PRACE Preparatory Access for HR ARPEGE-NEMIX porting on Bullx TGCC platform

E. Maisonnave, L. Coquart, S. Valcke

WN/CMGC/11/77

1. General information

Proposal ID	PA0490
Type of proposal granted	Preparatory Access (Class A)
Period of access to the PRACE facilities	from 18/08/2011 to 30/09/2011
Name of the PRACE facility assigned	Bullx Curie TGCC CEA (fr)

2. Project information

Project name : ARPEGE-OASIS-NEMO climate model set up for high resolution climate experiments

Research field : Earth Sciences and Environment

Institutions and research team members:

CERFACS, Global Change	Mr Maisonnave, Eric Mr Bretonnière, Pierre-Antoine Mr Jonville, Gabriel Dr Valcke, Sophie Dr Cassou, Christophe Miss Coquart, Laure
CNRS, Laboratoire Ecoulements Geophysiques et Industriels	Mr Molines, Jean-Marc
Météo-France, CNRM/GMGEC/EAC	Dr Déqué, Michel Mr Piédelièvre, Jean-Philippe

Summary of the project

This project deals with preparing the ARPEGE/OASIS/NEMO climate model, in order to complement the international inter-comparison exercise (CMIP5), which modeling will provide a basis for the Intergovernmental Panel on Climate Change (IPCC) 5th Report in 2013.

The preparatory project technical objectives are porting, set up on at least 1000 cores and optimizing the 3 coupled model different components ARPEGE,OASIS and NEMO, at high resolution (720 x 360 x 31 for atmosphere, 1442 x 1021 x 46 for ocean) on PRACE architectures.

Description of the results obtained from the scientific point of view, future perspectives, benefits to our society, and the benefits of using computer resources

The ARPEGE-NEMO climate model jointly developed by ECMWF, Météo-France, NEMO consortium and CERFACS, has been compiled and run on more than 1000 cores on the PRACE tier-0 “curie” Bullx supercomputer.

Our test configuration requires high resolutions components (50km-atmosphere, ¼ degree-ocean), to study regional scale / large scale interactions. The ocean model is used on a 1D mode, as a mixed layer model (this configuration is called NEMIX), to simplify and better understand coupled processes and validate porting more easily (comparing results with those of previously validated simulations on reference supercomputers).

As expected, performances of the coupled system reach 5 days per simulated decades (see figure 1), which corresponds to the best results observed on present supercomputers. Except data management problems, this speed allows to comfortably perform decadal (and even centennial) simulations.

The figure of 1024 cores has been reached without problems (atmosphere:500, ocean:512, coupler:12). A higher parallelism could be tested on ocean model, but atmosphere model scalability seems surprisingly limited to 256 cores (see figure 2).

Use of pseudo-parallelized version of our coupler (OASIS3) is still a bottleneck for the coupled system, as 20% of the elapsed time is spend to perform interpolations and ensure communications between coupled components (see figure 3).

To try to overcome this problem, 4th version of our coupler (fully parallel) should replace OASIS3, but still needs some development to be fully operational.

A simple coupled toy model that exchanges and interpolates an analytical field for one time step between the atmospheric and ocean models grid of our high resolution configuration, was tested on the machine to validate the OASIS4 user-defined regridding (this user-defined regridding is a necessary feature in our case).

To use this new functionality, it is necessary to describe links and weights associating specific source grid points with specific target grid points in a separate NetCDF file. The OASIS4 communication library reads these links and weights and automatically performs the multiplication of the source field values by the appropriate weights in parallel on the source side and the parallel redistribution of the results directly between the source and the target processes.

Bilinear and conservative interpolations were tested for different partitions (4 to 58 processes for the atmospheric and the ocean models, with only 1 process for OASIS4).

The results obtained can be found at the web addresses:
http://www.cerfacs.fr/oa4web/projet_test_weights_dev/RESULTS_PARA/projet_weights.html

As we observed a problem with the memory when increasing the number of processes, similar tests were done without using the user-defined regridding functionality to find out if the problem was linked to the user-defined or not and it seems to be the case. We will then follow our investigations on the user-defined regridding using “valgrind” debugging tool to try to find out what is happening.

Porting issues

Coupled model porting has been affected by various problems, preferentially on atmosphere part. As a consequence of the many porting realized for several years on scalar platforms (ECMWF machines, PC, Grids, GENCI supercomputers), our models are now well adapted to Intel compilers. This experience quickens compilation phase.

Nevertheless, problems can still occur when ARPEGE-Climat parallelism increases, due to a lack of extensive tests with high resolution configurations (particularly with 2 dimensional partitioning). Some debugging has been done in the code to overcome problems detected during first steps of execution. In particular, a problem in the parallelization of the mass correction routine cormass2.F90 drove a robust bias on surface pressure field over region covered by master processor.

This problem has been solved and a six months long simulation performed on the machine. Comparing it with similar test experience done on reference NEC SX8R and SGI Altix supercomputers, no important bias could be detected in the results (mean surface pressure and temperature, water/heat flux correction). The mean heat flux correction (mean on 6 months only) is similar to reference.

We consider that our ARPEGE-NEMIX-OASIS3 model is ready to be use at high resolution for a first scientific validation.

Expected future work in the area

Nevertheless, accessing the machine on such extremely short period prevents us:

- to fully optimize the various components: different compilation options have to be tested and best performance parameters of the models have to be found (like vectorization length NPROMA for ARPEGE or ideal partitioning for NEMO)
- to better balance coupling parameters: load balancing between components, mapping (if possible) of different executables on allocated nodes (and node cores)
- to reduce output data amount, organize data post-processing and migration to local storage disk. This work is crucial for Climate Modeling, considering size of produced data (4Gb/h in our test, much more on production phase).

Scientifically speaking, NEMIX mixed layer model, if efficient to quickly validate a porting and to lead interesting processes analysis, is not able to address most of the climate modeling community problems. Implementation of the fully advective NEMO model will then be necessary.

For this purpose, NEMO ORCA025-75 vertical levels own by Jean-Marc Molines (LEGI) would replace our NEMO model. Considering similarity of the codes (same NEMO 3.2 version, NEMIX configuration only differs with from NEMO in a few routines), this switch should not be, technically speaking, too problematic but could significantly delay the scientific validation, much difficult to reach because of the higher complexity of the represented coupled phenomena, inducing regional coupled biases, wrong heat balance, etc ...

To be able to address all those aspects of the porting, we will apply for a longer term access (on “curie”, at the same time than on “jade”), taking advantage of the yearly national access program (DARI). At this stage, no extra development is needed to enhance performances of our model but its complexity requires more time to find parameters that best fit machine characteristics.

Images of the results including description or caption

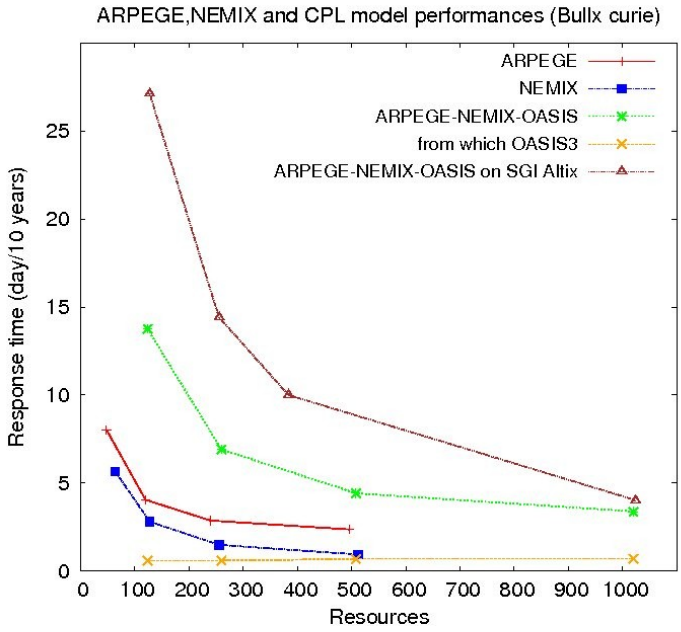

Figure 1: Performances of atmospheric (ARPEGE), oceanic (NEMO, mixed layer configuration NEMIX), coupler (OASIS3, always on 12 PE) and total coupled model (ARPEGE-NEMIX) + comparison with previous ARPEGE-NEMIX performances on SGI Altix

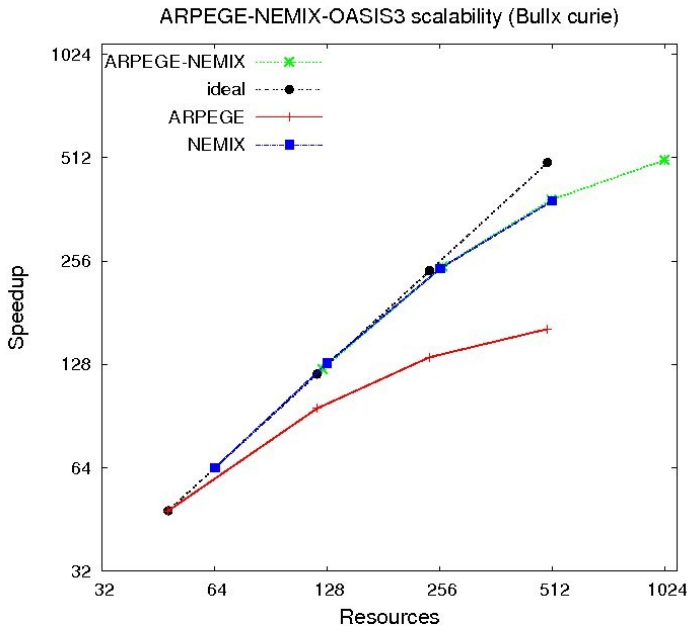

Figure 2: Scalability of atmospheric (ARPEGE), oceanic (NEMO, mixed layer configuration NEMIX) and total coupled model (ARPEGE-NEMIX)

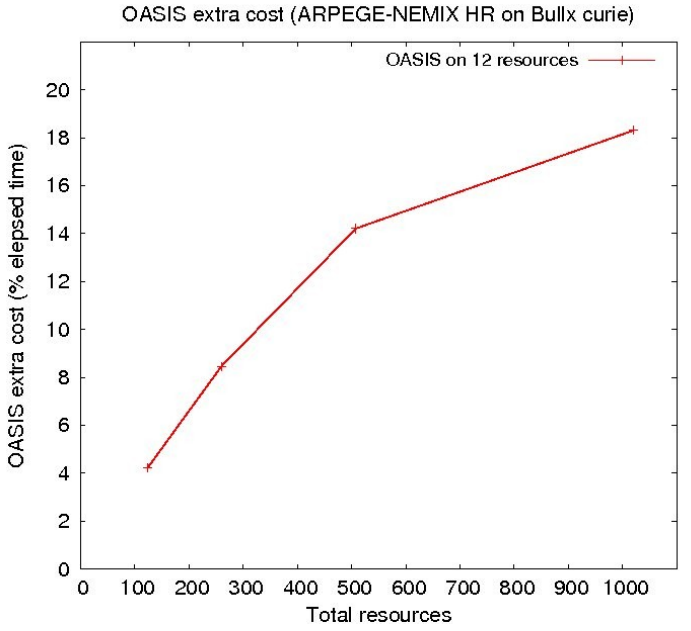

Figure 3: %age of OASIS coupling elapsed time (communications + interpolations) among total atmosphere + ocean + coupler elapsed time

3. Feedback and technical deployment

Feedback on the centres/PRACE mechanism

Preliminary porting of the model on TGCC Bullx “curie” occurs too quickly to reach the point when a particular support is necessary. A simple compilation/performance test was performed. During the second phase of the porting, a more accurate support will be necessary to better fit model parameters to machine characteristics.

Nevertheless, during this preliminary porting, we noticed that basic information (“curie.info” online users guide) was clear and easily accessible. In particular, information on MPI MPMD mode was documented (it is important for all OASIS users of Climate Community).

The “Module” system to load libraries and tools, now available on most of the European supercomputers, facilitated our work. Moreover, the “test” batch class, with sufficient resources for highly parallel problems, is open for debugging, which is particularly convenient, during this porting phase, to quickly localize problems occurring only at high parallelism level.

We also want to thanks TGCC hotline for their support when online information was not sufficient (or not read carefully enough !).

Explanation of how the computer time was used, compared with the work plan presented in the proposal. Justification of discrepancies, especially if the computer time was not completely used

About 30% of the allocated time has been used. Only a small amount of resources were necessary for debug, OASIS4 toy model tests and ARPEGE-NEMIX scalability measures. Most of the remaining CPU time was necessary to simulate 6 months of climate (needed for a first comparison with results coming from NEC SX8R and SGI Altix references).

Model porting on PRACE tier-0 supercomputers is part of IS-ENES FP7 infrastructure project and ANR SPADES (08-ANR-SEGI-025) project.