OASIS4 coupling interface implementation on ETHZ' land-atmosphere coupled model

E. Maisonnave

WN/GCCM/11/15

Abstract

During the 5th mission granted by the IS-ENES FP7 project, held from Nov 16 to Dec 9 2010 at ETH, Zürich(Switzerland), I implement and validate an OASIS4 interface for a regional atmosphere-land model, in collaboration with my host, Dr Edouard Davin.

A regional atmosphere (COSMO-CLM, DWD) and a land scheme (CLM, NCAR) model have been coupled with OASIS4, at low resolution on a MPP scalar machine (on 100 cores), to simplify version updates, allow the use of different time stepping and different grids at different resolutions, prepare a plug in of other model components.

Strong slowing down is observed in OASIS4 exchanges when core number increases. This problem, occuring on a highly particular configuration of grids and partitioning, has to be further investigated.

Table of Contents

Model / machine description:..4
Rationale...4
Implementation on models..5
Results/performances..6
OASIS developments..9
Needed improvements..12

Model / machine description:

COSMO-CLM (here called COSMO)
This regional atmosphere model (v4.8, and its climate version, v11) is used by a large community in several central Europe countries (in which ETHZ). DWD and several other meteorological agencies host the operational version of the model. Grid size: 109x121x32, 0.44 degrees. Parallelisation reaches 100 MPI tasks on the targeted supercomputer.

CLM
This land scheme is developed at NCAR (v3.5). It is used within the integrated CCLM climate model. Initially, CLM is launched on the same grid that COSMO.

The model is available on CRAY XT5 supercomputer, with 22,128 compute cores (2 six-core AMD Opteron 2.4 GHz Istanbul processors per node), CRAY SeaStar 2.2 interconnect. Peak performance of 212 Teraflop/s. The machine is located at CSCS, Manno, Ticino, Switzerland.

Rationale

This user support task proposes to upgrade an existing coupled system with the latest version of the OASIS coupler.

The previous coupling (called integrated coupling) gathers two models on one executable.

CLM reads input file describing its grid. This input file is written once by COSMO and contains COSMO grid specifications: CLM grid points are COSMO land points. CLM and COSMO take the same number of processes (used sequentially by one model and the other) which means that CLM and COSMO partitioning differs (CLM only process land points, COSMO both land and sea points). Exchanges between models do not need interpolation (CLM and COSMO land points are located at the same place), just communications between processors (a grid point could be located on different processes during CLM or COSMO computation). CLM is called as a subroutine of COSMO at each time step (sequential coupling).

COSMO could be launched on more processes than its number of subdomains to use independent IO processes (disabled with OASIS coupling). Subdomains can be re-organized on a cartesian grid (disabled with OASIS coupling).

Strong loss of scalability observed with COSMO-CLM coupled model, compared to COSMO stand alone mode, is probably due to too frequent calls to CLM (and coupling routines), on too many processes (same number than COSMO).

OASIS coupling is evaluated for its capacity to:

1. non intrusively be implemented on COSMO and CLM codes
2. let user choose the best time step for each model
3. launch models on a different number of processors (best number according to models distinct scalability)
4. investigate possibility of non sequential coupling

Implementation on models

To let users the choice of the coupling method, we keep the possibility to choose at compilation stage (by CPP key) between stand alone, existing coupling method (COSMO calls CLM as a subroutine) or OASIS4 coupling.

As previously implemented in several models (K.Mogensen - ECMWF), an as-separate-as-possible interface as been written. All the prism routines (psmile library) are called in newly created routines. These few routines (6) are called in the model:

- at the very beginning of the simulation (MPI communication setting)
- at the end of the initialization phase (partition and grid declaration to OASIS)
- at the beginning and at the end of the temporal loop (coupling field exchange, OASIS counter update)
- at the end of the simulation (MPI closing)

<u>CLM interface</u>

One of the interesting question of this user support was to test how difficult it could be to plug OASIS in an module originally designed for such integrated coupling as NCAR's one. We decided to start from the stand alone configuration of CLM. Original set up internal MPI communication is done with MPI_COMM_WORLD communicator: we simply changed this communicator with the local communicator provided by OASIS.

Due to a lack of appropriate option in present OASIS4 partitioning (see OASIS development paragraph), coupling fields have to be gathered on the master processor. Consequence on coupling performances has to be evaluated but is obviously an issue for further massively parallel configuration setting.

The possibility to define distinct grids for regional land and atmosphere models implies necessarily a geographic mismatch between global domains: some grid points of the larger grid cannot receive information from the narrowest grid.

A redefinition strategy for the largest subdomain (CLM) has to be designed. In the present User Support solution, CLM has to be coupled first with a domain larger than CLM's one (CLM latitude and longitude limits have to include COSMO limits). The first received field has to be saved and the simulation stopped. A new mask is deduced from this interpolated coupling field and CLM is restarted with this new mask.

Interface routines, driving exchanges with OASIS, are really non intrusive. As usual in such kind of implementation, the off-line mode routines which read forced fields in external files are switched off and replaced by our coupling fields receiving routines. Coupling fields sending is called as soon as coupling fields are available.

COSMO interface

The main originality of OASIS interface implementation lies in the possibility to involve a subset of model processes in the coupling, some of them providing no information to land model (all grid points are masked ocean grid points).

At definition stage, prior to any prism_def operation, unmasked point number is calculated and OASIS definition set only if this number is non zero.

COSMO grid point centre definition starts from a regular grid defined on both sides of equator, and then translated to the targeted region: a routine has been developed to transform identically grid corners.

OASIS exchanges with CLM replace a subroutine call in COSMO. prism_put and prism_get (in this order) are called one immediately after the other.

Gather/scatter operations needed in the previous integrated coupling are now disabled, and communication time saved at this stage.

Results/performances

Due to persistent bad mood of OASIS mpp_io unbedded output library, and thanks to Moritz Hanke (Max Planck Institute), an alternate netcdf based parallel output algorithm has been implemented in both interfaces. Coupling fields are written (and overwritten) at each coupling time step.

Those fields could be:

- used to re-built the CLM adjusted mask (see above)
- used to check their validity at implementation stage
- compared to arrays exchanged in the previous coupling at validation step.

Two examples of CLM and COSMO received coupled fields, produced after 17 days of simulation are shown on figure 1.

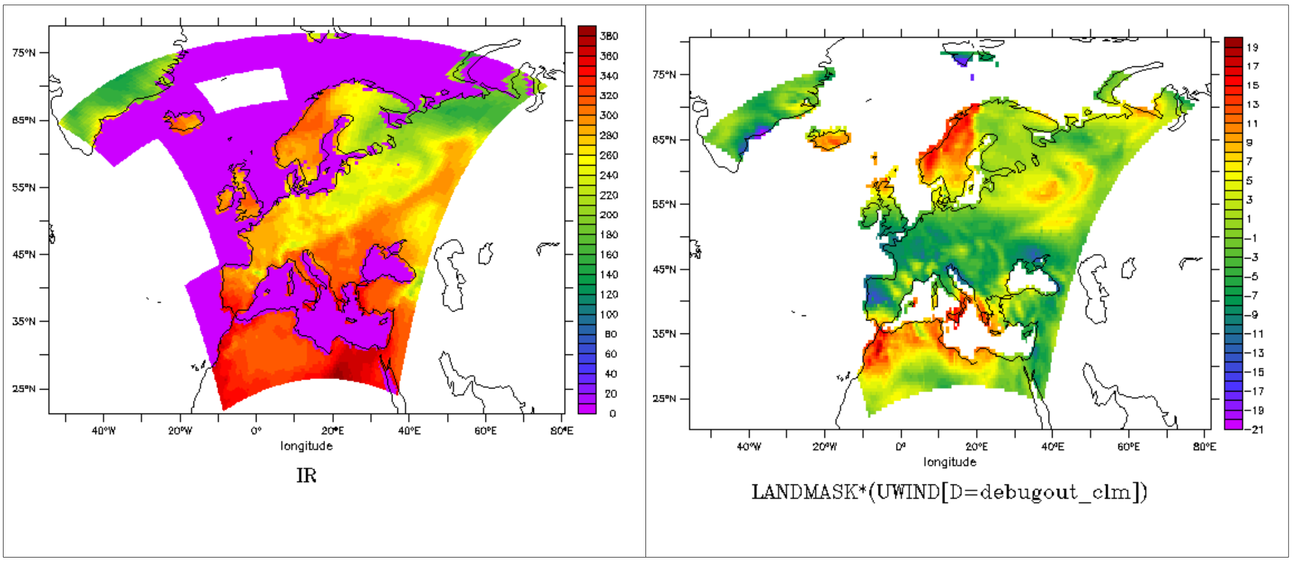

Figure 1: example of COSMO-CLM coupled fields: Infrared on COSMO grid (left) and zonal wind on CLM grid (right) after 17 days

Geographic discretization of CLM and COSMO can be now totally independent. OASIS performs an interpolation (bicubic) between those two grids. It means that a finer resolution can be used on one or the other model.

Also interesting for performance improvement: each model parallelism can be set at its own optimal level. Saved resources compensate the fact that both COSMO and CLM models need their own processors [1].

Performances also take benefit to the possibility to set different time step for each model (and coupling time step different from model time steps).

The newly OASIS4 interface is now able, with some additional development, to be plugged to other models used by COSMO and COSMO-Climate communities. NEMO or ECHAM are possible candidates to complement the regional system model. At the end of the user support period, Andy Dobler (Francfort University) starts with us an

1 On CSCS machine, or on every machine where number of processes on one core is limited to one, it is impossible to launch processes of the two executable on the same resources. If models are called sequentially, a lot of computing power is lost with processors employed by processes waiting the end of the other model calculations.

adaptation (following NEMO example) of the interface for OASIS3 coupling COSMO to a Mediterranean NEMO configuration.

Concerning performances, figure 2 shows scalability of previous integrated coupled model (red curb) and newly implemented OASIS coupled configuration, using the same time step (240s) on both models and the same coupling frequency (cyan curb), or decreasing until 1 hour land model and coupling frequencies (blue curb). For OASIS based configurations, the resources number is the total number of cores used for both models and coupler. 12 cores (1 node) are devoted to OASIS, 12 cores to CLM (land model calculations are much less expensive than the atmosphere ones).

Compared to previous integrated configuration, a strong slow down is observed, mainly due to coupling duration (about the same than needed for one time step calculations). The poor scalability is a consequence of the OASIS coupling duration, which increases with model parallelisation (even when model processes are not involved in the coupling, see OASIS development section).

Reducing the coupling time step, OASIS slow down is less visible and curb fits COSMO scalability. Response time is better than with the previous coupling. But CLM is called less often, and changes on model results have to be evaluated to consider this configuration equivalent to the existing one.

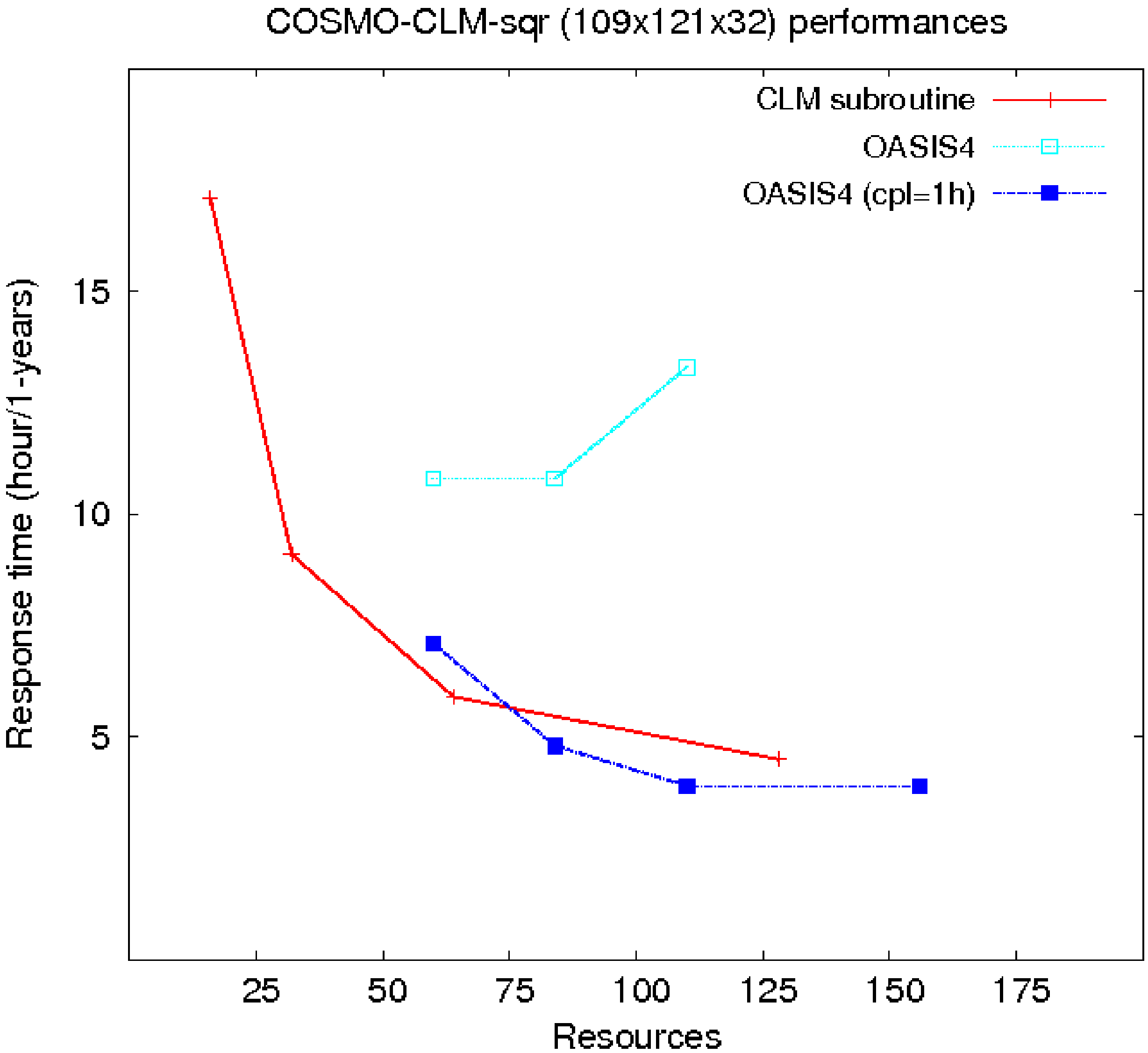

Figure 2: COSMO-CLM-OASIS4 coupled model performances on CRAY XT5

OASIS developments

User support tasks also explore capability of OASIS coupler to satisfy the complex community needs.

A strong day by day collaboration with OASIS developers allows us to identify and adresss or bypass related issues.

A first analysis of CLM partitioning let us identify a restriction on OASIS use for some rectangular grids. CLM calculations only occurs on land points and are independent: those points can be considered as a 1D vector, split onto processes without any geographic consideration (same kind of partitioning could be observed on models like LIM sea ice or SURFEX land scheme).

For the moment, OASIS assumes that (i plus 1) and (i minus 1) points are geographically neighbours, which is not the case on our partitioning. A non trivial development is needed to address this issue. The temporary solution implemented (exchange of gathered field on master processor) can be considered as a major bottleneck for further high resolution (and massively parallel) CLM configuration.

In our particular partitioning, some subdomains are not involved in the coupling because they don't intersect any unmasked points (COSMO) or also because they are not the master processor (CLM).

During coupling definition phase (halo detection), we identified some OASIS incapacity to consider separately MPI communications of process involved in the coupling from MPI waiting state in which non involved process lies. This problem has been addressed on-the-fly by OASIS developers.

Trying to validate coupling fields, OASIS outputs have been turn on without any success, because impossibility our mpp_io library implementation to manage processes not involved in the coupling.

This restriction, also concerning restart read/write operations, prevents to consider further test on parallel use of land and atmosphere model (instead of sequential calls of both models).

To let us visualize coupling fields, a workaround solution has been suggested by OASIS team and implemented, using parallel call to Netcdf within model interfaces (see above).

The last problem identified during this support task concerns abnormal coupling duration and its increase with number of model subdomains, even when those processes are not involved in the coupling (see table 1).

CRAY XT5 (12 cores/node)				SGI Altix (8 cores/node)			
OASIS	CLM	COSMO	Comm(s)	OASIS	CLM	COSMO	Comm(s)
12	12	36	0.20	8	8	32	0,01
12	12	60	0.30	8	8	64	0,02
12	12	96	0.37	8	8	128	0,05
12	12	132	0.43				
12	12	60	0.30	8	8	64	0,02
12	36	60	0.33	8	64	64	0,53
12	60	60	0.56	8	128	64	1,80

Table 1: mean duration of a total OASIS coupling sequence

Further investigations are needed to identify which service can be at the origin of the slowing down. A toy model is implemented to try to reproduce the problem, without success. This could suggest that an interaction between model calculations and MPI communication could occur in this configuration (memory ? MPI buffers ?)

The coupled model is ported on an SGI Altix platform and the same performance tests processed. If calculations are achieved during the same time with about the same number of processors, the time spent within OASIS calculations and communications is significantly higher (one order of magnitude, see the 4 right columns of the table 1).

But again, if we increase the number of CLM processors (not involved in the coupling), this duration amazingly increases. This identical behaviour on both machines suggests that the particular CSCS MPI installation on the CRAY machine (or some default MPI characteristics) is not at the origin of the noticed slowing down.

An instrumentation of every code is performed and a profiling can be checked with Paraprof tool.

Figure 3 describes the time spent on the main routines for OASIS (node 0 to 11), COSMO (node 12 to 47) and CLM (node 48 to 59).

With red, blue and green colours, MPI routines (Receive, Barrier, and Bcast) are much more represented than all other computation routines. Even though any model are waiting the calculation end of the other one, the total elapsed time of the simulation exceed the addition of CLM and COSMO routines calculation time.

At the same time, OASIS interpolation calculations appear negligible.

This suggest that the observed slowing down is mainly due to OASIS communication routines. A more accurate analysis should begin in collaboration with G. Piccinalli at CSCS, using Scalasca tracing tools to better understand the origin of the problem.

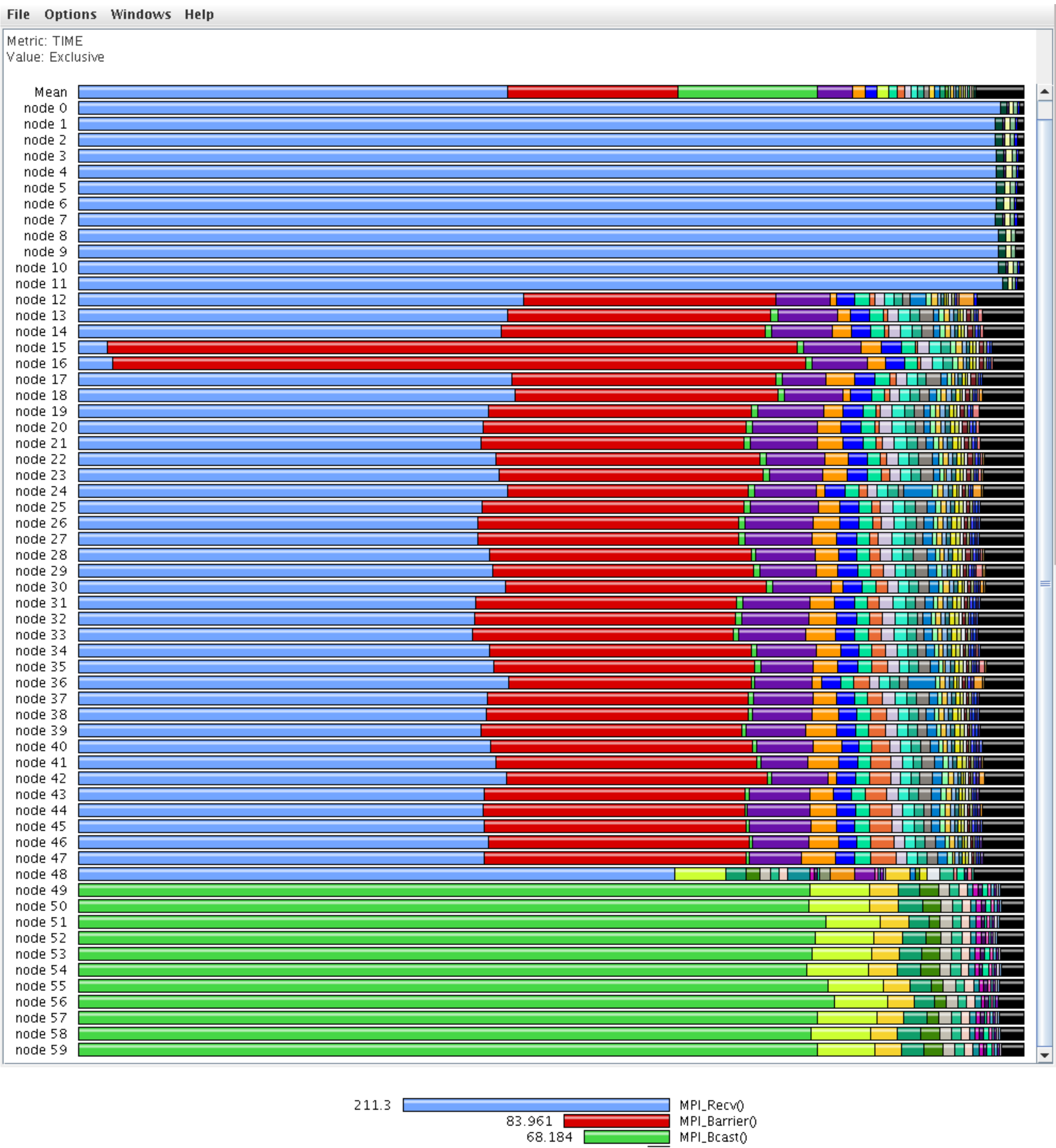

Figure 3: Paraprof profiling of a COSMO-Climate/CLM/OASIS4 run

Needed improvements

On top of those OASIS improvements, some additional tasks have to be tackled to be able to guarantee an iso-functional and efficient coupling between COSMO and CLM models.

Needed on every coupled system exchanging fluxes, conservative interpolation has to be tested for flux exchanges.

Then, coupled fields have to be compared with those of integrated coupling. Characteristics of those fields have to be checked when coupling field frequency decreases. A special care should be taken on the behaviour of exchange coefficient, recomputed in the atmosphere when fluxes are changed by land model. Models characteristics such as diurnal cycle or long term means should be compared.

An upgrade of CLM (v4) is planned. Even more integrated in an ESM structure, a new problem could occur during MPI communicator definition.

Finally, scalability tests with higher definition (operational) models could begin, to possibly investigate limitations of OASIS4 parallelism with such configuration.

Thanks to Edouard Davin, Sonia Seneviradne, Anne Roches, Olivier Fuhrer (MeteoSwiss), Jean-Guillaume Piccinalli (CSCS) for their strong support and the constant interest for our work. Once again, thanks to our four OASIS developers.