

DRAFT

**Ocean/sea-ice  
macro task parallelism in NEMO**

E. Maisonnave, S. Masson\*

\* LOCEAN-IPSL, Université Pierre et Marie Curie, CNRS, France

**TR-CMGC-15-54**

## Abstract

*In this document, we describe how to split NEMO into two components, ocean on one hand and sea-ice plus other surface processes on the other hand. The two separate executables produced with two separate FORTRAN routine sets are communicating via the OASIS coupling library, which ensures cross-component communications, and possibly interpolations, between sub domains of the MPI-parallel components. Bias in sea-ice volume is reduced compared to a standard NEMO configuration calling sea-ice model every 5 time steps. For some configurations (ORCA025), simulations can be performed faster, and at lower cost.*

Using NEMO [1], including sea-ice model, in any standard configuration implies that (i) ocean and sea-ice calculations are performed sequentially, (ii) on the same grid, (iii) following the same parallel decomposition and that (iv) both codes must be jointly maintained. In this document, we describe how to split NEMO into two pieces, ocean on one hand and sea-ice plus other surface processes on the other hand. The two separate executables produced with two separate FORTRAN routine sets are communicating via the OASIS [2] coupling library. It means that they are independently parametrized. The execution ordering (sequential/concurrent) is simply controlled via the OASIS parameter file, as well as, if needed, interpolation choice between different grids. OASIS automatically ensures cross-component communications between sub domains of the MPI-parallel components.

## 1- Implementation

### 1.1 Modification of existing configurations

The two new executables are built from two existing configurations of NEMO: a standard ORCA\_LIM3 (NEMO) and a stand-alone version of the surface module (SAS).

The LIM3 sea-ice model is removed from the original NEMO configuration (CPP `key_lim3` removal and routine list reduced to `OPA_SRC` in `cfg.txt` file). The single OPA model is then supposed to receive all fluxes from the other executable (SAS): a neutral forcing option (`jp_none`) is added in the `sbc` routine so that no forcing file or atmosphere coupling fields are provided. This new configuration is called *OPA* in this report.

The other starting configuration is less known. SAS is a testing tool, commonly used to force a sea-ice model with ocean surface fields and a set of atmospheric variables for flux calculations. SAS routine list includes `OPA_SRC`, `SAS_SRC` and `LIM_SRC_3` in `cfg.txt` file. In addition to bulk formulas, SAS can enable any component called in the `sbc` routine (icebergs, waves ...). In our implementation, the ocean surface field reading in files is disabled (no more call to `sbc_ssm`) because these surface fields are now transmitted by the OASIS coupler. This new configuration is called *SAS* in this report, the new double executable system *OPA-SAS*, in contrast with the existing single executable *NEMO* configuration.

OPA source files are taken from the standard NEMO configuration to NEMO. SAS also shares most of its source file with NEMO; only `nemogcm.F90`, `step.F90` and `sbc_ssm.F90` are different and modified by our implementation. Sea-ice namelists are only needed by SAS model. At the opposite, both SAS and OPA models require OPA reference and configuration namelists, which are duplicated and suffixed for SAS. OPA and SAS namelists mainly differs from the new parameter `nn_components`, an integer that can be set to (i) `i_am_nemo` for a standard use, (ii) `i_am_opa` or (iii) `i_am_sas` to communicate to the common OPA/SAS routines which component is executed. In addition, OPA namelist is defined with a `nn_ice` to 0 (SAS: 2) and `ln_blk_core` to FALSE (SAS: TRUE).

As any other OASIS coupled system, a parameter file, called `namcouple`, has to describe characteristics of coupling for each exchanged quantities, such as ID, coupling frequency, interpolation. This last topic do not have to be developed if OPA and SAS are sharing the same grid. In addition, coupled variables are saved in OASIS restart files. For the very first simulation, these quantities must be build (from NEMO and LIM restarts, for example) and provided at run start.

Finally, we mention that both SAS and OPA can use the XIOS output server. Two contexts must be declared in the same `iodef.xml` configuration file (as for any OASIS coupled components using XIOS concurrently<sup>1</sup>). SAS only outputs surface fields, that are no more output from OPA. XIOS executables are shared by the two components and perform alternatively SAS and OPA outputs.

## 1.2 Coupling interface

### 1.2.1 ocean/sea-ice coupled system

A new coupling between OPA and SAS components is highly facilitated by the existence of an OASIS based coupling interface, originally made for communication with atmosphere models. Our coding strategy consists in exchanging ocean variables from OPA to SAS and fluxes from SAS to OPA via the existing `sbc_cpl` routine. This routine, where exchanges with atmosphere are located, is already included in OPA but must be added in SAS. In both components, reading of initial coupling field selection in `namscb_cpl` namelist is disabled and a new set of coupling fields is automatically defined (see Table 1)

OPA to SAS	SAS to OPA
SST	Solar flux
SSS	Non solar flux
Currents (i,j components)	Evaporation-Precipitations
SSH	<i>Salt flux</i>
<i>First level thickness (opt)</i>	Wind stress (i,j components)
<i>Fraction of solar net radiation absorbed in the first ocean level (opt)</i>	Wind stress module
	Ice fraction

Table 1: List of OPA/SAS coupled variables. In italic, variables added to the initial atmosphere/ocean coupling field list. First level thickness (for variable volume VVL) and fraction of solar net radiation absorbed in the first ocean level (light penetration for ice-model) are optional

Within SAS, the incoming surface variable are also received at time step loop start. They fill the standard coupling arrays for surface variables (`sst_m`, `sss_m`, `ssh_m`, `ssu_m`, `ssv_m`). Bulk formulas produce the fluxes, that are used and transformed by the sea-ice model. The resulting fluxes are transmitted to OPA at time loop end (`sbc_cpl_send` routine).

Within OPA, the incoming fluxes are received at time step loop start. During this surface forcing reading phase, the variables directed to SAS are still filling the standard coupling arrays in `sbc_ssm`: they are needed in some routines (calculating runoff and damping)

<sup>1</sup> See for example IPSL-CM6 CGCM set up

that are not activated in SAS, but OPA. In addition, these quantities are transmitted to SAS at time loop end (*sbc\_cpl\_send* routine).

### 1.2.2. Ocean/sea-ice/atmosphere coupled system

In our implementation, bulk formulas forcing can be substituted by a full atmosphere model. In this case, a change in SAS namelist file is necessary: *ln\_cpl* is activated instead of *ln\_blk\_core* and *nam\_sbc\_cpl* has to be filled taking into account the atmosphere/SAS coupling specificity; this namelist won't be disabled like in ocean/sea-ice only configuration but the coupled fields exchanged with the atmosphere will be added to the ocean/sea-ice ones.

Consequently, a new *namcouple* must be designed to gather both SAS/OPA and SAS/atmosphere coupling specificity. Be careful that the atmosphere component is only exchanging information with SAS. The OPA/SAS exchanges remain the same in ocean/sea-ice or ocean/sea-ice/atmosphere configurations. Some quantities can be exchanged twice:

- fluxes are sent by the atmosphere and received by SAS, then send (after modification by sea-ice) by SAS and received by OPA. Compared to the ocean/sea-ice configuration, names of SAS sent fields and OPA received fields do not change, but SAS received field name prefixes 'O' are replaced by 'S' to let OASIS distinguish between them.
- SAS has to send twice the sea-ice cover to atmosphere and ocean. This coupling field definition is duplicated and the corresponding quantities are considered as different coupling fields by OASIS.

### 1.3 Coupling algorithm

In the original NEMO configuration (called "N"), bulks formulas, sea-ice and ocean calculations are performed sequentially. Our new OPA/SAS configuration is theoretically able to reproduce the same sequential algorithm. This configuration was used for validation purpose. It is not available in the configuration toolkit, but can easily derived from it by users familiar with OASIS<sup>2</sup>. It is called "OSS" for "Sequential OPA/SAS" in the document. N and OSS configurations are supposed to exactly share the same algorithm. However, our goal is a modification of the calculation sequence, in order to perform in parallel flux and sea-ice calculations in SAS on one hand and ocean calculations in OPA on the other hand. Speed gain of this macro task parallelism will be presented in § 3. This concurrent configuration of our OPA/SAS coupled system is called "OS" in the document.

---

<sup>2</sup> *namcouple* file must be adapted and "before" surface current exchanged instead of "now" values in *sbc\_cpl\_snd* routine

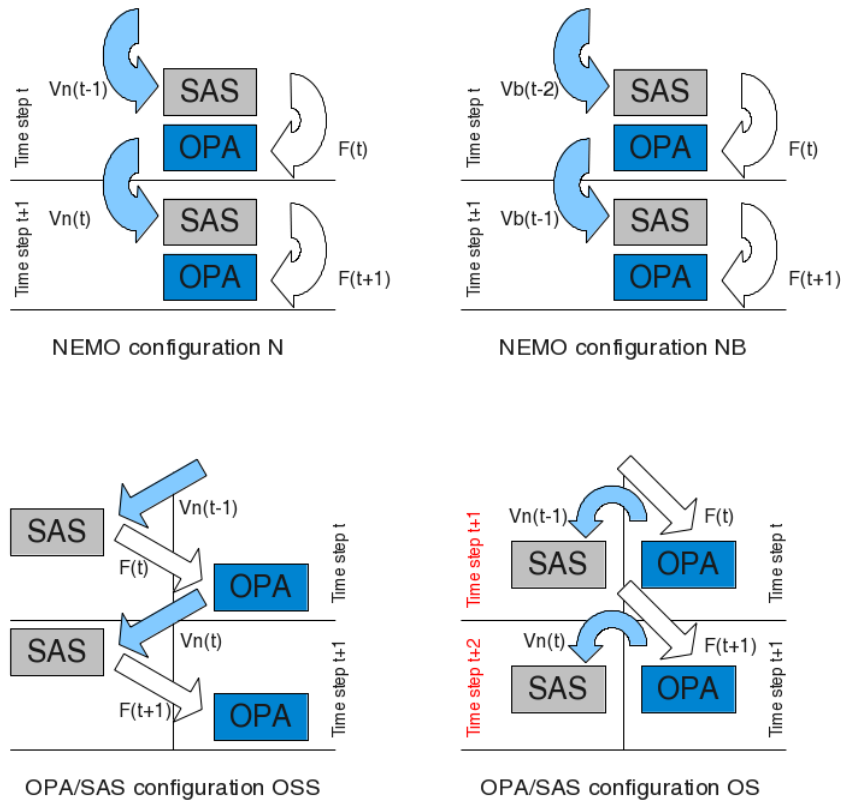


Figure 1: Algorithms, exchanged sea surface variables ( $V$ ) and fluxes ( $F$ ) for standard NEMO (N), shifted NEMO (NB), sequential OPA/SAS (OSS) and concurrent OPA/SAS (OS) configurations

When OPA and SAS are running concurrently, ocean is no more receiving fluxes calculated at time step  $t$  but  $t-1$ . To provide in OS the same fluxes than in N, SAS time stepping is shifted of one time step: when OPA is performing time step  $t$ , SAS is performing time step  $t+1$ . Thus, OPA is again receiving time step  $t$  fluxes at time step  $t$ . Practically, this time step shift operation does not need a change in SAS time step bounds, but all SAS forcing of time step  $t$  are read in files with a positive offset of 1. For their part, the ocean surface quantities received by SAS at time step  $t+2$  in OS are quantities calculated at time step  $t$ . There is a shift compared to N configuration, whose consequences are presented at § 2.

To help this analysis, the N coupling algorithm can be modified to reproduce, with a single executable, the OS algorithm. Two time step results of coupling surface variables are permanently available in the ocean code: results of previous time step  $t-1$  (“now” quantities) and results of time step  $t-2$  (“before” quantities). In N configuration, filtered  $t-1$  quantities are send to the sea-ice model<sup>3</sup> at time step  $t$ . In OS, at same time step,  $t-2$  quantities are received by SAS. Then, a shift in exchanged quantities is necessary and that is why “before” quantities are send by OPA instead of “now”. This configuration is called “NB” for “NEMO coupling before quantities”. OS and NB are supposed to exactly share the same algorithm. Again, this configuration is not include in the standard toolkit to avoid complicating source code and creating error possibilities.

3 Except for surface velocities (filtered  $t-2$ )

## 2. Validation

Three coupled systems are tested for validation. All are using 3.6 NEMO release. A first ocean/sea-ice configuration including LIM2 in SAS is not shown in this report. The same configuration including LIM3 and a LIM2+atmosphere coupled system are described in this paragraph. OPA and SAS are sharing the same ORCA2 grid. The ARPEGE atmosphere model grid is the Gaussian reduced T127.

An initial 10 year long N<sup>4</sup> simulation is performed to reach a first level balance in general circulation. The sea-ice model is called at each ocean time step. All presented simulations are starting from this spin-up phase. OASIS initial coupling field are built from ocean restart files.

We chose to base our validation on the unique sea-ice volume variable. It appears that this quantity is able to cumulate small errors in ocean/sea-ice coupling. Errors from different types are traced by Arctic and Antarctic: the sea-ice volume is shown on these two regions separately. A two year long test is found sufficient to reveal main errors and biases.

To avoid any of the coupling errors that our OPA/SAS configuration make possible, no runoff are prescribed, nor damping, and light penetration of solar flux is disabled. For an optimal use of these options, that are already provided by our implementation, please contact your NEMO system team support.

### 2.1. Ocean/sea-ice

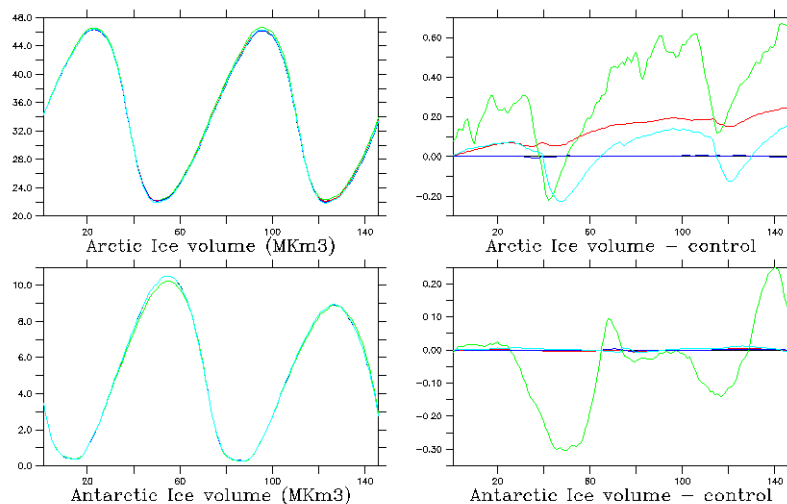


Figure 2: Two year long evolution (day pentads) of sea-ice volume (in million of Km<sup>3</sup>) for NEMO and OPA/SAS, including LIM3. The two upper (lower) plots are showing Arctic (Antarctic) volumes. The two left (right) plots are showing raw (difference with reference) quantities

In Figure 2, sea-ice volumes are presented for reference and tests configurations. The reference (N) is a two year long extension<sup>5</sup> of the spin-up phase (in black). In green, the

---

4 single executable using the released namelist

5 However, some namelist parameters are changed (runoff, damping), that can explain Antarctic drift in reference mode

same configuration (called N5) with a sea-ice model call every 5 time steps. In red, the reference configuration modified to provided 'before' quantities to the sea-ice model (NB). OSS, the sequential OPA/SAS configuration, is presented in dark blue and its concurrent version (OS) in light blue.

The error induced by the OASIS coupling is negligible for Antarctic, and for Arctic in sequential mode. In this region, the concurrent mode has a smaller impact than a 5 time step coupling.

## 2.2. Ocean/sea-ice/atmosphere

As shown in Figure 3, OS errors are significantly smaller than N5 when the atmosphere is added to the coupled system. Even if both are exhibiting larger errors (atmosphere/ocean initial bias) than in non-coupled-atmosphere configurations, one can estimate that our implementation (OS) does not change significantly or even enhance the standard atmosphere/ocean (N5) coupling quality.

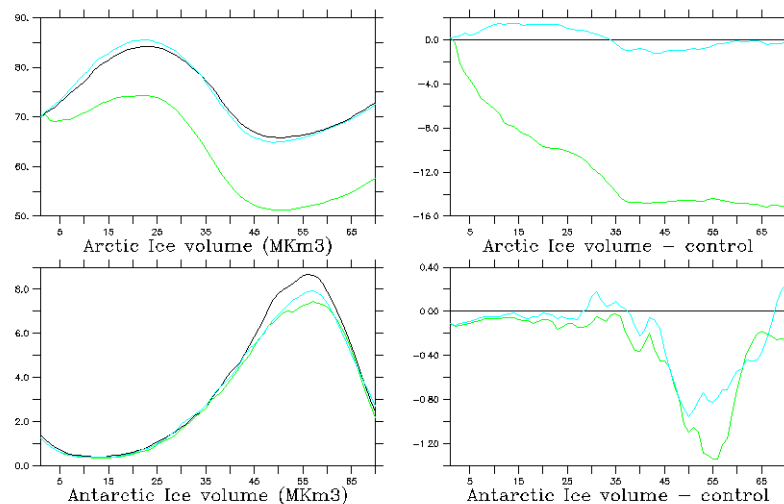


Figure 3: Same quantities than on Figure 2, for a one year long simulation, using LIM2 and coupled with ARPEGE atmosphere model. In this plot, NB and OSS results are omitted

## 3. Performances

When OPA and SAS are running concurrently, they are performing their operations in parallel and then, speeding up the simulation time to solution of a maximum factor 2. Nevertheless, more resources are needed to operate the extra executable, which can lead to additional costs.



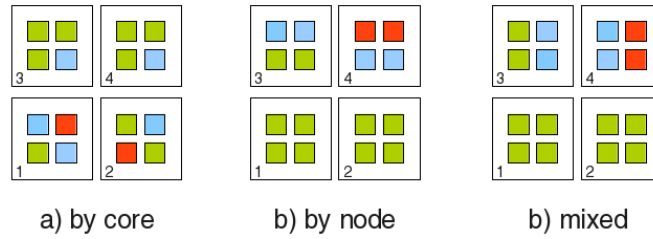


Figure 4: Different binding of OPA (green), SAS (blue) and XIOS (orange) processes on four four-core nodes

Measurements are done with two configurations of OPA/SAS without atmosphere, on two different grids: ORCA2 and ORCA025. Speed (in simulated years per day, SYPD) and cost (in computing hours per simulated years, CHSY) are evaluated in the TGCC Bullx supercomputer CURIE (fat nodes).

Reader must be warned that a non standard tuning of the processes binding can significantly change performances on some computers or using some schedulers. At the opposite of the standard N configuration, the coupled system MPI-processes host different calculations (OPA or SAS). All bindings presented in Figure 4 lead to different performances. The first one (default, using round robin distribution) exhibits the better on our supercomputer.

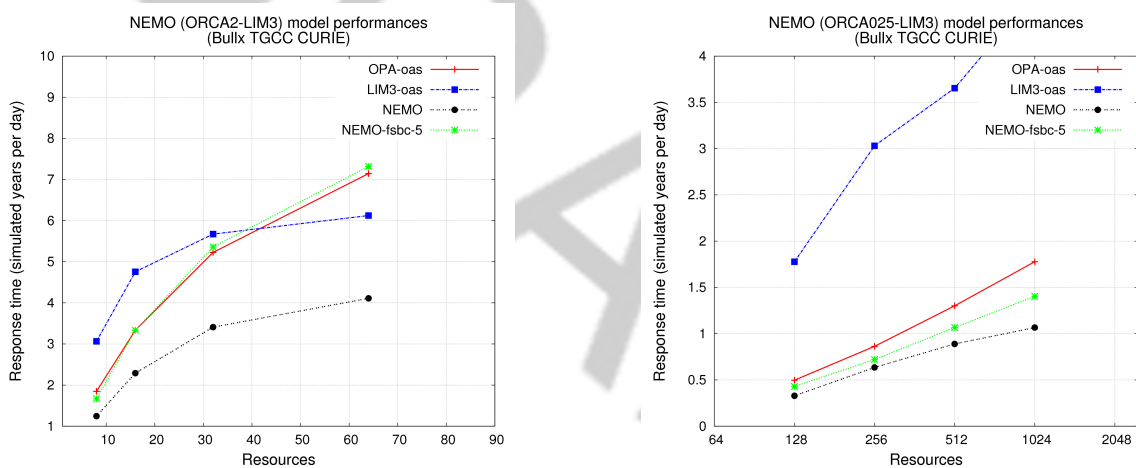


Figure 5: NEMO and OPA/SAS compared performances on ORCA2 (left) and ORCA025 (right) grids

For a each OS coupled simulation, separate performances are measured for OPA and SAS components thanks to the technique described in [3]. The reference simulation N with a LIM3 call frequency of 1 (5) is plotted in black (green). For OS, the separate performances of OPA (SAS) component is in red (blue). The maximum OS speed is given by the maximum speed of the slowest component (SAS for ORCA2, OPA for ORCA025). The optimum configuration is a compromise between speed and cost. It is deduced from Figure 5 and parallel efficiency considerations.

Configuration/ Grid	N	N5	OS (OPA or SAS)
ORCA2	4.1	7.1	6
ORCA025	1.1	1.4	1.7

Table 2: Maximum speed in SYPD

Configuration/ Grid	N	N5	OS (OPA+SAS)
ORCA2	225 (32)	143 (32)	282 (32+32)
ORCA025	960 (1024)	730 (1024)	648 (1024+128)

Table 3: Optimum cost in CHPSY (core #)

On ORCA2 grid, separate components are running slower if coupled with OASIS instead of being included in a single executable. We strongly suspect default binding not to be well optimized for MPMD mode. An optimum binding is still to found, with an important source of enhancement: a speedup of only 50% is observed with the ORCA2 configuration, where 100% could be expected.

On ORCA025 grid, default binding seems to be more efficient. Our OS configuration is 20% faster than N5 and 12% cheaper.

#### 4. Limitations

Parallel tasking is creating difficulties that any user not familiar with a tool like OASIS has to face:

- launching more than one executable at the time (MPMD mode of the MPI library) can be uneasy for some machines or schedulers. In any case, an efficient mapping of the different components on the available resources (complicated by XIOS memory requirements) can rapidly increase the difficulty for a novice
- debugging a multiple executable program is always complicating a developer task, since finding the failing component is a prerequisite before any correction
- maintaining this OASIS based interface in conjunction with the existing one (ocean/sea-ice in a single executable) implies additional care and tests

Additionally, since SAS is including both bulk formulas and sea-ice model calculations, OPA is necessarily receiving fluxes calculated on SAS grid. This has a strong impact, on free ocean grid points too, in case of SAS grid coarsening. The same problem occurs in atmosphere coupled mode: as mentioned at 1.2.2 §, SAS is playing a central role in ocean-atmosphere exchanges and no coupling fields are directly exchanged by the ocean-only component OPA and the atmosphere. All must transit by the surface module and adjusted to its grid. Which is necessarily covering the whole globe: it is not possible, with the current implementation, to process sea-ice calculations on meshes reduced to the high latitudes.

## 5. Conclusion, Perspectives

### 5.1. Applications

The OPA/SAS interface is proposed in the NEMO official release taking the occasion of CMIP6 developments. The international exercise should be the right occasion to familiarize a pioneer part of our community to the special requirements of our configuration. We hope that this first step will help a better diffusion of the “OPA/SAS best practice” to the whole community.

Some of the targeted CMIP6 configurations do not include the LIM3 sea-ice model, but CICE, GELATO, etc. Only LIM2 or LIM3 are fully tested in OPA/SAS configuration and additional work is required for other adaptations. Nevertheless, the clear OASIS interface can be easily cloned for a different use. Most of the additional work consists in identifying the coupled variables and disable the existing coupling implementation (copy in variables).

The PRIMAVERA H2020 project ultra-high resolution CGCM is one of the most promising configuration that can potentially take a large benefit from OPA/SAS. This model is including NEMO on the ORCA12 grid. In this case, a coarsening (to ORCA025) of the sea-ice model could significantly reduce its huge computing cost. We mention that a coarsening was successfully tested with OPA-ORCA025/SAS-ORCA2: a simple choice in appropriate OPA/SAS namelists and an automatic recomputing of OASIS interpolation weights were enough to perform a 1 month long run test.

### 5.2. Possible enhancements

New developments would be necessary with the objective to enhance the system performances. SAS domain decomposition on global grid follows the standard NEMO one, with a quasi-identical number of grid points per sub-domain. SAS sea-ice calculations mainly take place in high latitude regions. Thus, the standard decomposition leads to a high load imbalance between tropical-mid latitude and high latitude sub-domains. One possible solution would consist in defining larger latitude bounds for tropical and mid latitude sub-domains.

Another solution, unless requiring more developments, would split SAS into three different executables (one covering Arctic ocean, one Southern ocean and the last one the complementary region to the whole globe). The first two would include sea-ice model, and not the last one. The three sets of fluxes would be assembled in OPA, thanks to the newly developed forced/coupled mixed instructions located in `sbc_cp1` routines.

## Acknowledgment

This work is funded by the ANR project PULSATION. It is also supported by the IS-ENES2 FP7 infrastructure. We acknowledge that the results in this report have been achieved using the PRACE Research Infrastructure resource CURIE based in France at TGCC.

## References

- [1] Madec G., 2008: "NEMO ocean engine". Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No 1288-1619. <http://www.nemo-ocean.eu/>
- [2] Valcke, S., Craig, T. and Coquart, L., 2015: [OASIS3-MCT User Guide](#), OASIS3-MCT 3.0, Technical Report, TR/CMGC/15/38, CERFACS/CNRS SUC URA No 1875, Toulouse, France
- [3] 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

## Appendix

### OPA/SAS Quick Start Guide (June 2015)

#### Installing OPA/SAS using FCM (nemo\_v3\_6\_STABLE)

##### *OASIS installation: Step 1*

Register and extract OASIS3-MCT (v3) coupling library available here:

<https://verc.enes.org/oasis/>

Compile it with the appropriate compiler and libraries (modify make.inc accordingly). As an example, compile with ifort on linux:

```
cd oasis-mct/util/make_dir; cp make.intel_corail make.inc; make -f TopMakefileOasis3
```

##### *Extract and install XIOS 1.0: Step 2*

[Documented here](#) (and come back to this page afterwards to create NEMO executable!). If detached mode needed, XIOS executable must be compiled with `--use_oasis oasis3_mct` option.

##### *NEMO installation: Step 3*

Extract NEMOGCM (using your "my\_login" /pw registered on this web site):

```
svn --username "mylogin" co  
http://forge.ipsl.jussieu.fr/nemo/svn/branches/2015/nemo_v3_6_STABLE/NEMOGCM
```

The main script to compile and create executable is called **makenemo** and located in the CONFIG directories. First, these directories must be downloaded from an external ftp server:

```
cd NEMOGCM/CONFIG; wget  
ftp://ftp.cerfacs.fr/pub/globc/exchanges/maisonna/configs_opa_sas.tgz
```

This compile script must be called twice to build OPA and SAS executables. As an example, compile with ifort on linux, using LIM3 sea-ice model, the two configurations ORCA2\_OPA and ORCA2\_SAS\_LIM3:

```
cd NEMOGCM/CONFIG; ./makenemo -m ifort_linux -r ORCA2_OPA; ./makenemo -m ifort_linux -r ORCA2_SAS_LIM3
```

The WORK directory is created under each configuration directory.

#### Descriptions of files and directories

- `cfg.txt` : text list of configurations and source directories. ORCA2\_OPA only includes OPA\_SRC. ORCA2\_SAS\_LIM3 includes OPA\_SRC, SAS\_SRC and LIM\_SRC\_3
- `ARCH` : compilation option files, with format `arch_compiler.fcm`. This file must explicitly contains path of OASIS libraries
- `ORCA2_OPA` : the `cpp.fcm` file contains the list of cpp key including `key_oasis3` and `key_oa3mct_v3` but without `key_lim3`
- `ORCA2_SAS_LIM3`: the `cpp.fcm` file contains the list of cpp key including `key_oasis3`, `key_oa3mct_v3` and `key_lim3`

#### *Running: Step 4*

Once makenemo has run successfully, the OPA executable is available in `CONFIG/ORCA2_OPA/EXP00` and SAS executable (same name) in `CONFIG/ORCA2_SAS_LIM3/EXP00`

Input file are needed and are available via the NEMO website [Configuration pages](#) (Project configuration OPA/SAS). It consists in the same input tar file than in the ORCA2\_LIM reference configuration, with additional coupling information (OASIS restart files).

In `CONFIG/ORCA2_OPA/EXP00`, the `AA_job` file is given as a run script example. It includes the main launching instruction:

```
mpirun -np $NB_PROC_OPA ./opa.xx : -np $NB_PROC_SAS ./sas.xx
```