Today, our society is facing massive challenges: Climatechange and the biodiversity crisis are forcing a switch to a decarbonised and, more generally, circular and sustainable economy; growing threats for public health due to aging populations, pandemics and pollution; the need to adapt to geopolitical and economic shifts. Solutions have many faces: Switching to renewable forms of energy, developing novel materials reducing dependency on scarce resources, understanding and mitigating climate change, fast development of new drugs, personalized medicine, establishing more efficient forms of transport, under-standing and anticipating social and economic phenomena – the list could go on and on. At the same time, the European Union has clearly identified the strategic importance of reducing its dependency on the import of resources (both technological and energy resources) from other parts of the world. Therefore, an important long-term strategy for the European Union is to use its technological potential to gain more independence and create well-being for its citizens through the solution of the key societal challenges. The solution of many of the core technical challenges can be facilitated by the use of advanced computing: Starting from the small cases that can be run on a notebook or a workstation to extreme-scale computational analysis of the largest (entire climate system) to the tiniest (atomic structure of materials) structures, ranging from technical structures (like buildings, turbines, or airplanes) to geological and earth-scale (volcanoes, climate system) to societal (migration, economics) and biological (virology, biomechanics) systems. This involves many computational techniques, ranging from physics-based simulations to analysis of massive amounts of data, and increasingly a mix of these and other approaches.

The results enriched by the project aims at supporting community building around codes, provide trainings for users, and do a great job of putting the pieces together, transforming bunches of – at times - research-oriented codes used by a few highly specialised scientists on disparate platforms into something that runs efficiently on the most powerful computers, is easy to use, accesses data in an efficient way and bundles entire workflows, such that the whole technology is usable by a far broader community, both in industry and academia.

Meteorology
for Energy

DESCRIPTION

The increasing contribution of variable weather dependent renewable energy sources (RES) to the electricity grids of European countries challenges the existing energy system in many ways. Power grid and plant management as well as power stock exchange trading are the two major areas affected by insufficiently predicted wind and solar power.

Weather forecasts play a huge role in decision-making, and the most costly situations for all stakeholders are those extreme weather events missed by conventional forecasts, like stormy winds producing excess wind power or unexpected fog blocking all photovoltaic feed-in. Even probabilistic modern forecasts that simulate tens of possible scenarios can miss these events entirely.

Ultra-large ensemble sizes of O (1000) model runs can address the challenge of capturing all events using continuous probabilistic short-term forecasts, yielding probability density functions (pdfs) for wind and clouds respectively. At sufficient (1km) resolution, this ultimately requires exascale computing capability, which in turn means addressing a series of technical challenges, particularly in the areas of ensemble modelling, programming models, and big data analytics.

The flagship framework for accommodating these innovations will be the Ensemble for Stochastic Interpolation of Atmospheric Simulations (ESIAS), which will initially include the Weather Research and Forecast (WRF) model adopted to predict winds at rotor hub heights and cloud optical thickness (COT). The EURopean Air pollution Dispersion-Inverse Model (EURAD-IM) will further address the impact of aerosol-induced turbidity on solar power production. These tools will provide the meteorological data needed for wind and solar day-ahead power forecasting, as well as for short-term forecasting in confluence with satellite-based cloud-motion solvers.

RESULTS

Large-Scale Optimization of the Weather Research Forecasting Model

In order to capture all possible scenarios, one must properly model the underlying meteorology. Numerical weather forecasts consist of many underlying physical schemes and parameterizations based e.g. on typical conditions in the region of the model provider. Problems in the physical modelling can express themselves through an overall bias in the wind or COT output or in unrepresentative differences between scenarios due to the various models' sensitivities to perturbations of the initial condition. While post-processing techniques will improve the distribution in Meteo for Energy, the quality of the input is decisive.

In EoCoE II, we have conducted large-scale tests of an unprecedented 672 combination of WRF schemes for 12 simulated days, summarized for wind and solar in figure 1. We have verified the wind data against 1172 ground stations at the two-meter level, and the cloud coverage against satellite data [83].

For the physical scheme, PBL physics most affected the hourly wind speed, while the Mellor-Yamada-Nakanishi-Niino scheme (MYNN2 and MYNN3) and Asymmetric Convective Model (ACM2) performed best. For the microphysical scheme, CAM5.1, WSM3/5/6 and Goddard best detailed the cloud coverage (radiation schemes not shown here). CAM5.1 is the most accurate, but takes longer to simulate than the others. Among the faster schemes, Goddard appeared to be better in the later spring and early autumn. Finally, the cumulative schemes Tiedtke, Grell-3, and Simplied Arakawa-Schubert all achieved good scores with above-mentioned microphysics.

These results should be broadly useful to meteorological research in Europe using WRF, while in Meteo for Energy making the ultra-large ensembles of wind and photovoltaic power forecasts comparable to actual meter data. Large-Scale Optimization of the Weather Research Forecasting Model

In order to capture all possible scenarios, one must properly model the underlying meteorology. Numerical weather forecasts consist of many underlying physical schemes and parameterizations based e.g. on typical conditions in the region of the model provider. Problems in the physical modelling can express themselves through an overall bias in the wind or COT output or in unrepresentative differences between scenarios due to the various models' sensitivities to perturbations of the initial condition. While post-processing techniques will improve the distribution in Meteo for Energy, the quality of the input is decisive. In EoCoE II, we have conducted large-scale tests of an unprecedented 672 combination of WRF schemes for 12 simulated days, summarized for wind and solar in figure 1. We have verified the wind data against 1172 ground stations at the two-meter level, and the cloud coverage against satellite data [83]. For the physical scheme, PBL physics most affected the hourly wind speed, while the Mellor-Yamada-Nakanishi-Niino scheme (MYNN2 and MYNN3) and Asymmetric Convective Model (ACM2) performed best. For the microphysical scheme, CAM5.1, WSM3/5/6 and Goddard best detailed the cloud coverage (radiation schemes not shown here). CAM5.1 is the most accurate, but takes longer to simulate than the others. Among the faster schemes, Goddard appeared to be better in the later spring and early autumn. Finally, the cumulative schemes Tiedtke, Grell-3, and Simplied Arakawa-Schubert all achieved good scores with above-mentioned microphysics. These results should be broadly useful to meteorological research in Europe using WRF, while in Meteo for Energy making the ultra-large ensembles of wind and photovoltaic power forecasts comparable to actual meter data.

Post-Processing Solar Irradiation for PV

While a proper selection of schemes in WRF will provide a better basis, evee operational numerical weather predictions (NWP) can have very biased downward solar irradiation. It is both very difficult to model and has not had the attention of more broadly important variables in weather forecasting like wind, temperature, or precipitation. We have developed an effective calibration scheme for the irradiation data used in the solar power conversion model. Rather than correcting for bias only at the end of the forecast chain, comparing solar power calculations to German PV data, we use local satellite irradiance measurements to calibrate the solar input. This bias correction is possible local to each pixel of the high-resolution satellite images, identifying differences due e.g. to local topography not resolved by the weather model. While a correction of the aggregated German power works well on average, this correction does not differentiate local weather conditions across Germany, which we hope to capture with our O(1000) ensemble members. Figure 2 shows a calibration scaling based on one year, where each pixel is independently calibrated. The spatial correlation of the scaling indicates some robustness and apparent correlation between topography and NWP bias. We have so far been able to test this approach on operational day-ahead numerical weather forecasts, where we achieved a significant improvement in the root-mean-squared error between the raw power forecast and the German electric meter data from 3.5 to 3.2%.

CODES

Primary codes

ESIAS-Chem is a tool for generating and controlling ultra-large ensembles of chemistry transport models for stochastic integration, exploiting a two-level parallelism, combined with a particle filter data assimilation scheme.

ESIAS-Meteo is a tool for generating and controlling ultra-large ensembles of numerical weather forecast models for stochastic integration, exploiting a two-level parallelism, combined with a particle filter data assimilation scheme.

EURAD–IM (Current forecasts can be viewed here) simulates chemistry particle transportation in local atmospheres coupled with a weather forecast application WRF. An advection-diffusion-reaction equation, with multiple solvers for chemistry, is used.

ESIAS ensemble run part will be entirely refactored with proposed technologies in EoCoE-II.

Secondary Codes

The Wind Power Management System (WPMS), see Vogt et al., and the Solar Prediction System (SPS), Saint-Drenan et al. are empirical and physical wind and solar power models used at Fraunhofer IEE to calculate German power forecasts from meteorological forecasts and satellite measurements.

Scientific Leader: Garrett Good

garrett.good@iee.fraunhofer.de

Materials for Energy

DESCRIPTION

Advanced materials can contribute to the reduction in cost, increase in performance and extension of lifetime of the low-carbon energy technologies such as supercapacitors and solar cells. Thus, there is an urgent need for multi-functional and sustainable materials designed to provide a specific function in the final product. HPC can speed-up the entire process needed to identify new materials and to optimize them for the final use. In particular, the design of advanced materials needs to consider atomic-scale chemistry and how it affects the physical properties at larger scales close to the device itself. The Energy Materials objective in EoCoE-II will focus on three specific flagship applications in energy production respectively:

- libNEGF à neXGf (high efficiency silicon solar cells),
- Metalwalls (supercapacitor modelling),
- KMC-FMM/BoltMC (organic/perovskite photovoltaics)

In order to pursue this objective, the Materials for Energy Scientific challenge is divided in three main tasks:

Shedding light on carrier dynamics at hetero-interfaces in silicon solar cells

Amorphous-crystalline heterointerfaces play a crucial role in the photovoltaic operation of silicon heterojunction (SHJ) technology, but the microscopic mechanisms of transport and recombination mechanisms at the interface are still poorly understood. The purpose of the present task is to understand the transport mechanisms underlying photovoltaic devices based on SHJ technology by simulating at atomistic resolution amorphous crystalline heterointerfaces. Ab initio electronic properties of the c-Si/a-Si:H interfaces are calculated. Non-equilibrium Green's functions (NEGF) formalism is used to analyze the effect of interfaces on the carrier transport and dynamics in silicon solar cells.
libNEGF is a general library that calculates Equilibrium and Non Equilibrium Green's Function and related quantities in open systems, within an efficient sparse iterative scheme. By the end of the project the new code neXGf will be delivered.

Harvesting electricity from salinity or temperature gradient

This task focuses on optimizing capacitive blue energy electrodes. Blue energy corresponds to electric power production from salinity gradients. The idea is to harvest the free energy lost during the mixing of river with sea water in estuaries. The main technologies developed for this purpose to date exploit the electric potential differences applied across membranes. We analyze an alternative approach based on capacitive mixing, which involves carbon electrodes.
Metalwalls is a classical molecular dynamics code aiming at simulating electrochemical cells. It treats electrode as metallic systems held at constant potential and includes polarization effects for the liquids.

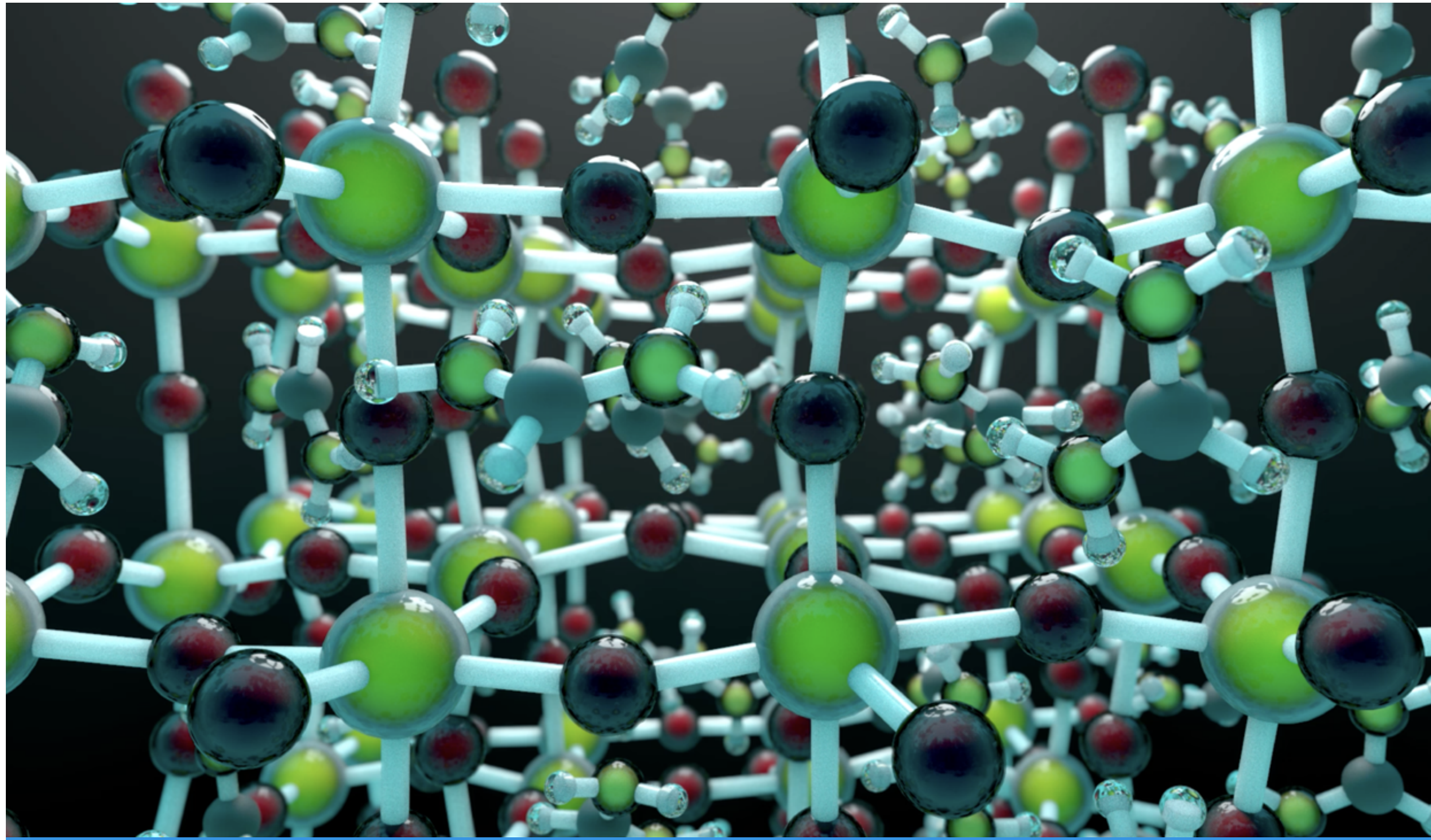
Organic and Perovskite solar cells

This task deals with the development of a flexible and modular scheme for the multiscale modelling of electronic and ionic transport in materials for next generation photovoltaic devices. The scientific goals of this project are: 1) simulate organic photovoltaic cells of 10 nm size and study interfaces on the nm length scale to refine models of charge generation and recombination; 2) Understand the complex processes of charge transport in a perovskite solar cell thanks to the implementation of a semiclassical approach based on solving the Boltzmann transport equation in submicron inorganic semiconductors.
KMC-FMM/BoltMC KMC-FMM simulates charge and energy transport in organic solar cells. BoltMC simulates charge transport in perovskite cells. The BoltMC code includes the effects of mobile ion motion that affects most perovskites. Both KMC-FMM and BoltMC codes are exascale flagship codes and are part of tasks in Scalable Solvers Technical Challenge of EoCoE-II

RESULTS

Four European institutions (CIEMAT, CNR, ENEA and FZJ) collaborated in understanding the transport mechanisms underlying photovoltaic devices based on the Silicon Heterojunction technology. The main purpose here is to design and develop cutting-edge computational methods and HPC-ready software for photovoltaic device simulations at atomistic resolution in the most insightful and predictive way. The results look very promising for further research since using libNEGF we have simulated the transport at the interface between passivated Silicon and amorphous Silicon by taking into account the presence of defects in a realistic structure, and increased the speed of specific algorithms used by the code more than six times. Further strategies include GPU acceleration and tensor products on GPU since it is clear that exascale machines will be heavily based on such devices.

In order to assess the best electrode/electrolyte combination to optimize the electricity production in electrochemical systems, collaborators from CEA and Maison de la Simulation use Metalwalls, a classical molecular dynamics code aiming at simulating electrochemical cells. It treats electrode as metallic systems held at constant potential and includes polarization effects for the liquids, which is coupled to MDFT, in which the solvent structure is simply calculated through the minimization of a density functional. The interactions are parameterized based on high-level quantum Monte-Carlo simulations. By testing a variety of carbon electrodes, the ones providing the largest capacitances will be selected and tested in future experiments.
In the research lead by the University of Bath (group of Prof. Alison Walker), we have been able to achieve good results on optoelectronic devices based on two classes of emerging photovoltaic materials, namely organic molecular semiconductors and metal-halide perovskite semiconductors. Both types of cells are cheap to make but have some stability concerns. Organic solar cells are only 18% efficient whereas perovskite solar cells are close to 26%efficiency. Algorithmic improvements and other optimization work on the KMC-FMM
(Kinetic Monte Carlo model with Fast Multipole Method) code enabled access to organic molecular semiconductors models that contain millions of charges with good parallel scaling. We demonstrated good performance scaling for 128 million charges up to 128 nodes on Intel and ARM architectures, making it possible to model doped organic semi-conductor containing 20,000 charges. Metal-halide perovskites are currently attracting intense interest for application in optoelectronic devices, due to their remarkable efficiencies in spite of low temperature synthesis methods. In order to test the influence of large polarons on the value of phenomenological parameters, we derived expressions for scattering of polarons by ionised impurities and acoustic phonons, and implemented them in the BoltMC simulator in order to compute and compare the temperature dependent mobilities of polarons and band electrons in three metal–halide perovskites. These studies are important for understanding the degradation mechanisms and for the design of solar cells based on metal-halide perovskites with enhanced stability in the performance of the device.



Scientific Challenge leader: Francesco Buonocore

francesco.buonocore@enea.it

WATER FOR ENERGY

DESCRIPTION

In hydropower and geothermal applications, modeling of shallow subsurface flow is of major importance in order to accurately simulate and predict the exchange of groundwater with streams under low-flow conditions, and the transport of energy. The major challenge is the representation of topographically driven groundwater convergence and streamflow generation, and of the geological heterogeneity across a number of space scales ranging from centimeters to thousands of kilometers in case of continental river systems. Constructing hydrological and geothermal models at this resolution over large spatial scales for scientific and operational applications constitutes a game changer, easily reaching up to 1012 degrees of freedom, where simulations must additionally assimilate observations to mitigate uncertainties in model data.

In EoCoE-II, the integration of hyper-resolved simulation of hydrological fluxes, routing along the river network, and management of storage reservoirs will be performed with a modernized version of ParFlow with adaptive mesh refinement (AMR) capability coupled with HYPERstreamHS model. The added values of these simulations will be shown by feeding ParFlow gridded runoff time series into the operational hydropower model embedded into HYPERstreamHS. The coupling will be specifically developed over the Italian Alpine region.

Previous work on geothermal reservoir characterization showed the successful application of optimal experimental design (OED) within the simulation code SHEMAT-Suite in order to identify optimal drilling locations for assessing uncertain reservoir parameters within a numerical reservoir model. However, the high computational cost has to date limited this approach to a numerical model with significantly reduced number of unknowns. Collaboration with experts in the EoCoE-II consortium will enable us to create a realistic geothermal reservoir model with vastly improved spatial resolution. Combining optimal experimental design for positioning boreholes with state-of-the art HPC techniques will improve the exploration and exploitation of geothermal reservoir systems, as it enables a sophisticated quantification of uncertainties in the subsurfac

RESULTS

Pan-European spatially and temporally consistent 3 km soil moisture reanalysis using land surface data assimilation

Soil moisture (SM) is a key state variable which controls the exchange of water, energy and carbon fluxes between the land surface and atmosphere (Seneviratne et al., 2010). As a result, it plays an important role in many regional-scale applications, including meteorology, hydrology, flood forecasting, drought monitoring, agriculture and climate change impact studies (Brocca et al., 2010). Because of its high spatiotemporal variability, it is difficult to monitor soil moisture at large spatial scales and remains the most difficult variable to obtain because there are no high-resolution soil moisture observations available at the continental scale. To produce pan-European spatially and temporally continuous information of soil moisture at 3 km resolution, the land surface data assimilation system CLM-PDAF consisting of the Community Land Model (CLM; Oleson et al., 2008) and the Parallel Data Assimilation Framework (PDAF; Nerger and Hiller, 2013) coupled CLM-PDAF framework (Kurtz et al., 2016) was used. We implemented CLM-PDAF to

assimilate the satellite based soil moisture dataset ESA CCI (the European Space Agency Climate Change Initiative; Wagner et al., 2012) into CLM, producing a high-resolution European SSM reanalysis (called ESSMRA hereafter) dataset.

This product overcome the shortcomings of sparse spatial and temporal distributions in observations and provides a better estimate of SM than obtained only by modeling or by sparse observations alone. To evaluate the quality of ESSMRA, we compared the daily SM from ESSMRA against in-situ data using the SM measurement acquired from the International Soil Moisture Network (ISMN; Dorigo et al., 2011) across Europe. The ESSMRA dataset shows overall good agreement with in-situ observations at daily time scale as shown in Figure 2 for two networks REMEDHUS and SMOSMANIA located in Spain and France, respectively. This comparison shows that at daily scale the model is able to reproduce the daily variations in soil moisture fairly well, with overall correlation above 0.60 and RMSE ranges 0.04 to 0.19. Overall, multiple validations revealed that ESSMRA data were consistent with the in-situ soil moisture data and other existing global reanalysis products. The relatively long time series and fine spatial resolution of this new European gridded ESSMRA dataset could provide a valuable data source for many hydrological applications. For example, it can be used as an initial input data for climate change analysis and for numerical weather prediction modelling to improve the model forecast in terms of location and amount of extreme precipitation events. This dataset will be also useful to understand the development and persistence of extreme weather events such as droughts, floods and heatwaves.

Experimental Design for Geothermal Modeling

Drilling boreholes during exploration and development of geothermal reservoirs not only involves high cost, but also bears significant risks of failure. In geothermal reservoir engineering, techniques of optimal experimental design (OED) can improve the decision making process. For instance, during exploration and production of geothermal fields, OED can be used to place additional slim holes, thus decreasing drilling costs and risk. Previous publications explained the formulation and implementation of this mathematical optimization problem and demonstrated its feasibility for finding borehole locations in two- and three-dimensional reservoir models that minimize the uncertainty of estimating hydraulic permeability of a model unit from temperature measurements (Seidler et al. 2014; Seidler et al. 2016). Subsequently, minimizing the uncertainty of the parameter estimation results in a more reliable parametrization of the reservoir simulation, improving the overall process in geothermal reservoir engineering. OED is a mathematical optimization method. The general approach is to find optimalexperimental conditions for constraining model parameters. In other words, OED gives answers to the question: How do I have to design an experiment in order to collect data from which I can predict model parameters with least uncertainty? This is a question of the sensitivity of the model with respect to the unknown parameters. This sensitivity is mathematically described by the Fisher Matrix. It contains the first order derivative of the model output towards the parameters. For evaluating the information contained in the Fisher matrix, OED criteria are formulated. One is the D-optimal design criterion which is based on the determinant of the Fisher Matrix (other criteria are based on the trace or eigenvalue). The minimum of the D-optimal criterion contains the maximum sensitivity. Various OED techniques are implemented in the Environment for Combining Optimization and Simulation Software (EFCOSS) (Seidler et al. 2014). This software framework links mathematical optimization software with SHEMAT-Suite, our geothermal simulation code for fluid flow and heat transport through porous media, for addressing problems arising from geothermal modeling. Within EoCoE-II we will extend this OED approach with further functionalities for geothermal reservoir modeling and aim at increasing its performance in order to apply it to reservoir scale production runs. As a first step, synthetic test models have been set up for studying certain aspects, such as sensitivity of the OED result to a priori data.

Synthetic OED study: 2D geothermal reservoir model above salt diapir[JF1]

Here we show first results of an OED study for a 2-dimensional synthetic model of a geothermal reservoir above a salt diapir (yellow unit), which is cross-cut by two high permeable faults (blue and red units) (modified after Rath et al. 2006). Each model unit is characterized by constant thermal and hydraulic properties (e.g., permeability, porosity, thermal conductivity). High heat conductivity of salt results in higher temperatures at shallower depth. This makes the sedimentary units above and close to salt diapirs interesting for direct use of geothermal energy. In addition, the permeable faults in our model provide pathways for advective heat transport, resulting in a heat transport towards the surface through the western fault (unit 11, blue). There is a borehole at location x0=7475 m with a certain depth z0=1475 m providing a temperature log Tlog0. The problem is to find the location x1 for an additional borehole of the same depth for measuring another temperature log Tlog1. These temperature data shall be used for estimating the fault permeability k10 and k11 with least uncertainty (i.e., optimally). The experimental condition is the horizontal position of the additional borehole. Overall, multiple validations revealed that ESSMRA data were consistent with the in-situ soil moisture data and other existing global reanalysis products. The relatively long time series and fine spatial resolution of this new European gridded ESSMRA dataset could provide a valuable data source for many hydrological applications. For example, it can be used as an initial input data for climate change analysis and for numerical weather prediction modelling to improve the model forecast in terms of location and amount of extreme precipitation events. This dataset will be also useful to understand the development and persistence of extreme weather events such as droughts, floods and heatwaves.

Experimental Design for Geothermal Modeling

Drilling boreholes during exploration and development of geothermal reservoirs not only involves high cost, but also bears significant risks of failure. In geothermal reservoir engineering, techniques of optimal experimental design (OED) can improve the decision making process. For instance, during exploration and production of geothermal fields, OED can be used to place additional slim holes, thus decreasing drilling costs and risk. Previous publications explained the formulation and implementation of this mathematical optimization problem and demonstrated its feasibility for finding borehole locations in two- and three-dimensional reservoir models that minimize the uncertainty of estimating hydraulic permeability of a model unit from temperature measurements (Seidler et al. 2014; Seidler et al. 2016). Subsequently, minimizing the uncertainty of the parameter estimation results in a more reliable parametrization of the reservoir simulation, improving the overall process in geothermal reservoir engineering. OED is a mathematical optimization method. The general approach is to find optimalexperimental conditions for constraining model parameters. The minimum of the D-optimal criterion contains the maximum sensitivity. Various OED techniques are implemented in the Environment for Combining Optimization and Simulation Software (EFCOSS) (Seidler et al. 2014). This software framework links mathematical optimization software with SHEMAT-Suite, our geothermal simulation code for fluid flow and heat transport through porous media, for addressing problems arising from geothermal modeling. Within EoCoE-II we will extend this OED approach with further functionalities for geothermal reservoir modeling and aim at increasing its performance in order to apply it to reservoir scale production runs. As a first step, synthetic test models have been set up for studying certain aspects, such as sensitivity of the OED result to a priori data.

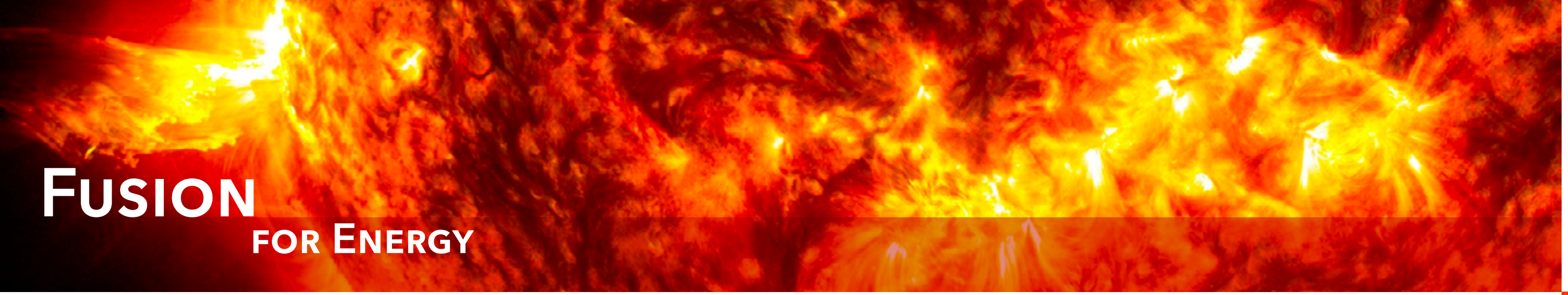
CODES

ParFlow (https://parflow.org/) is a physics-based 3D parallel hydrologic model, which simulates surface and 3D subsurface flow, based on Richard's and kinematic wave equations on a finite difference, finite volume grid with a preconditioned, implicit Newton-Krylov solver for the nonlinear PDEs. ParFlow will not be completely rewritten in EoCoE-II, but some code upscaling is planned as well as the activation of the Adaptive Mesh Refinement of the computing grids with the ad hoc numerical scheme and corresponding solver. It is also planned to have a strong effort on IO that will be coordinated with the effort on ensemble runs.

SHEMAT-Suite is a code for simulating single- or multi-phase heat and mass transport in porous media. It solves coupled problems comprising heat transfer, fluid flow, and species transport. SHEMAT-Suite can be applied to a range of hydrothermal or hydrogeological problems, be it forward or inverse problems.

HYPERstreamHS is distributed hydrological model based on the width function instantaneous unit hydrograph (WFIUH) theory, which is specifically designed to facilitate coupling with gridded climate datasets and climate model outputs. HYPERstreamHS inherits the computational grid from the overlaying meteorological forcing, still preserving geomorphological dispersion caused by the river network irrespectively of the grid resolution. In addition HYPERstreamHS is capable of simulating explicitly water transfers due to human activities (e.g., diversion channels and storage reservoirs).

Scientific Leader: Bibi Sarwat Naz



FUSION FOR ENERGY

DESCRIPTION

With the commissioning of large machines like the ITER tokamak, controlled fusion is poised to make huge step forward towards mastering the energy of the stars for a civil usage. The steady international progress regarding the achieved fusion performance gain – the ratio of the generated fusion power over the injected power – relies on our ability to understand, predict and possibly control the turbulent transport in view of optimizing the confinement properties of the plasma. Core transport studies in tokamak plasmas have now reached maturity with the development of state-of-the-art first-principle-based codes, using the gyrokinetic description. There, the distribution functions of plasma specie self-consistently evolve in time in a 5-dimensional phase-space under the action of electromagnetic fields governed by Maxwell's equations. However, despite their numerous successes to date, their predictive capabilities are still constrained with respect to the energy content in particular in optimized discharges.

Challenging this gap requires pushing gyrokinetic modelling towards the edge region of the container vessel – characterized by a colder and less dense plasma in interaction with solid materials, and as far as possible addressing edge and core transport on an equal footing, which makes nonlinear simulations mandatory. In EoCoE-II the goal for the fusion team is to bridge the gap between gyrokinetic core transport modelling and edge plasma physics for reliable predictions of fusion performance, which will require a number of numerical and physics bottlenecks to be overcome.

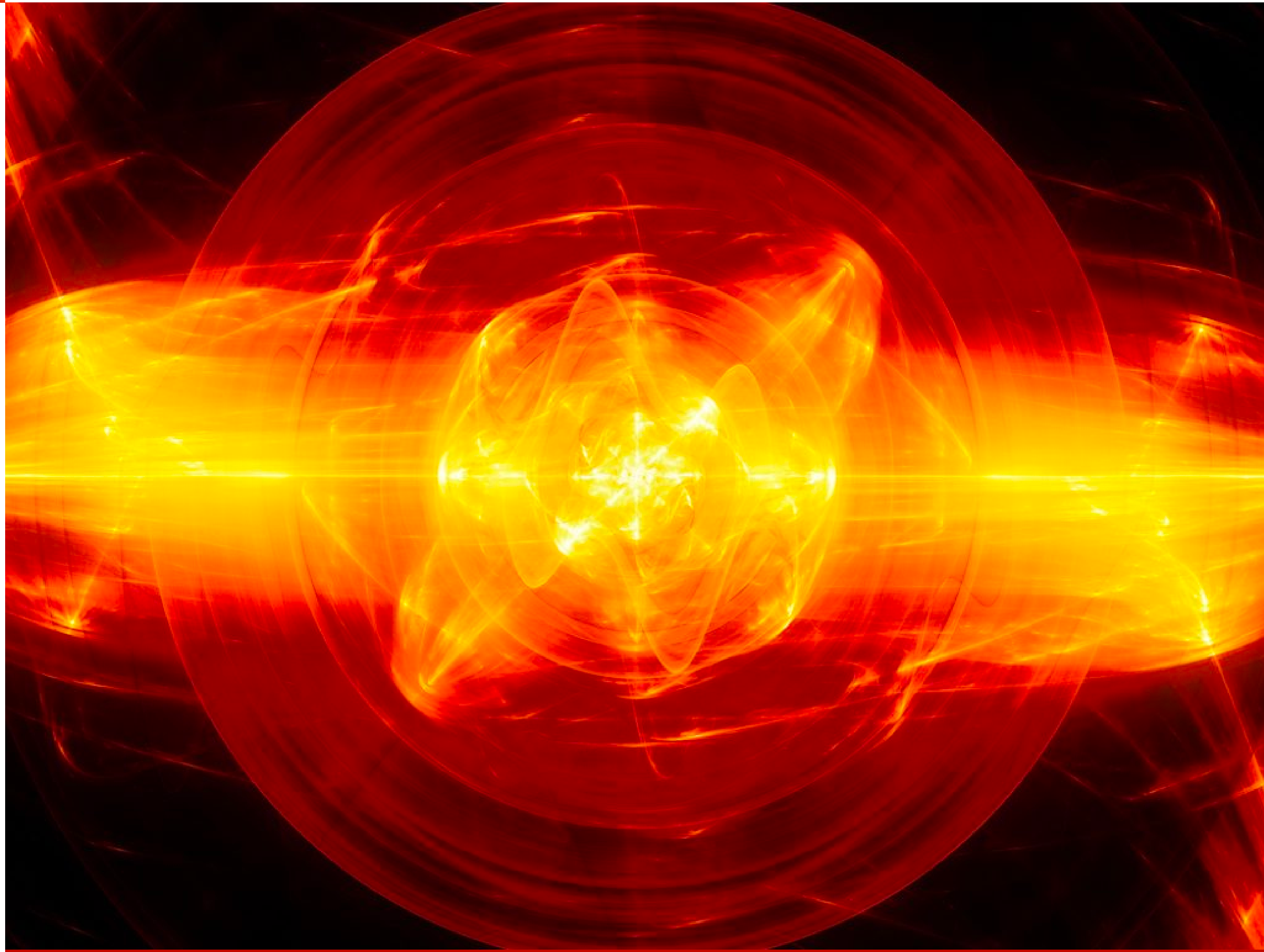
The objective is to develop a new numerical tool to address the core-edge issue, which will consist of refactoring and rewriting the flagship gyrokinetic code Gysela (its new name will be GyselaX), targeting the disruptive use of billions of computing cores expected in exascale-class supercomputers. This will be complemented by the upgrade of companion codes Tokam3X/Soledge2D (to be merged, leading to the new code Soledge3X), and GENE, to provide critical inputs both regarding numerical developments and physics issues.

RESULTS

1. The periphery of the confined plasma in tokamaks is characterized by plasma-wall interactions which govern many of its properties. This region, called the SOL (Scrape-Off Layer) is not a mere boundary condition for the core plasma: it is suspected – and partly observed in certain regimes – to impact fusion performance, and it controls the way power is conducted to and deposited on the target plates of the divertor (the element specifically designed to handle the large power out fluxes and particle exhaust).
   Immersed boundary conditions – using an adapted version of the penalization technique already operational in the companion fluid code Tokam3X – have been implemented and successfully tested in the version of GYSELA with adiabatic electrons. Both Vlasov and Poisson equations have been modified: mask functions have been implemented to account for the presence of a limiter – a perfect sink for the plasma – and for the expected electron response in the SOL. Preliminary tests especially show the buildup of a well of the radial electric field in the vicinity of the core-SOL boundary, as experimentally observed in experiments. This is very good news, since such a sheared electric field, if large enough, can lead to improved confinement regimes in tokamak plasmas, the so-called "H-mode."
2. Several Multigrid solvers have been tested on a reduced problem – using a 2-dimensional multi-grid solver – in view of providing efficient solutions for the GyselaX field solver when using very large grids: (i) a flux surface aligned polar mesh and (ii) a locally refined Cartesian mesh. Both options have advantages and drawbacks: while the former is well adapted to the intrinsically anisotropic turbulence in tokamak plasmas, there is no singularity issue at the magnetic axis with the latter.
   AMReX has revealed the most appropriate adaptive mesh refinement library for our purpose. In case (i), a mapping from a block-structured logical grid refined at the edge to the circular physical domain has been implemented, yet leaving the magnetic axis unsolved. Good numerical performance in terms of scalability and accuracy for both approaches has been obtained with OpenMP on a single node, although the multigrid solver required considerably fewer iterations on the Cartesian domain. The MPI parallelization on several nodes is ongoing.
   In parallel, a polar multigrid solver has been developed from scratch on a mesh that can be strongly refined in the radial direction (collaboration MPG-IPP / CERFACS). This solver, specifically adapted to our problem, should overcome the above issues encountered when using "on-the-shelves" libraries. Promising results have already been obtained in circular and shaped geometries. The solver will now be coupled to the GyselaX code.
3. The use of non-equidistant meshes also impacts the treatment of the Vlasov equation governing the time evolution of the distribution functions. To this end, preliminary coupling tests of a non-equidistant spline module have been performed on prototype codes – the VOICE code and the SELALIB library (collaboration CEA-IRFM / MPG-IPP) – and reveal conclusive. The coupling to the GyselaX code is in progress.

CODES

Gysela (http://gyseladoc.gforge.inria.fr/) is a 5D full-f (regarding Vlasov equations) and flux-driven gyrokinetic Fortran parallel code that solves Vlasov (ions and electrons) and Poisson (electric potential) equations to simulate plasma turbulence and transport in Tokamak devices. GyselaX developments include the replacement of the current Gysela code will be initiated to consider the whole tokamak from the core to the edge with advanced geometric constraints would require so deep refactoring of the existing Fortan90 code aims at increasing its modularity for a performant use of exascale supercomputers, with a special effort on the efficient numerical treatment of the large heterogeneities of physical quantities characterizing core and edge plasmas, Input-Output and a scalable solver for the Poisson like equation.



Scientific Leader: Yanick Sarazin
Yanick.sarazin@cea.fr

WIND

FOR ENERGY

DESCRIPTION

As Europe moves towards a decarbonized energy ecosystem, we need precise and efficient methods to produce, store and manage clean energy. Wind power is the renewable source with the most successful deployment over the past decade; if you were setting up a windfarm, wouldn't you want to ensure your turbines are optimally placed? Simulating your entire windfarm would give you a way to, ultimately, maximize your installation's energy output. But doing so requires the use of cutting-edge computational methods to run HPC simulations of a wind farm in complex terrain and up to 100 turbines. Such a modelization is one of EoCoE's core use cases, which also include simulations applied to perovskite solar cells, hydrology, meteorology and tokamak fusion reactors, all crucial domains a successful the European energy transition..

RESULTS

The wind scientific challenge has two main tasks. The first one is related to the CFD simulation of flow over a wind farm, commonly known as microscale The simulation of a full rotor, with rotor blades passing very closely to the mast, causes enormous difficulty for generating a computational mesh, which was solved by the EoCoE team. Also, the turbulent flow past a rotor is very complex and a challenge for numerical solvers. The EoCoE team improved the linear system solver with new algorithmic components. In order to ensure wide uptake and provide standardized feedback loops to developers, a portable version of the Alya code with all its dependencies, designed to run on any HPC platform, was created. Thanks to this work, a simulation of a (mid-to-large scale) turbine with 70m rotor diameter and wind velocity of 16 m/s was run on the MareNostrum IV supercomputer. The turbine's solid mesh consists of over 4.3 million elements, and its fluid mesh of close to 18 million elements.

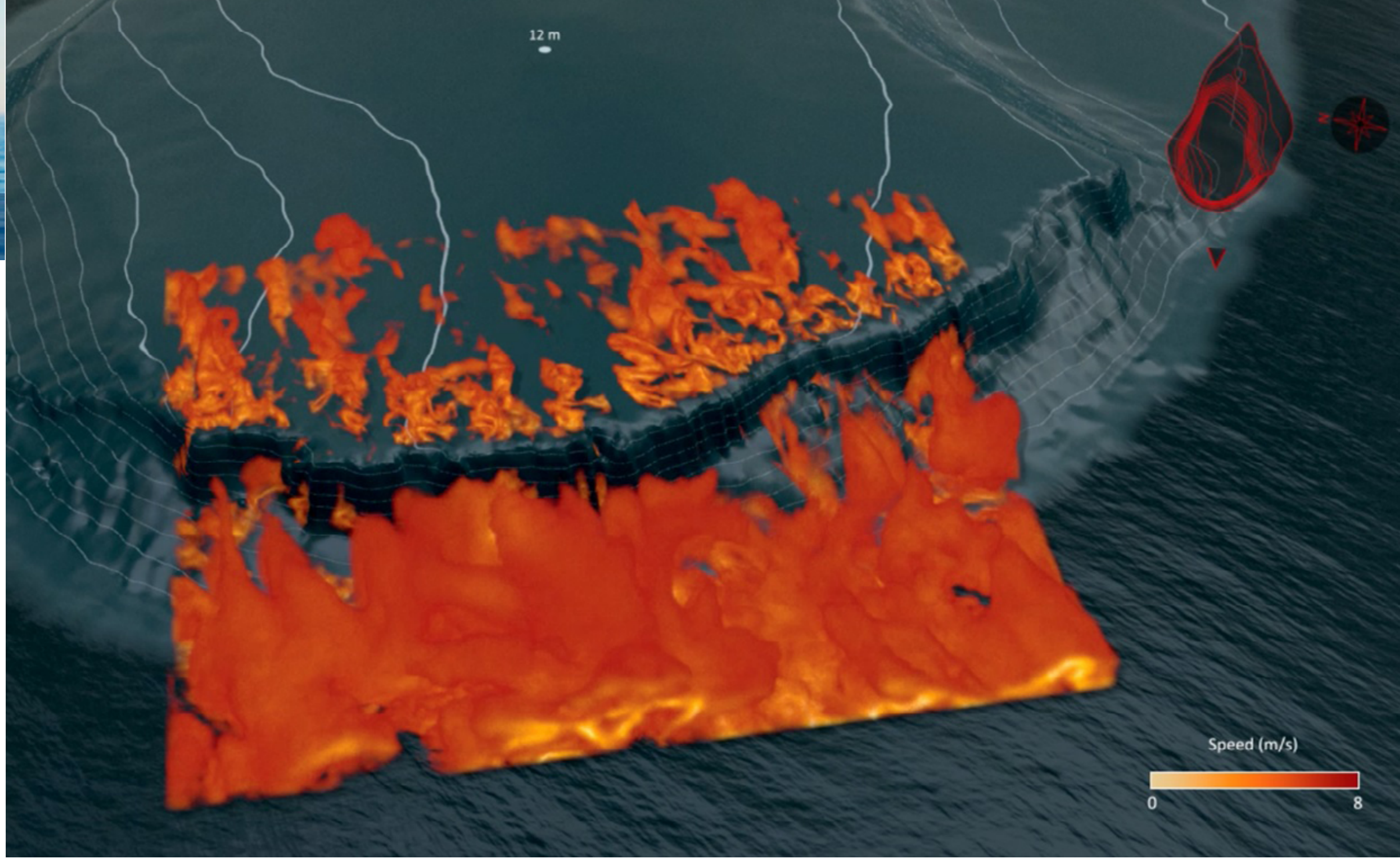
As the second phase of the EoCoE project comes to a close, its teams can run full rotor simulations and air flow over complex terrain, and plan to keep strengthening their collaborations and improving their codes in the future, making exascale HPC simulations a fundamental tool for improving the European energy ecosystem as a whole. Beyond the windfarm challenge, and as mentioned above, the EoCoE project applies simulations to several crucial domains for the European energy transition. In all of them, significant progress was achieved by the consortium, in line with what was detailed here for the wind simulations.

CODES

The increased understanding of turbulent flow in the wind farms has played a key role in the spectacular advance in wind power. To fully exploit these advances and move from a single wind turbine to the simulation of a large windfarm and its terrain requires increases in the size of the problems to be solved and the amount of computational resources of at least two orders of magnitude. EoCoE used the Alya code for performing these simulations (a flagship code for showcasing the power of Europe's next generation supercomputer).

To reach its performance objectives for complex terrain simulations, the EoCoE team ported the Alya code on GPUs, using only OpenACC. Tremendous results were achieved; for instance the team got Alya running 68 times faster for a core part (element assembly) compared to a CPU. Porting Alya meant ensuring the code would fully exploit modern GPUs, and EoCoE reached 50% of the maximum floating-point performance on an NVIDIA A100 GPU – a very good ratio. In a cooperation between Barcelona Supercomputing Centre (BSC) and the National Research Council of Italy, a fast linear solver (AMG4PSBLAS) was integrated, running 3 times faster than the standard solver used before. Thanks to work on dynamic load balancing taking full advantage of the GPU computing power, Alya GPU-only simulations now run 4.3 times faster than the previous CPU-only simulations.

In addition to GPU porting, the EoCoE team tackled the co-execution of CPU/GPU fuel simulations. The joint task force between BSC and Friedrich-Alexander University achieved a 20% improvement on CPUs - a significant result that validates the consortium's collaborative code improvement strategy. EoCoE believes that there is still more potential in the CPU/GPU combination for this type of application; a study ("mini-app") built using CUDA, an alternate framework for GPU programming, obtained a massive speed-up, as it ran 11 times faster than the OpenACC GPU-only runs.



Detailed simulation of irregular wind velocity over complex terrain (using a mesh with 358 million nodes and 2 billion volume elements)

Scientific Leader: Herbert Owen

herbert.owen@bsc.es

Technical Challenges

DESCRIPTION

The emergence of deep memory hierarchies and the need for extreme scalability are key issues applications must face for their arrival in the coming pre-exascale and exascale systems.

EoCoE-II proposes a flexible approach that embraces Input/Output (IO) methods, ensemble runs and programming models in order to contribute to the necessary paradigm shift.

From a software engineering point of view, the current implementation of physics diagnostics and post-processing in production codes is mixed with the numerical core implementation. The resulting codebase is then monolithic and intricate. The development and the introduction of a Parallel Data Interface in production codes separate clearly these two concerns and will improve its modularity, its maintainability and its adaptability to new coming hardware. The integration of the Parallel Data Interface with the code enables for different associated plug-ins to be implemented dealing with general IO issues, in-situ/in-transit post-processing capabilities or fault tolerance mechanisms.

EoCoE-II project partners offer expertise in the following domains:

- Experts on **Programming Models** focus on how to handle efficiently complex computing nodes having a deep memory hierarchy and possibly accelerators, how to address more operation concurrencies and to minimize development effort that maximizes performance portability. There will not be new research in this area but the application of existing developments and of established standards or emerging technologies, like task-based programming models and Kokkos.
- Experts on **Scalable Solvers** focus on algorithmic issues that are strongly linked to linear algebra. Refactoring existing solver packages will not be enough to reach Exascale. New algorithms are necessary to reach the required level of scalability. The range of the expertise is wide and will be applied to at least seven codes where linear algebra has been identified as one of the main bottlenecks.
- Experts in **Input/Output & Data Flow** methods aim to leverage efficiently the capacities of the coming hardware that will further extend the depth of memory hierarchy (NVRAM, SSD…). Efficiently writing large amounts of data to the parallel file system will be the expertise on offer. Deep memory hierarchies call for paradigm shift in the way fault tolerance is handled and post processing is implemented with in-situ and in-transit capabilities.
- Experts in **Ensemble Runs** propose the integration of software technologies to run efficiently simulation ensembles and more generally workflows on coming pre-exascale and exascale systems. This will potentially include data assimilation. The objective is to provide a flexible and maintainable way of executing simulation ensembles in multiple jobs on a given supercomputer while enabling communication between distinct ensemble members in ensemble management and data assimilation processes.

PROGRAMMING MODELS

Choosing the right programming model(s) is an important decision that can strongly affect the development, the maintenance, the readability, the portability and the optimization of a code. Programming model challenges are tackled in the Work Package 2 of EoCoE. Here, the term programming models mainly refers to as programming, optimization and parallelism languages, frameworks or libraries.It could be rephrased High Performance Computing (HPC) challenges. Our goal consists in helping, advising and guiding scientist code developers non-expert in HPC to refactor and optimize their applications to reach the best performance and scalability on massively-parallel super-computers. Targeted super-computers can be composed of CPUs with accelerators such as GPUs. We currently test our improvements on Tier-0 European machines but our work is not limited to present technologies. Experts in HPC are preparing applications to run on future pre-exascale machines with disruptive technologies such as ARM-based processors and accelerators. We aim at prepare selected codes in the domain of energy to the Exascale ecosystem.
To achieve this goal, the programming model technical challenge has experts in performance evaluation. Having a systematic and continuous evaluation of the code performance enables to guide optimization developments all along the project and to monitor improvements. It helps to determine the best options to solve performance bottlenecks.

SCALABLE SOLVERS

Solving Linear Algebra (LA) problems is a core task in four out of five EoCoE II Scientific Challenges (SC) and thus the availability of exascale-enabled LA solvers is fundamental in preparing the SC applications for the new exascale ecosystem. The goal of WP3 is to design and implement exascale-enabled LA solvers for the selected applications and to integrate them into the application codes. A co-design approach between LA and SC experts will be used; nevertheless, the solvers will be also developed in a more general perspective, to obtain LA tools useful for a wider range of applications. The LA experts involved in WP3 have a long-standing experience in developing solvers for HPC platforms, and the planned activities will build upon software and methodologies developed by them and tested during EoCoE I. New algorithms and disruptive technologies will be also considered, to tackle the new challenges posed by the envisioned exascale systems.
In order to achieve the previous goal, the following steps will be performed:

- analysis of the LA kernels of the applications, to clearly identify the needs of the applications in terms of LA solvers, and to select the best LA methodologies and software to work with. Actually, this work has been triggered during EoCoE I, providing a sound basis for EoCoE II
- extension, modification, and re-factoring of the selected LA solvers, based on a co-design approach between LA experts and application experts, to ensure that solvers and applications evolve accordingly in their route toward exascale
- design and implementation of novel solvers, for applications where modifying and re-factoring the available LA solvers does not appear satisfactory
- Integration of the LA solvers into the applications, in strict collaboration between LA and application experts, and testing and tuning on problems of interest
- The work will address the Material, Water, Fusion and Wind Scientific Challenges, where strong needs for exascale-enabled LA solvers have emerged during EoCoE I. A different task for each SC has been planned, plus a task concerning transversal activities. The HPC packages developed by the LA experts participating in EoCoE II will provide a sound basis.

I/O & DATA FLOW

Four different I/O & Data Flow releated technical challenges are targeted within EoCoE-II:

- Improvement of I/O accessibility: Different I/O libraries support a variety of different configuration options. Depending on the situation these options must be continually updated, or a complete new library must be adapted. We want to introduce a generic interface with the help of the Portable Data Interface (PDI), which decouples the I/O API and the application to allow easier switching between different I/O subsystems. These interface should either serve standard I/O operation but should also be useful in context of ensemble or in-situ visualisation data movement.
- I/O performance: The data writing and reading time can consume a significant part of the overall application runtime and should be minimized. For this we want to leverage the optimization options of different I/O libraries in use as well by adapting intermediate storage elements such as flash storage devices.
- Resiliency: Running an application on a large scale increases the chance of hard- or software problems if more and more computing elements are involved in the calculation. Additional I/O techniques can be used to reduce the effort needed to restart a broken run or even avoid an overall crash, by storing intermediate snapshots to the storage elements. In particular we want to focus on resiliency for ensemble calculations.
- Data size reduction: Running an application on a larger scale often implies an increasing data size, which can become unmanageable and consume too much resources. Within this task we want to reduce the overall data size without losing necessary information via in-situ and in-transit processing, moving post-processing elements directly into the frame of the running application.

ENSEMBLE RUNS

The goal is to efficiently run large simulation ensembles, commonly called “ensemble runs”, on coming pre-exascale and exascale systems, an approach used when a sample of simulation runs is required for getting a statistical estimation of the application behavior on some given parameter ranges. Ensemble run are used for uncertainty quantification and data assimilation for instance. EoCoE-II is targets developing an innovative ensemble management infrastructure by enabling elasticity, resilience and modularity to enable the execution of very large ensembles of advanced parallel applications. The standard approach for data assimilation relies on a single large monolithic MPI execution that makes it sensitive to faults, load balancing issues and require to reserve up-front large fractions of supercomputers.
Two out of five EoCoE-II Scientific Challenges (Weather and Hydrology) integrate support for ensemble runs for data assimilation or sensitivity analysis. The goal is to develop this novel framework and demonstrates that it enables to run efficiently run ensemble of more than 1024 members for the Weather and Hydrology Scientific Challenges.

DISSEMINATION

Transversal Joint Programme

Having assessed the potential of digitalization as an asset for energy research, and studied the current successes and shortcomings of the application of digitalization to energy domains, EoCoE and EERA have been working hand-in-hand to set up a transversal Joint Programme to promote and streamline how digitalization as a whole, including but not limited to HPC and simulations, is used. Users are, first and foremost, scientists working on energy domains, but also industrial actors in a position to exploit the scientific results having reached their full potential. This transversal EERA Joint Programme 'Digitalisation for Energy', launched in October 2020, was designed as a cross-cutting structure to the other EERA Joint Programmes with the aim of leveraging pre-existing expertise within the established EERA structure and of complementing it with leading edge knowledge on the latest digital concepts and technologies. As the focus on transdisciplinarity and community building are mainstays of the programme, it promotes and fosters transdisciplinary collaborations across all EERA activities. EERA and EoCoE are confident that this Joint Programme will be an important asset for research and industry energy communities. It will provide the tools and expertise to run energy focused simulations on large scale systems, thus fully exploiting the immense potential of high performance computing; ultimately, this Joint Programme will play an important role in changing Europe's energy ecosystem.

EoCoE School

EoCoE School offers collaborative online learning opportunities on Energy-oriented Centre of Excellence project tools and codes. EoCoE Yuotube channels presents the 5 scientific Exascale challenges in the low-carbon sectors of Energy Meteorology, Materials, Water, Wind and Fusion, coming from a world-class consortium of 18 complementary partners from 7 countries related into a unique network of expertise in energy science, scientific computing and HPC, including 3 leading European supercomputing centres.

School participants are instructed for free and supported by experts in High Performance Computing and in Sustainable Energies. The school creates a great opportunity to share experience and learn more about EoCoE technical and industrial work. During the hand-on sessions attendees use the EoCoE SaaS Portal that benefits from the Altair supercomputer (Top 500). Target audience are scientists and researchers from academia and industry from across Europe.

EOCOE.EU

European Commission

Horizon 2020
European Union funding
for Research & Innovation