Proposal n°2010PA4064

Submitted for the assessment cut-off date : Friday 1 September 2017 at 11:00:00 CEST

General information

Type of proposal: Preparatory type **C – Code development with support from PRACE experts.**
Start date: 01/11/2017
The start date should be either one and half or two months after the announced date for evaluation cut-off.

Continuation of : Previous type D project: undefined

Project name: Optimisation of EC-Earth 3.2 model
Research field: Earth Sciences and Environment

Template documents for the mandatory final report :
(as described at **http://www.prace-ri.eu/Information-for-PRACE-Awardees**)

- **PRACE_-_FRF_PC_TypeA.doc**
- **PRACE_-_FRF_PC_TypeB.doc**
- **PRACE_-_FRF_PC_TypeC.doc**
- **PRACE_-_FRF_PC_TypeD.doc**

Project leader (personal data and contact)

Gender: Female
Title: Dr
First name: Virginie
Last name: Guemas
Initials: VG
Date of birth: 21/11/1983
Nationality: French
e-mail : virginie.guemas@bsc.es
Please give your professional e-mail address. E-mail addresses such as Gmail and Hotmail are not accepted.
Phone number: +34 934137679
Fax number:

Project leader (organisation and job title)

Job title: Climate Prediction group manager
Website: https://www.bsc.es/guemas-virginie
Organisation name: Barcelona Supercomputing Center
Department: Earth Sciences
Group: Climate Prediction
Address: Edifici Nexus II c/ Jordi Girona 29
Postal code: 08034
City: Barcelona
Country: SPAIN

Organisation with a research activity: Yes

Employment contract of the project leader is valid at least 3 months after the end of the allocation period: –

For commercial companies,

- **Is the head office of the organisation in Europe?** –
- **% of R&D activity in Europe as compared to total R&D activity :** –

Contact person for all correspondence

Name: Kim Serradell Maronda
E-mail: kim.serradell@bsc.es

Collaborators

Organisation name: Barcelona Supercomputing Center
Department: Earth Sciences
Group: Climate Prediction
Address: Nexus II Carrer Jordi Girona 29
Postal code: 08034
City: Barcelona
Country: SPAIN

Mr. Kim Serradell Maronda

Job title: Computational Earth Sciences group manager
Website: https://www.bsc.es/serradell-maronda-kim
Gender: Male
Title: Mr.
First name: Kim
Last name: Serradell Maronda
Initials: KSM
Date of birth: 19/12/1980
Nationality: Spanish
e-mail : kim.serradell@bsc.es
Phone number: +34934134051
Fax number:

Requested computer systems

Computing center	Machine
BSC *Barcelona Supercomputing Center*	☑ **MareNostrum**
CINECA	☐ **Marconi – Broadwell**
CINECA	☐ **Marconi – KNL**
CSCS *Swiss National Supercomputing Centre*	☐ **Piz Daint**
Gauss/HLRS *High Performance Computing Center Stuttgart (HLRS)*	☐ **Hazel Hen**
Gauss/LRZ *Leibniz-Rechenzentrum*	☐ **SuperMUC**
GENCI/CEA *Commissariat à l'Énergie Atomique*	☐ **Curie**
Tier-1 system **(for PA Type D)** *A Tier-1 system will be selected based on your given information*	☐
Specific Tier-1 system **(for PA Type D)** *Name of the preferred Tier-1 system*	

1. Summary of the project *(for Types A, B, C, D)*

To be published in the PRACE website. Maximum 500 words.

EC-Earth is the community European global climate model, based on the world-leading weather forecast model of the ECMWF (European Centre of Medium Range Weather Forecast) in its seasonal prediction configuration, along with NEMO, a state-of-the-art modelling framework for oceanographic, forecasting and climate studies which is developed by the NEMO European Consortium.

BSC has developed a coupled version of EC-Earth 3.2 at a groundbreaking resolution. In the atmosphere the horizontal domain is based on a spectral truncation of the atmospheric model (IFS) at T1279 (approx. 15 km globally, i.e. the highest resolution we can use with the standard IFS - higher resolutions would require e.g. non-hydrostatic parameterizations) together with 91 vertical levels. The ocean component (NEMO), is run on the so-called ORCA12 tripolar (cartesian) grid, at a horizontal resolution of about 1/12° (approximately 16 km); with 75 vertical levels which thickness increases from 1m below surface up to 500m in the deep ocean.

2. Scientific case of the project *(for Types A, B, C, D)*

Explain the scientific case for which you intend to use the code(s). Maximum 500 words.

This code will be used to perform Climate Prediction Group scientific activities in the next years as well as next PRACE allocations. These activities involve historical experiments, which consists in simulations of 50 and 100 years, and seasonal prediction experiments, which are commonly performed in ensemble mode, using many independent simulations with different initial conditions. The number of members is determined by the computational resources available. These means that increasing the performance of the model, will allow to run simulations with more members, so getting better scientific results.

On the other hand, the historical experiments (simulating up to 100 years) for this ultra-high resolution involve a

computational cost huge in sequential (simulating from the first to the last year). A proper computational and energy efficiency of these experiments will be mandatory, in order to not waste the parallel resources needed.

3. Computer resources requested *(for Types A, B, C), or expected for Type D*

Total storage required (Gbyte) (only available during the duration of the preparatory access project)	2 000.00
Maximum amount of memory per core (Mbyte)	2 048.00

4. Please provide the details listed below for the main simulation application *(for Types A, B, C, D)*

Name and version	EC-Earth 3.2
Webpage or other reference	http://www.ec-earth.org/
License If the code is open source please, fill out *open source* for this query.	The EC-Earth model has a license via ECMWF that is not open source, but it allows for free 'Academic use' for users in member states of ECMWF (most European countries). The NEMO ocean model is open source (French CeCILL license) and as such there is no co

5. Describe the main algorithms and how they have been implemented and parallelized *(for Types A, B, C, D ; Maximum 250 words)*

EC-Earth 3 comprises two major components: IFS as atmosphere model, and NEMO to simulate the behaviour of the ocean. Both components run coupled using OASIS3, which in its more recent version is used as a library. IFS and NEMO are built as separate executables, the same as XIOS, which is the I/O server used by NEMO, and the runoff mapper, which is a small program that distributes runoff water from land to the Ocean. OASIS synchronises all this components and is responsible for the exchange of information.

The resolution proposed (T1279: ~2140702 horizontal grid points, ORCA12: 13220998 horizontal grid points) will help efficiently share calculations between 2,521 processes, increasing the range of efficient compute-core usage per model executable.

IFS and NEMO fully support a parallel environment, while OASIS3 supports a pseudo-parallel environment. IFS supports both distributed (MPI) and shared memory parallelism (OpenMP), while NEMO (and XIOS I/O server) has been only parallelized through MPI, so in EC-Earth only MPI is used. OASIS3 requires Cray pointers.

6. Current and target performance *(for Types A, B, C, D ; including the points below. Maximum 250 words)*

- Describe the scalability of the application and performance of the application
- What is the target for scalability and performance? (i.e. what performance is needed to reach the envisaged scientific goals)

This resolution has been tested only in MareNostrum 3. Currently, we are at 0.16 SYPD (Simulated Years Per Day). This value is very low compare to other resolution results (up to 30 and 15 SYPD using T255/ORCA1 and T511/ORCA025 respectively) Our goal is reaching at least 6 SYPD (threshold to use this

configuration in production). This target performance is mandatory to achieve if we want to obtain simulation results for scientific activities in a proper time.

7. Confidentiality *(for Types A, B, C, D)*

Is any part of the project covered by confidentiality? No

If YES, please specify which aspect is confidential and justify:

8. Describe the I/O strategy regarding the parameters indicated below *(for Types A, B, C, D)*

8.a) Is I/O expected to be a bottleneck?

Yes, at the resolution targeted, I/O will be a bottleneck, mainly for the atmosphere component (IFS). IFS produces the outputs in a sequential process using a standard known as GRIB. All the MPI processes send the output information to a master process, which do the writing process.

8.b) Implementation: I/O libraries, MPI I/O, netcdf, HDF5 or other approaches:

NEMO's I/Os are handled by a parallel I/O library called XIOS that reads and writes NetCDF3 and NetCDF4 files in NEMO. IFS uses a master process to manage the output.

8.c) Frequency and size of data output and input:

Frequency and size can be easily managed in IFS and NEMO using namelists. This can be changed accordingly to scientific goals of the experiment. Typically, the historical and seasonal experiments require outputs for some variables hourly (6h and 12h) and most of them daily.

8.d) Number of files and size of each file in a typical production run:

In a typical production run, 1 year of simulation of T1279-ORCA12 would occupy ~ 1.5 TB.

9. Main performance bottlenecks *(for Types B, C, D. Maximum 250 words)*

We assume bottlenecks related to IFS I/O due to the exclusive master writing approach used. With the increase of the resolution, this behaviour will have a strong impact on performance, due to the need of increasing the number of processes to achieve a good parallel efficiency of the model, increasing the number of communications from all the slave to the master process.

Another bottleneck is related to the solver used by NEMO model. Currently, in EC-Earth NEMO uses a filtered solver which has a not very good performance when increasing the resolution. This is due to the domain decomposition process, some works prove that the particular solver used is useful when the number od subdomains is not large. For the ultra-high resolution, where the number of processes needed increase, the solver used is not the best choice.

Finally, NEMO runs the ocean and ice govern equations inside the same executable. However, both of them are independents and their computational efficiency is not the same. This ice model (LIM3) is more computational intensive and it's only run each N timesteps of the ocean model (OPA). These means that though both of them are using the same parallel resources, the ideal number of processes should be different for the ocean and ice models.

10. Describe possible solutions you have considered to improve the performance of the project

(for Types B, C, D. Maximum 250 words)

For the solver issue, a possible solution would be to enable another approach for the resolution of the linear system to solve, NEMO includes other possible solvers, depending on the each particular situation. We think that a time splitting solver approach, including in NEMO which be a better option. This solver has already been coded but never implemented in EC-Earth.

A solution for the ice limitation would be to remove LIM3 from NEMO binary and run it as a coupled component, using the OASIS coupler. This means that the ice model would be an independent component (similar to IFS and NEMO), where the number of parallel resources could be set independently, according to the needs of the ocean and ice models.

To avoid node contention we have been considering changing the affinity of the processes. Instead of having node only with atmosphere or ocean tasks, we could mix them and avoid possible hardware contentions. Morever, the NUMA affinity could be also studied for these MPMD applications, in order to increase the efficiency of EC-Earth.

11. Describe the application enabling/optimization work that needs to be performed to achieve the target performance

(for Types B, C, D. Maximum 250 words)

For the solver issue, time splitting solver should be enabled through namelist and analyse the performance of the model using this solver. Different parameter configuration should be evaluated, in order to test a different number of parallel resources. Once the model runs smoothly with this change, a numerical evaluation of the results has to be performed to validate them.

For the ice issue, a new configuration has to be developed where LIM3 is an independent model and OASIS manages the coupling between ice and ocean. Once the model runs smoothly with this change, a numerical evaluation of the results has to be performed to validate them.

For the affinity tests, a large benchmark of combinations, exchanging NEMO and IFS processes should be tested, in order to evaluate which is the best combination.

12. Which computational performance limitations do you wish to solve with this project?

(for Types B, C, D. Maximum 250 words)

We wish to solve the current limitation when running EC-Earth model at the ultra high resolution proposed. We want to run the model faster than it's done currently and with the best balanced configuration. The main goal is to increase the computational efficiency of EC-Earth, this will allow to not waste the parallel resources of the supercomputers and obtain the simulation results in an affordable time.

13. Describe the impact of the optimization work proposed? *(for Types B, C, D. Maximum 250 words)*

- Is the code widely used?
- Would the code be used only within this original research project?
- Would the code be used for other similar research projects with minor modifications?
- Would the code be used in many research projects of the research field indicated in the proposal?
- Would the modification be easy to add to the main release of the software?

The code is widely used in more by more than 22 partners and in many international scientific projects. All the developments will be reported to the repository of the model to be used by other scientists. BSC is an active member of the Technical Working Work of EC-Earth model so deeply involved in the development of the model. There are other partners which are using EC-Earth are part of PRACE projects, all these partners will benefit from the optimizations proposed here, which will be available for the community intermediately, since all the developments are done using common projects.

14. Describe the request plans for work with support from PRACE experts *(for Type C,D)*

14.a. Describe the level of collaboration with PRACE experts you have planned for and how much effort (person months) have you reserved for this?

We plan to have a full collaboration with PRACE experts to tackle the problems described in prior sections. EC-Earth is a combination on several models so the knowledge needed to improve the performance is deep. In this situation, BSC experts can be a valuable help.

Specify a rough estimate for the amount of person months this work entails : **6.00**

14.b. Describe the optimization work you expect to be done with the support of a PRACE expert for your project

We plan to apply the developments described above but also work with the PRACE experts to evaluate the impact of these developments. For this task, we need experts with high knowledge in HPC performance tools like extrae or paraver. These are profiling analysis tools which are needed in order to evaluate the main bottlenecks of parallel executions when the number of processes increase, this is the case for our project, where the overhead produced by the parallel implementation is difficult to evaluate.

14.c. Please specify the amount of PRACE experts person months required to support your project (1-6 PMs): 6.00

15. Please describe your hardware architecture currently in use to allow an optimal selection of a feasible Tier-1 system *(for Type D only, Maximum 250 words)*

☑ **I certify that I have read, understand, accept and comply with the terms and**

conditions of *PRACE Preparatory access – Call for proposals* available at http://prace-ri.eu/PRACE-Preparatory-Access

Those terms include the ones reproduced hereinafter for the sake of clarity:

The users commit to:

a. Provide to PRACE within the period established in the guide for applicants a final report, using the proper PRACE template, with the results obtained through the access to the PRACE Research Infrastructure, as well as a qualitative feedback on the use of the resources.
b. Acknowledge the role of the HPC Centre and PRACE in all publications which include the results above mentioned. Users shall use the following (or equivalent) wording in such acknowledgement in all such papers and other publications:

   « We acknowledge PRACE for awarding us access to resource [machine name] based in [country] at [site] »

   Where technical support has been received the following additional text should also be used:

   « The support of [name of person/people] from [organisation name], [country] to the technical work is gratefully acknowledged. »

c. Allow PRACE to publish the mentioned report as of one year from the termination of the allocation period.
d. Commit to collaborate with PRACE, upon its request, in the preparation of dissemination material.
e. The applicant commits to not use the project results for military purposes.