In Need of Partnerships
An Essay about the Collaboration between Computational Sciences and IT Services

Dieter Kranzlmüller

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

Leibniz Rechenzentrum
Germany

SuperMUC - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR / 2012 IBM

147456  2897.00  3185.05  3422.7

www.top500.org, June 2012

Video: SuperMUC rendered on SuperMUC by LRZ
http://youtu.be/OlAS6iigWrQ
LRZ Supercomputers

next to come (2014): SuperMUC Phase II 6.4 PFlop/s
SuperMUC and its predecessors
SuperMUC and its predecessors
SuperMUC and its predecessors
LRZ Building Extension

Picture: Horst-Dieter Steinhöfer

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

Picture: Ernst A. Graf
## 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>2014</td>
<td>SuperMUC Phase II</td>
<td>3.2 + 3.2 Pflop/s</td>
<td>229960</td>
</tr>
</tbody>
</table>
SuperMUC Architecture

Snapshots/Replika
1.5 PB
(separate fire section)

18 Thin node islands
(each >8000 cores)

1 Fat node island
(8200 cores) ➔ SuperMIG

I/O
nodes

NAS
80 Gbit/s

Spine Infiniband switches

$HOME
1.5 PB / 10 GB/s

Internet

10GbE
access

GPFS for
$WORK
$SCRATCH

Parallel Storage

I/O

nodes

Login Support
nodes

Desaster Recovery Site

1.5 PB
(separate fire section)

Snapshots/Replika

Achive and Backup ~ 30 PB

16 cores/node
2 GB/core

SB-EP

non blocking

Compute nodes

10 PB
200 GB/s

WM-EX
40 cores/node
6.4 GB/core

non blocking

Compute nodes

pruned tree (4:1)

D. Kranzlmüller

In Need of Partnerships
Questions

- How to use today’s supercomputers?
- How to cope with the complexity?
- How to use these machines efficiently?
- How to scale applications?
- How to do I/O?
- How about resilience?
- ...
July 2013:

First SuperMUC Extreme Scale Workshop

Participants:
- 15 international projects

Prerequisites:
- Successful run on 4 islands (32768 cores)

Participating Groups (Software packages):
- LAMMPS, VERTEX, GADGET, WaLBerla, BQCD, Gromacs, APES, SeisSol, CIAO

Successful results (> 64000 Cores):
- Invited to participate in PARCO Conference (Sept. 2013) including a publication of their approach
Regular SuperMUC operation
- 4 Islands maximum
- Batch scheduling system

Entire SuperMUC reserved 2,5 days for challenge:
- 0,5 Days for testing
- 2 Days for executing
- 16 (of 19) Islands available

Consumed computing time for all groups:
- 1 hour of runtime = 130,000 CPU hours
- 1 year in total
## Results (Sustained TFlop/s on 128000 cores)

<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, Intel</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>
5 Software packages were running on max 16 islands:
- LAMMPS
- VERTEX
- GADGET
- WaLBerla
- BQCD

VERTEX reached 245 TFlop/s on 16 islands (A. Marek)
Lessons learned – Technical Perspective

- Hybrid (MPI+OpenMP) on SuperMUC still slower than pure MPI (e.g. GROMACS), but applications scale to larger core counts (e.g. VERTEX).

- Core pinning needs a lot of experience by the programmer.

- Parallel IO still remains a challenge for many applications, both with regard to stability and speed.

- Several stability issues with GPFS were observed for very large jobs due to writing thousands of files in a single directory. This will be improved in the upcoming versions of the application codes.
Next Steps

- LRZ Extreme Scale Benchmark Suite (LESS) will be available in two versions: public and internal
- All teams will have the opportunity to run performance benchmarks after upcoming SuperMUC maintenances
- Second LRZ Extreme Scaling Workshop ➔ 2-5 June 2014
- Initiation of the LRZ Partnership Initiative πCS
Slices through the three-dimensional gas density (top panels) and vorticity (bottom panels) for fully developed, highly compressible, supersonic turbulence, generated by solenoidal driving (left-hand column) and compressive driving (right-hand column), and a grid resolution of $4096^3$ cells.

Federrath C MNRAS 2013;mnras.stt1644
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)
Effective usage of High Performance Computing infrastructures requires substantial amount of knowledge and expertise.

Collaboration between Computational Sciences and IT Services leads to new research results.

A partnership between CS and IT is preferred over a provider-user relationship.

Incentives are needed to ensure fruitful collaboration.

LRZ has established the Partnership Initiative Computational Sciences $\pi$CS.
In Need of Partnerships
An Essay about the Collaboration between Computational Sciences and IT Services

Dieter Kranzlmüller
kranzlmueller@lrz.de