Newsletter 12/2017

To subscribe to/unsubscribe from this newsletter, please visit:
https://lists.dkrz.de/mailman/listinfo/esiwace-external

You find further details on ESiWACE at www.esiwace.eu. All newsletters are also available at www.esiwace.eu/newsletter.

Contents

Introduction 1

Dissemination and Upcoming Events: Meet Us! 1

Scalability and Optimisation: Towards the High-Resolution Demonstrators 2

Usability: Handling Complex Earth System Model Software Stacks and Workflows 3

Exploitability: Software Solutions for Efficient Storage 4

Collaborations and Industry Interactions 4

Governance and Engagement 5

Introduction

ESiWACE focuses on preparing Earth system models and tools for the exa-scale, and on supporting the weather and climate modelling communities to operate these models efficiently on a variety of HPC systems. The key target is to prepare global cloud and ocean eddy resolving simulations at 1km scale and at a rate of 1 simulated year per calendar day. This will allow breaking down some of the greatest scientific challenges thanks to better predictions of climate trend, and it will result in much more fidelity in the representation of high-impact regional weather and climate events. This is clearly not achievable today and will require exa-scale computing and data handling capability.

Dissemination and Upcoming Events: Meet Us!

In 2018, ESiWACE will participate in

- EGU, 8-13 April, Vienna/Austria, www.egu2018.eu
- 5th ENES HPC workshop, 17-18 May, Lecce/Italy, www.esiwace.eu/events/5th-enes-hpc-workshop (organised by ESiWACE)
- PASC 2018, 2-4 July, Basel/Switzerland, pasc18.pasc-conference.org
- 18th ECMWF Workshop on High Performance Computing in Meteorology, 24-28 Sep, Reading/UK, www.ecmwf.int/en/learning/workshops

ESiWACE has participated in various application-specific and HPC events, amongst others:

- At EGU 2017 in Vienna, we presented ESiWACE in the session "Recent developments in numerical atmospheric, oceanic and sea-ice models: towards global cloud and eddy resolving simulations on exascale supercomputers".
- At the European HPC summit week 2017 in Barcelona, ESiWACE organised a half day workshop.
- A BOF session with international speakers from USA, Japan, China and Europe on "Cloud Resolving Global Earth System Models: HPC at its Extreme" was held at ISC 2017 in Frankfurt. Additionally, ESiWACE co-coordinator Peter Bauer delivered a keynote talk at ISC.
- At PASC 2017 in Lugano, we organised several minisymposia within the Climate & Weather track.
- At TeraTec 2017, a booth was reserved to present the work of Bull/Atos.

Scalability and Optimisation: Towards the High-Resolution Demonstrators

One key achievement has been the swift inclusion of a new task on very high-resolution global model simulations—so-called demonstrators—to investigate the "computability" of the target resolution and target production rate. This will determine the discrepancy between current and required code performance and will feed into a roadmap proposing the steps to close this gap. Already, a first set of 1km- to 5km-resolution runs with both IFS and ICON global atmospheric models has been performed. Scalability was found to be good for IFS and ICON in the high-resolution simulations. A performance benchmark has been extracted from aqua-planet experiments with ICON and is available at https://redmine.dkrz.de/projects/icon-benchmark/wiki/Instructions_on_download_execution_and_analysis_ICON_Benchmark_v160. See https://www.esiwace.eu/results/achievements for parallel performance data of both IFS and ICON and a 3-week forecast animation using the global 5km-NWP ICON model. Besides, preparations for coupled ocean-atmosphere simulations with EC-Earth have been made and a 10km coupled ocean-atmosphere run has already been conducted, producing a few years of output.
Long-term strategies do not only require research and code optimisation to meet the target criteria, but they particularly introduce the need of code support and usability (for the latter, see following section). In this scope, ESiWACE offers user support (including on-line forum, user guide, tutorials, etc.) for the software tools XIOS and OASIS.

Various code optimisation techniques have been employed for different models:

- IFS: Single-precision tests with IFS have shown up to 40% efficiency gains while retaining good accuracy. Concurrent execution of radiation with other model components within MPI tasks produced up to 13% run time gains. The assessment of different programming models (Fortran co-array, MPI, GPI) also started for the IFS spectral transforms. The usage of GPI was enabled through the link with the EPiGRAM project as this link provided free access to the software license.
- ICON: In 2016 we started a collaboration with the centre of excellence POP to analyse an R2B4 ICON AMIP configuration. It was shown that the main factors affecting the parallel performance are load balance and the latency of the communications. As these issues are directly influenced by the number of used MPI tasks, we recently started another POP audit that will analyse the OpenMP scalability of ICON. By an improved hybrid version of the code, we seek to minimise the MPI communication and hence get better overall performance.
- NEMO: We performed an analysis of the memory requirements for version 2 of the XIOS I/O server. We further started analysing the impact of running different parts of NEMO in single precision. Bull realised diverse optimisations following several tracks: vectorisation, computation

reduction, memory and communication schemes. Further activities focussed on the development of a NEMO hybrid version with the introduction of a second level of parallelism on the vertical levels based on OpenMP. Biogeochemistry, sea-ice and almost the whole physics of the NEMO code have been parallelised using the hybrid approach. About 25% of the code has been analysed and modified with a gain of 5% on the total execution time. These activities are coordinated within the NEMO consortium and, more specifically, within the NEMO HPC Working Group.

- EC-Earth T255L91-ORCA1L75: Activation of OASIS3-MCT options gathering coupling fields in the coupling exchanges yielded 40% coupling cost reduction. Optimising in OASIS3-MCT the mass conservation calculation with local sums to reduce the global communication resulted in 90% coupling cost reduction. Finally, a collaboration with the Montblanc EU project team has been established to produce energy profiling of EC-Earth 3.2 on a ThunderX ARM platform, which has been communicated to Montblanc developers in a report delivered beginning 2017.

However, the gains from all the optimisations alone will not be sufficient to achieve the required runtime targets. ESiWACE is set to implement a long-term strategy including incremental as well as potentially disruptive measures to support the weather and climate community to achieve those aims beyond the end of the project.

Usability: Handling Complex Earth System Model Software Stacks and Workflows

In addition to making Earth system models “exa-scale ready”, ESiWACE aims at improving their usability especially in view of their exa-scale application. The aim is to identify, design and support the end-to-end workflow for climate modelling and weather forecasting applications in both research and production mode. This will provide the increased human efficiency that will be required to exploit increasingly complex exa-scale systems.

ESiWACE handbooks, which will constantly evolve throughout the project, support creation of the environment and system software stack which are necessary for the communities to run their models. Spack has been investigated and is to be used to easily port and create a common software stack. This has already been successfully tested on five supercomputing platforms. One success story consisted in porting a daily air quality forecast system from BSC to another machine within 2 days; a task which had initially been expected to take at least a week. Details for system administrators can be found in https://www.esiwace.eu/results/deliverables/d3-5-how-to-select-configure-and-install-esm-software-stacks/view.

Further, we foster the automation of scientific workflows by delivering development and support services to the community for the meta-scheduler Cylc. This will also facilitate exploitation of results. Various optimisations and enhancements have been performed in Cylc, amongst others

- Client-server interaction enhancements, including moving from the obsolete Pyro3 protocol to industry standard HTTPS protocol,
- new built-in support for parameterised tasks for a cleaner workflow,
- restart robustness and runtime database improvements,
- performance and feature enhancement to the task/job management subsystem,
- performance and efficiency improvements to the suite validation and runtime.
- The latter implied, for example, that for a large ensemble suite which is used at Met Office, memory could be reduced by more than 50% going from Cylc 6.1.2 to 6.11.2.

Exploitability: Software Solutions for Efficient Storage

The amount of data that will be generated by high-resolution Earth system model runs will be unprecedented, and efficient exploitation of data storage will become a major challenge. One of the ESiWACE objectives is therefore to develop tools to enhance and support exploitability of data sets arising from next generation weather and climate models—and to use prototypes of the software system components to handle data from the demonstrator simulations.

First, we addressed requirements and basic (coarse-grained) modelling of performance, cost and resilience for weather and climate data storage systems, introducing several alternative scenarios for potential and existing architectures of data centres. We have already started to adopt these architectures. Details on the alternative scenarios can be found in https://www.esiwace.eu/results/deliverables/d4-1-business-model-with-alternative-scenarios/view.

Second, we worked on designs for the architecture of a software "middleware" layer, which can be implemented between user applications and storage systems aiming at increased I/O and data handling performance by local customisation. Requirements for the community were discussed. We introduced the semantic perspective, i.e., the impact on end-users and developers, when switching to the middleware, and provided a draft for the implementation of a low-level layer for file systems that utilises concepts from log-structured file systems to optimise performance for random write workloads. Additionally, it was explored how the middleware can be integrated into NetCDF and HDF-libraries from the NWP/climate community. An in-memory plugin for HDF5 was implemented to study the behaviour of interfaces and to act as performance baseline for any subsequent implementation that utilises persistent storage. Together with Seagate, the APIs for data were explored.
Requirements and design concepts around exploiting an HDF server architecture were investigated, in particular to support high-performance data access between a secure data environment and an insecure private cloud environment available to external users with root access to their systems. The HDF Group (unfunded) provided input on their plans and architectures for their new HDF to S3 work. Further strands of work focussed on understanding scientific library storage requirements for Mero usage and collaborative exchange on domain specific I/O libraries, as well as investigations of both in-memory support and object store technology as backend storages for scientific data analytics.

A third task is the development of a prototype for fine-grained simulations of hierarchical storage systems with a focus on tape and modelling both cost and resilience aspects. In the scope of a master's thesis, work has been acccomplished to simulate hierarchical storage systems which benefits the modelling effort and will be used to understand the impact of alternative tape strategies.

Collaborations and Industry Interactions

Interactions with the HPC industry were intensified at the 4th ENES HPC workshop 2016 co-organised by ESiWACE and IS-ENES2. Contacts with the centres of excellence EoCoE and PoP, several FET projects, namely ESCAPE and NextGenIO, as well as with the European Extreme Data & Computing Initiative (EXDCI) were established. The coordinators were involved in various discussions with ETP4HPC, in particular with respect to the rolling revision of the Strategic Research Agenda and the implementation of the envisaged Extreme scale Demonstrators (EsD) in Europe. The ENES HPC task force has been extended to include experts from the domain of numerical weather prediction. In particular, the task force contributes to the EXDCI working group on “Weather, Climate and Solid Earth Sciences”. DKRZ and MPI-M

organised a series of workshops under NDA with leading HPC hardware vendors, which were also attended by representatives of the ENES HPC task force. This led to an intense exchange on requirements by the weather and climate communities on the one hand, and expectations with regard to upcoming (exa-scale) hardware on the other hand.

Governance and Engagement

With regard to software governance, we collected information on the purpose, features and governance of community software packages. Concrete steps have been taken with regard to governance for certain central and established packages, including OASIS, YAXT, Cylc and CDO. Amongst others, governance developed for OASIS under the umbrella of IS-ENES2 has been launched in the ESiWACE context. Based on feedback received by the extended user community, including partner and supporting institutions, suggestions for governance rules for new software developments in ESiWACE were recently internally formulated; upon request of direct feedback we could moreover assess the interest of the extended user community, including partner and supporting institutions, in ESiWACE activities and collect elements on how to inform it.

As Centre of Excellence, ESiWACE has been granted access to PRACE compute resources. In total, ESiWACE partners have been granted 1.5M core hours on the supercomputers Mare Nostrum 3, Marconi Broadwell and Marconi KNL.

ESiWACE further participates in the European Commission Open Access Data Pilot for Research Data and thus provided a Data Management Plan in compliance with the requirements set by this pilot. A kick-off meeting and first meeting of the General Assembly was held in November-December 2015. The second meeting of was held in April 2016 in conjunction with the 4th ENES HPC workshop. The third general assembly took place on 12-13 December in Berlin.

Participants of the ESiWACE General Assembly, 12-13 Dec 2017, Berlin

The ESiWACE project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 675191