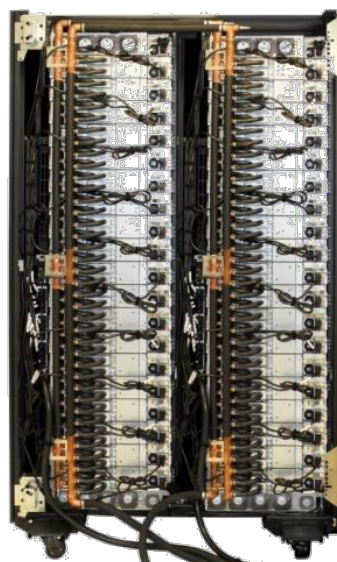
Picture: Ernst A. Graf

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
2014	SuperMUC Phase II	3.2 + 3.2 Pflop/s	229960



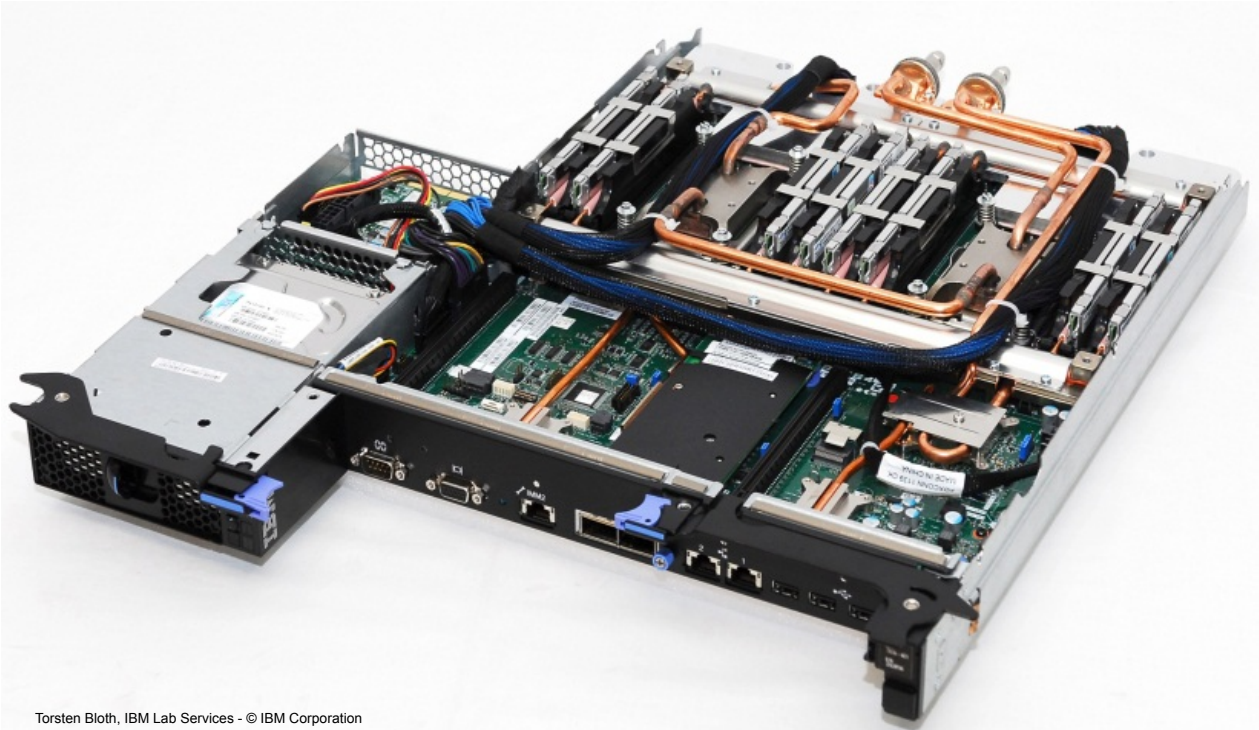


iDataPlex DWC Rack
w/ water cooled nodes
(front view)



iDataPlex DWC Rack
w/ water cooled nodes
(rear view of water manifolds)

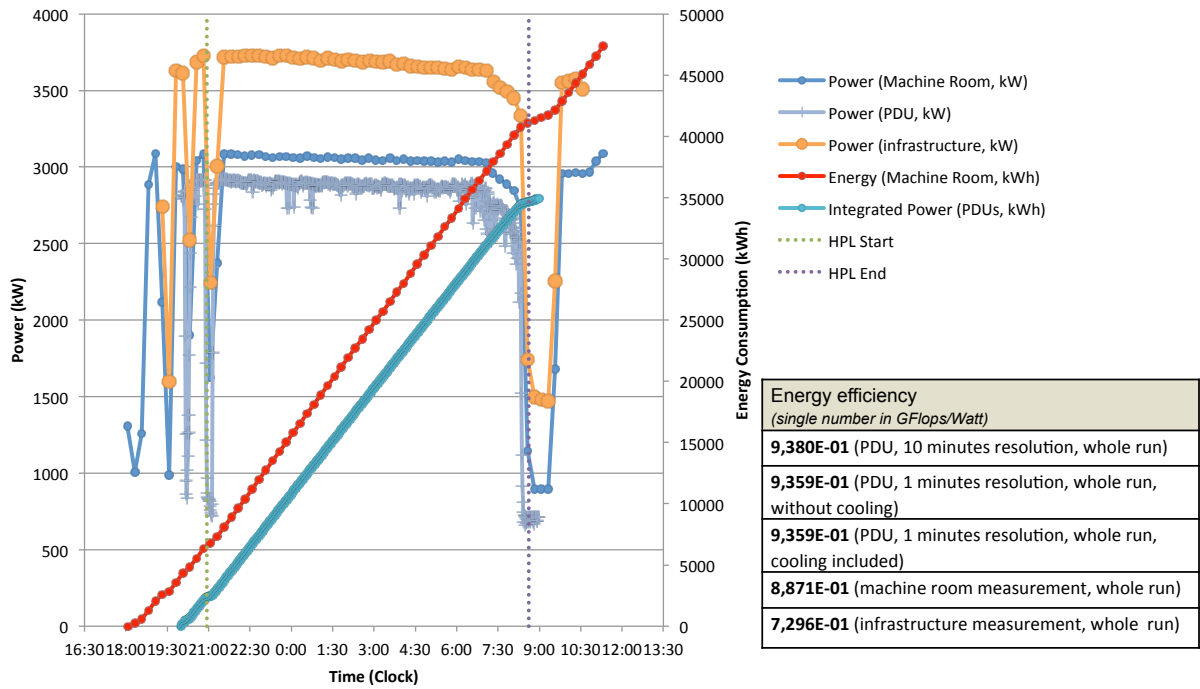
Torsten Bloth, IBM Lab Services - © IBM Corporation



Torsten Bloth, IBM Lab Services - © IBM Corporation



Photos: StBAM2 (staatl. Hochbauamt München 2)

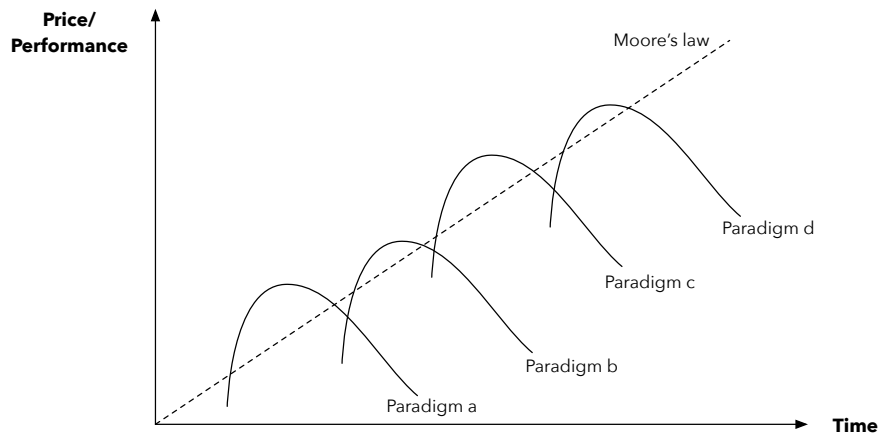


Name	MPI	# cores	Description	TFlop/s/island	TFlop/s max
Linpack	IBM	★ 128000	TOP500	161	2560
Vertex	IBM	★ 128000	Plasma Physics	15	245
GROMACS	IBM, Intel	★ 64000	Molecular Modelling	40	110
Seissol	IBM	★ 64000	Geophysics	31	95
waLBerla	IBM	★ 128000	Lattice Boltzmann	5.6	90
LAMMPS	IBM	★ 128000	Molecular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



Consequences for Data-intensive Sciences

- Moore's law is deceptive
 - Medium term based on paradigm shifts
 - Most recent: "MHz race" -> "core race"
 - Drop in app performance, unless code is adapted/rewritten



- Nevertheless, technology is getting cheaper
 - Eventually...
- Key: Time-to-solution!
- Corresponding organizational changes are getting more expensive
 - More complex collaborations: international and interdisciplinary
 - Standards, legal requirements
 - More and more tight QoS demands that need to be maintained through the change
- Working with e-Infrastructure makes it possible to
 - Adapt technology to "surf on next wave" (instead of inhaling it)
 - Prepare organization and processes to maximize benefits
 - Make accurate predictions of system-level capacities and capabilities of tomorrow

- Big Data
 - New Big Data: Open datasets, mashups, new discovery tools – Fashion/Hype?
 - Old (really) Big Data: meteo, climate, bio,...
- DRIHM is a perfect example of Data-Intensive Science
- Strong links with practical/operational uses
- e-Infrastructure needs are intrinsic to DRIHM
- General purpose characteristic needs to be adapted to specific needs of DRIHM applications
- “Not just for (single) experts anymore”
 - DRIHM model chain: hard for single individuals to be an expert on all components
 - Link with Civil Protection (time pressure), policy formation (lobbying pressure), Citizen scientists

- Key challenges
 - Understanding e-Infrastructures: scale, complexity
 - Shaping future e-Infrastructures:
 - Expectations from users (interfaces, integration,...)
 - Characteristics from applications
 - Limitations:
 - Budgetary constraints (capital and operational)
 - Capability constraints (hardware features)
 - Scalability constraints (cores, memory, bandwidth)
 - Metrics and incentives need to be adjusted (open data?)
 - Patent vs. well-curated open data?
- Opportunities
 - DRIHM shows: We can do more and better
 - Others to follow: more visible and understandable → more impact

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

- **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 optimized IT infrastructures
 - Early access to latest IT infrastructure (hard- and software) developments and specification of future requirements
 - Access to IT competence network and expertise at Computer Science and Mathematics departments
- **Partner contribution**
 - Embedding IT experts in user groups
 - Joint research projects (including funding)
 - Scientific partnership – joint publications
- **LRZ benefits**
 - Understanding the (current and future) needs and requirements of the respective scientific domain
 - Developing future services for all user groups

Goals for LRZ:

- Thematic focusing – **Environmental Computing**
- Strengthening science through innovative, high performance IT technologies and modern IT infrastructures and IT services
- Interdisciplinary integration (technical and personnel) of scientists and (international) research groups
- Novel requirements and research results at the interface of scientific computing and computer-based sciences
- Increased prospects for attracting research funding through established IT expertise as contribution to application projects
- Outreach and exploitation

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)

Trends in Computation, Communication and Storage: Consequences for Data-Intensive Science

Dieter Kranzlmüller
kranzlmue@lrz.de

