HPC AND LARGE SCALE INSTRUMENTS – FZ JÜLICH‘S VIEW

6th Scientific Computing Forum, CERN

FEBRUARY 25, 2019 | THOMAS EICKERMANN, JSC

Mitglied der Helmholtz-Gemeinschaft
FZJ AND JSC AT A GLANCE
FORSCHUNGSZENTRUM JÜLICH BY NUMBERS

- **Budget**: 610 Mio €, including 245 Mio € third party funding
  100 Horizon 2020 projects, 420 national projects
- **Employees**: 5,900
  incl. 1,950 scientists including PhD students
  800 guest scientists from 75 countries
- **Publications**: 2,450
  (source: fact sheet 2017)
JULICH SUPERCOMPUTING CENTRE (JSC)

Facts and Figures

Staff:
- 220 Total (185 FTE)
- 160 Scientists
- 13 PhD Students (+13 external)

Budget:
- 30 Mio. € Institutional Funding (PoF)
- 15 Mio. € Third Party Funding
JÜLICH SUPERCOMPUTING CENTRE AT A GLANCE

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

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

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

Education and training
JSC’S ARCHITECTURE STRATEGY - RATIONALE

Dilemma #1

- Grand Challenge applications require extreme performance
- Not achievable with general purpose architectures (x86 Clusters): Cost, Energy
- Highly scalable architectures not suitable for applications requiring high single node performance, large memory per core

Solution: Dual architecture approach

- JUQUEEN - highly scalable system
- JUROPA / JURECA Cluster – general purpose system
- Common Storage
COMPUTER ROADMAP
DUAL ARCHITECTURE

IBM Power 6
JUMP, 9 TFlop/s

IBM Power 4+
JUMP (2004), 9 TFlop/s

IBM Blue Gene/L
JUBL, 45 TFlop/s

IBM Blue Gene/P
JUGENE, 1 PFlop/s

IBM Blue Gene/Q
JUQUEEN (2012)
5.9 PFlop/s

JURECA Cluster
(2015) 2.2 PFlop/s

JUST Gen 4:
20 PB raw

JUROPA
200 TFlop/s
HPC-FF
100 TFlop/s

General Purpose Cluster

Highly scalable
JSC’S ARCHITECTURE STRATEGY - RATIONALE

Dilemma #1
- Grand Challenge applications require extreme performance
- Not achievable with general purpose architectures (x86 Clusters): Cost, Energy
- Highly scalable architectures not suitable for applications requiring high single node performance, large memory per core

Solution: Dual architecture approach
- JUQUEEN - highly scalable system
- JUROPA / JURECA Cluster – general purpose system
- Common Storage

Dilemma #2
- Parts of complex applications or workflows often have different requirements and scalability properties
- Heterogeneous architectures / clusters with accelerators: static ratio of CPU / Accelerator performance potentially wastes resources and energy

Solution: Modular computing
- Separation of CPU (Cluster) and highly scalable Accelerator (Booster, e.g. GPU-centric system) allows dynamic resource allocation to optimize resource utilization
- Requires substantial application modifications
A EUROPEAN ARCHITECTURE

The DEEP Projects

- 20 Partners
- 45 Mio € EU
- 30 Mio € EU funded by EU
MODULAR SUPERCOMPUTING

Smart Exascale

HBP Data Analytics workflow

Module 1 Central Cluster

Module 2 ExaScale Booster

Module 3 Neuromorphic System

Module 4 Quantum Annealer

Module 5 Data Analytics Module

Module 6 Extreme Storage

Climatology workflow

Deep Learning workflow
COMPUTER ROADMAP
TOWARDS MODULARITY

IBM Power 6
JUMP, 9 TFlop/s

IBM Power 4+
JUMP (2004), 9 TFlop/s

IBM Blue Gene/L
JUBL, 45 TFlop/s

IBM Blue Gene/P
JUGENE, 1 PFlop/s

IBM Blue Gene/Q
JUQUEEN (2012)
5.9 PFlop/s

JURECA Cluster
(2015) 2.2 PFlop/s

JURECA Booster
(2017) 5 PFlop/s

JUWELS Cluster
Module (2018)
12 PFlop/s

JUWELS_Scalable
Module (2020)
50+ PFlop/s

Position 23
in June 2018,
Module 1
HPC AND LARGE SCALE INSTRUMENTS
OPPORTUNITIES AND CHALLENGES

Opportunities

- HPC systems provide vast amounts of compute capability and capacity
- HPC centres usually operate high-performance and large capacity storage resources
- Tier-0 / Tier-1 HPC centres typically have excellent NREN / GÉANT connectivity

Challenges

- HPC Systems are optimized for HPC not HTC
  - Applications / workflows may need adaptation
- Existing systems are fully committed and utilized
- Deployment and mode of operation may not be well adapted to requirements of large scale instruments:
  - 1 year project allocations based on Peer Review
    - long-term programme requirements
  - Batch mode with varying turnaround times
    - QoS / near-realtime requirements
  - Compute nodes without Internet Access
    - Access to remote data during runtime
ADDRESSING THE CHALLENGES #1: PILOTS

Demonstrators / Pilot Projects
- Address technical challenges and demonstrate feasibility on a small scale
- Postpone policy and financial questions until added value has been proven

PRACE link with Large Scale Scientific Instruments (F. Berberich at CERN on Oct. 22\textsuperscript{nd}, 2018)
- PRACE-5IP Pilot with European Synchrotron Radiation Facility (ESRF)
  - Application Use Cases: Simulation and Data Processing
  - Moderate resource requirement: 256 cores x 2-3 month ~ 500 T core-hours
  - Executed at PSNC (CPU cluster) and MPCDF (GPU cluster)
- First results
  - CPU cluster well suited for use case
  - PRACE’s user support beneficial for porting codes and adapting environment to the codes
- Another Pilot in preparation between CERN and PRACE-6IP
COSMIC-RAY RESEARCH ON THE ISS – AMS-02

A Production-Level Use-Case at JSC: Alpha Magnetic Spectrometer

- Multi-purpose cosmic-ray detector installed on the ISS
- Indirect search for dark matter and detection of cosmic antimatter as main objectives
- JSC provides compute and data facilities to support the data analysis of the AMS partner RWTH Aachen (Prof. Dr. S. Schael)
DATA FLOW AND DATA PROCESSING

- Raw Data
  - 40 TB p.a.
- ROOT Files
  - ca. 160 TB p.a.
- ACQ Files
  - ca. 20 TB p.a.
ADDRESSING THE CHALLENGES #2: PARTIALLY DEDICATED RESOURCES

Use Case Characteristics
- Multi-year project: AMS in operation since 2011; new cooling system planned for 2019
- Plannable IT-resources: data transfer, compute and storage requirements, campaign-oriented mode of operation

JUAMS: dedicated 2000-core CPU Cluster at JSC
- Jointly funded by FZ Jülich and AMS-researchers at RWTH
- Access to JSC’s HPC-storage, currently using 2.5 PB
- Leveraging:
  - High-performance Filesystem Access via MPI
  - Plannable (dedicated) CPU capacity, 100% utilized: competitive turnaround times
  - JUAMS will be replaced in 2019, new system will be tailored for AMS requirements (same compute, 5 PB)

CLAIX: Tier-2 HPC system at RWTH Aachen
- Relying on project resources granted in competitive Peer Review process via JARA-HPC (11.4 Mio. Core-h recently)
- Leveraging direct access of compute nodes to Internet (Data at CERN)
HPC Trends

- Single core performance gain is slowing down – need to increase level of parallelism
- Accelerators such as GPUs help improving performance / Watt & performance / Euro
- Modular Computing is JSC’s approach to address scalability limitations

HPC and Large Scientific Instruments

- HPC can help to address compute demand of large scale instruments
- Challenges
  - Technical: adaptation of applications and workflows
  - When leveraging existing resources at HPC centres: deployment, mode of operation, access policy, …
- There are success stories showing that these challenges can be overcome