Urgent Computing for Disaster Response Mitigation –
State-of-the-Art and Challenges

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

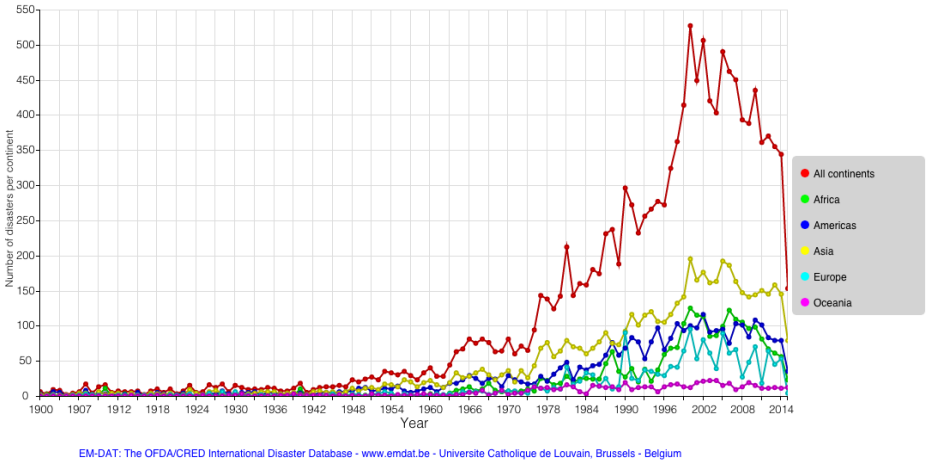
Foreword | At a glance | Preface | Introduction | Part I | Part II | Part III | Backmatter

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)

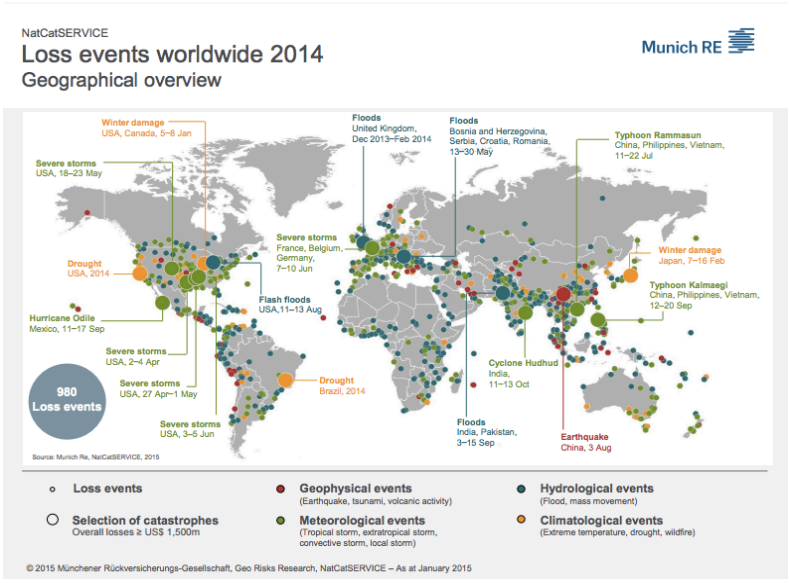
http://www.preventionweb.net/english/hyogo/gar/2015/en/home/GAR_2015/GAR_2015_6.html

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

Everybody is concerned

Philippine Congress commends former UNISDR chief for her advocacy on disaster risk reduction

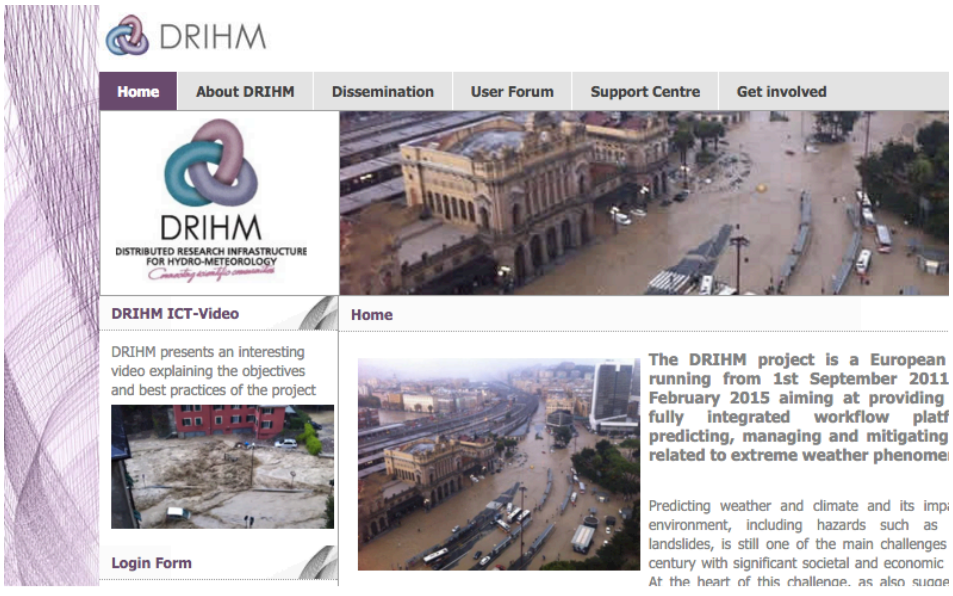
Ms. Margareta Wahlström (centre), former head of the UN Office for Disaster Risk Reduction (UNISDR), receives from Philippines Senate President Franklin Drilon the Senate Resolution authored by Senator Loren Legarda, UNISDR Global Champion for Resilience. Also pictured are Secretary Emmanuel de Guzman (first left), Vice Chairperson of the Climate Change Commission, and San Francisco, Camotes Island Councilor Al Arquillano Jr. (first right), UNISDR Asia Pacific Regional Champion for making cities resilient (Photo: Office of Senator Legarda)

https://www.unisdr.org/archive/47429

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*

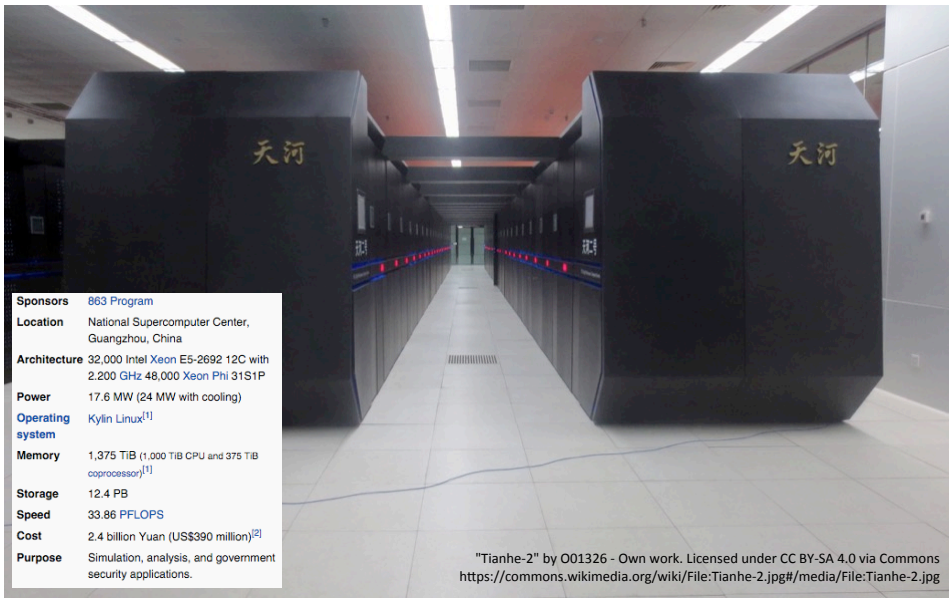


Possible Solution

- 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**

Tinahe-2 – Fastest Computer of the World since June 2013 (www.top500.org)



Sponsors	863 Program
Location	National Supercomputer Center, Guangzhou, China
Architecture	32,000 Intel Xeon E5-2692 12C with 2.200 GHz 48,000 Xeon Phi 31S1P
Power	17.6 MW (24 MW with cooling)
Operating system	Kylin Linux[1]
Memory	1,375 TiB (1,000 TiB CPU and 375 TiB coprocessor)[1]
Storage	12.4 PB
Speed	33.86 PFLOPS
Cost	2.4 billion Yuan (US$390 million)[2]
Purpose	Simulation, analysis, and government security applications.

Problems

- How to get allocations on high-performance computing infrastructures?
- How to start jobs on HPC on demand (instead of batch queue)?
- How to access multiple, widely distributed sources of data?
- How to execute with high performance to finish before a certain deadline?
- How to analyze the simulation results to initiate mitigation activities?

Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

With 156 employees + 38 extra staff 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

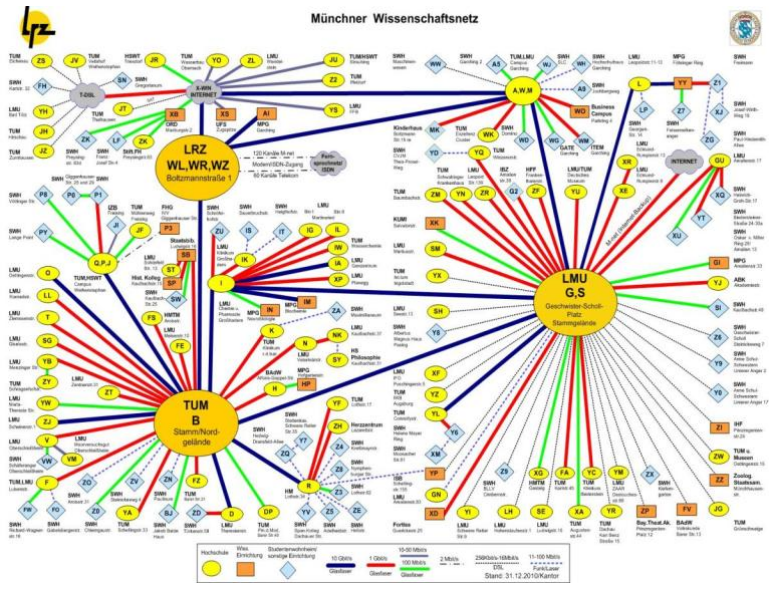
Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

■Computer Centre for all Munich Universities

IT Service Provider:
- Munich Scientific Network (MWN)
- Web servers
- e-Learning
- E-Mail
- Groupware
- Special equipment:
  - Virtual Reality Laboratory
  - Video Conference
  - Scanners for slides and large documents
  - Large scale plotters

IT Competence Centre:
- Hotline and support
- Consulting (security, networking, scientifc computing, …)
- Courses (text editing, image processing, UNIX, Linux, HPC, …)

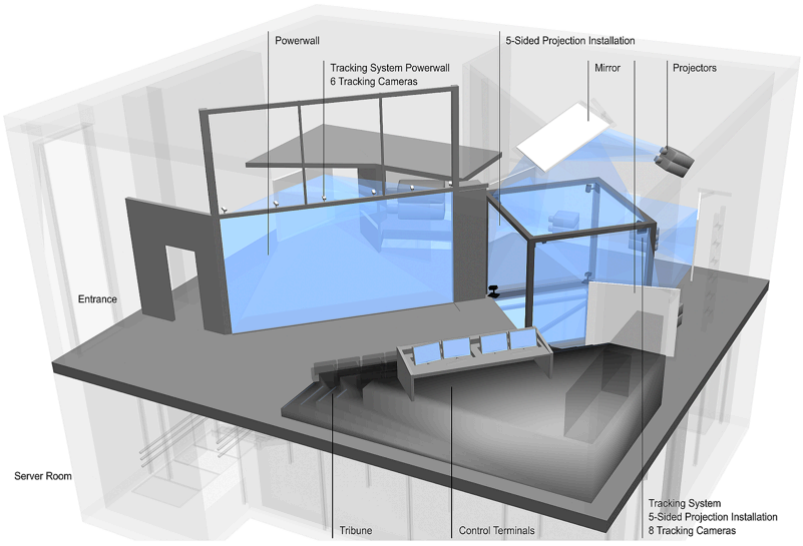
The Munich Scientific Network (MWN)



Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

- Regional Computer Centre for all Bavarian Universities
- Computer Centre for all Munich Universities

Virtual Reality & Visualization Centre (LRZ)



Examples from the V2C

Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

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

Gauss Centre for Supercomputing (GCS)

- 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)

PRACE Research Infrastructure Created

- Establishment of the legal framework
  - PRACE AISBL created with seat in Brussels in April
    (Association Internationale Sans But Lucratif)
  - 20 members representing 20 European countries
  - Inauguration in Barcelona on June 9

- Funding secured for 2010 - 2015
  - 400 Million € from France, Germany, Italy, Spain
    Provided as Tier-0 services on TCO basis
  - Funding decision for 100 Million € in The Netherlands
    expected soon
  - 70+ Million € from EC FP7 for preparatory and implementation
    Grants INFSO-RI-211528 and 261557
    Complemented by ~ 60 Million € from PRACE members

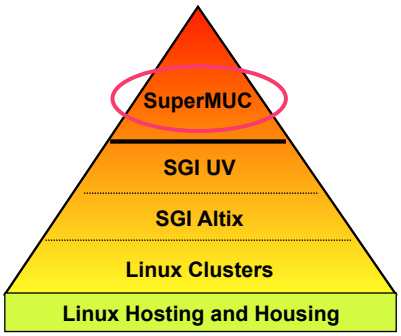
PRACE

PRACE Tier-0 Systems

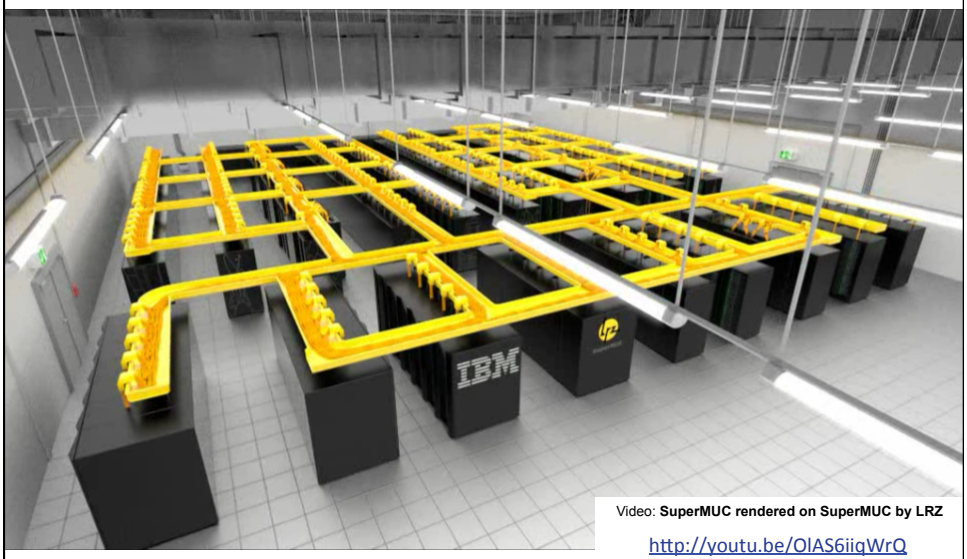
- **Curie @ GENCI**:
  Bull Cluster, 1.7 PFlop/s
- **FERMI @ CINECA**:
  IBM BG/Q, 2.1 PFlop/s
- **Hermit @ HLRS**:
  Cray XE6, 1 Pflop/s
- **JUQUEEN @ FZJ**:
  IBM Blue Gene/Q, 5.9 PFlop/s
- **MareNostrum @ BSC**:
  IBM System X iDataPlex, 1 PFlop/s
- **SuperMUC @ LRZ**:
  IBM System X iDataPlex, 3.2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

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**

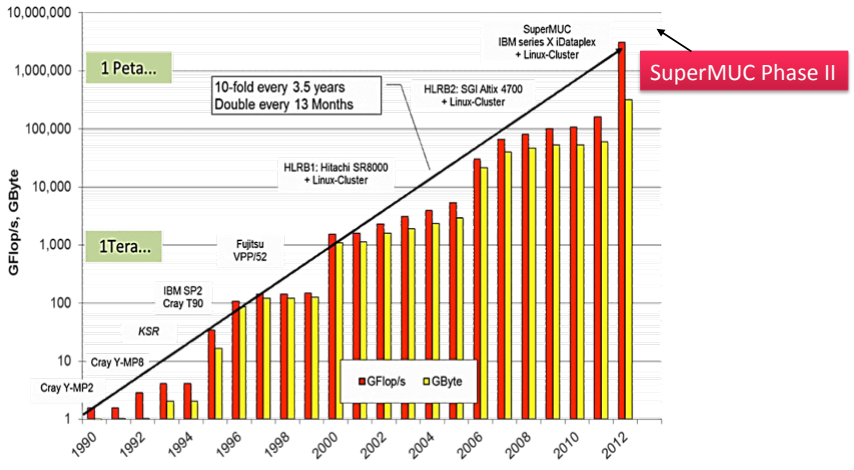
http://youtu.be/OlAS6iiqWrQ

Top 500 Supercomputer List (June 2012)

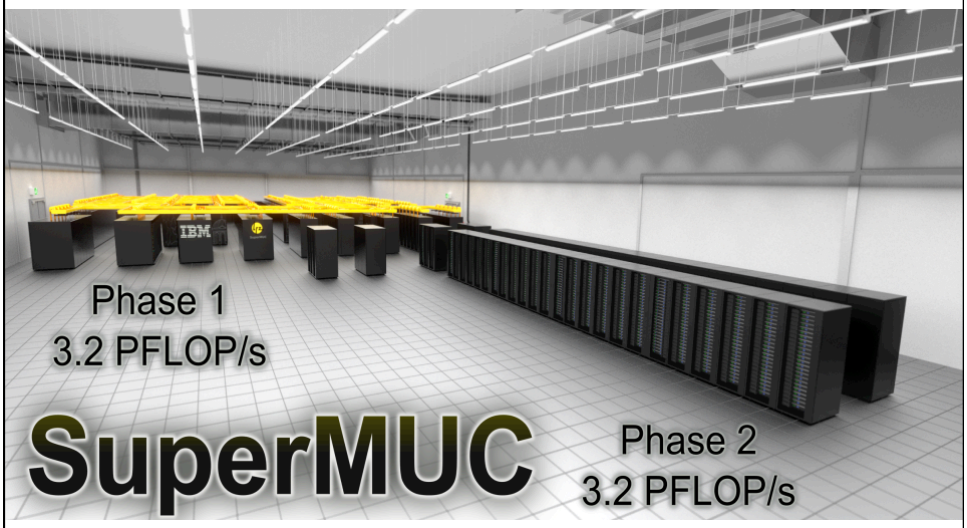
Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom / 2011 IBM	1572864	16324.75	20132.66	7890.0
2	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect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

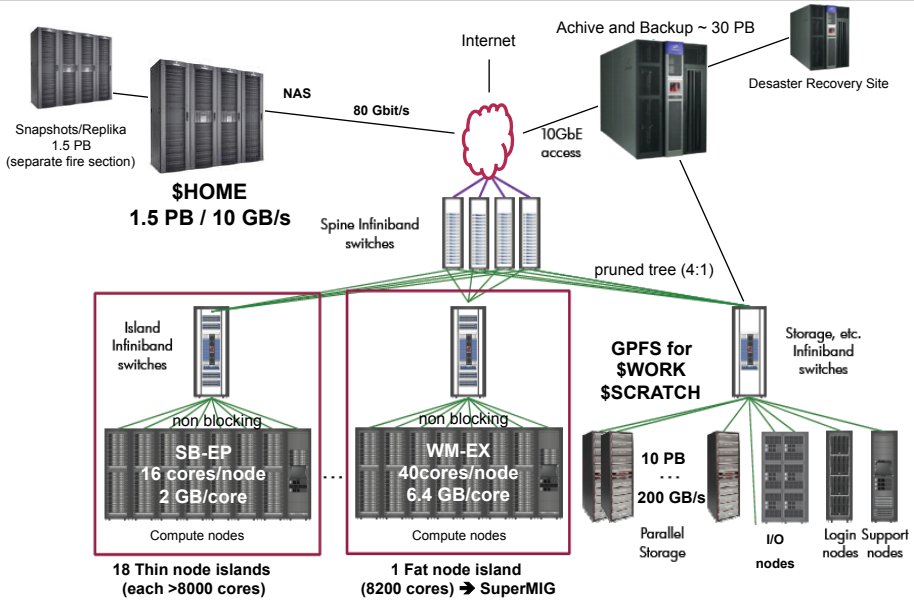
LRZ Supercomputers



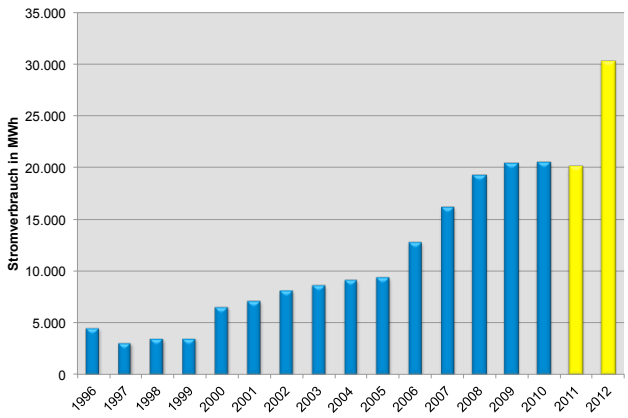
SuperMUC Phase 1 + 2



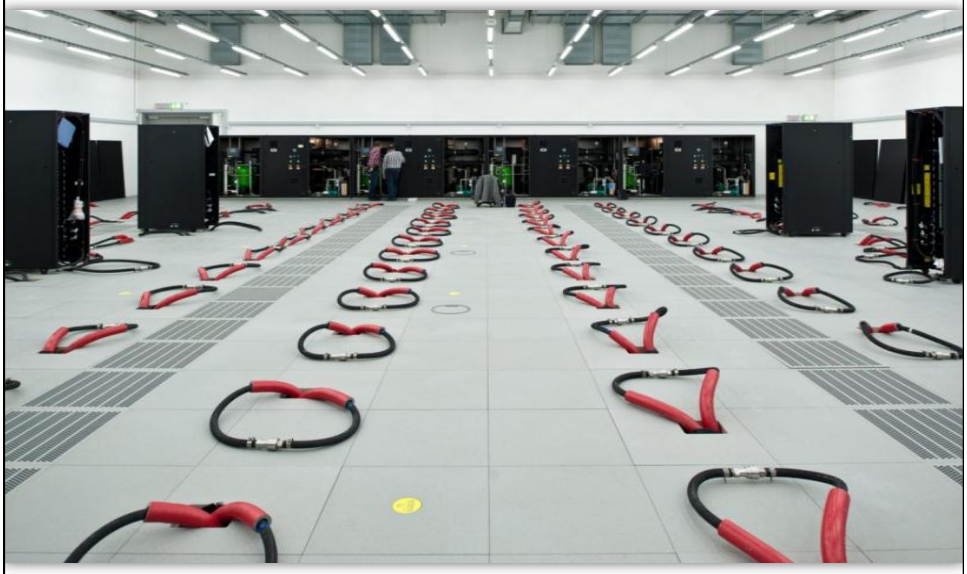
SuperMUC Architecture



Power Consumption at LRZ



Cooling SuperMUC



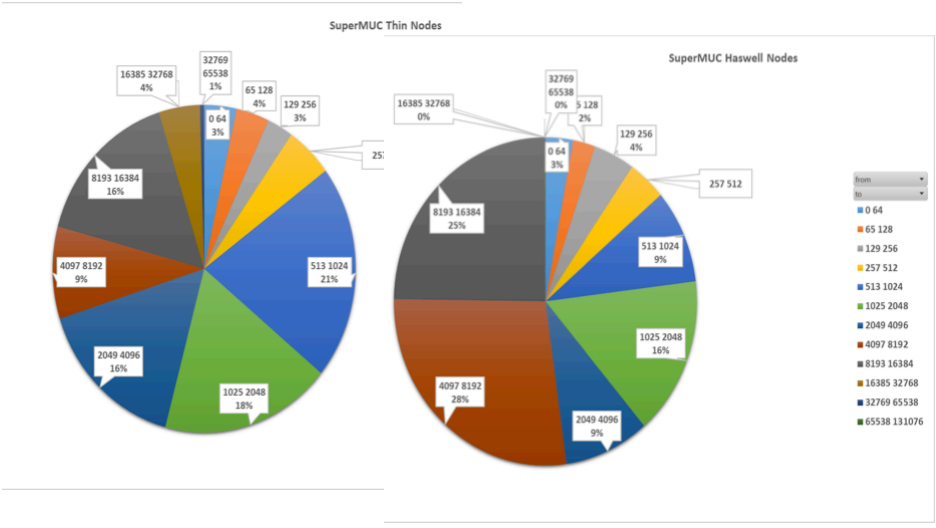
LRZ Application Mix

- 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

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

SuperMUC Jobsize 2015 (in Cores)



Problems

- How to get allocations on high-performance computing infrastructures?
- How to start jobs on HPC on demand (instead of batch queue)?
- How to access multiple, widely distributed sources of data?
- How to execute with high performance to finish before a certain deadline?
- How to analyze the simulation results to initiate mitigation activities?

How to get allocations on high-performance computing infrastructures?

- Apply for allocations
  - At your local center
  - At your regional center (Gauss Alliance)
  - At your national center (GCS)
  - At international centers (PRACE)
- Keep the performance pyramid in mind – Not every code needs to run on the fastest machine
- Discuss new policies – Preemptive scheduling in terms of crisis – Who takes the decision?
- Improve training and education – Provide information on infrastructure availability and access policies to (future) computational scientists

How to start jobs on HPC on demand (instead of batch queue)?

- Technical solution: Urgent computing
  - SPRUCE Framework
  - Novel UC Framework by Cerlane Leong
- Policy issues:
  - Who takes the decision?
  - Which jobs are stopped?
  - What happens to the allocations of interrupted jobs?
- Costs:
  - What costs are associated with Urgent Computing?

    S. H. Leong and D. Kranzlmüller, "A Case Study - Cost of Preemption for Urgent Computing on SuperMUC", in: High Performance Computing, Volume 9137 of Lecture Notes in Computer Science (LNCS), pages 422–433. Springer International Publishing, 2015.

How to access multiple, widely distributed sources of data?

- Network connectivity and data transfer are important – NRENs/ GEANT ➔ Take full e-Infrastructure landscape into consideration (Networks, Grids, Clouds, …)
- Bandwidth in critical situations possibly limited
- Access mechanisms: Federated identity management
- Access policies: Public data should be publicly accessible
- Different applications often need similar data, e.g. terrain height map ➔ **Environmental Computing**

How to execute with high performance to finish before a certain deadline?

- Potentially lots of performance needed – Depends on crisis situation and development over time – Flexibility needed in access to computational power (Grids, Clouds, …)
- Performance optimization needs corresponding software development

Results (Sustained TFlop/s on 128000 cores)

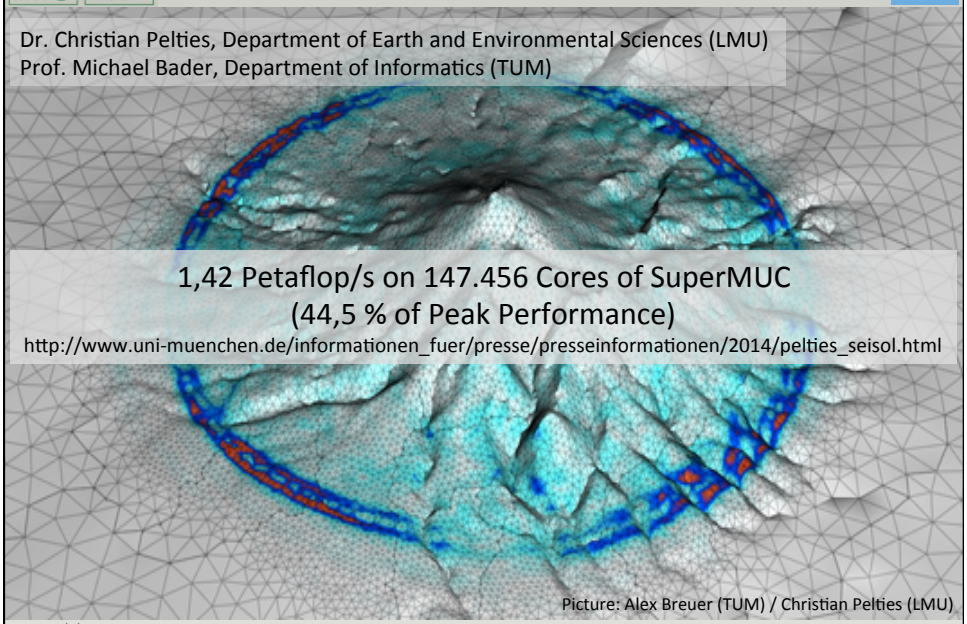
Name	MPI	# cores	Description	TFlop/s/island	TFlop/s max
Linpack	IBM	128000	TOP500	161	2560
Vertex	IBM	128000	Plasma Physics	15	245
GROMACS	IBM, Intel	64000	Molecular Modelling	40	110
Seissol	IBM	64000	Geophysics	31	95
waLBerla	IBM	128000	Lattice Boltzmann	5.6	90
LAMMPS	IBM	128000	Molecular Modelling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

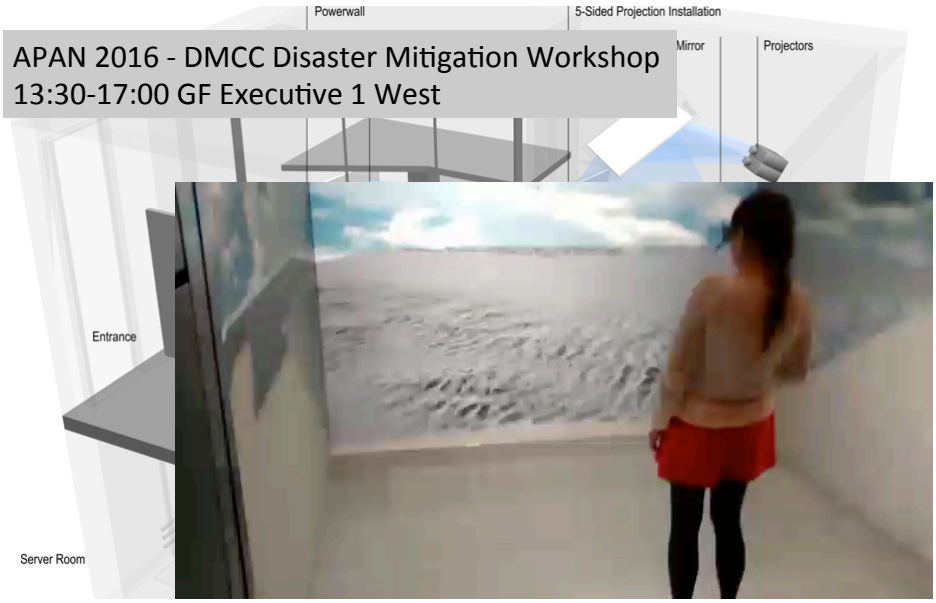


1,42 Petaflop/s on 147.456 Cores of SuperMUC
(44,5 % of Peak Performance)
http://www.uni-muenchen.de/informationen_fuer/presse/presseinformationen/2014/pelties_seisol.html

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

How to analyze the simulation results to initiate mitigation activities?

APAN 2016 - DMCC Disaster Mitigation Workshop
13:30-17:00 GF Executive 1 West



Conclusions

- HPC centers need to establish policies for Urgent Computing
  - Interrupt jobs with lower priority to provide cycles for ad hoc situations
- Users need to possibility to reliably execute (and optimize) their codes on the full size machines with more than 100.000 cores
- 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 – portability, efficiency, reliability
- The LRZ Partnership Initiative Computational Science (piCS) tries to improve user support

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


Urgent Computing for Disaster Response
Mitigation – State-of-the-Art and Challenges
Dieter Kranzlmüller
kranzlmueller@lrz.de