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)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>Computer/Year Vendor</th>
<th>Cores</th>
<th>Rmax</th>
<th>Rpeak</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DOE/NASA/LLNL United States</td>
<td>Sequoia - BlueGene/Q, Power BCC 19C</td>
<td>1.39 Gflop, Current 2011 IBM</td>
<td>1572904</td>
<td>10324.75</td>
<td>20132.05</td>
</tr>
<tr>
<td>2</td>
<td>JPSR Advanced Institute for Computational Science (AICS) Japan</td>
<td>K Computer: SPHINX 2.0/40, Tera Intelligenz / 2011 Fujitsu</td>
<td>760624</td>
<td>10251.00</td>
<td>12080.38</td>
<td>12659.0</td>
</tr>
<tr>
<td>3</td>
<td>DOE/Argonne National Laboratory United States</td>
<td>BlueGene/Q, Power BCC 19C</td>
<td>1.39 Gflop, Current 2012 IBM</td>
<td>785352</td>
<td>8192.38</td>
<td>10000.38</td>
</tr>
<tr>
<td>4</td>
<td>LRZ Rechenzentrum Germany</td>
<td>SuperMUC -拍View DK10,000,000 cores, in 2).70GHz, Seflom TFLOPS / 2012 IBM</td>
<td>147456</td>
<td>2560.76</td>
<td>3105.05</td>
<td>3422.7</td>
</tr>
<tr>
<td>5</td>
<td>National Supercomputing Center in Tokyo, China</td>
<td>Tsuru MA - Intel XMPi, Xeon, 5.3GHz</td>
<td>2.71 GHz, XEON 2015</td>
<td>185160</td>
<td>2560.76</td>
<td>4701.00</td>
</tr>
<tr>
<td>6</td>
<td>DOE/Argonne National Laboratory United States</td>
<td>Jaguar - Cray X3K, Opteron XT4 19C</td>
<td>2.00GHz, Cray X3K, Opteron XT4 19C</td>
<td>289092</td>
<td>1941.00</td>
<td>2627.01</td>
</tr>
<tr>
<td>7</td>
<td>CINEA Italy</td>
<td>Fermi - BlueGene/Q, Power BCC 19C</td>
<td>1.39 Gflop, Current 2012 IBM</td>
<td>164546</td>
<td>1725.49</td>
<td>2097.15</td>
</tr>
<tr>
<td>8</td>
<td>Peking University Shanghai, China</td>
<td>Aurtex 10 - Intel XMPi, Xeon, 5.3GHz</td>
<td>2.71 GHz, XEON 2015</td>
<td>185160</td>
<td>2560.76</td>
<td>4701.00</td>
</tr>
<tr>
<td>9</td>
<td>CERN/GENO Switzerland</td>
<td>GigaP/SuperMUC</td>
<td>130172</td>
<td>1380.39</td>
<td>1677.72</td>
<td>657.5</td>
</tr>
<tr>
<td>10</td>
<td>National Supercomputing Centre in Beijing, China</td>
<td>Nebula - Power BCC 19C</td>
<td>1.39 Gflop, Current 2012 IBM</td>
<td>129840</td>
<td>1271.06</td>
<td>2504.93</td>
</tr>
</tbody>
</table>

---

### 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
- ...

---

www.top500.org
SOFIA - Stratospheric Observatory For Infrared Astronomy

**Plattform:** Boeing 747 SP
- Gewicht: 300.000 kg
- Spannweite: 60m
- Max. Entfernung: 15.000km

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

**Phylogenetic Tree Computation**

Sequencing → Alignment → Phylogenetic Tree

Alexandros Stamatakis, H-ITS
# of possible trees for 150 species

- \( 4226517247809112252219618802377042809718932383449 \)
- \( 8822942857479880831434032178759024536798491951168 \)
- \( 3076494692867414802738570221298292428457687814873 \)
- \( 4552121861861600804474608426626044448936698500560 \)
- \( 2468116186441264227425440726676614927906540649360 \)
- \( 2976397461917469326750931190889241406694054603576 \)
- \( 66015625 \)

- \( n \approx 4.22 \times 10^{301} \)

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
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

https://natureecoevo.nature.com/content/521/7497/beh/behresearchpaperposts/1606002-micah-dunthorn-perspective-of-tropical-rainforests/
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

DOI: 10.1038/s41559-017-0091


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
### Increasing numbers

<table>
<thead>
<tr>
<th>Date</th>
<th>System</th>
<th>Flop/s</th>
<th>Cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>HLRB-I</td>
<td>2 Tflop/s</td>
<td>1512</td>
</tr>
<tr>
<td>2006</td>
<td>HLRB-II</td>
<td>62 Tflop/s</td>
<td>9728</td>
</tr>
<tr>
<td>2012</td>
<td>SuperMUC</td>
<td>3200 Tflop/s</td>
<td>155656</td>
</tr>
<tr>
<td>2015</td>
<td>SuperMUC Phase II</td>
<td>3.2 + 3.6 Pflop/s</td>
<td>229960</td>
</tr>
</tbody>
</table>

D. Kranzlmüller

---

### LRZ Extreme Scaling Workshop Series

- **Results:**
  
<table>
<thead>
<tr>
<th>Name</th>
<th>MPI</th>
<th># cores</th>
<th>Description</th>
<th>TFlop/s/island</th>
<th>TFlop/s max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linpack</td>
<td>IBM</td>
<td>128000</td>
<td>TOP500</td>
<td>161</td>
<td>2560</td>
</tr>
<tr>
<td>Vertex</td>
<td>IBM</td>
<td>128000</td>
<td>Plasma Physics</td>
<td>15</td>
<td>245</td>
</tr>
<tr>
<td>GROMACS</td>
<td>IBM,</td>
<td>64000</td>
<td>Molecular Modelling</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td>Seissol</td>
<td>IBM</td>
<td>64000</td>
<td>Geophysics</td>
<td>31</td>
<td>95</td>
</tr>
<tr>
<td>waLBerla</td>
<td>IBM</td>
<td>128000</td>
<td>Lattice Boltzmann</td>
<td>5.6</td>
<td>90</td>
</tr>
<tr>
<td>LAMMPS</td>
<td>IBM</td>
<td>128000</td>
<td>Molecular Modelling</td>
<td>5.6</td>
<td>90</td>
</tr>
<tr>
<td>APES</td>
<td>IBM</td>
<td>64000</td>
<td>CFD</td>
<td>6</td>
<td>47</td>
</tr>
<tr>
<td>BQCD</td>
<td>Intel</td>
<td>128000</td>
<td>Quantum Physics</td>
<td>10</td>
<td>27</td>
</tr>
</tbody>
</table>

- **Sustained TFlop/s on 64000/128000 cores**
SuperMUC Jobsizes 2015 (in Cores)

Statistics from www.top500.org
SuperMUC System @ LRZ

Phase 1 (IBM System x iDataPlex):
- 3.2 PFlops peak performance
- 9126 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

- **I/O nodes**
- **NAS** 80 Gbit/s
- 18 Thin node islands (each >8000 cores)
- 1 Fat node island (8200 cores) => SuperMIG
- Spine Infiniband switches
- $HOME 1.5 PB / 10 GB/s
- Pruned tree (4:1)
- SB - EP
- WM - EX
- GPFS for $WORK $SCRATCH
- 10 PB...200 GB/s
- Parallel Storage
- I/O nodes
- Login Support nodes
- Internet
- Active and Backup ~ 30 PB
- Disaster Recovery Site
- Snaps/(Replika 1.5 PB (separate fire section))

Power Consumption at LRZ

- **Stromverbrauch in MWh**
- **Power Consumption**
  - 0
  - 5,000
  - 10,000
  - 15,000
  - 20,000
  - 25,000
  - 30,000
  - 35,000
- **Years**
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
LRZ as IT Competence Centre: Operating Cutting-edge IT Infrastructure

- High Speed Networking: Munich Scientific Network
- High Performance Computing: SuperMUC, LinuxCluster
- Big Data: Bavarian State Library Digital Archive
- Virtual Reality & Visualisation: V2C (CAVE, Powerwall)

LRZ as IT Competence Centre: Providing Comprehensive IT Services for Science

- Service Desk
- Application Software Support
- Personal Consulting
- Tailored Solutions
- Backup & Archive
- Authentication & Authorization
- Trainings & Workshops
- IT Security
- Big Data: Bavarian State Library Digital Archive
- High Speed Networking: Munich Scientific Network
- Performance Computing: SuperMUC, LinuxCluster
- Virtual Reality & Visualisation: V2C (CAVE, Powerwall)
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 (hardware & 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

---

**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
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, ...)

Internet

Slide courtesy Herbert Huber
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
kranzmueller@lrz.de