Taking on Exascale Challenges: Key Lessons & International Collaboration Opportunities

Birds-of-a-Feather Session at SC15
Jointly organised by European Exascale Projects and ETP4HPC
Objectives of this BoF

• Present the European landscape of HPC technology research

• Discuss collaboration opportunities
| Introduction 15 min | Overview on the European Exascale Landscape  
By Jean-François Lavignon, ETP4HPC Chairman and Atos |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Talks 30 min</td>
<td><strong>Presentation of Focus Technology Research Areas</strong></td>
</tr>
</tbody>
</table>
| **Area 1: Architecture & Compute**  
By Filippo Mantovani, Technical Project Coordinator Mont-Blanc, Barcelona Supercomputing Centre | **Area 2: Interconnect, Memory & Storage and Data-intensive Real-Time**  
By Prof. Jesus Carretero, Computer Architecture Professor, Computer Science and Engineering Dep. University Carlos III of Madrid |
| **Area 3: Programming Tools, Algorithms & Mathematics**  
By Stefano Markidis, Assistant Professor, KTH Royal Institute of Technology |
<table>
<thead>
<tr>
<th>Panel Discussion</th>
<th>45 min</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>International Collaboration Opportunities arising and mechanisms needed</strong></td>
<td></td>
</tr>
<tr>
<td>Featuring distinguished international guests:</td>
<td></td>
</tr>
<tr>
<td><strong>Mitsuhisa Sato</strong>, Co-project leader of Post-K, University of Tsukuba &amp; RIKEN AICS</td>
<td></td>
</tr>
<tr>
<td><strong>Franck Capello</strong>, Argonne National Laboratory, Senior Computer Scientist, Director of the INRIA, UIUC, ANL, BSC, JSC and Riken Joint Laboratory on Extreme Scale Computing</td>
<td></td>
</tr>
<tr>
<td><strong>Eric Van Hensbergen</strong>, Senior Principal Research Engineer, ARM</td>
<td></td>
</tr>
<tr>
<td><strong>Moderated by Sai Narasimhamurthy</strong>, Staff Engineer, Research, Seagate</td>
<td></td>
</tr>
</tbody>
</table>
1. A **European HPC Technology Project Handbook** detailing all the projects is available at the event (Please collect one from the organisers now).

2. A **Report** from this BOF.

Please leave your **business card** and both documents will be sent to you via email.

These documents will also be published on the website of ETP4HPC (the European Technology Platform for HPC), [www.etp4hpc.eu](http://www.etp4hpc.eu), and the **individual project websites**.

You can also contact the Office of ETP4HPC at office@etp4hpc.eu with a request to receive a copy of either.
Presentation of Focus Technology Research Areas
A bit of history

• In 2004 start thinking about a HPC research infrastructure (HPCEUR, HET) ending with the creation of PRACE in 2010
• In 2010, expert group meetings between EC and European stake holders for a more global policy
• In 2011 creation of ETP4HPC
• Actions in FP7
  – PRACE Preparatory Phase and first Integrated Projects
  – Supports action : EESI and EESI2
  – calls for Exascale computing in WP2011 and WP2013
The EC communication

• High-Performance Computing: Europe's place in a Global Race issued in Feb 2012

• Policy with 3 pillars

Access to best HPC for industry and academia (PRACE)

- Collaboration of HPC Centres and application CoEs
- Provision of HPC capabilities and expertise

FET/HPC: EU development of Exascale technologies

- Specifications of exascale prototypes
- Technological options for future systems

Excellence in HPC applications (Centres of Excellence)

- Identify applications for co-design of exascale systems
- Innovative methods and algorithms for extreme parallelism of traditional/emerging applications
The HPC Public Private Partnership

- Mutual commitment
  - European Commission
    - HPC as a priority of Horizon2020
    - Funding of 700 M€
  - ETP4HPC
    - Investment to match EC funding in R&D
    - Effort to maximize impact on European industry

- Partnership board
  - strategy setting
  - impact monitoring
ETP4HPC

• Open association for organizations with HPC research activities in Europe
  – industry led with more than 70 members
• The Strategic Research Agenda

www.etp4hpc.eu
Support actions

- **EESI and EESI2 (EU FP7)**
  - run by European experts from 2008 to 2015
  - important recommendations covering technology, applications, algorithms
  - [http://www.eesi-project.eu/](http://www.eesi-project.eu/)

- **NESUS (COST action)**
  - sustainability in ultrascale systems, started in April 2014
  - programming models, resilience, runtime systems, data management, and energy efficiency
  - [http://www.nesus.eu/](http://www.nesus.eu/)

- **Eurolab4HPC (H2020)**
  - To build the foundation for a European Research Center of Excellence in High-Performance Computing (HPC) Systems
  - [http://eurolab4hpc.eu/](http://eurolab4hpc.eu/)

- **EXDCI: ecosystem development (H2020)**
  - managed by PRACE and ETP4HPC, started in Sept 2015
  - support of application and technology roadmap
  - training and education; support to SMEs; international cooperation; monitoring and impact assessment of the H2020 programme
  - [http://www.exdci.eu/](http://www.exdci.eu/)
FP7 Projects

• Started in 2011

• EU funding: 45 Mio € + in-kind contributions by the partners (budget: more than 90 Mio € in total)

• 8 projects: CRESTA, DEEP & DEEP-ER, EPIGRAM, EXA2CT, Mont-Blanc (I+II), Numexas
The Horizon 2020 HPC projects

• First call of Horizon2020
  – 19 research projects and 2 support actions
  – Most projects start in Sept-Oct 2015 for 3 years
  – Global effort: 94 M€ for R&D projects

• Global facts
  – 170 organizations involved in this effort
  – Project repartition
    • 9 HPC core technologies and architectures
    • 5 Programming methodologies, environments, languages and tools
    • 0 APIs and system software
    • 5 New mathematical and algorithmic approaches

<table>
<thead>
<tr>
<th>Industry</th>
<th>Non EU</th>
<th>Research</th>
<th>SME</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.7%</td>
<td>0.2%</td>
<td>68.9%</td>
<td>15.3%</td>
</tr>
</tbody>
</table>
More will come

• 2 calls in Work Programme 2016-2017
  – FET HPC 1 2016: Co-design of HPC systems and applications
    • budget: 41 M€ - deadline: 27 September 2016
  – FET HPC 2 2017: Transition to Exascale Computing
    • 5 subtopics:
      – High productivity programming environments for exascale
      – Exascale system software and management
      – Exascale I/O and storage in the presence of multiple tiers of data storage
      – Supercomputing for Extreme Data and emerging HPC use modes
      – Mathematics and algorithms for extreme scale HPC systems and applications working with extreme data
    • budget: 40 M€ - deadline: 26 September 2017

• On-going discussion on additional Horizon2020 programme

• New HPC initiative by the European Commission
Area 1: Architecture and Compute

Filippo Mantovani
Barcelona Supercomputing Center
Topics covered by this area

Architecture of future HPC platforms will deal with:

• Energy efficiency
  – 1 ExaFLOP in 20 MW
• Heterogeneity
  – Name your device, please.
• Reconfigurability
  – If you have named FPGA in the previous point...
    You most probably want to reconfigure it, right?
• Resource balance
  – Mostly balance compute throughput, memory and network bandwidth
• Co-design driven
  – And now you want to balance also for CFD, QCD, MD, ...
• Integration and reliability
  – 50K+ compute nodes... Heterogeneous...
  – How to detect failures? How to survive with them?
Energy Efficiency

• Low power compute components
• Advanced nanotechnologies
• Extreme resource efficiency
• Power monitoring
• Power aware scheduling and programming
• Hot-water cooling

• and of course heterogeneity (coming soon)
Heterogeneity

• Classical heterogeneity
  – CPU + accelerator

• Fine grained heterogeneity
  – Heterogeneous compute cores on chip
  – Mobile on chip GPUs

• System-level modularity
  – Cluster-Booster approach
  – Custom interconnection

• Hierarchical system partitioning
  – “Workers” grouped by address spaces
Reconfigurability

• Support for specific HPC applications
  – Mapping of functions
  – Mapping of algorithms

• Hardware runtime support

• Leveraging low reconfig. overhead

• Aiming reducing data traffic
Balanced / Co-design driven

• Heterogeneous hardware platform with matching software stack and optimized grand-challenge HPC applications

• Analysis of requirements HPC applications, mini-apps and kernels via performance analysis tools

• Programming models enabling hw resources with minimal impact on app
Integration and Reliability

- High density
  - 2.5D technology / Interposer
  - System on Package
- Advanced cooling
  - Air/Liquid
  - Single/double loop
  - Heat reuse
- Fault tolerance
  - Error detection mechanisms
  - Reliability of large systems
  - Quality of Service
Collaboration Opportunities

• Liaison with international technology companies (not only EU)
• Workshops and trainings
• Lessons learned in prototyping
• Lessons learned in co-design and software modernization
• Accessing prototype platforms
  – DEEP → 500 TFlop/s peak performance prototype installed at JSC (pmt@deep-project.eu)
  – Mont-Blanc → 1000+ nodes ARM-based cluster installed at BSC (http://montblanc-project.eu/industrial-user-group)
Area 2: Interconnect, Memory & Storage
Data-Intensive Real Time

Jesus Carretero
University Carlos III of Madrid
Topics covered by this area

• System architecture for Exascale and data-centric HPC
  – Very Tightly Coupled Data & Computation
  – Codesign using accelerators and FPGAs

• Develop a new server architecture using next generation processors and memory advances
  – Integration of NVRAM technologies in the I/O stack
  – Extreme compute power density
  – Develop the systemware to support their use at the Exascale

• New applications and use cases emerging for HPC
  – Real-time, data-intensive, and energy efficiency
Develop a new server architecture using next generation interconnection, and memory advances

- Integration of NVRAM technologies in the I/O stack
- Fast, distributed in-node non-volatile and network attached memory Storage
- Extreme compute power density
- Low-latency unified Interconnect for compute & storage traffic
- Codesign using accelerators and FPGAs
- Liquid cooling technologies
Storage

• Develop the systemware to support new architectures used at the Exascale
  • Data Centric Computing System based on object-storage
  • Leveraging the I/O stack to enhance data-intensive computing
  • Computation off-loading to I/O system

• Model different I/O workloads and use this understanding in a co-design process
  – Very Tightly Coupled Data & Computation
  – API for massive data ingest and extreme I/O
  – Extreme data management and analysis
Exploring New Heterogeneous Architectures For HPC Systems

- Deeply heterogeneous manycore architectures
- Real-time support providing a unified access to the systems via a smart interconnect
- Adaptive programming models and compiler support to the new architectures

New applications and use cases emerging for HPC arena

- Real-time, data-intensive, and energy efficiency
- Applications used to identify system requirements
- Real scientific and data center applications (e.g. MAORY RTC system).
Collaboration Opportunities

• Deeply heterogeneous system architecture for Exascale and data-centric HPC
  • Integration of NVRAM technologies in the I/O stack
  • Develop the systemware to support their use at the Exascale
• Leveraging the I/O stack for Exascale systems
• Data staging coordination and control for Exascale applications
  • API for massive data ingest and extreme I/O
• New applications and use cases needing Real-time, data-intensive, and energy efficiency
Area 3: Programming Models, Algorithms and Mathematics

Stefano Markidis

KTH Royal Institute of Technology, Sweden
8 Programming Models/tools Projects:

- AllScale, Antarex, CRESTA, Exa2CT, EPiGRAM, ESCAPE, Intertwine, READEX

6 Algorithms and Mathematics Projects:

- ComPAT, ExCAPE, ExaFLOW, ExaHYPE, NLAFT, NUMEXAS
• Innovative Programming Models for Exascale
  – **AllScale**: An exascale programming, multi-objective optimization and resilience management system supporting nested recursive parallelism.

• Enhanced MPI and PGAS: the incremental approach
  – **CRESTA**: Enhancing programming models and system software by co-design. Large-scale real-world applications to guide the development of the software stack for exascale.
  – **EPIGRAM**: MPI and GPI for exascale. Combining best features of Message-Passing and PGAS programming model.
  – **Exa2CT**: cutting edge of the development of solvers, related algorithmic techniques, and HPC software architects for GASPI communication.
• Interoperability of programming models: the ‚‘+‘‘ issue
  – **Intertwine**: enhanced programming and runtime systems for effective interoperability.

• Autotuning for energy efficiency and green HPC:
  – **READEX**: developing a tools-aided methodology for dynamic auto-tuning of HPC applications to exploit the dynamically changing resource requirements for improved energy-efficiency.
  – **ANTAREX**: providing a breakthrough approach to express by a DSL the application self-adaptivity and to runtime manage and autotune applications for green and heterogeneous HPC
ALGORITHMS

• Computational Fluid Dynamics for Exascale
  – **ExaFLOW**: addressing algorithmic challenges to enable the use of accurate simulation models on exascale with focus on error control, AMR in complex geometries, heterogenous modeling, energy efficiency, and in-site I/O.
  – **NUMEXAS**: develop, implement and demonstrate the next generation of numerical simulation methods with focus on industrial applications and on pre- and post-processing.
  – **ESCAPE**: developing next generation IFS numerical blocks for weather forecast.

• Multiscale Applications
  – **ComPAT**: development of the High Performance Multiscale Computing paradigm.
ALGORITHMS AND MATHEMATICS

• Machine Learning:
  – **ExCAPE**: better machine learning algorithms for predicting biological activity of drugs and their deployment on HPC systems.

• Solvers:
  – **NLAFET**: linear solvers for exascale with novel algorithms, communication avoiding, advanced scheduling strategies and autotuning
  – **ExaHyPE**: High-order Discontinuous Galerkin hyperbolic PDE engine for geo- and astrophysics.
Opportunities for International Collaboration

• Programming and Autotuning Systems
  • Stellar group on HPX
  • MPI Forum and main MPI implementors (MPICH, OPENMPI)
  • OpenMP ARB
  • PNNL (Auto-tuning)

• Algorithms
  – University of Tennessee for magma and plasma libraries
Panel discussion

**Mitsuhisa Sato**, Co-project leader of Post-K, University of Tsukuba & RIKEN AICS

**Franck Capello**, Argonne National Laboratory, Senior Computer Scientist, Director of the INRIA, UIUC, ANL, BSC, JSC and Riken Joint Laboratory on Extreme Scale Computing

**Eric Van Hensbergen**, Senior Principal Research Engineer, ARM

Moderated by **Sai Narasimhamurthy**, Staff Engineer, Research, Seagate