# Energy Efficient Environmental Computing on SuperMUC

Dieter Kranzlmüller

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

## Flash Flood Genoa, Italy, 2011

http://www.drihm.eu/images/video/DRIHM_final.mp4

## Flash Floods

- Form swiftly due to (extremely) high rainfall rates
- Little or no prior warning
- Devastating consequences (casualties, economic losses, ...)



## UNISDR – The United Nations Office for Disaster Risk Reduction

https://www.unisdr.org/

## GAR – Global Assessment Report on Disaster Risk Reduktion 2015

Global Assessment Report on Disaster Risk Reduction 2015

Making development sustainable: The future of disaster risk management

Home | Pocket GAR | **GAR 2015 Main Report** | Documents | Data | Download | Press

Foreword | At a glance | Preface | Introduction | Part I | Part II | Part III | Backmatter

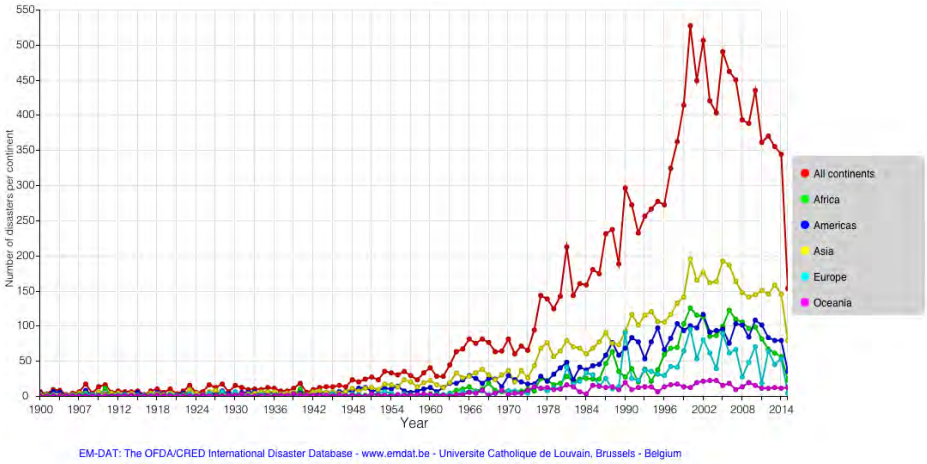
**Most disasters that could happen have not happened yet.**

Economic losses from disasters such as earthquakes, tsunamis, cyclones and flooding are now reaching an average of US$250 billion to US$300 billion each year. **Future losses** (expected annual losses) are now estimated at US$314 billion in the built environment alone. **This is the amount that countries should set aside each year to cover future disaster losses.** ( → Chapter 3 )

http://www.preventionweb.net/english/hyogo/gar/2015/en/home/GAR_2015/GAR_2015_6.html

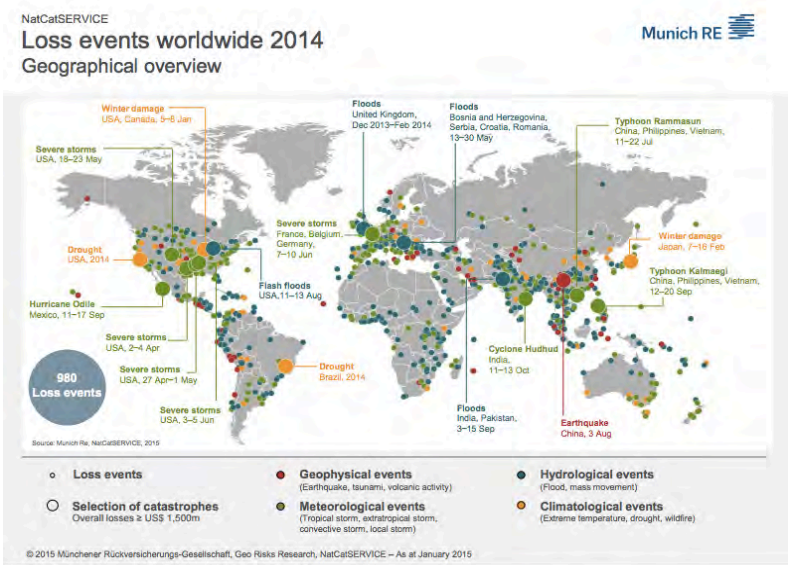


## Number of Disasters per Region




http://www.emdat.be/disaster_trends/index.html

## Munich Re – Loss Events Worldwide 2014



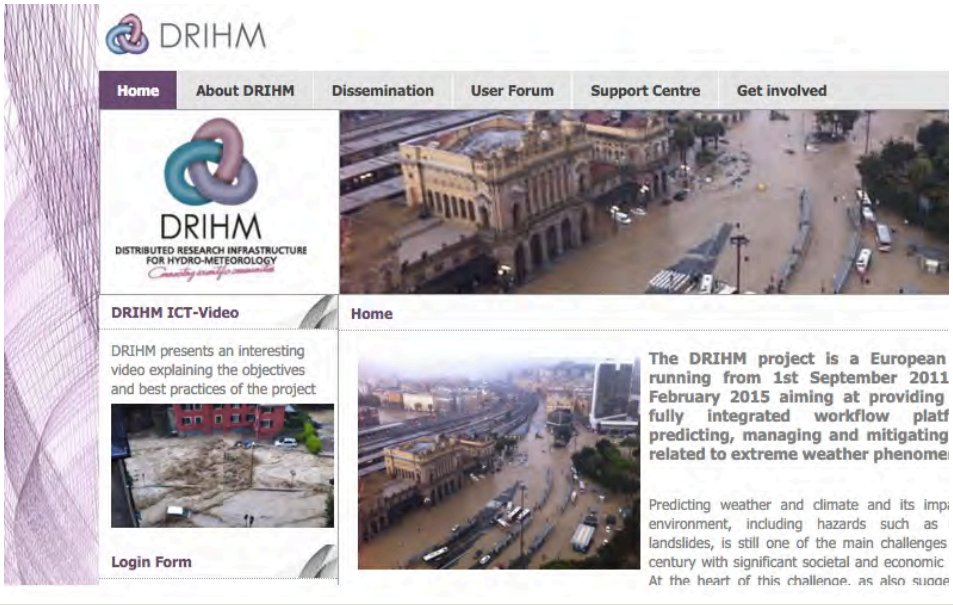

http://www.preventionweb.net/files/41773_munichreworldmapnaturalcatastrophes.pdf



## Flash Floods

- Form swiftly due to (extremely) high rainfall rates
- Little or no prior warning
- Devastating consequences (casualties, economic losses, ...)
- Monitoring and forecasting of floods:
  - European Flood Awareness System (EFAS)
  - Global Flood Detection System (GFDS)
  - Global Flood Awareness System (GloFAS)
- Problem: spatial resolution 50-100 km
  ➔ Flash floods remain undetected

## The EU Project Series DRIHM*






## Possible Solution – Environmental Computing

- Combine meteorology, hydrology, hydraulics through computer science
- Increase spatial and temporal resolution (data quality)
  - Regional Climate Models (RCM)
- Compute ensembles of forecasts to cover all potential outcomes
- Start and finish computation in time to provide lead time for evacuation measures

➔ **Simulate ensembles of forecasts with high-resolution on high-performance computing (HPC) infrastructures ~~on demand when triggered by increased rainfall rates~~**

Not in this talk

## Leibniz Supercomputing Centre
### of the Bavarian Academy of Sciences and Humanities

With approx. 230 employees
for more than 100.000 students and
for more than 30.000 employees
including 8.500 scientists

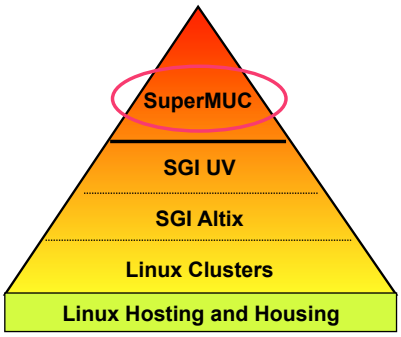
- European Supercomputing Centre
- National Supercomputing Centre
- Regional Computer Centre for all Bavarian Universities
- Computer Centre for all Munich Universities

Photo: Ernst Graf

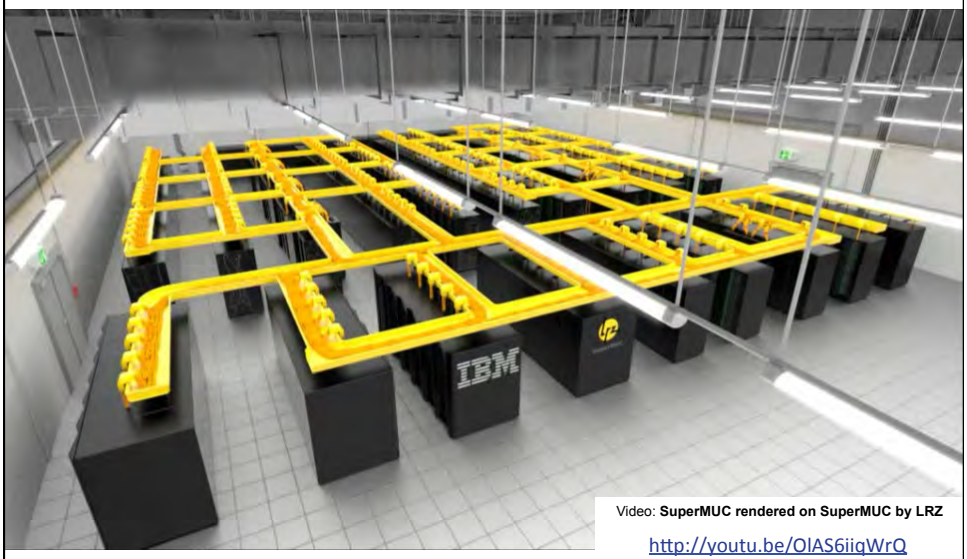


## Leibniz Supercomputing Centre
### of the Bavarian Academy of Sciences and Humanities

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



## SuperMUC @ LRZ




Video: **SuperMUC rendered on SuperMUC by LRZ**

http://youtu.be/OlAS6iiqWrQ

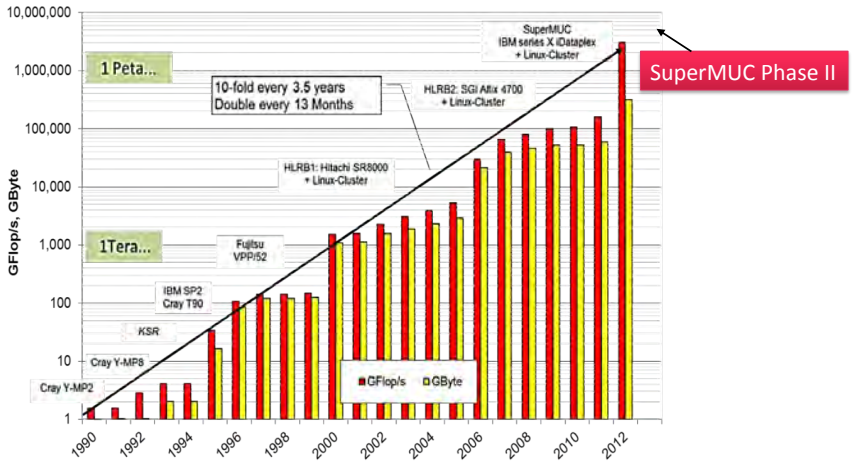


## Top 500 Supercomputer List (June 2012)

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

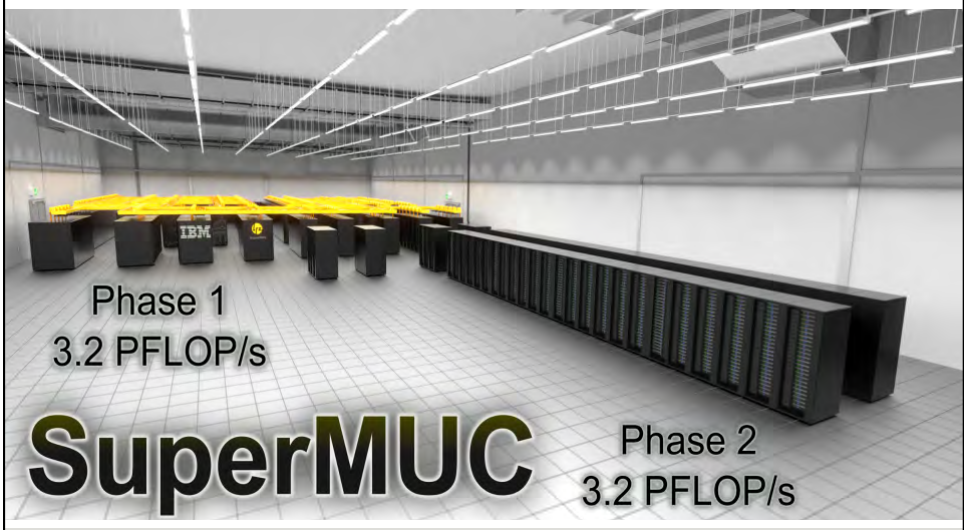
www.top500.org

## LRZ Supercomputers






## SuperMUC Phase 1 + 2



## 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)

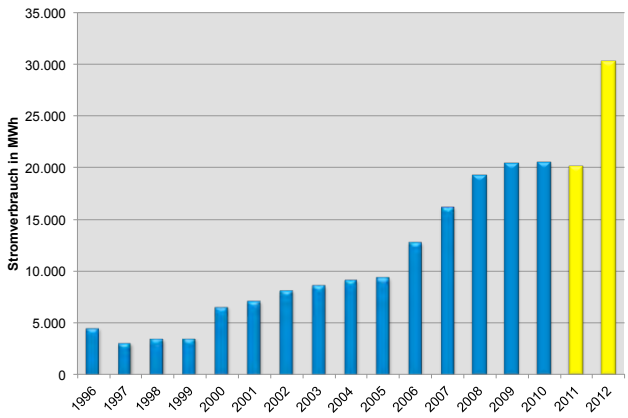
Phase 2 (Lenovo NeXtScale WCT):

- 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)

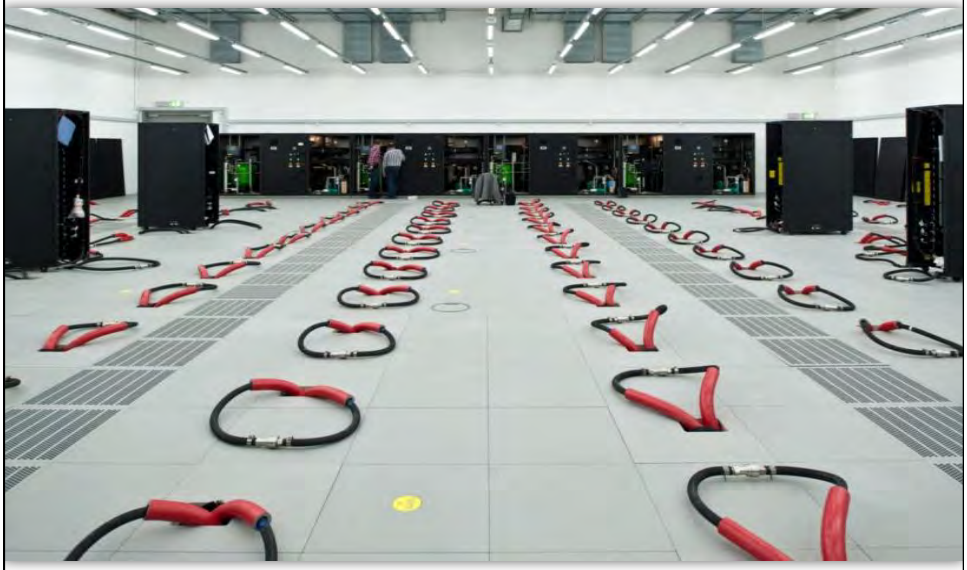
Common GPFS file systems with 10 PB and 5 PB usable storage size respectively
Common programming environment
Direct warm-water cooled system technology



## Power Consumption at LRZ



## Cooling SuperMUC





## SuperMUC Phase 2 @ LRZ

Photos: Torsten Bloth, Lenovo

High Energy Efficiency

- ✓ Usage of Intel Xeon E5 2697v3 processors
- ✓ Direct liquid cooling
  - 10% power advantage over air cooled system
  - 25% power advantage due to chiller-less cooling
- ✓ Energy-aware scheduling
  - 6% power advantage
  - ~40% power advantage
  - Total annual savings of ~2 Mio. € for SuperMUC Phase 1 and 2

## 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



## Results (Sustained TFlop/s on 128000 cores)

| 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 |

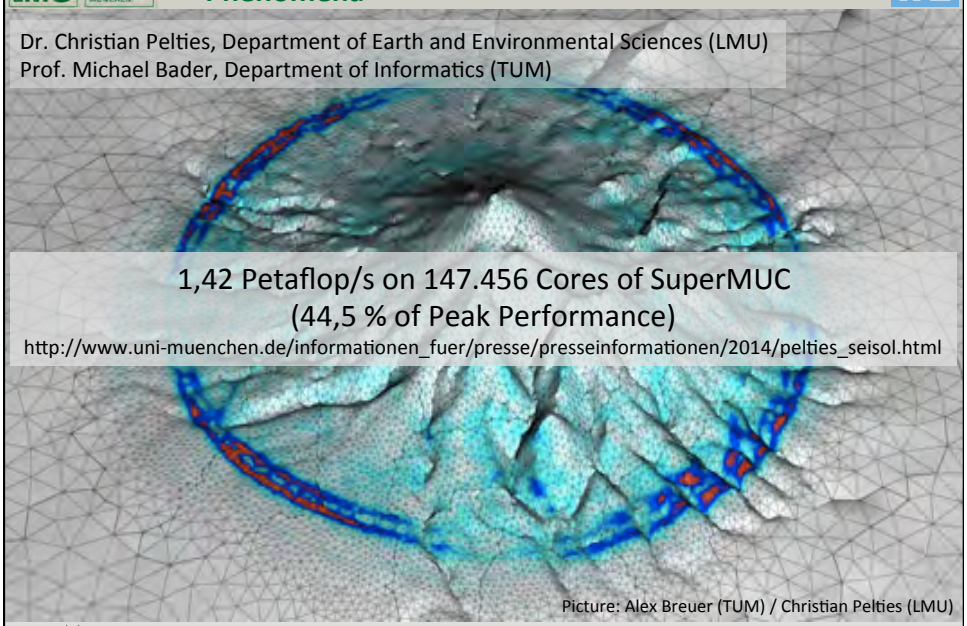
## Partnership Initiative Computational Sciences πCS

- **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 CS and Math departments
- **Partner contribution**
  - Embedding IT experts in user groups
  - Joint research projects (including funding)
  - Scientific partnership – equal footing – joint publications
- **LRZ benefits**
  - Understanding the (current and future) needs and requirements of the respective scientific domain
  - Developing future services for all user groups
  - Thematic focusing: **Environmental Computing**



## 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)

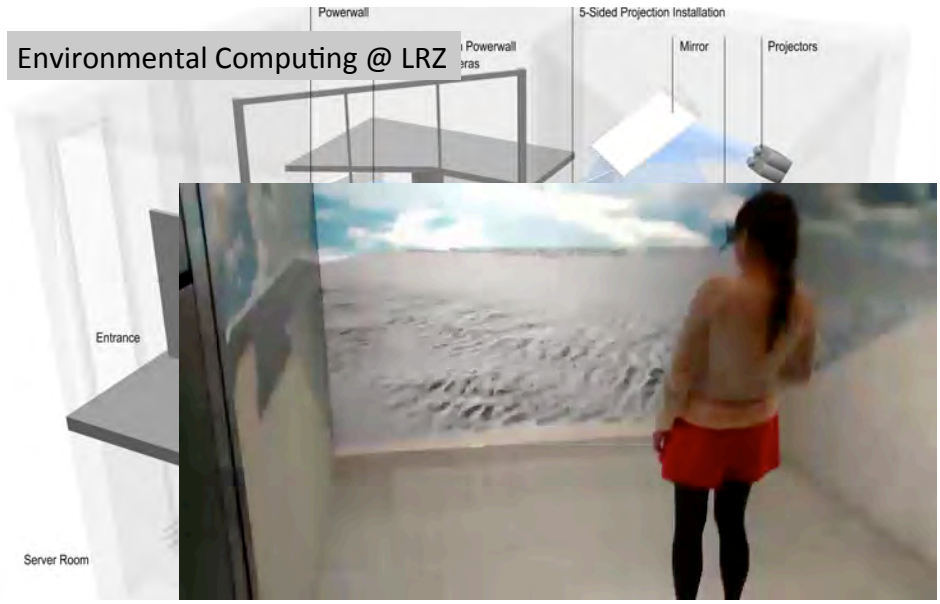
## Conclusions

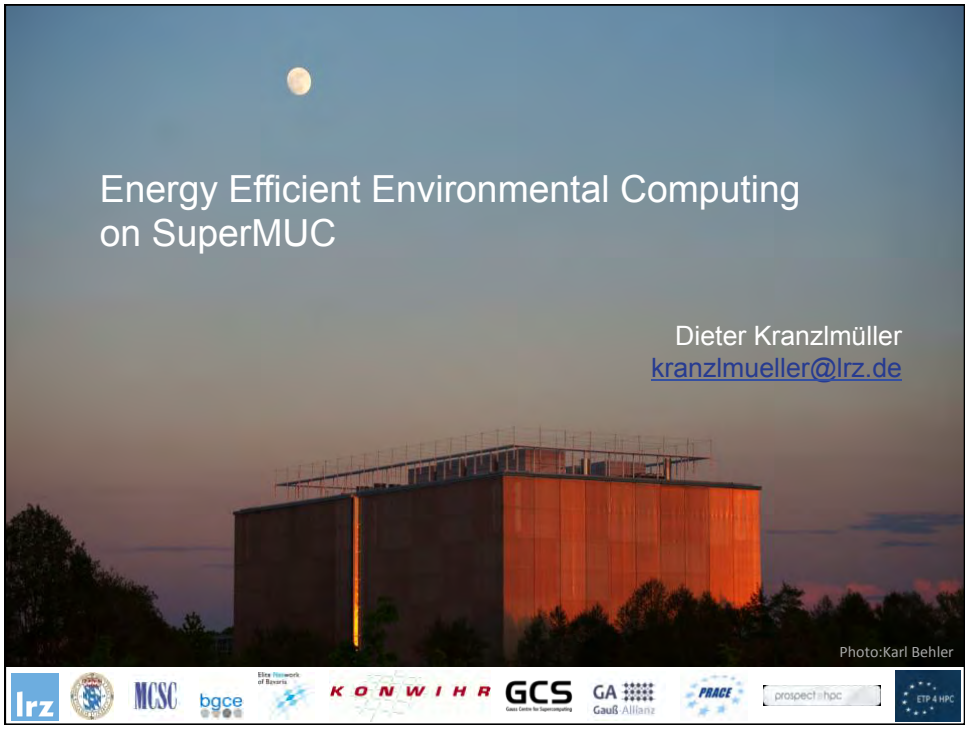
- **Environmental Computing** needs IT-Infrastructures (including HPC)
- **Energy Efficiency** is an important part to maximize scientific throughput
- Computational science needs to be an integral part of teaching domain scientists
  - Learn how to get access to HPC infrastructures
  - Learn how to program HPC infrastructures with increasing complexity, heterogeneity and scalability – efficiency, reliability, portabiliy
- The LRZ Partnership Initiative Computational Science (piCS) tries to improve user support

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



## Analyze Simulation Results




Energy Efficient Environmental Computing
on SuperMUC
Dieter Kranzlmüller
kranzlmueller@lrz.de
Photo:Karl Behler