

TerrSysMP coupled model:
load balancing and further improvements
E. Maisonnave
WN/CMGC/18/33

The TerraSysMP model [1], shared by an important community in Europe, relies on the OASIS3-MCT [2] coupling library to synchronise and perform the on-line exchanges between its three components (COSMO regional atmosphere, CLM soil and ParFlow hydrology model). In the framework of the European modelling infrastructure ENES [3], CERFACS regularly provides support to help the setup and the optimisation of OASIS based coupled system, from which TerraSysMP [4,5].

In the following note, we detail the modifications we made on the Research Training Day (RTD) tutorial package (“idealRTD”), on the JURECA super computer at Jülich Supercomputing Center. We show how to simply identify the load imbalance between components, optimise the resources and speedup the model. In a second step, we propose some possible enhancements, taking benefit of the next OASIS release (OASIS3-MCT_V4.0) scheduled on April 2018.

Load balance analysis

The OASIS version currently used by TerraSysMP was modified to include the internal LUCIA instrumentation [6]. Detail of this implementation is given in Annex 1. On runtime¹, the OASIS library produces ASCII files that describe and date each OASIS exchanges. On a post-processing step, an independent FORTRAN tool processes these files and gives a summary on between which components and on how fast the exchanges were performed. In particular, it is possible to (i) determine the sequence of exchanges between the three components (see Illustration 1), (ii) estimate for each component the total time spent to wait an information from another component (coarse grain load imbalance, see Table 1) and (iii) measure the time spent by each component to perform their OASIS interpolation(s) (see also Table 1).

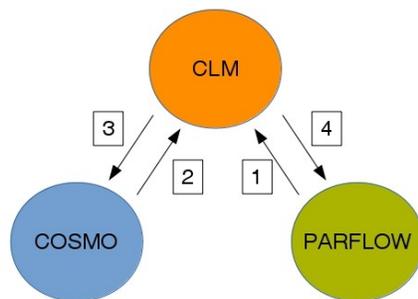


Illustration 1: Coupling exchange sequence

¹ when the specific OASIS parameter 'NLOGPRT(2)' is set to -1

A first rough analysis of these results would be that CLM is the fastest model and is waiting the results of COSMO and ParFlow computations (that are performed concurrently). Consequently, it is possible to re-estimate the distribution of resources between components, in order to better synchronise COSMO and ParFlow. The new distribution and its results are described in Table 2. The total restitution time for the whole TerraSysMP coupled model is reduced from 93 to 73 seconds (using the same amount of resources – 1 node).

Component	# MPI process	Calculations (s)	Waiting time (s)	Interpolations (s)	# coupled time steps
COSMO	30	51.33	37.78	3.03	598
PARFLOW	4	81.00	8.11	2.51	598
CLM	4	8.38	80.72	2.14	598

Table 1: Per-component performances (total calculation time, from which OASIS interpolations, and waiting time) before optimisation

Component	# MPI process	Calculations (s)	Waiting time (s)	Interpolations (s)	# coupled time steps
COSMO	18	57.40	12.86	4.63	598
PARFLOW	16	59.92	10.34	3.04	598
CLM	4	5.73	64.53	0.86	598

Table 2: Per-component performances (total calculation time, from which OASIS interpolations, and waiting time) after optimisation

A finer analysis would be necessary to find the distribution that minimize restitution time (or maximize the parallel efficiency of the component). Detailed strategies are described in [6]. This analysis would also need to take into account the mapping of each MPI process, a parameter which is particularly important when hyper-threading mode is on (as it is the case on JURECA).

The aim of this working note is not to provide this optimal number, but to demonstrate that a simple instrumentation exists in OASIS, that allows to speedup any TerraSysMP, without any advance knowledge on computing or any extra instrumentation of the code (using i.e. SCALASCA). This result would be useful for beginners (the tool is already included in the RTD package) but can also contribute to save a significant amount of computing time if used routinely by scientists dealing with high resolution configurations or long simulations.

Proposed enhancements

Interpolations from grids that are identical are not mandatory in OASIS3-MCT, and the chosen

nearest-neighbours and bilinear can even introduce an error. The corresponding cost, which is non negligible (see Tables), could be avoided, at least between COSMO and ParFlow 2D grids. For CLM grids (1D), a single nearest-neighbour interpolation would be the most efficient solution.

The current OASIS version of the RTD package can possibly be updated. With the current v3.0, it is already possible to simplify the namcouple header: the SEQMODE, CHANNEL, JOBNAME, NBMODEL, INIDATE, MODINFO, CALTYPE are no more taken into account. In addition, it is no more necessary to force exchange sequentiality, which means that the SEQ variable doesn't need to be specified in the namcouple. In addition, with the future OASIS v4.0 (to be released in spring 2018), it will be possible to define a 3D exchange with only one coupling field entry + the number of levels of the 3D variable.

References

- [1] Shrestha, P., M. Sulis, M. Masbou, S. J. Kollet, and C. Simmer, 2014: "A scale-consistent terrestrial systems modeling platform based on COSMO, CLM, and ParFlow", *Mon. Wea. Rev.*,142,3466–3483, doi:10.1175/MWR-D-14-00029.1
- [2] Craig A., Valcke S., Coquart L., 2017: "Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0", *Geoscientific Model Development*, 10, pp. 3297-3308, doi:10.5194/gmd-10-3297-2017
- [3] Mitchell J., Budich R., Jousaume S., Lawrence B. and Marotzke J., 2012: "Infrastructure strategy for the European Earth System Modelling community 2012-2022", *ENES Report Series 1*, 33 pp. <https://doi.org/10.5285/ca90b281d6ff4cffb9a9bbdeb5fa63f3> .
- [4] Maisonnave, E. and Valcke, S., 2012: "[OASIS Dedicated User Support 2011, Annual report](#)", Technical Report, **TR/CMGC/12/18**, SUC/CERFACS, URA CERFACS/CNRS No1875, France
- [5] Maisonnave, E. and Valcke, S., 2017: "OASIS3-MCT, a coupling software for climate modeling", *International Fall School on Terrestrial Modeling and High-Performance Scientific Computing*, 25-29 September, Bonn, Germany. Meteorological Institute, Bonn University
- [6] Maisonnave, E., Caubel, A., 2014: "[LUCIA, load balancing tool for OASIS coupled systems](#)", Technical Report, **TR/CMGC/14/63**, SUC/CERFACS, URA CERFACS/CNRS No1875, France

Annex 1: LUCIA implementation on RTD framework

A - Original routines copied from

slts26@jureca.fz-juelich.de:\$HOME/RTD/terrsysmp/oasis3-mct/lib/psmile/src and

slts26@jureca.fz-juelich.de:\$HOME/RTD/terrsysmp/blsava/intf_oas3/oasis3-mct/tsmp :

mod_oasis_advance.F90, mod_oasis_data.F90, mod_oasis_method.F90 ,

mod_oasis_coupler.F90 & mod_oasis_grid.F90

to

\$HOME/RTD/terrsysmp/blsava/intf_oas3/oasis3-mct/tsmp

B - Post-processing tool 'lucia' directory added to

slts26@jureca.fz-juelich.de: \$HOME/RTD/terrsysmp/oasis3-mct/util/

- Compiling of the FORTRAN program with default ifort compiler and executing:
\$HOME/RTD/terrsysmp/oasis3-mct_JURECA_1.2.0MCT_clm-cos-pfl/util/lucia -c

- the LUCIA post-processing analysis is performed on the running directory

\$WORK/run_pft15_sID2_Sv0.35 :

launch \$HOME/RTD/terrsysmp/oasis3-mct_JURECA_1.2.0MCT_clm-cos-pfl/util/lucia