aXXLs:
Application Extreme-scaling Experience of Leading Supercomputing Centres

2015-07-16 | Dirk Brömmel, Wolfgang Frings & Brian J. N. Wylie
Jülich Supercomputing Centre

b.wylie @ fz-juelich.de

http://www.fz-juelich.de/ias/jsc/aXXLs
Extreme scaling
# Diversity of current leadership HPC computer systems

<table>
<thead>
<tr>
<th>Cores</th>
<th>Name</th>
<th>System</th>
<th>Processor</th>
<th>(cores)</th>
<th>Accelerator</th>
<th>(cores)</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 120 000</td>
<td>Tianhe-2</td>
<td>NUDT IVB-FEP</td>
<td>Xeon(12C)</td>
<td>384 000</td>
<td>31S1P(57M)</td>
<td>2 736 000</td>
<td>NSCC-GZ, China</td>
</tr>
<tr>
<td>1 572 864</td>
<td>Sequoia</td>
<td>IBM Blue Gene/Q</td>
<td>PowerPC(16C)</td>
<td>1 572 864</td>
<td></td>
<td></td>
<td>LLNL, USA</td>
</tr>
<tr>
<td>786 432</td>
<td>Mira</td>
<td>IBM Blue Gene/Q</td>
<td>PowerPC(16C)</td>
<td>786 432</td>
<td></td>
<td></td>
<td>ANL, USA</td>
</tr>
<tr>
<td>705 024</td>
<td>Kei</td>
<td>K computer</td>
<td>SPARC64(8C)</td>
<td>705 024</td>
<td></td>
<td></td>
<td>RIKEN AICS, Japan</td>
</tr>
<tr>
<td>560 640</td>
<td>Titan</td>
<td>Cray XK7</td>
<td>Opteron(16C)</td>
<td>299 008</td>
<td>K20x(14S)</td>
<td>261 632</td>
<td>ORNL, USA</td>
</tr>
<tr>
<td>462 462</td>
<td>Stampede</td>
<td>Dell PowerEdge</td>
<td>Xeon(16C)</td>
<td>102 400</td>
<td>SE10P(61M)</td>
<td>390 400</td>
<td>TACC, USA</td>
</tr>
<tr>
<td>458 752</td>
<td>JUQUEEN</td>
<td>IBM Blue Gene/Q</td>
<td>PowerPC(16C)</td>
<td>458 752</td>
<td></td>
<td></td>
<td>JSC, Germany</td>
</tr>
<tr>
<td>452 400</td>
<td>Blue Waters</td>
<td>Cray XE6+XK7</td>
<td>Opteron(16C)</td>
<td>393 600</td>
<td>K20x(14S)</td>
<td>58 800</td>
<td>NCSA, USA</td>
</tr>
<tr>
<td>234 312</td>
<td>SuperMUC</td>
<td>IBM iDataPlex</td>
<td>Xeon(8-14C)</td>
<td>230 472</td>
<td>5110P(60M)</td>
<td>3 840</td>
<td>LRZ, Germany</td>
</tr>
<tr>
<td>138 368</td>
<td>Tera-100</td>
<td>Bull bullx S6010</td>
<td>Xeon(8C)</td>
<td>138 368</td>
<td></td>
<td></td>
<td>CEA-DAM, France</td>
</tr>
<tr>
<td>115 984</td>
<td>Piz Daint</td>
<td>Cray XC30</td>
<td>Xeon(8C)</td>
<td>42 176</td>
<td>K20x(14S)</td>
<td>73 808</td>
<td>CSCS, Switzerland</td>
</tr>
<tr>
<td>76 032</td>
<td>Tsubame2.5</td>
<td>NEC/HP SL390s</td>
<td>Xeon(6C)</td>
<td>16 896</td>
<td>K20x(14S)</td>
<td>59 136</td>
<td>TITech, Japan</td>
</tr>
<tr>
<td>35 712</td>
<td>Magnus</td>
<td>Cray XC40</td>
<td>Xeon(12C)</td>
<td>35 712</td>
<td></td>
<td></td>
<td>PSC, Australia</td>
</tr>
</tbody>
</table>
Background

Technology trends

- energy-efficient processors with many relatively weak cores
- organised in compute nodes with restricted shared memory
- combined with co-processors and accelerators with separate address space

Current generation of leadership supercomputers have many thousands of processors/cores
- and exascale computer systems are expected to have many more

Applications need to be extremely scalable and adaptable to effectively exploit such systems
- hybrid parallelisation combining message-passing, multi-threading, vectorisation
Motivation

Extreme-scaling applications is highly challenging

- correctness and performance issues only seen at large scale
  - *new issues encountered with each doubling of scale (or perhaps factor of ten)*

Application code teams need good support from supercomputing centres

- expert advice and training regarding most suitable algorithms, compilers, libraries, tools
  - *often specific to a particular code, but sometimes more generally applicable*
- access to large system configurations for testing and validation
  - *often only for short periods, but potentially disruptive*

Supercomputing centres have a variety of support approaches

- favouring capability versus capacity usage
- organising dedicated workshops and documenting successes
This workshop

Invited leadership supercomputing centres to share and discuss their extreme-scaling experience

- with real-world HPC applications (or proxies)
- with a variety of the largest scale homogeneous & heterogeneous computer systems
- accepted by presenters from Europe, America, Asia & Australia

addressing

- policies for using large (capability) compute node configurations
- porting and tuning support to demonstrate and exploit extreme-scaling
- characteristics of extreme-scaling applications and their execution configurations
- observed gaps between current and an 'ideal' solution
  - technology limitations, programming models, file I/O, debugging/analysis tools, ...
- application readiness in respect of envisaged exascale computer systems
Agenda

09:00-11:00 Session 1: Non-accelerated HPC systems
(break)
11:30-13:00 Session 2: MIC-accelerated HPC systems
(lunch break)
14:00-16:00 Session 3: GPU-accelerated HPC systems
(break)
16:30-18:00 Session 4: Open discussion and wrap-up
Session 1: Non-accelerated HPC systems

09:00-11:00

- Dirk Brömmel, Jülich Supercomputing Centre (JSC)
  - Extreme-scaling applications 24/7 on JUQUEEN Blue Gene/Q
- Miwako Tsuji, RIKEN Advanced Institute for Computational Science (AICS)
  - “K-scale” applications on the K computer and co-design effort for “post-K”
- Jean-Philippe Nominé (CEA)
  - Application Extreme-scaling Experience on Bullx Curie, Helios & Tera100
- Bronis de Supinski, Lawrence Livermore National Laboratory (LLNL)
  - Programming Future Large Scale Systems
Session 2: MIC-accelerated HPC systems

11:30-13:00

- Yunfei Du, National Supercomputing Center Guangzhou (NSCC-GZ)
  - Scaling applications from six application domains on Tianhe-2
- Carlos Rosales-Fernandez, Texas Advanced Computing Center (TACC)
  - Application Extreme-scaling Experience on Stampede
- Ferdinand Jamitzky, Leibniz Supercomputing Centre (LRZ)
  - SuperMUC Extreme Scale-out Phase2, lessons learned
Session 3: GPU-accelerated HPC systems

14:00-16:00

- Stefan Andersson, Cray Inc.
  - A Strategy for Developing a Performance Portable Highly Scalable Application
- William Sawyer, Swiss National Supercomputing Centre (CSCS)
  - Can large clusters of GPUs lead us to ExaScale?
- Akihiro Nomura, Tokyo Institute of Technology (TITech)
  - Performance measurement and monitoring in TSUBAME2.5 towards next generation supercomputers
- Neil Stringfellow, Pawsey Supercomputing Centre
  - From Tailored Systems to Scalable Machines, Multiple Architectures to Meet Competing Demands in a Data-Intensive World
Session 4: Open discussion

16:30-18:00

- Other experience
- Additional topics
- Follow-up
- Wrap-up
Workshop organisers

Dirk Brömmel
d.broemmel
@fz-juelich.de
Simulation Laboratory:
Plasma Physics

Wolfgang Frings
w.frings
@fz-juelich.de
Cross-sectional team:
Application Optimization

Brian Wylie
b.wylie
@fz-juelich.de
Cross-sectional team:
Performance Analysis