Enhancing Energy Production with Exascale HPC Methods

José J. Camata[1], José M. Cela[2], Danilo Costa[1], Alvaro LGA Coutinho[1], Daniel Fernández-Galisteo[3], Carmen Jiménez[3], Vadim Kourdioumov[3], Marta Mattoso[1], Rafael Mayo-García[3], Thomas Miras[1], José A. Moríñigo[3], Jorge Navarro[3], Philippe O.A. Navaux[4], Daniel de Oliveira[5], Manuel Rodríguez-Pascual[3], Vítor Silva[1], Renan Souza[1], Patrick Valduriez[6]

(On behalf of the project)

[1] COPPE / Federal University of Rio de Janeiro,
Rio de Janeiro, Brazil

[2] Barcelona Supercomputing Center-Centro Nacional de Supercomputación,
Barcelona, Spain

[3] Centro de Investigaciones Energéticas Medioambientales y Tecnológicas,
Madrid, Spain

[4] Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

[5] Fluminense Federal University, Niterói, Brazil

[6] Zenith Team, Inria and LIRMM, Montpellier, France

Abstract. High Performance Computing (HPC) resources have become the key actor for achieving more ambitious challenges in many disciplines. In this step beyond, an explosion on the available parallelism and the use of special purpose processors are crucial. With such a goal, the HPC4E project applies new exascale HPC techniques to energy industry simulations, customizing them if necessary, and going beyond the state-of-the-art in the required HPC exascale simulations for different energy sources. In this paper, a general overview of these methods is presented as well as some specific preliminary results.

1 Introduction

New energy sources, if untapped, might become crucial in the mid-term. Intensive numerical simulations and prototyping are needed to assess their real value and improve their throughput. The impact of exascale HPC and data intensive algorithms in the energy industry is well established in the U.S. Department of Energy document "Synergistic Challenges in Data-Intensive Science and Exascale Computing" [1], for example.

The High Performance Computing for Energy (HPC4E) project, whose kick-off was in February, 2016, aims to provide these new exascale HPC and data intensive algorithms to three energy sources: wind energy, biomass, and oil. To do so, several European and Brazilian institutions are closely working, fostering in this way a collaboration that can be extended to other countries in Latin America as the computational solutions provided will be useful to many scientific and industrial fields as well as to other software and middleware developers.

Moreover, the current collaboration embraces both public and private companies who are leaders in their respective sector. The industrial sector is formed by REPSOL, Iberdrola Renovables Energía, TOTAL, and PETROBAS. Regarding the public sector, Brazilian (COPPE/UFRJ, LNCC, ITA, UFRGS, UFPE) and European (BSC-CNS, INRIA, UNLAC, CIEMAT) institutions belonging to the HPC and the energy fields are represented.

As previously stated, the main objective of HPC4E is to develop beyond-the-state-of-the-art high performance simulation tools that can help the energy industry to respond future energy demands and also to carbon-related environmental issues using the state-of-the-art HPC systems. This Brazilian-European collaboration also aims at improving the usage of energy using HPC tools by acting at many levels of the energy chain for different energy sources:

- Exploitation: In wind energy (respond to demand peaks, output prediction)
- Efficiency: In biomass-derived fuels (develop more efficient and renewable fuels, reduce green-house gas emissions, reduce hydrocarbon dependency and fuel cost)
- Exploration: In wind energy (resource assessment) and in hydrocarbons (improve available reserves, explore with less financial and environmental risk).

Another main objective is to improve the cooperation between energy industries, guaranteeing that the Technology Readiness Levels (TRL) of the particular project technologies will be very high.

2   A New Architecture for the exascale challenge

The project will set up a disruptive exascale computer architecture 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. In order to properly test this infrastructure, porting, tuning, and testing efforts of the different simulations codes will be previously carried out.
Specifically, four lines of action will be pursued.

The selected computing kernels of the codes coming from the energy sector will be optimized for architectures based on accelerators. The goal will be 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. To guarantee the maximum portability of the codes we will use programing models and tools like openCL, ompSs and BOAST.

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, AMD and SGI, but also new platforms based on ARM processors will be analyzed. Thus, the key point will be the load balancing and data placement, takin into account new scheduling algorithms able to improve locality too.

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 multiphysic problems; parallel I/O optimization; design of efficient check-pointing strategies; and, fault tolerance strategies at MPI level.

Last, 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. Also 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.

2.1 Some results: integrating checkpointing techniques into Slurm

Slurm is a workload manager designed for Linux clusters of all sizes. Among the plethora of available LRMSs, SLURM is currently one of the most advanced ones, and the most commonly used by the Top500 supercomputers. Thus, any improvement on its behavior will deeply impact the HPC community.

DMTCP (Distributed Multi Threaded Checkpointing) transparently checkpoints a single-host or distributed computation in user-space. It supports the commonly used OFED API for InfiniBand, as well as its integration with various implementations of MPI. It is specially suitable to be used in environments with legacy applications, as it does not oblige to modify their code to include checkpoint directives (as application-level checkpoint libraries) or to re-compile them (as most other system-level libraries). Instead, it acts as an additional software layer between the application and the Operating System, being able to save its status and restore it in the same -or on a different- computational resource.

Both Slurm and DMTCP are Open Source solutions. This allows examining their insights and perform enhancement required for this work.

SLURM is designed to be easily extended with third party plugins. In particular, it already counts with plugins to support checkpoint libraries and perform some basic operations like checkpoint and restart. The performed work consists on the development a driver serving as a bridge between SLURM and DMTCP APIs. This way Slurm scheduling algorithms and user-level commands will be able to transparently perform these operations.

As the number of cores composing HPC resources grows, applications are being required to run with an increasing degree of parallelism. However, the Mean Time To Failures (MTTF) of the hardware components (from coolers to memories or random issues) does not grow as fast as the number of resources, so the probability of one or more tasks being affected by a failure increases. In addition, the higher number of tasks composing a job, the higher computational and economic loss.

Thus, in order to achieve the objectives of this project, it is mandatory to increase the robustness and fault tolerance of the developed software. With the proposed software stack based on Slurm + DMTCP plus their coupling any job can be checkpointed and restarted transparently to the user and the job itself.

3   Algorithms and Solvers for Exascale Computations

Innovative computational algorithms well suited to the numerical simulation of complex phenomena on exascale architectures will be introduced. These refer to numerical schemes for Partial Differential Equations (PDE), sparse linear solvers, adaptivity, and data management.

Particularly the first topic is concerned with scalable implementations of high order schemes for wave propagation models. The second topic will develop and demonstrate the benefits of generic (i.e. algebraic) parallel solvers for large sparse linear systems of equations. The third topic addresses mesh and (local) time-step adaptive algorithms in order to optimize the use of computational resources. The fourth topic focuses on leveraging techniques to support simulation data management as required by the pre- and pos-processing steps involved in highly complex simulations.

Regarding scalable 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. First, we will implement and demonstrate the benefits of a recently designed class of high order multiscale methods. The common core approximation framework is the Multiscale Hybrid-Mixed (MHM) methods combined with Discontinuous Galerkin (DG) or Stabilized Continuous Galerkin (SCG). For frequency-domain problems, we will perform the same analysis with the so called hybridized DG formulations that drastically reduce the number of globally coupled degrees of freedom. Both types of solvers (i.e. time-domain and frequency-domain) are linked to the simulation/inversion framework for subsurface imaging proposed in the geophysics domain.

With respect to scalable sparse linear solvers, the goal is to provide state of the art parallel solvers for spare linear systems of equations or numerical schemes adopted in the simulation software associated to the applications coming from the energy field. Both direct and hybrid direct/iterative solvers will be considered. Regarding the former solver type, the PaStiX software [2] will be adopted. It is based on a supernodal approach and has been implemented on top of various runtime systems enabling an efficient use of heterogeneous manycore platforms. The hybrid iterative/direct strategy will be made available through the MaPHyS software [3] that implements algebraic domain decomposition ideas and relies in parallel on parallel sparse direct solvers such as PaStiX for each subproblem. On top of those two solvers, Krylov subspace methods are implemented either for the iterative refinement steps of PaStiX or to solve the reduced Schur complement system in MaPHyS.

Finally, the high performance dense linear algebra kernels on which those solvers are relying will also be part of the project but no specific action is foreseen on it. This software stack will be made available through a coherent and flexible API where the matrices can be provided by the application in various format such as centralized or distributed, assembled or unassembled.

Optimal numerical schemes for PDEs involve adapting the grids in space and time to minimize errors in the simulation. The activity on this topic will then be to explore libraries to support adaptivity such as the libMesh library [4]. libMesh provides a framework for the numerical simulation of partial differential equations using arbitrary unstructured discretizations on serial and parallel platforms. Adaptive time stepping controlling strategies will be also studied. The objective here is to demonstrate the applicability of such strategies to large-scale parallel computations of the simulation of polydisperse mixtures typically found in the geological processes.

Big Data management and analysis of numerical simulations will be explored by the use of three systems: SimDB, UpsilonDB and Chiron. The first is being designed to manage spatial-temporal time series predictive data from numerical simulations, represented as a multidimensional array. The second system, UpsilonDB, is currently an early stage prototype aiming at managing the uncertainty on numerical simulation data, integrated with a probabilistic database system, MayBMS. UpsilonDB supports simulation post-processing analysis. Chiron is a scientific workflow management system focused on managing scientific dataflow with provenance data support. Chiron strong support in data analytics at runtime, allowing for dynamic configuration fine-tuning, including uncertainty quantification data steering.

4 Atmosphere for Energy

The fundamental knowledge barriers to further progress in wind energy are defined as scientists' understanding of atmospheric flows, unsteady aerodynamics and stall, turbine dynamics and stability, and turbine wake flows and related array effects. The use of computational fluid dynamics (CFD) large-eddy simulation (LES) models to analyze atmospheric flow in a wind farm capturing turbine wakes and array effects requires exascale HPC systems.

In this way, microscale atmospheric models are based on CFD solvers adapted to simulate the Atmospheric Boundary Layer (ABL) in order to approach two fundamental wind energy problems: analysis (mainly focused on wind resource assessment and wind farm design) and forecast (mainly focused on short-term prediction for wind farm dispatch to the electricity network). 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. All models, methods and techniques developed for analysis will be tuned to produce short-term online forecasts of the wind farms output.

The objective is then to have the CFD models ready to exascale systems in order to overcome the present limitations and increase the accuracy on the evaluation of technical and economic feasibility of wind farms.

Regarding dynamical downscaling in order to assess wind resource, CFD models must account for the coupled effects of complex terrain, Coriolis forces, thermal stability, presence of forests, and wind turbines. Modifications need to be made to the RANS/LES CFD models including turbulent closures for ABL. The objective is to characterize the accuracy of the different ABL-CFD models, the numerical stability (robustness) and the convergence behavior of their HPC implementation.

The statistical downscaling approach will be complementary to the dynamical downscaling. Statistical downscaling models will be developed using local observations and large scale circulation and wind fields in the wind farm region.

A compilation of the available data and an evaluation of the quality issues that might affect the succeeding analyses will be accomplished, for this purpose is necessary to identify an appropriate metrics that account for the deviations in the wind power production predictions. Transfer functions between wind and wind power should be determined and serve as reference to translate the wind into wind power estimates for the rest of analyses.

On the other side, in order to efficiently plug wind farm power production to a distribution electricity network it is mandatory a forecast of the power production that allows the network operator to manage the electricity resources. In this sense, wind power short-term prediction within hourly to daily time scales is of fundamental importance. The models, methods and techniques developed for wind farm modelling will be tuned to produce short-term online forecasts of the wind farms output. Efficient use of HPC resources is critical to have these forecasts online. We will develop a forecast based on dynamical and statistical downscaling strategies.

4.1 Improving CFD microscale models

In order to achieve such an improvement, the following actions will be made:
- For RANS/LES models, study the turbulence closure models for ABL simulations as an alternative to the existing parametrizations.
- Implementation of a canopy model [5]
- Validate the HPC implementation using experimental data from the New European Wind Atlas project (ERA-Net)
- Wind farm modelling: Currently, farm models simulate downwind effects of rotors by extracting axial momentum at the turbines. In this simplistic approach, the rotor characteristics are incorporated trough velocity-independent drag coefficients obtained from tunnel experiments. This tasks aims at characterizing numerically the downwind effects of rotors by solving complex turbulent rotating flows.

4.2 Dynamical Downscaling strategies

Boundary conditions for solving wind flow on microscale domains are typically assumed steady and homogeneous over the computational inflow. These limitations can be overcome by dynamically coupling microscale CFD models with mesoscale

simulations furnishing initial and time-dependent boundary conditions at the computational boundaries. The following will be made:

  - − Blending between mesoscale (WRF) and CFD computational meshes to have consistent terrain information (topography and roughness) at the computational margins. The topography will be interpolated in such a way that it will be coincident with the CFD resolution in the inner zone and coincident with the mesoscale (WRF) over the boundaries;
  - − Initial condition and time-dependent CFD boundary conditions consistent with the mesoscale outputs that will drive the CFD model through boundary conditions;
  - − Study the use of nudging strategies based on introducing a force term over the momentum equation in the CFD model close to the boundary in order to enhance consistency between models near the boundaries;
  - − Validate the methodologies for different site conditions, onshore and offshore, considering benchmark validation cases from other projects such as IEA-Task 31 [6] Wakebench and FP7-NEWA [7].

4.3 Some results: WRF model coupled to Large-Eddy Simulations

The Weather Research and Forecasting Model (WRF) is a tool for multiscale atmospheric simulations that can be coupled to other methods, such as CFDs. In this case, turbulence-resolving Large-Eddy Simulations (LES) have been executed with WRF in real cases with a mesoscale resolution (grid cell size ~10 km). Specifically, 6 different domains have been simulated with a 3D unit-cell composed of 121x121x95 inner points through nested simulations in which the mesoscale domains drive the LES domains.

In order to determine how the system computationally evolves, a fixed total wall-time of 2 h has been set. Thus, the WRF-LES calculi with the WRF MPI-based 3.6.1 version have been executed on an increasing number of processors in order to determine how much time of an atmospheric phenomenon can be simulated in such a period of 2 h. As can be seen in Table 2, moving from 8 to 24 processors a linear behavior is obtained as the total of simulated phenomenon is multiplied by a factor of 3, but the simulation looses this behavior as the number of processors increases, getting a time of simulation multiplied by ~6.4 when the number of processors has been increased by a factor of 12.

This 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 1. WRF-LES computational behavior.

Number of processors	8.	24	48	96
Time of the simulation obtained	00:07:41	00:25:50	00:35:29	00:45:40

5 Biomass for Energy

Another important challenge is to develop a validated, predictive, multi-scale, combustion modeling capability to optimize the design and operation of evolving fuels. The next exascale HPC systems will be able to run combustion simulations in parameter regimes relevant to industrial applications using alternative fuels, which is required to design efficient furnaces, engines, clean burning vehicles and power plants

Thus, in order to obtain a thorough understanding of the effects of fuel variability on energy utilization of biomass-derived gaseous fuels, a coupled approach which covers three distinct areas of development will be employed: generation of chemical kinetic mechanisms for biomass-derived fuels, integration of the schemes into a CFD code, and creation of efficient algorithms for data exchange that can run efficiently in HPC platforms.

The activities include analyzing the physical characteristics of bio-syngas flames, assessment of performance in practical systems and providing an optimized industrial guideline for biomass derived gaseous fuel compositions and performance. The application scenario corresponds to industrial devices of the energy sector: stationary gas turbines, furnaces and portable combustion devices.

Generation of detailed chemical schemes that reproduce accurately the oxidation of biomass-derived gaseous fuels will be developed and assessed via detailed chemical kinetics mechanisms that can predict the oxidation process and species formation with certain level of accuracy. Several well-established mechanisms (GRI 3.0, San Diego, Leeds, etc.) will be examined and compared for different fuel compositions and the accuracy at predicting reference species and radicals will be provided for different operating conditions. These mechanisms are the starting point for reduction or tabulation techniques.

Also, development of skeleton and reduced chemical schemes for biomass combustion for engine operating conditions will be carried out. The use of detailed chemical mechanisms imposes an important limitation for practical applications of turbulent combustion. These mechanisms include both slow and fast chemical reactions involving a large number of species leading to a highly costly numerical problem. Besides, the effects of turbulence and flow strain also contribute to a complex interaction between chemistry and fluid mechanics that has to be accurately reproduced by CFD codes. In order to reduce the stiffness of the chemical problem, skeleton and reduced mechanisms will be developed so it can be integrated into a multiphysics code. The reduction technique will be based on flame-generated manifolds (FGM) and quasy-steady state (QSS) approximation for operating conditions of interest.

For studying the combustion dynamics of laboratory flames and comparing it with available data an integration of the chemical schemes developed in the previous points and the corresponding validation using benchmarking cases will be performed. Several cases using experimental data of laboratory flames will be investigated and the effects of fuel variability on the flame dynamics will be investigated.

Regarding industrial applications of biomass fuels in practical systems, numerical simulation of an industrial engine burning biomass will be carried out. Different operating conditions and fuel compositions will be examined. The activities will be

focused on providing details of the system performance to develop industrial guidelines for the use of biomass derived gaseous fuels.

Last, applications of biomass-derived gaseous fuel combustion in portable reformers for hydrogen production will be explored as well. Limitations to the miniaturization of hydrogen production reformers are linked to the large surface to-volume ratio, which enhances heat losses through the walls. Additional measures, such as heat recirculation, catalytic combustion, reactant preheating etc., are needed to sustain their proper operation. The activities of this Task will consist in numerical studies of the stability of combustion in small size reformers, with the objective of improving the understanding of their operation and determining the stable and more efficient regimes as a function of the fuel composition. The reduced chemical mechanisms developed in the aforementioned paragraphs as well as the use of HPC shall prove essential to this large parametric numerical investigation.

5.1 Some results: dynamics of combustion regimes in small confined chambers

A thorough parametric study of the combustion regimes associated to the intrinsic instabilities of flames in confined chambers is addressed in the context of the small size reformers. Direct Numerical Simulation (DNS) of a reactive fuel mixture with Arrhenius kinetics is carried out in a classical configuration known as Hele-Shaw cell (i.e., two parallel plates separated by a narrow gap). 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.

6 Geophysics for Energy

The third energy sector is related to oil as an energy source. Huge computational requirements arise from full wave-form modelling and inversion of seismic and electromagnetic data. By taking into account the complete physics of waves in the subsurface, imaging tools are able to reveal information about the Earth's interior with unprecedented quality. Nevertheless, actual wave physics has a high cost in terms of computational intensity, which can only be matched by using the exascale HPC systems.

In this sense, the capacity for imaging accurately the Earth's subsurface, on land and below the sea floor is a challenging problem that has significant economic implications in terms of resource management, identification of new energy reservoirs and storage sites as well as their monitoring through time. As recoverable deposits of

petroleum become harder to find the costs of drilling and extraction increase accordingly. Thus the oil and gas industry needs more detailed imaging of underground geological structures in order to find the best representation of the subsurface in terms of which model sticks better to the data recorded during acquisition surveys. This involves research based on advanced methods combining mathematics, geophysics and scientific computing. Such multidisciplinary collaboration is essential to the design of numerical simulation codes capable of delivering the clearest possible picture of the subsurface.

The data types involved in geophysical imaging are mostly seismic (acoustic or elastic) and electromagnetic. Modern imaging techniques (RTM, FWI …) rely on intensive usage of full 3D physical modeling engines. Hence, in order to attain results in a reasonable time, these engines must use, as efficiently as possible, the fastest hardware architectures in a massively parallel way. On top of that, the larger and more complex the scenarios become, techniques which attain results with low computational complexity or few degrees of freedom become preferable.

The main goal is, then, attaining the sharpest possible images of the subsurface with the best possible quantitative content (i.e. parameter estimation, uncertainty analysis) in the shortest possible time. In addition, in the exascale era, power efficiency is becoming ever more a crucial factor in establishing the usability of HPC in industrial applications. The main developments will involve geophysical inversion of elastic and electromagnetic waves using high-order structured and unstructured computational grid types. Test-driven code development will allow us to put a special focus on having detailed comparisons and benchmarks between all possible approaches in the most realistic scenarios. Architecture-oriented programming optimizations are expected to play a crucial role in establishing cost/accuracy/complexity relationships which will help delineate the future directions of geophysical imaging in the exascale era.

Specific sub-objectives will be: 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; and, Industry validation.

6.1 Some results: Uncertainty Quantification in Seismic Imaging using Chiron SWfMS

A computational simulation may be defined as the execution of complex programs chained in a coherent flow. This flow of simulation programs may be modeled as a scientific workflow [9]. A scientific workflow is an abstraction that allows for scientists to specify a set of activities and a data flow between them [9]. Each activity is associated to a simulation program, which is responsible for the consumption of a set of input data and the production of another set of output data. Many of the experiments modeled as scientific workflows have to process a large volume of data, thus requiring the effective usage of HPC environments allied to parallelization techniques, such as data parallelism or parameter sweep [10]. In order to support the modeling and execution of scientific workflows in HPC environments, parallel Scientific Workflow Management Systems (SWfMS) [9] were developed, such as

Chiron [11]. To foster the workflow parallelism, the activities of workflows can be subdivided in smaller tasks, known as activations [11].

Although these SWfMS execute on large amounts of HPC resources, as the workflow becomes increasingly complex, they tend to execute for weeks or even several months depending on the amount of input data and the availability of computational resources. This way, it is fundamental for scientists of any domain to be aware of the execution status in order to analyze if the current execution complies with some pre-defined quality and performance criteria. Based on this information steering [12], scientists may decide if they have to interfere in the execution (also known as dynamic workflows) or pause the workflow execution to change parameter values or even change the original workflow specification. These facilities allow for scientists to perform debugging, partial results analysis or to identify failures as early as they happen as online analysis of scientific workflows. This paper presents a real workflow modeled for Uncertainty Quantification (UQ) in seismic imaging using Chiron.

Reverse Time Migration (RTM) is a standard algorithm for producing accurate images of subsurfaces that help to improve the decision process in the Oil and Gas Industry. However, the use of this algorithm becomes challenging with high dimensional uncertain input due to computational costs. In this context, we present in this study a workflow using Chiron adapted for Uncertainty Quantification (UQ) in seismic imaging. Taking advantage of an optimized RTM code previously developed [13], we consider a framework which allows the uncertainty quantification on the output of a large scale computation managed by Chiron. This approach is non-intrusive, since we use the deterministic simulation of our optimized RTM code, and couple dimension reduction with sparse grid stochastic collocation.

6.1 1. Chiron SWfMS for managing provenance in UQ domain

Chiron [14] is a parallel SWfMS that aims at supporting users to model, execute, and monitor scientific workflows in cluster environments. With this purpose, Chiron uses workflow algebra SciWfA [12] to rule all workflow activities and represents consumed and produced data as relations. This workflow algebra uses a set of operators (Map, Filter, Reduce, SplitMap, SRQuery, and MRQuery) to be associated with workflow activities. Workflow activities (i.e., invocation of scientific programs) are considered operands, while consumed and produced data are also operands.
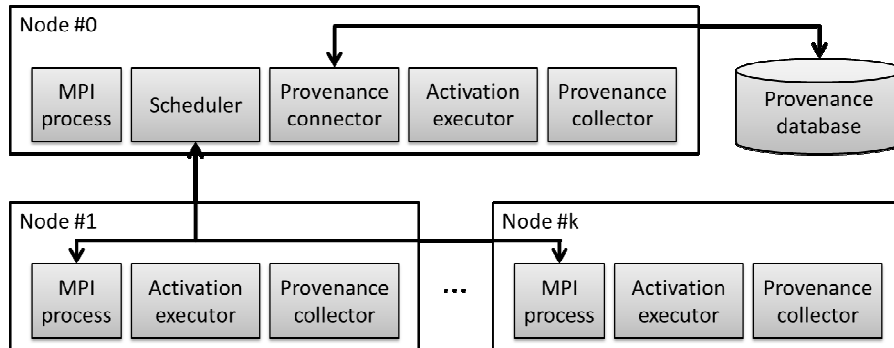
Fig. 1. Architecture for the Chiron's engine.

Therefore, each activity is associated to an operator, which is composed of input (data consumption) and output (data production) relations. Following this workflow algebra, Chiron is able to manage the dataflow generation at the physical (i.e., file flow) and logical (i.e., data element flow) levels from parallel execution of scientific workflows, as presented in [15]. Chiron captures and stores provenance and performance data at runtime on a database. Then, users can run queries in provenance database to analyze domain-specific data and performance execution data at runtime. Chiron presents a provenance data model that follows W3C PROV recommendations, known as PROV-Df [15]. Chiron's parallel engine distributes activations to the available computing resources using a client-server architecture. Figure 1 presents the architecture for Chiron's engine. To enable the communication between available computing resources, Chiron employs Message Protocol Interface (MPI) for Java [16] to exchange messages between those computing resources. To access performance, we coupled the Tuning and Analysis Utilities (TAU) to Chiron. With this, we can collect and visualize the usual performance execution data (MPI operations, scientific programs) plus transactions in the provenance database.

6.1 2. Preliminary Results

Considering the application for UQ in seismic imaging, we modeled and executed a scientific workflow for this domain using Chiron SWfMS (Figure 2). As an experimental evaluation, this workflow was executed in the Endeavour cluster, an Intel Linux Cluster based on 51,000+ cores with Intel Xeon E5-2697v2 (Ivy Bridge) processors and an Intel Xeon Phi Coprocessor 7110 (MIC Architecture).
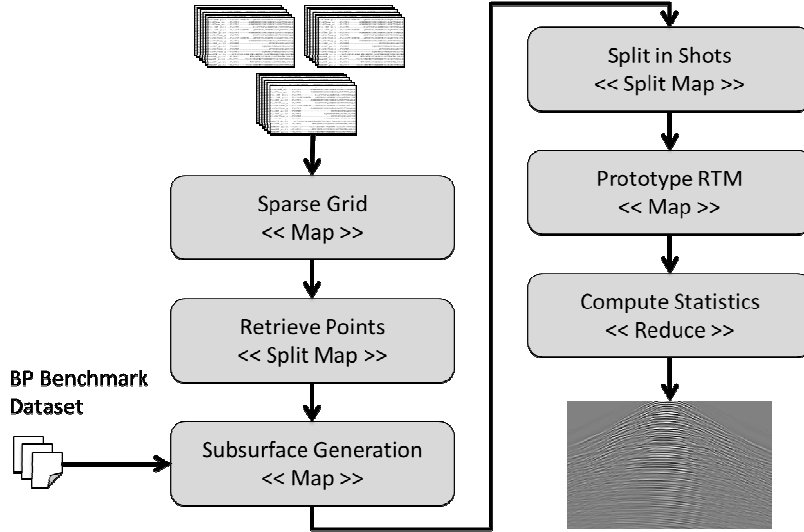
Fig. 2. Scientific workflow for UQ in seismic imaging using Chiron.

Then, as preliminary results, we obtained the workflow elapsed times presented in Table 2, when we vary the number of cores and the stochastic collocation interpolation level (consequently, the number of solver invocations or tasks) to evaluate the scalability of this workflow using Chiron. With those results, we observed that the sequential workflow elapsed time does not increase in the same proportion as we vary the number of stochastic collocation (SC) points. Further, TAU helped to identify potential performance bottlenecks associated to the parallel workflow execution. Thus, we observed that Chiron presents a non-negligible overhead due to the MPI communication and the query processing using a centralized DBMS. For this reason, we are improving Chiron to support provenance data management using a decentralized DBMS or modifying the query execution plan of some provenance transactions.

Table 2. Experimental results for executing the UQ in seismic imaging workflow.

Number of nodes	Number of cores	Interpolation level	SC points	Workflow elapsed time (minutes)
1	16	1	17	47.18
2	32	2	145	305.15
4	64	3	849	1758.70

References

1. Synergistic Challenges in Data-Intensive Science and Exascale Computing, DOE ASCAC Data Subcommittee Report, March 2013
2. The PaStiX software, available at http://pastix.gforge.inria.fr
3. The MaPHyS software, available at http://maphys.gforge.inria.fr
4. The libMesh library, available at http://libmesh.github.io/
5. A. Sogachev, M. Kelly, and M.Y. Leclerc, "Consistent Two-Equation Closure Modelling for Atmospheric Research: Buoyancy and Vegetation Implementations". Boundary-Layer Meteorology, vol 145 (2), 2012, pp..307–327.
6. The IEA-Task 31 Wakebench, http://www.ieawind.org/task_31.html
7. The NEWA project, http://euwindatlas.eu/
8. G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
9. E. Deelman, D. Gannon, M. Shields, and I. Taylor, "Workflows and e-Science: An overview of workflow system features and capabilities," Future Generation Computer Systems, vol. 25, no. 5, pp. 528–540, 2009.
10. E. Walker and C. Guiang, "Challenges in executing large parameter sweep studies across widely distributed computing environments," in Workshop on Challenges of large applications in distributed environments, Monterey, California, USA, 2007, pp. 11–18.
11. E. Ogasawara, J. Dias, D. Oliveira, F. Porto, P. Valduriez, and M. Mattoso, "An Algebraic Approach for Data-Centric Scientific Workflows," in Proceedings of the VLDB Endowment, 2011, vol. 4, pp. 1328–1339.
12. M. Mattoso, J. Dias, K. A. C. S. Ocaña, E. Ogasawara, F. Costa, F. Horta, V. Silva, and D. de Oliveira, "Dynamic steering of HPC scientific workflows: A survey," Future Generation Computer Systems, vol. 46, pp. 100–113, May 2015.
13. D. L. Costa, A. L. Coutinho, B. S. Silva, J. J. Silva, and L. Borges, "A trade-off analysis between high-order seismic RTM and computational performance tuning."
14. E. Ogasawara, J. Dias, V. Silva, F. Chirigati, D. Oliveira, F. Porto, P. Valduriez, and M. Mattoso, "Chiron: A Parallel Engine for Algebraic Scientific Workflows," Concurrency and Computation, vol. 25, no. 16, pp. 2327–2341, 2013.
15. V. Silva, D. de Oliveira, P. Valduriez, and M. Mattoso, "Analyzing related raw data files through dataflows," Concurrency and Computation: Practice and Experience, vol. 28, no. 8, pp. 2528–2545, 2016.
16. B. Carpenter, V. Getov, G. Judd, A. Skjellum, and G. Fox, "MPJ: MPI-like message passing for Java," Concurrency: Practice and Experience, vol. 12, no. 11, pp. 1019–1038, 2000