Jülich Supercomputing Centre

Introduction

E. Suarez
Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich
Outline

- JSC in a Nutshell
- Production systems
  - System use
  - How to get access
  - Training
- Architecture evolution
  - Dual approach
  - Cluster-Booster
  - Modular Supercomputing
- JSC vision
JSC in a Nutshell

Supercomputer operation for:
- Centre – FZJ
- Region – RWTH Aachen University
- Germany – Gauss Centre for Supercomputing
  John von Neumann Institute for Computing
- Europe – PRACE, EU projects

Application support
- Unique support & research environment at JSC
- Peer review support and coordination

Education and training

R&D work
- Methods and algorithms, computational science, performance analysis and tools
- Scientific Big Data Analytics with HPC
- Computer architectures, Co-Design
  **Exascale Labs** together with IBM, Intel, NVIDIA
Domain-specific User support
PRODUCTION SYSTEMS
JURECA Cluster – Hardware

- Dual-socket Intel Haswell (E5-2680 v3)
  - 12× cores/socket
  - 2.5 GHz
  - ≥ 128 GB main memory

- 1,884 compute nodes (45,216 cores)
  - 75 nodes: 2× K80 NVIDIA GPUs
  - 12 nodes: 2 × K40 NVIDIA GPUs
    - 512 GB main memory

- Peak performance: **2.2 Petaflop/s** (1.7 w/o GPUs)
- Mellanox InfiniBand EDR
- Connected to the GPFS file system on JUST
  - ~15 PByte online disk and
  - 100 PByte offline tape capacity
JUQUEEN - Hardware

- IBM Blue Gene/Q
- PowerPC® A2
  - 16× cores per node
  - 1.6 GHz
  - 16 Gbyte main memory
- 28,672 nodes
  - 458,752 cores
  - 28 racks
- Peak performance: 5,9 PFlop/s
- Connected to a GPFS:
- 5D network

June 2017: #7 in Europe
#21 worldwide
#94 in Green500
JUST: Storage server

- **8 PB**
  - **$WORK**
  - **$DATA**

- **$HOME**
  - 3 x 600 TB

- **$ARCH**
  - 2 x 600 TB

- 220 GB/sec

JUQUEEN | JURECA | JUROPA3 | JUDAC Jülich Data Access | TSM Server
JURECA Cluster – Software

- **Operating system**: Linux CentOS 7.X
- **Scheduler**: SLURM
- **Filesystem**: GPFS ($HOME, $WORK, $ARCH)
- **Compilers** (C/C++, Fortran, CUDA): Intel, GNU, PGI, CUDA
- **Debuggers**: TotalView, DDT, MUST
- **Programming**: Intel MPI, ParaStation MPI, OpenMP, CUDA
- **Performance analysis tools**: Score-P, Scalasca, Vampir, TAU, NVIDIA Visual Profiler, Darshan…
- **Libraries** (modules): MKL, SIONlib, HDF5, netcdf, PETSc …
- **Domain specific packages**: NAMD, QuantumExpresso,…

Support: sc@fz-juelich.de
On-line documentation: [http://www.fz-juelich.de/ias/jsc/jureca](http://www.fz-juelich.de/ias/jsc/jureca)
JUQUEEN – Software

- **Operating system**: Linux
- **Scheduler**: SLURM
- **Filesystem**: GPFS ($HOME, $WORK, $ARCH)
- **Compilers** (C/C++, Fortran): Intel, GNU, PGI
- **Debuggers**: TotalView, STAT, MUST
- **Programming**: MPICH2, OpenMP, CUDA
- **Performance analysis tools**: Score-P, Scalasca, Vampir, TAU, Darshan,…
- **Libraries** (modules): MKL, SIONlib, HDF5, netcdf, PETSc …

Support: sc@fz-juelich.de

On-line documentation: http://www.fz-juelich.de/ias/jsc/juqueen
SYSTEMS USE
System Usage

JURECA Cluster

Launch of JURECA, phase 1: 260 nodes: Jul 13, 2015
Launch of JURECA, phase 2: 1,884 nodes: Nov 2, 2015
Research fields – Current projects

General-Purpose Cluster

Massively Parallel system

JURECA
ca. 170 Projects

JUQUEEN
ca. 95 Projects

Granting periods
11/2016 – 10/2017
05/2017 – 04/2018

1 Earth & Environment
2 Biophysics
3 Particle Physics
4 Soft Matter
5 Condensed Matter
6 Plasma Physics
7 Chemistry
8 Fluid Dynamics
9 Materials Science
10 Computer Science
11 Astrophysics
How to get access – peer review process

- To JURECA via:
  - JARA-HPC: for FZJ + RWTH staff members only
    - Vergabegremium (VGG) and/or Kommission zur Vergabe von SC Ressourcen (VSR)
    - John von Neumann Institute for Computing (NIC)

- To JUQUEEN via:
  - JARA-HPC (for FZJ + RWTH staff only): VGG and/or VSR
  - Gauss Centre for Supercomputing (GCS)
    - Proposals evaluated by NIC
  - PRACE: European Research Infrastructure
    - Project Access: Biannual CfPs since June 2010
    - Call for preparatory access open, no closing dates
How to use the systems

- Introductory courses every 6 months
- Detailed documentation (slides) on JSC website:

Further training events

- Very varied training events:  jsc-events-join@fz-juelich.de

---

**Events at JSC**

If you would like to receive regular information on our events per e-mail, please send an e-mail to: jsc-events-join@fz-juelich.de.

### More events

<table>
<thead>
<tr>
<th>Date</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 Sep 2017 - 29 Sep 2017</td>
<td>International HPSC TerrSys Fall School 2017</td>
</tr>
<tr>
<td>09 Oct 2017 - 11 Oct 2017</td>
<td>Training course &quot;Einführung in Python&quot;</td>
</tr>
<tr>
<td>09 Oct 2017 - 10 Oct 2017</td>
<td>Training course &quot;Porting code from Matlab to Python&quot;</td>
</tr>
<tr>
<td>16 Oct 2017 - 17 Oct 2017</td>
<td>Training course &quot;Introduction to GPU programming using OpenACC&quot;</td>
</tr>
<tr>
<td>19 Oct 2017 - 20 Oct 2017</td>
<td>GAMM CSE Workshop 2017</td>
</tr>
<tr>
<td>06 Nov 2017 - 15 Nov 2017</td>
<td>Training course &quot;Programmierung in C&quot;</td>
</tr>
<tr>
<td>15 Nov 2017</td>
<td>Training course &quot;JusD - JusNet/Internet Security Day&quot;</td>
</tr>
<tr>
<td>20 Nov 2017</td>
<td>Training course &quot;Software Development in Science&quot;</td>
</tr>
<tr>
<td>21 Nov 2017 - 22 Nov 2017</td>
<td>Training course &quot;Vectorisation and portable programming using OpenCL&quot;</td>
</tr>
<tr>
<td>23 Nov 2017 - 24 Nov 2017</td>
<td>Training course &quot;Introduction to the Programming and usage of the supercomputer resources at Jülich&quot;</td>
</tr>
<tr>
<td>27 Nov 2017 - 28 Nov 2017</td>
<td>Training course &quot;Advanced Parallel Programming with MPI and OpenMP&quot;</td>
</tr>
<tr>
<td>29 Nov 2017</td>
<td>Training course &quot;Das Programmierwerkzeug make&quot;</td>
</tr>
</tbody>
</table>

http://www.fz-juelich.de/ias/jsc/EN/News/Events
High-Q-Club: >30 generic codes at scale!

dynQCD  Gysela  JusPIC  MP2C  μφ

PEPC  PMG+PFASST  TeraNEO  WalBerla

Gyrokinetic code
Laser-Plasma
CFD with Particles
Water in Porous Media
Particle Tree Code
ODE-Solver
MG for Geophysics
Lattice Boltzmann etc.

Non-trivial kernels only!
From dual to Modular Supercomputing

ARCHITECTURE EVOLUTION
Dual Architecture

- **JUMP**
  - IBM Power p690
  - 9 TFlop/s

- **JUROPA**
  - Intel Nehalem
  - 300 TFlop/s

- **JURECA Cluster**
  - Intel Haswell
  - ~2.2 PFlop/s

- **JUBL**
  - IBM Blue Gene/L
  - 5 TFlop/s

- **JUGENE**
  - IBM Blue Gene/P
  - 1 PFlop/s

- **JUQUEEN**
  - IBM Blue Gene/Q
  - 5.9 PFlop/s

General purpose clusters

- Deeply parallel architecture
Both combine:
- Hardware
- Software
- Applications
in a strong co-design

EU-Exascale projects
27 partners
Total budget: 44 M€
EU-funding: 30 M€
Nov 2011 – Jun 2020

www.deep-projects.eu
Homogeneous cluster
Heterogeneous cluster
Cluster-Booster architecture
DEEP Prototype

- Installed at JSC
- 1.5 racks
- 500 TFlop/s peak perf.
- 3.5 GFlop/s/W
- Water cooled

Cluster (128 Xeon)

Booster (384 Xeon Phi KNC)
Software environment

- **Scheduler**: Torque/Maui → future moving to SLURM
- **Filesystem**: BeeGFS
- **Compilers**: Intel, gcc, PGI
- **Debuggers**: Intel Inspector (threading, memory), TotalView (source code, memory debugger)
- **Programming**: ParaStation MPI (mpich), OpenMP, OmpSs
- **Performance analysis tools**: Extrae/Paraver, Scalasca, Intel Advisor, Intel, VTune…
- **Libraries**: SIOnlib, SCR, E10, HDF5, netCDF, PETSc…
Application-driven approach

**DEEP projects applications (15):**

- Brain simulation (EPFL + NMBU)
- Space weather simulation (KULeuven)
- Climate simulation (Cyprus Institute)
- Computational fluid engineering (CERFACS)
- High temperature superconductivity (CINECA)
- Seismic imaging (CGG + BSC)
- Human exposure to electromagnetic fields (INRIA)
- Geoscience (LRZ)
- Radio astronomy (Astron)
- Lattice QCD (University of Regensburg)
- Molecular dynamics (NCSA)
- Data analytics in Earth Science (UoI)
- High Energy Physics (CERN)

*CO-DESIGN*
Cluster-Booster architecture advantages

- **Full user flexibility** – many different use modes
  - Dynamic ratio of processors/coprocessors
  - Use Booster as pool of accelerators (globally shared)
  - Discrete use of the Booster
  - Discrete use + I/O offload
  - Specialized symmetric mode

- **More efficient use of system resources**
  - Only resources really needed are blocked by applications
  - Dynamic allocation further increases system utilization

- Better I/O performance and resiliency
Dual Architecture

**JUMP**
IBM Power p690
9 TFlop/s

**JUROPA**
Intel Nehalem
300 TFlop/s

**JURECA Cluster**
Intel Haswell
~ 2,2 PFlop/s

**JUBL**
IBM Blue Gene/L
5 TFlop/s

**JUGENE**
IBM Blue Gene/P
1 PFlop/s

**JUQUEEN**
IBM Blue Gene/Q
5.9 PFlop/s

**JURECA Booster**
DELL/Intel Xeon Phi
*Fictive picture

**General purpose cluster**

**Massively parallel architecture**
Dual Architecture

**JUMP**
IBM Power p690
9 TFlop/s

**JUROPA**
Intel Nehalem
300 TFlop/s

**JURECA Cluster**
Intel Haswell
~ 2.2 PFlop/s

**JUBL**
IBM Blue Gene/L
5 TFlop/s

**JUGENE**
IBM Blue Gene/P
1 PFlop/s

**JUQUEEN**
IBM Blue Gene/Q
5.9 PFlop/s

**JURECA Booster**
DELL/Intel Xeon Phi
5 PFlop/s
* Fictive picture

---

25 September 2017  
TerrSys Workshop – JSC Introduction  
37
Dual Architecture

- **JUMP**
  - IBM Power p690
  - 9 TFlop/s

- **JUROPA**
  - Intel Nehalem
  - 300 TFlop/s

- **JURECA Cluster**
  - Intel Haswell
  - ~2.2 PFlop/s

- **JUBL**
  - IBM Blue Gene/L
  - 5 TFlop/s

- **JUGENE**
  - IBM Blue Gene/P
  - 1 PFlop/s

- **JUQUEEN**
  - IBM Blue Gene/Q
  - 5.9 PFlop/s

- **JURECA Booster**
  - DELL/Intel Xeon Phi
  - 5 PFlop/s

*Fictive picture*

General purpose cluster

Massively parallel architecture
MODULAR SUPERCOMPUTING
Cluster – Booster architecture

Module 0: Storage

Module 1: Cluster

Module 2: Many core Booster

Disk

CN

Disk

BN

BN

BN

BN

BN

BN

BN

CN
Modular Supercomputing

Generalization of the Cluster-Booster concept
Modular Supercomputing

Module 0
Exascale Storage

Module 1
HPC Cluster

Module 2
Extreme Scale Booster

Module 3
Data Analytics Module

Module 4
Graphics Module

Module 5
Neuromorphic

Module n:
Modular Supercomputing

Module 1
HPC Cluster

Module 2
Extreme Scale Booster

Module 0
Exascale Storage

Module 3
Data Analytics Module

Module 4
Graphics Module

Module 5
Neuromorphic

Workload 1

Workload 2

Workload 3

Module n:
JSC Vision

Neuromorphic

Cluster

Quantum Computer

Data Analytics Module

Booster