

WN/CMGC/07/16

**LEGO
Grid Compliant Climate Model
Analysis**

Eric Maisonnave

Table of Contents

General purpose.....	4
Numerical simulation.....	5
Algorithm.....	5
Components characteristics.....	7
Files.....	7
Functions	7
Pre-processing tasks:.....	7
Task 1.1: modify_parameters.....	7
Task 1.2: concatenate_atmospheric_input_files.....	8
Main computing task: the climate model.....	8
Task 2: process_coupled_run	8
General presentation.....	8
Climate model components.....	8
Post processing tasks.....	10
Task 3.1: convert_output_format.....	10
Task 3.2: extract_minimum_information and Task 3.3: compress_diags.....	10
Task 3.4: validate_run.....	11

The aim of this document is to describe the application provided by Cerfacs within the LEGO ANR funded project. This application mainly consists on ocean-atmosphere numerical simulations. Adapting this tool, new Grid technologies must help us to run embarrassingly parallel scenarios of XXI th century climate.

The world climate is currently changing due to the human-induced increase in the atmospheric concentrations of CO₂ and other greenhouse gases. Exacerbated by man's impact, seasonal-to-decadal climate fluctuations are predicted to become increasingly extreme in the coming years.

In order to study these changes, we need to produce numerical simulations using general circulation models of climate system main components (atmosphere, ocean, sea ice, continental surfaces) on forced mode or, like here, on coupled mode, that is to say allowing information exchange between each component during simulation.

General purpose

Climatologists have recourse to numerical simulation and particularly coupled models in several occasions: for example, to estimate natural variability (thousand of simulated years), for seasonal forecasting (only a few simulated months) or to study global warming characteristics (some simulated decades).

To take advantage of Grid5000 platform, we choose to launch parallel simulations (ensemble) on several nodes, approximatively 10 or more, according to platform loading. “Scenario” simulation ensemble, that is to say simulations of present climate followed by XXI st century ones, perfectly match with Grid requirement. Each simulation will differ from each other in physical parametrization of atmospheric model. In this way, we expect to better estimate global warming prediction sensibility to model parametrization.

Practically, a 150 years long scenario combines 1800 simulations of one month each, launched one after the other. This partition eases workflow and implements checkpoint. The initial state of monthly simulation “n” is the ending state of monthly simulation “n-1”.

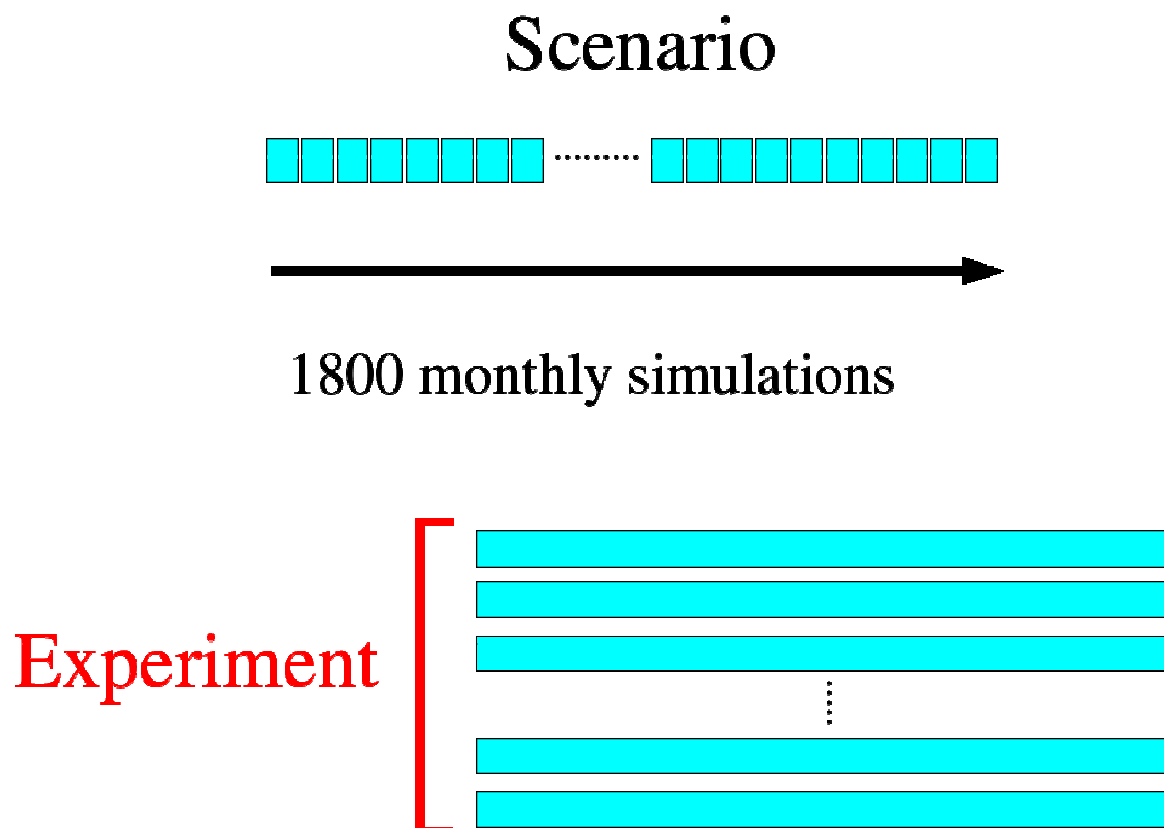


Fig. 1: Monthly simulation, scenario and whole experiment

Numerical simulation

We hereby details characteristics of each task making part of a scenario simulation, as observed on a CRAY XD1 platform. This Linux cluster is the main calculator at Cerfacs. Performance on its 58 nodes of AMD Opteron bi-processors seems a good approximation of most of the Grid5000 platform clusters (characteristics per node: 4.8 peak Gflops, with 2Gb of memory and 6.4 Gb/s bandwidth).

Algorithm

The following algorithm give details of one experiment called “scenario”. We expect to launch such experiment 10 times or more during WP5 validation.

Some precisions concerning syntax:

- In argument calls, bold characters mark input arguments and underline output.
- Directory names refers to local node home directories (**WORKDIR** and current directory) and CERFACS data server (**EXTERNAL_DIR**).
- Square brackets are used to define elements in an array, parenthesis to describe function arguments.

```

cp $WORKDIR/forcing_files .
cp $WORKDIR/standard_parameter_files .
cp $WORKDIR/executables process_coupled_run

do month=1,1800

    if ( month eq 1 ) then
        cp $WORKDIR/restart_files_in .
    else
        mv restart_files_out restart_files_in
    end if

    modify_parameters(      standard_parameter_files,
                          appropriate_parameter_files)

    concatenate_atmospheric_input_files( atmospheric_restart_file_in,
                                         atmospheric_forcing_files)

    process_coupled_run(    restart_files_in,
                          forcing_files,
                          appropriate_parameter_files
                          monthly_diags,
                          restart_files_out)

    convert_output_format( monthly_diags,
                          converted_diags[month])
    rm monthly_diags

    extract_minimum_information(      converted_diags[month],
                                     minimum_diags[month])

    if ( modulo (month, 180 ) == 0 ) then
        compress_diags(      converted_diags[month-179, month],
                            compressed_diags)
        rm converted_diags[month-179, month]
        mv compressed_diags $EXTERNAL_DIR
    end if

    if ( month == 1 ) then validate_run(converted_diags[1])

end do

mv minimum_diags[:] $EXTERNAL_DIR

```

Components characteristics

Files

Files	Size
forcing_files	270 Mb
standard_parameter_files or appropriate_parameter_files	40 Kb
executables	32 Mb
restart_files_in	120 Mb
atmospheric_restart_file_in	4 Mb
atmospheric_forcing_files	1.7 Mb
monthly_diags	150 Mb
restart_files_out	120 Mb
converted_diags[1] (array size: 180)	50 Mb (18+32)
minimum_diags[1] (array size: 10)	100 Kb
compressed_diags	1.2 Gb

Fig. 2: Synthetic presentation of scenario IOs characteristics.

Functions

Functions	Memory	Elapsed time	Processors
modify_parameters	1 Kb	1 sec	mono
concatenate_atmospheric_input_files	200 Mb	1 sec	mono
process_coupled_run	600 Mb	25 min	8 (MPI)
convert_output_format	5 Mb	1 min	mono
extract_minimum_information	20 Mb	1 min	2 (OPEN-MP)
compress_diags			2 (OPEN-MP)
validate_run	5 Mb	1 min	mono

Fig. 3: Synthetic presentation of scenario functions characteristics.

Pre-processing tasks:

Task 1.1: modify_parameters

The purpose of this function is to update parametrization of each model, accordingly to time

coordinate position of the monthly run within whole scenario experiment. For example, green house gas concentration changes each month during scenario and must be updated at this stage.

Coding characteristics: Unix sed command.

Task 1.2: concatenate_atmospheric_input_files

Before simulation, we modify atmospheric model input files merging initial state and limit conditions (soil characteristics, green house effect gazes, ozone ...).

“updcli” executable read initial state file (also named restart file) and et limit conditions file and rewrite only one file compliant with ARPEGE IO format. “updozo” executable add ozone concentration.

Coding characteristics: Fortran 90 code (2 files, 500 lines code).

Main computing task: the climate model

Task 2: process_coupled_run

General presentation

Climate model often flows from specific adaptation of a weather forecast numeric model. This model no longer head for operational short term forecasting but to study of climate mechanisms and consequently, to simulate its variability and changes during period from months to decades.

The climate global approach need coupling atmospheric model with several other models describing evolutions of ocean, soil, sea-ice, biosphere ...

Within LEGO project, chosen numerical model (OCC17) comes from the ultimate version of one of the two french community. Cerfacs is just ending OCC17 assembling procedure.

OCC17 includes atmospheric model (ARPEGE), ocean and sea-ice (OPA) and runoff pathway (TRIP). OASIS coupler ensures their simultaneous running and synchronic exchange of information at component interfaces.

This 22 fields coupling does not need flux correction. It allows oceanic current coupling to correct usual wind stress biases. With help of CNRM and LOCEAN research teams, it includes state-of-the-art physical and dynamical parametrization in both atmospheric and oceanic models.

Climate model components

Atmosphere

The model is the 4.6 version of the ARPEGE [DEQ94] climate model, which is designed for use by the French climate community. (ARPEGE is the acronym for Action de Recherche Petite Echelle Grande Echelle: Research Project on Small and Large Scales.)

The model has a spectral horizontal representation (spherical harmonic basis functions) with transformation to a Gaussian grid for calculation of nonlinear quantities and some physics. Primitive-equation dynamics are expressed in terms of vorticity and divergence, temperature, and specific humidity, and the natural logarithm of surface pressure (or, on option, the surface pressure itself).

Cerfacs implemented this version for the Dynamite European Project (FP6). Next IPCC exercise plan to use it in order to achieve necessary simulations

ARPEGE code is fully parallel. At low resolution use, a 4 processors configuration has a profitable computation/cost ratio.

Coding characteristics: mostly Fortran 90, 2.500 files, 600.000 lines code.

Libraries: BLAS, LAPACK, LAM-MPI, PSMILE-OASIS, NETCDF

Optimized parallelization: 4 processors.

Ocean

The OPA system is an Ocean General Circulation modeling System [MAD06] shared by projects (research and operational) in oceanography and Climate change studies. It is developed at the Laboratoire d'Océanographie et du Climat Expérimentations et Approches Numériques (LOCEAN). OPA is the acronym for « Océan Parallélisé ».

The core of the system is a primitive equation model applied to both regional and global ocean circulation.

Prognostic variables are the three-dimensional velocity field and the thermohaline variables. They are distributed on a three-dimensional Arakawa-C-type grid using prescribed z- or s-levels. Various physical choices are available to describe ocean physics, including a 1.5 turbulent closure for the vertical mixing, geopotential or isopycnal mixing, eddy induced velocity parameterization, simple bottom boundary layer representation, etc ...

NEMO (Nucleus for European Modeling of the Ocean), last version of the model, includes a sea-ice model (LIM2 module , implemented at Université de Louvain la Neuve University).

Coding characteristics: Fortran 90, 230 files, 100.000 lines code.

Libraries: NETCDF, PSMILE-OASIS, LAM-MPI

Runoff pathways model

In addition to those general circulation models, TRIP (Total Runoff Integrating Pathways) [OKI] allows us to simulate river flows over continental surfaces. This module come from Industrial Sciences Institute of Tokyo University.

Coding characteristics: Fortran 77, 50 files, 3.200 lines code.

Libraries: PSMILE-OASIS, NETCDF, LAM-MPI

Coupler

OASIS [VAL04] is a software allowing synchronized exchanges of coupling information between numerical models representing different components of the climate system. OASIS is the acronym for Ocean Atmosphere Sea Ice Soil. It has been developed at Cerfacs since 1991.

OASIS ensures simultaneous launch of executables on allocated nodes, exchange of coupling fields at model interfaces and suitable transformation (spherical interpolations) between model grids.

Due to interpolations contingency from t63 gaussian reduced grid of ARPEGE to unstructured ORCA2 (OPA) grid, the 3rd version of OASIS interpolation toolkit was chosen.

Coding characteristics: mostly Fortran 90, 220 files, 45.000 lines code.

Libraries: NETCDF, LAM-MPI

	CPU (sec)	Elapse (min)	Max Memory (Mb)	IO (Mb)	MPI IO (Mb)
OASIS	200	3	350	170	240
ARPEGE	1500/proc	25	200/proc	190	60
OPA	1500	25	600	320	140
TRIP	200	3	20	2	40
Total	7900	25	600	680	420

Fig. 4: Synthetic presentation of coupled model components characteristics.

Post processing tasks

Task 3.1: convert_output_format

The ARPEGE->NETCDF format converter SAFO « Sorties ARPEGE au Format des Océanographe » was developed at Cerfacs to standardize climate model components output. An ASCII file permit to choose fields to convert between raw ARPEGE outputs. Saving frequencies could also be tune at this same stage.

Coding characteristics: Fortran 90, 66 files, 8.500 lines code.

Libraries: NETCDF

Task 3.2: extract_minimum_information and Task 3.3: compress_diags

The STATPACK statistical analysis toolkit, developed at LOCEAN by Pascal Terray, will help us to reduce oversized data, calculating global or regional means with `comp_serie` tool (`extract_minimum_information`) and decomposing temporal series to orthogonal vectors using Empirical Orthogonal Function method (`compress_diags`).

Coding characteristics: Fortran 90, 250 files, 130.000 lines code.
Libraries: NETCDF

Task 3.4: `validate_run`

In a first stage, to allow all LEGO developer validate their implementations, a simple analysis tool will be called each month. Matching some basic physical quantities with standard outputs, various criterion of validity will be produced to validate coding modifications.

Coding characteristics: Fortran 90, 1 file, 500 lines code.
Libraries: NETCDF

Bibliography

DEQ94: Déqué M, Drevet C, Braun A, Cariolle D (1994) The ARPEGE/IFS atmosphere model: a contribution to the french community climate modeling ,Clim Dyn, **10**, 249-266

MAD06: Madec G. (2006) NEMO Reference manual, ocean dynamic component: NEMO-OPA, Institut Pierre Simon Laplace (IPSL), **27**, ISSN 1288-1619

OKI: Oki T and Sud Y C (1998) Design of Total Runoff Integrating Pathways (TRIP), Earth Integration, **2**

VAL04: Valcke S, Caubel A, Vogelsang R and Déclat D (2004) Oasis 3, User Guide, PRISM Report Serie no 2 (5th edition), CERFACS, Toulouse, France, 60 pp