Enhancing Energy Production with Exascale HPC Methods

The research leading to these results has received funding from the European Union's Horizon 2020 Programme (2014-2020) and from Brazilian Ministry of Science, Technology and Innovation through Rede Nacional de Pesquisa (RNP) under the HPC4E Project (www.hpc4e.eu), grant agreement n° 689772.

CARLA 2016 Conference
Aug 29-Sep 02, 2016 – Mexico DF
High-Performance Computing For Energy
OUTLINE

- The consortium
- A Disruptive Exascale Computer Architecture
- Simulators for Exascale Computations
- Energy fields approached
- General and Latin American Impact
- Conclusions
The Consortium
THE HPC4E CONSORTIUM

- HPC4E (http://www.hpc4e.eu) is a project co-funded by the European Commission under its H2020 programme and the Brazilian Rede Nacional de Pesquisa.
- Its lifetime is of 2 years starting on December 1, 2015.
- This project aims to apply the new exascale HPC techniques to energy industry simulations, customizing them, and going beyond the state-of-the-art in the required HPC exascale simulations for different energy sources:
  - Wind energy production
  - Efficient combustion systems for biomass-derived fuels (biogas)
  - Exploration geophysics for hydrocarbon reservoirs.
- As a result of these intensive prototyping and numerical simulations, a more accurate assessment of their real value and improvement their throughput will be achieved.
THE HPC4E CONSORTIUM

- European Partners
  - Barcelona Supercomputing Center (BSC) – Coord.
  - Institut National de Recherche en Informatique et Automatique (INRIA)
  - Lancaster University (ULANC)
  - Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT)
  - Repsol S.A. (REPSOL)
  - Iberdrola Renovables Energía S.A. (IBR)
  - Total S.A. (TOTAL)
THE HPC4E CONSORTIUM

- Brazilian Partners
  - Fundação Coordenação de Projetos, Pesquisas e Estudos Tecnológicos (COPPE) – Coord.
  - National Laboratory for Scientific Computation (LNCC)
  - Instituto Tecnológico de Aeronáutica (ITA)
  - Petroleo Brasileiro S. A. (PETROBRAS)
  - Universidade Federal do Rio Grande do Sul (INF-UFRGS)
  - Universidade Federal de Pernambuco (CER-UFPE)
HPC4E (HPC for Energy)

- 2 Million Euro for European partners; 6 Million Reais for Brazilian partners

"Apply Exascale HPC technology to energy industry simulations"  
Wind + Oil and Gas + Biomass
CONSORTIUM AS A WHOLE

- The HPC4E project partners provide their facilities and expertise
  - **Exascale systems**: LNCC, BSC, INF-UFRGS
  - **Numerical methods**: BSC, CIEMAT, INRIA, ULANC, COPPE, INF-UFRGS, ITA, LNCC
  - **Wind Modeling**: BSC, CIEMAT, CER-UFPE, IBR
  - **Combustion Simulations**: BSC, ULANC, CIEMAT
  - **Exploration Geophysics**: REPSOL, TOTAL, PETROBRAS, BSC, INRIA, ITA, COPPE, LNCC
Project Structure

WP 1 MANAGEMENT

WP2 DISRUPTIVE EXASCALE COMPUTER ARCHITECTURES

WP3 SIMULATORS FOR EXASCALE COMPUTATIONS

WP4 ATMOSPHERE FOR ENERGY

WP5 BIOMASS FOR ENERGY

WP6 GEOPHYSICS FOR ENERGY

WP7 DISSEMINATION & EXPLOITATION
A Disruptive Exascale Computer

Architecture
A DISRUPTIVE EXASCALE COMPUTER ARCHITECTURE WILL BE SET UP

- To study the mapping and optimization of the codes proposed for each energy domain on novel architectures for exascale, as well as developments in the underlying software infrastructure.

- Porting, tuning, and testing efforts of the different simulations codes will be carried out.

- Specifically, four lines of action will be pursued.
LINE OF ACTION #1

- The selected computing kernels of the codes coming from the energy sector will be optimized for architectures based on accelerators.
- The goal is to optimize the performance but keeping a high degree of portability.
  - The ratio flops/watt obtained in each platform will be analyzed.
- The main target architecture platforms are those based on Xeon Phi and NVIDIA GPUs, but other platforms based on embedded processors will be also analyzed.
LINE OF ACTION #2

- The selected kernels will be also ported to architectures based on symmetric multicore processors with NUMA memory
  - The goal will be to optimize the performance
- The main target architectures will be Intel and AMD, but also new platforms based on ARM processors will be analyzed
- The key point will be the load balancing and data placement
  - New scheduling algorithms able to improve locality will be taken into account
LINE OF ACTION #3

- The management of the MPI level parallelism in the codes coming from the energy sector will be guaranteed for achieving a high scalability of the applications in HPC clusters with millions of cores.

- The main topics to be analyzed will be:
  - Creation of tools for migration of running parallel tasks inside clusters
  - Hierarchical MPI structures to manage coupled multiphysics problems
  - Parallel I/O optimization
  - Design of efficient check-pointing strategies
  - Fault tolerance strategies at MPI level
LINE OF ACTION #4

- Performance analysis will be focused on the performance analysis of the different applications and kernels
  - The proper environments and tools (Paraver, Triva, Ocelotl, TAU, etc.) will be deployed to analyze all the parallel levels in the applications
- Inside a computational node roof-line analyses will be done to understand the bottlenecks of the architectures
- At the cluster level, network traffic, I/O traffic, and load balancing will be analyzed to guarantee the application scalability
- Performance prediction tools will be used to analyze the potential benefits of architecture or algorithm modifications
- Different proposal of exascale architectures will be studied for the selected applications as well.
SOME RESULTS: INTEGRATING CHECKPOINTING TECHNIQUES INTO SLURM

- Slurm is the workload manager designed for Linux clusters and the most commonly used by the TOP500 supercomputers
- Checkpointing can be done at both
  - Application level $\rightarrow$ FTI library
  - System level $\rightarrow$ DMTCP library
- Slurm is being extended with third parties plugins to these APIs
  - Resilience is increased
  - New scheduling algorithms can be designed
Simulators for Exascale Computations
COMPUTATIONAL ALGORITHMS FOR EXASCALE ARCHITECTURES

- Numerical schemes for Partial Differential Equations (PDE)
  - Scalable implementations of high order schemes for wave propagation models
- Sparse linear solvers
  - Generic (i.e. algebraic) parallel solvers for large sparse linear systems of equations
- Adaptivity
  - Mesh and (local) time-step adaptive algorithms in order to optimize the use of computational resources.
- Data management
COMPUTATIONAL ALGORITHMS: HIGH ORDER NUMERICAL SCHEMES

- Two families of innovative high order finite element methods and a family of (standard and mimetic) finite difference schemes will be considered for both time-domain and frequency domain.

- These numerical schemes exhibit a high level of parallelism. In particular, they are well suited to a mixed coarse grain/fine grain (MIMD/SIMD) parallelization targeting many-core (Xeon Phi/GPU) systems.
  - Multiscale Hybrid-Mixed (MHM) methods combined with Discontinuous Galerkin (DG) or Stabilized Continuous Galerkin (SCG) for time-domain.
  - For frequency-domain problems, the same analysis with the so called hybridized DG formulations that drastically reduce the number of globally coupled degrees of freedom will be performed.

- Both types of solvers (i.e. time-domain and frequency-domain) are linked to the simulation/inversion framework for sub-surface imaging proposed in the geophysics domain.
COMPUTATIONAL ALGORITHMS: SCALABLE SPARSE LINEAR SOLVERS

- Both direct and hybrid direct/iterative solvers will be considered
- This hybrid iterative/direct strategy will be made available through the MaPHyS\(^1\) software that implements algebraic domain decomposition ideas and relies in parallel on parallel sparse direct solvers such as PaStiX\(^2\) for each sub-problem
- On top of those two solvers, Krylov subspace methods are implemented

\(^1\) http://maphys.gforge.inria.fr
\(^2\) http://pastix.gforge.inria.fr
COMPUTATIONAL ALGORITHMS: NUMERICAL SCHEMES FOR PDES

- They involve adapting the grids in space and time to minimize errors in the simulation.
- The numerical simulation of partial differential equations will be performed with arbitrary unstructured discretizations on serial and parallel platforms.
- Adaptive time stepping controlling strategies will be also studied.
- The objective is then to demonstrate the applicability of such strategies to large-scale parallel computations of the simulation of polydisperse mixtures typically found in the geological processes.
COMPUTATIONAL ALGORITHMS: BIG DATA MANAGEMENT

- Big Data management and analysis of numerical simulations will be explored by the use of three systems
  - **SimDB**
    - Designed to manage spatial-temporal time series predictive data from numerical simulations, represented as a multidimensional array.
  - **UpsilonDB**
    - Early stage prototype aiming at managing the uncertainty on numerical simulation data, integrated with a probabilistic database System
    - Simulation post-processing analysis is also supported.
  - **Chiron**
    - Scientific workflow management system focused on managing scientific data flow with provenance data support
    - Data analytic at runtime is also supported, allowing for dynamic configuration fine-tuning, including uncertainty quantification data steering.
Energy Fields Approached
HPC4E: Wind farm design and optimization

- Design the best possible spatial design for onshore and offshore wind farms.
- Predict energy throughputs with local weather forecasts.
- Enable using larger turbines.
WIND ENERGY

○ The use of computational fluid dynamics (CFD) large-eddy simulation (LES) models are used to analyze atmospheric flow in a wind farm capturing turbine wakes and array effects
  ○ Wind resource assessment and wind farm design (analysis)
  ○ Short-term prediction for wind farm dispatch to the electricity network (forecast)

○ Both problems will study key aspects concerning microscale modelling simulations, as standalone CFD models or in connection with mesoscale models, by developing dynamical and statistical downscaling strategies
WIND ENERGY: SOME RESULTS ON COUPLING WRF TO LES

- The Weather Research and Forecasting Model (WRF) is a tool for multiscale atmospheric simulations that can be coupled to other methods.
- Turbulence-resolving Large-Eddy Simulations (LES) have been executed with WRF.
- An experiment has been performed in a cluster composed of Blade nodes Dual Xeon quad-core 3.0 GHz (2 GB per core) and clearly demonstrates how the solutions proposed as part of HPC4E are needed in order to better achieve a performance capable of exploiting the coming Exascale supercomputers.

<table>
<thead>
<tr>
<th>Number of processors</th>
<th>8</th>
<th>24</th>
<th>48</th>
<th>96</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of the simulation obtained</td>
<td>00:07:41</td>
<td>00:25:50</td>
<td>00:35:29</td>
<td>00:45:40</td>
</tr>
</tbody>
</table>
HPC4E: Biofuels

- Describe the performance of biomass fuel blends on practical combustion systems.
- Improve the design of combustion engines.
COMBUSTION FOR BIOMASS-DERIVED FUELS

- The objective is to develop a validated, predictive, multi-scale, combustion modeling capability to optimize the design and operation of evolving fuels.
- The effects of fuel variability on energy utilization of biomass derived gaseous fuels will be studied.
  - Generation of chemical kinetic mechanisms for biomass derived fuels.
  - Integration of the schemes into a CFD code.
  - Creation of efficient algorithms for data exchange that can run efficiently on HPC platforms.
COMBUSTION: SOME RESULTS ON SMALL CONFINED CHAMBERS

- Direct Numerical Simulation (DNS) of a reactive fuel mixture with Arrhenius kinetics is carried out in a classical configuration known as Hele-Shaw cell.
- The stability of the solution depends on:
  - The ratio of the thermal to molecular diffusivity of the fuel, through the Lewis number.
  - The thermal expansion, through the heat release parameter.
  - The buoyant convection term.
  - The heat losses through the walls.
- Biomass-derived gaseous fuel combustion in portable reformers suffers for the same instabilities, depending on the characteristic parameters of the mixture.
- A complete parametric study requires the use of the HPC techniques proposed above.
HPC4E: Hydrocarbon Exploration

- Exascale-level computational kernels.
- Exploration risk reduction through uncertainty quantification.
- Industry-driven benchmarks for geophysical imaging.
HYDROCARBON RESERVOIRS (OIL)

- Full wave-form modelling and inversion of seismic and electromagnetic data will be pursued
- Numerical simulation codes capable of delivering the clearest possible picture of the subsurface will be designed
  - Development and optimization of high-order finite-element schemes for 3D elastodynamics
  - Development and optimization of classical extrapolation schemes in 3D
  - Uncertainty estimation of petrophysical quantities
  - Synthetic benchmarking of exascale geophysical problems
  - Industry validation.
HYDROCARBON RESERVOIRS: SOME RESULTS ON UNCERTAINTY QUANTIFICATION IN SEISMIC IMAGING

- The Parallel Scientific Workflow Management System Chiron and the Reverse Time Migration (RTM) algorithm are used.
- The scalability of this system is evaluated in terms of the number of cores and the stochastic collocation (SC) interpolation level (i.e., the number of solver invocations or tasks).

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>Number of cores</th>
<th>Interpolation level</th>
<th>SC points</th>
<th>Workflow elapsed time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>1</td>
<td>17</td>
<td>47.18</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>2</td>
<td>145</td>
<td>305.15</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>3</td>
<td>849</td>
<td>1758.70</td>
</tr>
</tbody>
</table>
General Impact and Conclusions
**IMPACT - GENERAL**

- Vast improvement in simulation efficiency in terms of Watts needed per execution and reduced time-to-solution.
  - This will be applied to critical aspects of the energy value chain, with rapid deployment in the partner’s current production systems

- Establishing transnational “numerical laboratories”, which are cheaper, safer and faster than real-life experiments
CONCLUSIONS

- The HPC4E is a two-year project and has recently started in Dec 2015
- Several scientific codes are being improved in order to get them efficiently running on current and future supercomputers
- Its impact goes beyond the three selected energy sources as the computing solutions proposed will be able to be applied to other scientific and engineering domains
Solutions Envisioned

- **HPC**: Efficient use of the future 100 Petaflops and Exaflop systems.
- **WIND ENERGY**: Wind farm design and short-term micro-scale wind simulations to forecast the daily power production and reduce CO$_2$ targets.
- **BIOMASS**: Predict the performance of different biomass-derived fuels in practical systems.
- **GEOPHYSICS**: Obtain detailed images of deep hydrocarbon reservoirs.
Thank you for your attention.

The research leading to these results has received funding from the European Union’s Horizon 2020 Programme (2014-2020) and from Brazilian Ministry of Science, Technology and Innovation through Rede Nacional de Pesquisa (RNP) under the HPC4E Project (www.hpc4e.eu), grant agreement n° 689772.