European Exascale Software Initiative EESI2
Towards exascale roadmap implementation

Quick Overview
Fukuoka, February 27, 2014

Philippe RICOUX (TOTAL), EESI2 Coordinator
Jean Yves BERTHOU, Thierry BIDOT, all WP leaders
EESI objectives

Build and consolidate a vision and roadmap at the European level, including applications, both from academia and industry to address the challenge of performing scientific computing on the new generation of supercomputers, hundreds of Petaflop/PBytes in 2017 and Exaflop/ExaBytes in 2020/2022.

Propose & Initiate International collaborations

In order to tackle the key issues
What is at stake?

- Velocity Model
- Snap Shot
- Seismic interpretation
- Geological Study

B – Seismic Depth Imaging
C – Interpretation structurale
Seismic interpretation – Geological Study

- Human Brain Project

MODELING THE CLIMATE SYSTEM
Includes the Atmosphere, Land, Oceans, Ice, and Biosphere

Cellular Neurosurgical Colonies (in vivo and ex vivo)
**Why Exascale?**

**EXAFLOP (capability and capacity)**

*Drug design*: realistic cell membrane models, including **drug permeation and binding**

*Oil&Gas*: huge 3D seismic wave inversion, reservoir modelling, robust optimization

*Materials*: material properties identification

*Aeronautics*: Greening the aircraft

*Astronomy*: **Square Kilometre Array**, Earth Sciences: Natural hazard mitigation

*Fundamental Sciences, Life Sciences, Engineering sciences* (Turbulence, Combustion, acoustics, Mechanical, chemical engineering, …)

*Climate*: satellite sensors create floods of data (x 1000), leading to Exa-scale archives, ex. projected frequency of intense tropical cyclones in some region of the globe

*Quantum Chemistry*: discover a material and properties: ab initio databases of materials and molecular properties connected to existing databases of experimental properties

*Industrial applications*: management of data generated from micro/macro sensors and automated measurement devices
EXAFLOP also means a Petaflop in a box … and 20 KW

EXAFLOP also means a Petaflop in a box … and 20 KW

**Huge impact** for those, academic, industrial, large and small structures, including SMEs, that will be able to take advantage of “Exascale” technology, not just for few heroes applications

---

**Exascale, a technological breakthrough**

- Compare to Petascale computers: memory per core 1/10, CPU heterogeneity, total node interconnect BW & node memory 1/10, concurrency *10
  
  => concurrency/load balancing, data locality/Memory management, resilience/fault tolerance, energy efficiency

- Software layer and applications need to exploit these new hardware trends that cannot be handled by existing software stack

- Community codes unprepared for sea change in architecture while: designing and developing a new generation of Scientific Applications takes 5 to 10 years, lifetime of Scientific Applications are several decades

---

### System attributes

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>“2015”</th>
<th>“2018”</th>
<th>Difference Today &amp; 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>System peak</td>
<td>2 Piflop/s</td>
<td>200 Piflop/s</td>
<td>1 Eflop/sec</td>
<td>O(1000)</td>
</tr>
<tr>
<td>Power</td>
<td>6 MW</td>
<td>15 MW</td>
<td>~20 MW</td>
<td>O(100)</td>
</tr>
<tr>
<td>System memory</td>
<td>0.3 PB</td>
<td>5 PB</td>
<td>32-64 PB</td>
<td>O(100)</td>
</tr>
<tr>
<td>Node performance</td>
<td>125 GF</td>
<td>0.5 TF</td>
<td>7 TF</td>
<td>O(10) – O(100)</td>
</tr>
<tr>
<td>Node memory BW</td>
<td>25 GB/s</td>
<td>0.1 TB/sec</td>
<td>1 TB/sec</td>
<td>O(100)</td>
</tr>
<tr>
<td>Node concurrency</td>
<td>12</td>
<td>O(100)</td>
<td>O(1,000)</td>
<td>O(10,000) – O(1000)</td>
</tr>
<tr>
<td>Total Concurrency</td>
<td>225,000</td>
<td>O(10^9)</td>
<td>O(10^9)</td>
<td>O(10,000)</td>
</tr>
<tr>
<td>Total Node Interconnect BW</td>
<td>1.5 GB/s</td>
<td>20 GB/sec</td>
<td>200 GB/sec</td>
<td>O(100)</td>
</tr>
<tr>
<td>MTTI</td>
<td>days</td>
<td>O(1 day)</td>
<td>O(1 day)</td>
<td>O(10)</td>
</tr>
</tbody>
</table>
Exascale, a technological breakthrough, imposes

To think different and differently

Exascale cannot be justified only if we are just planning to do the usual thing but bigger

Exascale needs breakthroughs in several domains (Algorithms, Algebra, Uncertainties, Couplers, Meshing …)
Towards Exascale: Main issues to be addressed (EESI1)

**At the level of simulation environment:**
- **Unified Simulation Framework** and associated services: CAD, mesh generation, data setting tools, computational scheme editing aids, visualization, etc.
- **Multi-physics simulation**: establishment of standard coupling interfaces and software tools, mixing legacy and new generation codes
- Common (jointly developed) **mesh-generation tool**, automatic and adaptive meshing, highly parallel
- **Standardized efficient parallel IO and data management** (sorting memory for fast access, allocating new memory as needed in smaller chunks, treat parts of memory that are rarely/never needed based on heuristic algorithms, …)

**At the level of codes/applications:**
- **New numerical methods**, algorithms, solvers/libraries, improved efficiency
- Coupling between stochastic and deterministic methods: Numerical scheme involving Stochastic HPC computing for uncertainty and risk quantification
- **Meshless methods and particle simulation**
- **Scalable program, strong and weak scalability**, load balancing, fault-tolerance techniques, multi-level parallelism (issues identified with multi-core with reduced memory bandwidth per core, Collective communications, Efficient parallel IO)
- Development of standards programming models (MPI, OpenMP, C++, Fortran, …) handling multi-level parallelism and heterogeneous architecture
- **Human resources, training** (what level?)

50% Computer Power for Data movement

Synchronization and Communication reducing algorithms
EESI roadmaps, vision and recommendations need to be monitored, updated, on a dynamical way

Key issues to be addressed are pointed out in EESI1 … Now EESI2 must recommend R&D actions how to tackle them

- Extend, refine, and update Exascale cartography (directly in the dedicated WG for better analysis of each WG) and roadmaps from HPC community, on software, tools, methods, R&D and industrial applications, ...

  With a Gap Analysis.

  Including WG on disruptive technologies

- Address “Cross Cutting issues”: Data management and exploration, Uncertainties - UQ&VQ, Power & Performance, Resilience, Disruptive technologies

- Investigation on funding scheme and opportunities, education, co-design centres, international coordination

- Operational Software maturity level methodology, evaluation
Investigation on funding, education, co-design centres, international coordination

WG Application WGs

WG Enabling Technologies WGs

WG Cross cutting issues WGs

Operational software maturity level methodology

Management

Dynamic Vision and Roadmap

Dissemination

Updated Cartography

One by year

Contractual partners: TOTAL (coordinator), PRACE AISBL (acting for third parties LZR, GENCI, BSC, CINECA, EPCC, SARA…)

Contributing partners, involved in the management of EESI2 tasks but not associated to PRACE AISBL: INTEL, DLR, EDF, ANR, CERFACS, …

Supporting partners: more than 50 letters of Support

EESI2 proposal submitted in November to INFRA-2012-3.3: Coordination actions, conferences and studies supporting policy development, including international cooperation, for e-Infrastructures.

Requested funding: 1.5 M€ → 1.36 M€ accepted by EC

Duration: 30 months, Start 1st September 2012 - kick off 18th September 2012 (extended to 34 months)
EESI2 Working Groups

Enabling technologies for Exaflop computing
- Hardware roadmap, links with vendors
- Numerical Analysis
- Scientific software engineering, software eco-system and programmability
- Disruptive technologies

Application Grand Challenges
- Industrial and Engineering Applications
- Weather, Climatology and Earth Sciences
- Fundamental Sciences
- Life science & Health
- Disruptive technologies

Cross cutting issues
- Data management and exploration
- Uncertainties - UQ&VQ
- Power & Performance
- Resilience
- Disruptive technologies

10 EESI2 Overview January 2014
Disruptive Technologiest (Iain Duff, EASC 2013)

Disruptive technologies in algorithms

Potential Breakthroughs ?:

Tensor Algebra

- Primitive functions (Sum, Product, \textbf{Eigen values}, ...), to be defined (works of Pierre Comon)
- coupled with \textbf{hierarchical embedded structures} (Dickson Algebra?)

Surrogate Model

- (multi scale discretization, splitting equations, big data compression, ...)

Model reduction

Communication

Computation

Simulation

Mixing

Low dimensional transforms

Universal

Model reduction

Chaotic relaxation

Hybrid

Storage
EESI2 FOCUS

- Scientific key issues to be tackled
  - Fundamental, Industrial, Engineering, eco software, ...
  - Cross Cutting
    - Big Data in extreme computing
    - Uncertainties (Quantification, propagation, ...)
  - *Not only challenges list but technical propositions (funding specific experts)!*

- Detailed periodic roadmaps on the key challenges, including:
  - Gap analysis,
  - *Breakthroughs*,
  - Identification of priority actions,
  - *Recommendations to R&D project*

- Define educational programs for Exascale

- Contributing to build an international Exascale software Organization
The EESI2 Vision

The vision which is presented here is also inspired by some worldwide recent installations in Europe, Asia and in USA of 10 (and more) Petaflops computers and by the feedback of several applications and tests running on full configurations of these systems. These tests have shown the extreme difficulty to get some acceptable results in term of performance on these computers. In particular the following points appear to be critical:

- Resilience
- Error propagation
- Reproductibility
- Data transfert, communication
- Task synchronization

As a consequence, Exascale applicative software appear to be a very difficult challenge and most worldwide experts consider that this challenge will not be solved with existing algorithms.
What appears presently, shared by US, Japanese and European experts, is that:

- Exascale technology will *trickle down to every scale* (architecture system as well physics and time)
- Exascale cannot be justified only if we are just planning to do the usual thing but bigger
- Exascale machines will be useless without algorithms that use their specific features
- Exascale imposes to do something different and differently

The following points are on the critical path to Exascale Computing:

- The use of hierarchical algorithms which reduce communications and tasks synchronizations
- The use of multi-physics methods which do not need or minimize data transfers and include multi scaling and parallel space-time methods
- The reshaping of operating systems and management tools such as MPI and OpenMP and mesh generation tools to the new developed algorithms
- The use of in situ data processing
In term of organization, the EESI experts agree that multi disciplinary research teams approach enable the emergence of significant progress toward the implementation of Exascale applications. The best coupling of Architecture, Algorithm and Application (AAA) is the challenge of efficient Exascale software. That leads to increase international collaboration, international working teams.

But, even if all agree on these points, there is an on going debate on co-design centre structure. EESI2 will go on deeper on this item. The related EESI2 working group just begins its works and will give recommendation within the next deliverable in 2014. As a first recommendation, EESI claims that co-design centers/centers of excellence should conciliates scientific multi-disciplinarity, international dimension, critical mass of researchers working at the same place, the balance of vertical (specialty) and horizontal (transverse) scientific domain and the need to do things differently.
2013 EESI2 recommendations

- R1. Ultra scalable algorithms
- R2. Resilience
- R3. Big data
- R4. Couplers
- R5. High productivity programming models
- R6. Mini apps
- R7. Software Engineering Methods for High-Performance Computing
- R8. Uncertainty Quantification Verification, Validation
R1. Ultra scalable algorithms

Motivations:
- architectural features: highly parallel, hierarchical and heterogeneous processor layout
- energy and resilient aware algorithms
- increasing complexity of computations that application scientists and engineers

Type of recommendation/Proposal for the next step
An IP over 4 years could host this project. It should mean a 15 million Euros budget, 6M on SP1, 6M on SP2, 3M on SP3. Deliverables will come mostly in the form of Open Source software, ready for use for the scientific and industrial community.

- Maturation of existing algorithms and software tools to optimize them for Exascale computing
- Development of new numerical techniques for handling very large problems
- Target large scale applications/real life problems
- International coopération (EMWG DOE WG)
R2. Resilience: an holistic approach

Motivation:
- dealing the resilience challenge,
- global solution from hardware/OS, algorithms, software

Sources of faults
- tolerance to fail stop errors (Process crashes)
- tolerance to silent data corruptions (Data corruptions)

Type of recommendation/Proposal for the next step
An IP over 4 years could host this project. It should mean a 15 million Euros budget, 3M on SP1, 2M on SP2, 3M on SP3, 1M on SP4, 3M on SP5, 3M on SP6. Deliverables will come mostly in the form of Open Source software, ready for use for the scientific and industrial community.

Proposal:
- tolerance to fail stop errors (Process crashes)
- tolerance to silent data corruptions (Data corruptions)

- Resilient OS (make sure that OS confine faults) - No leader for OS
- Resilient Algorithms - Avoid overlap with USA and Japan, focus on complementarities (may need to establish a forum/working group)
- An international effort: take a complementary position to USA and Japan
R4. Couplers: an holistic approach

Motivations:
- Legacy codes & models (X10 century.man, validated models, x10 millions.line)
- Context of Exascale: massively parallel, big data, data locality (meshing), resilient, energy aware (informatics)

Proposal:
- Improving/adapting to Exascale specificities
- MiniApps, coupling-aware performance modeling
- Couplers: standard coupling API, data and communication management
- Codes to be coupled: component-based approach, new programming model (PGAS, hybrid, ...)
- Coupling methods/algorithms (maths): data and communication management
- Pre and Post treatment: mesh management (generation, adaptation, conformity verification, ...), Visualization

Type of recommendation/Proposal for the next step:
- An IP over 4 years could host this project. It should mean an approximative 15 million Euros budget.
EESI Conclusions

- Europe need for a **sustainable, long term and coordinated** effort
- Europe is still well positioned to be part of the few player worldwide deploying and exploiting Exascale technology but action is needed **now**
- A 2,5 to 3,5 billions euros total budget over 10 years, supported by EC, National European funding agencies, industry, … a **several decades** engagement
- Scientific Computing at Exascale, from a computing and data intensive point of view are **strategic** for maintaining and developing both **European Scientific Excellence and Industry Competitiveness**
- **International** collaboration is required
- Beside legacy codes, Europe should encourage the development of **Open Source solutions** to foster international collaborations and the emergence of international **de facto** standards, enabling commercial exploitation
http://www.eesi-project.eu