Project Number GA:	**228203**
Project Acronym:	**IS-ENES**
Project Title:	**InfraStructure for the European Network for Earth System Modelling**
Programme:	**SEVENTH FRAMEWORK PROGRAMME Capacities Specific Programme Research Infrastructures**

D8.1 - Report on the Definition of the Evaluation Suite

WP8/JRA2: European ESM: Performance Enhancement

Due Date:	M6	
Submission Date:	-	
Start Date of Project:	01/03/2009	
Duration of Project:	42 months	
Organisation Responsible for the Deliverable:	CMCC (9)	
Version:	1.0	
Status	Final	
Author(s):	G. Aloisio - I. Epicoco - S. Mocavero	CMCC (9)

Version History

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	04/08/2009	First draft: definition of the outline	G. Aloisio, I.Epicoco, S.Mocavero
0.2	28/08/2009	Comments and improvements	G. Riley
0.3	01/09/2009	Second draft: ready for the distribution on the internal mailing list	G. Riley, G. Aloisio
0.4	17/12/2009	Contribution on models description	E. Maisonnave, C. Lévy, M.A. Foujols, A. Caubel, S. Mullerworth, M.Giorgetta, R. Budich
0.5	19/01/2010	Final Draft: ready for mailing list comments	G. Aloisio, I.Epicoco, S.Mocavero
1.0	30/09/2010	Final V1	G. Aloisio, G. Riley

Table of Contents

1. INTRODUCTION **5**

1.1 PURPOSE 5

1.2 GLOSSARY OF ACRONYMS 6

2. EVALUATION SUITE **7**

2.1 RELEVANCE FOR IS-ENES 7

2.2 SELECTION CRITERIA 7

2.3 SUITE OVERVIEW 8

2.4 MODELS DESCRIPTION AND PRELIMINARY PERFORMANCE RESULTS 8

2.4.1 Coupled models 9

2.4.1.1 CMCC-MED 9

2.4.1.2 ARPEGE-NEMO 13

2.4.1.3 IPSLCM5 Model 17

2.4.1.4 HadGEM2 22

2.4.1.5 MPI-M 23

2.4.2 Stand-alone models 28

2.4.2.1 ECHAM 28

2.4.2.2 NEMO 31

Index of Tables

Table 1: Glossary of Acronyms 6
Table 2: CMCC-MED Model 13
Table 3: ARPEGE-NEMO Model 17
Table 4: IPSLCM5 Model 21
Table 5: HadGEM2 Model 23
Table 6: MPI-M Model 28
Table 7: NEMO Model 33

1. Introduction

The work package WP8/JRA2 is mainly focused on the analysis of the performance aspects concerning the configuration, deploying and running of Earth System Models (ESMs). It covers a number of key areas relating to model performance including: researching portability and performance of key models on a range of platforms to provide improved models; to improve the links between model support initiatives and to ensure the European climate science community is better prepared to work together on the petascale machines that are imminent and on the machines available today; to develop tools to ease the composition of new ESMs from existing model components and coupler technologies which will lower the technical hurdle for small climate research organisations.

Within these objectives, task 8.1 is a first step in the definition of the ESMs initially brought to IS-ENES by the project partners. These ESMs constitute the basis of the evaluation suite. Applications in the suite will be used for benchmarking and evaluating the performance of the emerging and existing hardware architectures and to highlight the limits and the advantages of the current ESMs' software infrastructures. This work will be undertaken as a series of well defined tasks in which application owners and selected performance specialists work together on a specific ESM. A task may involve, for example, porting, performance analysis and improvement of ESM model code and/or of software infrastructure, etc. Results from work on the evaluation suite will inform the design and development of future Earth System Models optimized for peta and exascale architectures, considering that performance and portability may bring to conflicting strategies.

The report is organized as follows: the next section describes the main objectives and methodology used to select the climate models belonging to the suite; this is followed by a description of each model.

1.1 Purpose

The main goal of this report is to define and describe the initial contents of the IS-ENES Evaluation Suite, which gathers key existing and developing ESMs belonging to the IS-ENES partners. The models in these ESMs are considered as examples of ESM infrastructure. The evaluation suite will provide an application-focussed lead to the other tasks of WP8/JRA2, ensuring cross-fertilisation of effort and knowledge. Even though the selection of the applications belonging to the Evaluation Suite has been initially based on the ESMs developed by IS-ENES project partners, it must not be considered a closed list. The list can be extended, as the project evolves, by the IS-ENES management group.

The main purpose of this report regards the definition of a suite of climate models to be used:

- for evaluating the performances of the models themselves on existing parallel architectures
- for benchmarking the parallel architectures and defining the critical aspects that climate models stress
- for profiling and for optimizing current ESMs and related infrastructure on the available architectures
- for highlighting the bottlenecks, the strengths and weaknesses of each of the models to guide the design and development of future optimized ESMs for the upcoming peta and exascale architectures

1.2 Glossary of Acronyms

Acronym	Definition
ESM	Earth System Model
NEMO	Nucleus for European Modelling of the Ocean
IPSLCM5	IPSL coupled model
PRACE	

Table 1: Glossary of Acronyms

2. Evaluation Suite

2.1 Relevance for IS-ENES

Earth system models are the only analytical tools available for prediction of the future climate evolution either under natural conditions or under the influence of humankind. The development and use of realistic climate models requires a sophisticated software infrastructure and access to the most powerful supercomputers and data handling systems.
Several scientific institutions, universities, governmental organizations and industrial partners in Europe have developed a world class expertise in different aspects of Earth System Modeling and have contributed to international assessments of climate change.
The complex component processes of a high-resolution climate system models, that better simulate the interactions and feedbacks among the physical and biological processes, include: atmosphere, ocean, land, soils, permafrost, vegetation cover, sea ice, land ice, carbon and other biogeochemical cycle, clouds and related microphysics, hydrology, atmospheric chemistry, aerosols, ice sheets, human systems.
Current models already provide improved simulation and prediction of changes in temperature and precipitation, and extreme weather events.

2.2 Selection Criteria

A huge amount of climate models are actually available worldwide. They differ each other both for computational requirements and scientific reliability of the prediction and modelization of the climate behavior and its impacts. In order to define the evaluation suite we have discussed some relevant selection criteria, described below:

1. **Interest for the model**. The main criteria adopted is based on the estimation about how much the model is of interest for IS-ENES partners. From a pragmatic point of view, all of the models, belonging to the Evaluation Suite, should be maintained, ported and optimized on the major parallel architectures. More a model is of interest more it will be maintained, optimized and analyzed.
The interest for a model have be measured evaluating the plans that a partner has for porting, optimizing, providing support on it

2. **Model spread**. This criteria measures how much the model can be relevant for a wider climate community. It consist in the number of climate centers the regularly use the model for production. More

3. **Supported platforms**. This criteria is strictly related to the availability of the model on different parallel architectures. More are the platform where the model is already ported and analyzed, less will be the effort to support new architectures

4. **Performance requirements**. A qualitative evaluation of memory, cpu, I/O, network use. The metric could be low/medium/high

5. **Model extension.** The criteria refers to the number of model components integrated within the couple model.

6. **Potentiality for peta-exascaling.** This is a qualitative criteria that defines how much a model could scale up to exaflop architectures and how difficult is to modify the model for the exascale platforms.

8. **Level of model documentation.** The availability of a good documentation related to the model is relevant for maintaining the model and for running the evaluation suite

2.3 Suite Overview

The Evaluation Suite gathers existing models used by partners to provide other tasks with a an application-focused lead for understanding ESMs performance and improving them for current state-of-the-art computing systems and for future, such as for peta and exascale architectures. An initial set of applications has been defined, taking into account both partners experience and knowledge and selection criteria described in the previous section. The Evaluation Suite will include two different categories of applications: (i) coupled models, often used by climatologists to evaluate the complexity of the system as the range of model components involved (i.e., physical, chemical, and biological) and their interactions, and (ii) stand-alone models, generally used for performing high-resolution simulations in the short period. The suite overview synthesizes models features. In particular:

- The configuration referring to a test case run. This configuration is used for testing/benchmarking the model.
- A description of the model (for coupled models each component has to be analyzed).
- A description of platforms on which the model has been tested and the execution details, such as libraries, compilation flags, etc.
- Details about scalability such as the execution time, the I/O time and the Communication time at different number of cores.
- Recommendations about potential for Peta/Exascaling and optimization and related expected effort.

In section 2.4.1 the following coupled models will be described:

- CMCC-MED from CMCC.
- ARPEGE-NEMO from CERFACS.
- IPSLCM5 from CNRS-IPSL.
- HadGEM3 from METOFFICE.
- MPI-M from MPI.

In section 2.4.2 the following stand-alone models will be detailed:

• NEMO from IPSL.

2.4 Models description and preliminary performance results

The information about models have been organized in the following tables in the following. Two sections 3.4.1 and 3.4.2 have been introduced since in order to differentiate the information regarding the Coupled models (3.4.1) or Stand alone models (3.4.2)

2.4.1 Coupled models

2.4.1.1 CMCC-MED

General	
Name of the model	CMCC-MED
Reference person	Silvio Gualdi, gualdi@bo.ingv.it, CMCC
Version	2.0
Brief description	Three components fully coupled climate model with mediterranean sea region
Infrastructure (e.g. coupler)	OASIS3 ver. 2.5 (CMCC parallel version)
Download	For information about the code download, please contact the reference person
Reference for detailed information	
Configuration	
Input configuration	The coupled model is made of: Echam5 T159L31 OPA 8.2 global Nemo 1/16° for Mediterranean sea Length of model run: 1 month Oasis configuration: total number of exchanged fields 35 exchanged fields 17 (Echam -> Oasis -> OPA) exchanged fields 9 (Echam -> Oasis -> Nemo) exchanged fields 6 (OPA -> Oasis -> Echam) exchanged fields 3 (Nemo -> Oasis -> Echam) coupling period: 2h 40' Echam5

	Region: Global Resolution: T159L31 (480 x 240) time step: 240 sec OPA Region: Global Resolution: 2° (182 x 149) time step: Nemo: Region: Mediterranean Sea Resolution: 1/16° (871 x 253) time step:
Input files dimension	Echam5: 912MB OPA: 96MB NEMO: 1.6GB OASIS: 1.1GB restart files (1 month): Echam5: 1.23GB OPA: 102MB Nemo: 1.5GB OASIS3: 30,5MB
Output files dimension	Echam5: 9.2GB Nemo: 18GB OPA: 1.5GB

Component Model	
Component model's name	OASIS
Version	2.0.1 CMCC parallel version based on version 3_prism_2-5
License policy	Lesser GNU General Public License (LGPL)
Programming Language(s)	C, Fortran

Libraries	NetCDF, MPI, OpenMP,
Parallelization method supported	MPI1, MPI2, OpenMP
Component Model	
Component model's name	Echam
Version	v. 5
License policy	"MPI-M Software Licence Agreement", which must be signed by each user
Programming Language(s)	Fortran
Libraries	NetCDF, MPI, OpenMP, Lapack, Blas
Parallelization method supported	MPI2, OpenMP
Component Model	
Component model's name	OPA
Version	v. 8.2
License policy	CeCILL license (public license)
Programming Language(s)	F95
Libraries	NetCDF, MPI
Parallelization method supported	MPI1, MPI2
Component Model	
Component model's name	Nemo
Version	v. 9
License policy	CeCILL license (public license)
Programming Language(s)	F95
Libraries	NetCDF, MPI
Parallelization method supported	MPI

Platform	

Execution platform	Ulysses / NEC-SX9
Details of the execution platform	CPU Type: SX9 CPU Speed: 3.2 GHz # of nodes: 7 SMP size: 16 vector CPUs Memory: 512 GB per node Interconnection: IXS Teoretical Peak Performance: 11.2 TFLOPS Installed Libraries: Blas, Lapack, MPI, OpenMP Development tools: C compiler: sxcc Fortran compiler: sxf90 profiler: SXFtrace
Libraries	**Oasis**: -lnetcdf **Echam**: -llapack -lblas -lnetcdf **OPA**: -lnetcdf **Nemo**: -lnetcdf
Compilation flags for each component	**Oasis**: -Pstack -pi auto nest=3 line=10000 exp=iminim, rmaxim, rminim, grid_search_bilin -Ep -sx9 -Wf,"-P nh" -Wf,"-pvctl noassume loopcnt=5000000 vworksz=100M" -Wf,"-A idbl4" -Wf,"-msg o" -Wf,"-pvctl fullmsg" -Wf,"-L fmtlist transform map summary noinclist" -Chopt -Wf,"-ptr byte" **Echam**: -Popenmp -Chopt -sx9 -Wf,-init heap=zero stack=zero -Wf,-K a -Wf"-L source fmtlist noinclist mrgmsg summary transform map objlist" -Ep -pi line=1000 -Wf,-pvctl chgpwr fullmsg,-msg o,-ptr byte **OPA**: -Ep -Pstack -sx9 -Chopt -Wf,-P nh -Wf,-pvctl noassume loopcnt=200000 vworksz=4M -Wf,-A idbl4 -Wf,-msg o -Wf,-pvctl fullmsg -Wf,-L fmtlist transform map summary noinclist **Nemo**: -size_t64 -dw -Wf\"-A dbl4\" -sx9 -pi auto -P stack -C vopt -Wf"-init stack=nan" -Wl"-f nan" -Wf"-P nh" -Wf,-pvctl noassume loopcnt=10000 -L transform -Wf,"-msg o" -Wf,"-L fmtlist transform map summary noinclist"
General comments	

Scalability details	
The performance, reported below, refers to 1 month run. In the experiment, the number of processes for OPA, Nemo and Echam are constant: OPA -> 1 proc Nemo -> 6 procs Echam -> 8 procs	

#cores (OASIS)	**#cores (total)**	**Wall clock (secs)**
1	16	4320.95
2	17	4032.18
3	18	3963.80
7	22	3898.64
9	24	3856.03
11	26	3864.05
13	28	3790.30
15	30	3786.89
17	32	3774.50
26	41	3793.12
33	48	3822.48

Recommendations	
Expected potential for Peta/Exascaling	low
Expected effort to reach potential Peta/Exascaling	Improve kernels to be able to scale to high number of processes. Increase resolution of the models.
Expected potential for optimization	medium
Expected effort to reach the optimization potential	Optimize I/O operations in order to improve simulation time. Reduce bank conflicts on vector machines.

Table 2: CMCC-MED Model

2.4.1.2 ARPEGE-NEMO

General	
Name of the model	ARPEGE-NEMO

Reference person	Eric Maisonnave, eric.maisonnave@cerfacs.fr, CERFACS
Version	ARPEGE-Climat v5.2 – NEMO v3.2
Brief description	CGCM high definition
Infrastructure (e.g. coupler)	OASIS v3 (pseudo parallel) – OASIS v4 (work in progress)
Download	Due to component licence restriction, CERFACS does not provide any version of the coupled model
Reference for detailed information	For general information on ARPEGE-NEMO workflow, see LEGO: Grid Compliant Climate Model Analysis. For information on porting and tuning the high-resolution configuration, document to come.
Configuration	
Input configuration	The coupled model consists on: ARPEGE-5 T359L31 (possibly higher resolution) NEMO-3 ¼ degree (possibly 1/12 degree – 1D configuration also available) Length of model run: 1 month Oasis configuration: OASIS-3 Pseudo-parallel mode (possibly OASIS-4) Coupling period: 3h Coupling fields (reduced configuration) O2A: surface temperatures, albedo A2O: heat fluxes, water fluxes (without calving and runoff), non solar heat flux derivative ARPEGE-5 Region: Global Resolution: T359L31 (360 x 180) time step: 900 sec NEMO-3 Region: Global Resolution: 1/4 (1442 x 1021)

	time step: 1080 sec
Input files dimension	ARPEGE: 87 MB NEMO: 384 GB OASIS: 8 GB restart files (1 month): ARPEGE: 48 MB NEMO: 8.4 GB OASIS3: 50 MB
Output files dimension	Monthly means: ARPEGE: 4.5GB NEMO: 745 MB
Component Model	
Component model's name	ARPEGE-Climat
Version	v5
License policy	"ARPEGE-Climat Software Licence Agreement", which must be signed by each user (contact CNRM, Météo-France)
Programming Language(s)	Fortran
Libraries	BLAS, LAPACK
Parallelization method supported	MPI1 (MPICH, OpenMPI, LAM tested), possibly OpenMP
Component Model	
Component model's name	NEMO
Version	v3
License policy	CECIL licence, see http://www.nemo-ocean.eu/user/register

Programming Language(s)	Fortran90
Libraries	Netcdf, xml
Parallelization method supported	MPI1 (MPICH, OpenMPI, LAM tested)
Component Model	
Component model's name	OASIS
Version	V3
License policy	LGPL, see https://oasistrac.cerfacs.fr/
Programming Language(s)	Fortran/C
Libraries	Netcdf
Parallelization method supported	Pseudo parallelism IPSL/CERFACS method

Platform	
Execution platform	Météo-France NEC-SX9 (but also SGI Altix, IBM JS21/BG-L/BG-P at lower resolution)
Details of the execution platform	CPU Type: SX9 CPU Speed: 3.2 GHz # of nodes: 6 SMP size: 16 vector CPUs Memory: 1 TB per node Teoretical Peak Performance: 102 GFLOPS/proc Installed Libraries: Blas, Lapack, MPI, OpenMP, Netcdf Development tools: C compiler: sxcc Fortran compiler: sxf90 profiler: SXFtrace
Libraries	**Oasis**: -lnetcdf **Echam**: -llapack -lblas -lnetcdf

	OPA: -lnetcdf **Nemo**: -lnetcdf
Compilation flags for each component	**Oasis/NEMO/ARPEGE:** -Ep -Pstack -dwW -Wf\"-A idbl4 -pvctl vwork=stack fullmsg\" -sx9 -Wf,-P nh
General comments	Slow-down at GPFS file copy stage Multi-core communications slower than intra-core About 1.7 faster than SX8R configuration
Scalability details	
1 month long simulation, NEMO on 4 cores, one OASIS without additional resource. Cores # below for ARPEGE (but NEMO always faster)	

#cores	Wall clock (secs)
4	7416
5	6912
7	6516

Recommendations	
Expected potential for Peta/Exascaling	medium
Expected effort to reach potential Peta/Exascaling	Weak-scaling: increase of the resolution for Terascale Strong-scaling: code rewriting for Peta/Exascale
Expected potential for optimization	medium
Expected effort to reach the optimization potential	Porting on MPP platforms: fully parallel coupling (OASIS4) and IO improvements

Table 3: ARPEGE-NEMO Model

2.4.1.3 IPSLCM5 Model

General	
Name of the model	IPSLCM5
Reference person	Arnaud Caubel and Marie-Alice Foujols, IPSL, arnaud.caubel@lsce.ipsl.fr foujols@ipsl.jussieu.fr

Version	IPSLCM5 v3
Brief description	Ocean-sea ice-atmosphere-land climate model, ready for Earth System Model (chemistry, marine biogeochemistry and carbon cycle added)
Infrastructure (e.g. coupler)	OASIS v3 (pseudo parallel)
Download	For information about the code download, please contact the reference person
Reference for detailed information	For information on porting and tuning the high-resolution configuration, look at www.
Configuration	
Input configuration	The IPSLCM5 model consists on: LMDZ4+ORCHIDEE (CMIP5) NEMO v3_2 Length of model run: 5 days Region: Global Oasis configuration: Total number of exchanged fields 21 Exchanged fields 17 (LMDZ4 -> Oasis -> NEMO) Exchanged fields 4 (NEMO -> Oasis -> LMDZ4) Coupling period: 24h LMDZ4 Resolution: 280x280x19 (280x280x39 also possible if more memory available) Time step: 72 s (dynamics) and 30 mn (physics) NEMO Region: Global Resolution: ORCA05 ½° (511x722x31) time step: 40 mn
Input files dimension	LMDZ4/ORCHIDEE: 140 MB NEMO: 1.1GB

	OASIS: 900 MB (9GB with calculated weights for runoff) restart files (5 days): LMDZ4/ORCHIDEE: 600 MB NEMO : 1.5 GB OASIS: 23 MB
Output files dimension	LMDZ4/ORCHIDEE: 315 MB (histday.nc and histhf.nc) NEMO: 650 MB

Component Model	
Component model's name	LMDZ4 - ORCHIDEE
Version	CMIP5
License policy	CeCILL licenses
Programming Language(s)	Fortran
Libraries	NetCDF, MPI, OpenMP, Lapack, Blas
Parallelization method supported	MPI1, OpenMP
Component Model	
Component model's name	NEMO
Version	v. 3.2
License policy	CeCILL license
Programming Language(s)	Fortran
Libraries	NetCDF, MPI
Parallelization method supported	MPI1
Component Model	
Component model's name	OASIS
Version	V3
License policy	LGPL, see https://oasistrac.cerfacs.fr/
Programming Language(s)	C, Fortran
Libraries	NetCDF

Parallelization method supported	MPI1, Pseudo parallelism IPSL/CERFACS method

Platform	
Execution platform	BSC MareNostrum, IBM PPC for this resolution. All performances and options below come from this computer. Bull, NEC SX-8/9, IBM Power6, SGI for lower resolutions
Details of the execution platform	CPU Type: IBM PPC CPU Speed: 2.3 GHz # of cores: 72 (LMDZ) + 20 (NEMO) +2 (Oasis) SMP size: 4 cores per nodes Memory: 8 GB per node, ie a maximum of 3 LMDZ MPI process per node at this resolution (280x280x19) Teoretical Peak Performance: 94,2 TFLOPS Installed Libraries: NetCDF, Blas, Lapack, MPI 1, OpenMP Development tools: C compiler: mpicc Fortran compiler: xlf90_r Profiler: Paraver
Libraries	**Oasis**: -lnetcdf **LMDZ4**: -llapack -lblas -lnetcdf **Nemo**: -lnetcdf
Compilation flags for each component	**Oasis**: -O2 -qextname=flush -q64 -qarch=ppc970 -qtune=ppc970 -qrealsize=8 **LMDZ4**: -O3 -qstrict -qarch=ppc970 -qtune=ppc970 -qcache=auto -q64 -qautodbl=dbl4 **Nemo**: -O2 -qsave -qstrict -qrealsize=8 -qsuffix=cpp=F90 -qextname=flush -qsource -qlargepage -qmaxmem=-1
General comments	
Scalability details	
5 days run for LMDZ (MPI only), 1 day run for NEMO (on MareNostrum). Performances depend on the repartition of MPI process on nodes (1/4, 2/4 or 3/4).	

#cores (LMDZ)	Wall clock (secs) BSC		#cores (NEMO)	Wall clock (secs) BSC
1	50000 s		5	760 s
2	18805 s		10	382 s
4	12524 s		20	201 s
8	8610 s		27	141 s
16	4389 s		32	125 s
24	3213 s			
32	2830 s			
40	2329 s			
45	1881 s			
48	1700 s			
56	1561 s			
60	1484 s			
72	1350 s			
80	1316 s			

5 days run

#cores (LMDZ)	#cores (NEMO)	#cores (OASIS)	#cores (total)	Wall clock (secs) BSC
72	20	2	94	1270 s

Recommendations	
Expected potential for Peta/Exascaling	High
Expected effort to reach potential Peta/Exascaling	Weak-scaling: increase of the resolution Strong-scaling: dynamical core code rewriting for Peta/Exascale from scratch or from a shared kernel
Expected potential for optimization	High, Use OpenMP to increase the scalability and the number of cores used by LMDZ4 (OpenMP is available in LMDZ4 sources but running with it needs investigation on BSC). IO server allows asynchronous outputs and increases scalability
Expected effort to reach the optimization potential	Hybrid parallelisation based on MPI/OPenMP need more investigation on all platforms Use of IO server in standard configuration

Table 4: IPSLCM5 Model

2.4.1.4 HadGEM2

General	
Name of the model	Unified Model: HadGEM3A configuration
Reference person	Steve.Mullerworth@metoffice.gov.uk
Version	7.4
Brief description	Atmosphere model
Infrastructure (e.g. coupler)	Input set up through user interface. No coupler (can be used coupled to NEMO through OASIS)
Download	Provided on CD. License is required
Reference for detailed information	Documentation provided with installation disk
Configuration	
Input configuration	Configuration is defined by a User Interface, and different configurations can be supplied. A typical configuration is a resolution of 192 points E-W by 145 points N-S on 63 atmospheric levels, running with a 20 minute time-step for 1 month.
Input files dimension	Typically 4Gb
Output files dimension	Configurable – minimum 2Gb
Component Model	
Component model's name	HadGEM3A
Version	7.4
License policy	License available for academic/research users
Programming Language(s)	FORTRAN + some C
Libraries	GCOM
Parallelization method supported	MPI. Some OpenMP currently being developed

Platform	

Execution platform	IBM Power 6
Details of the execution platform	100 nodes, 32 processors per node.
Libraries	GCOM, mass
Compilation flags for each component	Typical choice: -qextname -qsuffix=f=f90 -qarch=pwr6 -qtune=pwr6 -qrealsize=8 -qintsize=8 -NS32768
General comments	
Scalability details	

One month run

#cores	Wall clock (secs)
32	5800
64	3080
96	2240
128	1800

Recommendations	
Expected potential for Peta/Exascaling	low
Expected effort to reach potential Peta/Exascaling	
Expected potential for optimization	medium
Expected effort to reach the optimization potential	

Table 5: HadGEM2 Model

2.4.1.5 MPI-M

General	
Name of the model	ECHAM5/MPIOM
Reference person	Marco Giorgetta, marco.giorgetta@zmaw.de, MPI-M
Version	COSMOS-1.2.1.1

Brief description	Earth system model consisting of coupled atmosphere, ocean, and land, includes carbon cycle
Infrastructure (e.g. coupler)	IMDI SCE/SRE, OASIS3
Download	Available for download upon request from the following URL: http://www.mpimet.mpg.de/en/wissenschaft/modelle/model-distribution.html
Reference for detailed information	The MPI-M special section in J. Climate Vol.19 No.16
Configuration	
Input configuration	"ASOB" configuration, i.e. coupled climate carbon cycle model ECHAM5J/MPIOM Setup like Control run for preindustrial conditions, as done for ENSEMBLES stream 2. Length of integration in this case = 20 years Oasis configuration: exchanged fields 17 (ECHAM5J → Oasis → MPIOM) exchanged fields 8 (MPIOM → Oasis → ECHAM5J) coupling period: 1 day ECHAM5 Region: Global Resolution: T31L19 (96 x 48) time step: 2400 sec MPIOM Region: Global Resolution: 3° time step: 8640 sec
Input files dimension	specify the size of the input files: 58 Mbyte

Output files dimension	size of the output files: o restart: 36 Mbyte o diagnostic:
Component Models (including coupler)	repeat this section for each of the model components
Component model's name	**ECHAM5**
Version	ECHAM5.4+JSBACH
License policy	"MPI-M Software Licence Agreement", to be signed by each user, See http://www.mpimet.mpg.de/en/wissenschaft/modelle/model-distribution.html
Programming Language(s)	Fortran, C, (inkl OpenMP)
Libraries	NetCDF, CDI, MPI (to name a few)
Parallelization method supported	MPI1, MPI2, OpenMP
Component model's name	**MPIOM**
Version	
License policy	"MPI-M Software Licence Agreement", to be signed by each user, see http://www.mpimet.mpg.de/en/wissenschaft/modelle/model-distribution.html
Programming Language(s)	Fortran, C(inkl OpenMP)
Libraries	NetCDF, CDI, MPI (to name a few)
Parallelization method supported	MPI1, MPI2, OpenMP
Component model's name	**OASIS3**
Version	prism_2_3
License policy	Lesser GNU General Public License (LGPL)
Programming Language(s)	Fortran, C (inkl OpenMP)
Libraries	NetCDF, CDI, MPI (to name a few)
Parallelization method	MPI1, MPI2, OpenMP

supported	

Platform	on platform #1 (repeat the table for all the platforms the code is available on)
Execution platform	Blizzard, IBM P6 parallel SMP
Details of the execution platform	Thomas Jahns, jahns@dkrz.de

Libraries	specify the actual linked libraries
	`for all components:`
	
	`IBM MPI, lapack, essl, blas, mass, netcdf-3.6.3, hdf5-1.8.2, szip-2.1 zlib-1.2.3`

Compilation flags for each component	specify the CFLAGS or FFLAGS used for compiling and for optimizing the execution on the target platform
	`MPIOM:`
	`-q64 -qsuffix=cpp=f90 -O3 -qsuppress=1518-061:1518-128`
	`-qstrict -qarch=pwr6 -qtune=balanced -qzerosize –qessl`
	`-qhot -qxflag=nvecvter -qxflag=nsmine -qfloat=fltint`
	`-qextname -qdpc=e -qrealsize=8`
	
	`ECHAM5:`
	`-q64 -qsuffix=cpp=f90 -qsuppress=1500-036 -O3 –qlist`
	`-qreport -qxflag=nvectver -qxflag=nsmine -qarch=auto`
	`-qtune=auto -qcache=auto -qfloat=fltint –qzerosize`
	`-qessl -bdatapsize:64k -bstackpsize:64k -qextname`
	
	`OASIS3:`
	`-q64 -qfixed=72 -qsuffix=cpp=F -qtbtable=full -O3`
	`-qstrict -qMAXMEM=-1 -Q -qarch=auto -qtune=auto`
	`-qcache=auto -qfloat=fltint -qzerosize -qessl`
	`-qextname -qdpc=e -qrealsize=8`

General comments	The selected model configuration is typical for millennia time scale simulations. The primary goal is to achieve a high turnover rate (many years per 24 hr wall clock time). Hence the model runs at low resolution. This is a special challenge on computers like Blizzard (IBM Power 6), which have a large number of relatively slow processors. Gaining speed by parallelization is difficult for low resolution models. But high resolution models would still have a lower turnover rate.
Scalability details	

No detailed numbers on scalability are currently available for this version of the model since a) the platform is still rather new to us and b) most of the investigations have been directed to the CMIP5-Version of our ESM, which is currently not yet public.

In general we know that the T31/L19-GROB30 version of the coupled model runs at 60 forecast years per day on one node of the IBM.

The version, which is distributed, has been extensively tested on our linux cluster in the beginning of this year, see below.

New numers on these issues can be provided if need arises.

Platform	on platform #2 (repeat the table for all the platforms the code is available on)
Execution platform	tornado, AMD Operon dual cpu based linux custre
Details of the execution platform	Carsten Schmitt, carsten.schmitt@zmaw.de
Libraries	specify the actual linked libraries for all components e.g.: OpenMPI, acml, acml_mv, netcdf-4.0.1-without-hdf5, rdmacm, libverbs, numa, dl, nsl, util, pthread
Compilation flags for each component	specify the CFLAGS or FFLAGS used for compiling and for optimizing the execution on the target platform for all components: -Mpreprocess -O2 -Kieee -fastsse –Mnorecursive -Mextend -tp amd64e -byteswapio -r8

General comments	
Scalability details	

Insert the table reporting the execution time, I/O time and Communication time at different number of cores

The numbers we have do not specify the details you asked for, sorry:

Atm-oce coupled version:

Resolution	CPUs	nproca_ atm	nprocb_ atm	nproca_ oce	nprocb_ coe	Real Time/s
T31GR30	12	8	1	4	1	339
T31GR30	32	6	3	7	2	175

T31GR30	64	6	6	3	9	135
T63GR15	12	1	8	3	1	4069
T63GR15	32	4	4	3	5	1955
T63GR15	64	6	6	3	9	996

Atm-oce-lnd coupled version

CPUs	Prca_a	Prcb_a	Thrd_a	Prca_o	Prcb_o	Thrd_o	Real Time / s
12	1	7	1	2	2	1	675.90
12	1	5	1	2	3	1	910.78
12	1	5	1	3	2	1	915.50
12	5	1	1	2	3	1	548.20
12	1	3	1	2	4	1	874.56
24	12	1	1	11	1	1	389.83
24	12	1	1	1	11	1	344.35
24	6	2	1	11	1	1	383.18
24	6	2	1	1	11	1	361.48
24	2	6	1	11	1	1	394.38

Recommendations	All
Expected potential for Peta/Exascaling	low
Expected effort to reach potential Peta/Exascaling	Switch to a new dynamical core, and software infrastructure
Expected potential for optimization	high
Expected effort to reach the optimization potential	Need to invest time for these items. But time normally is spend on improving the physics….

Table 6: MPI-M Model

2.4.2 Stand-alone models

2.4.2.1 ECHAM

General	
Name of the model	ECHAM5 ECHAM6 in 2nd half 2010

Reference person	Marco Giorgetta, marco.giorgetta@zmaw.de, MPI-M, name, email and institution
Version	ECHAM5.4.02 ECHAM6.0.nn in 2nd half 2010
License policy	See http://www.mpimet.mpg.de/en/wissenschaft/modelle/model-distribution.html
Programming Language(s)	Fortran, C (inkl OpenMP)
Libraries	
Parallelization method supported	MPI, OpenMP, pthreads, hybrid, SHMEM, ...
Download	ECHAM5: yes ECHAM6: later
Reference for detailed information	Roeckner et al., The atmospheric general circulation model ECHAM 5. PART I: Model description, MPI report 349, 2003. http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf
Configuration	ECHAM5 T63L31
Input configuration	192 x 96 horizontal points, 31 levels, Δt = 720 sec. ;6 hrly output, yrly restart files
Input files dimension	19,8 MB
Output files dimension	Output: 15,5 GB / yr; Restart: 200 MB / yr; Monthly means: 141 MB / yr; $\sum$ = 17,8 GB / yr

Platform	on platform #1 (repeat the table for all the platforms the code is available on) See above
Execution platform	name and type of the architecture (SMP, MPP, ...)
Details of the execution platform	specify the HW and SW configuration of the platform
Libraries	specify the actual linked libraries

Compilation flags	specify the CFLAGS or FFLAGS used for compiling and for optimizing the execution on the target platform
General comments	specify know issues on that platform, limitations, comments, ...
Scalability details	

Insert the table reporting the execution time, I/O time and Communication time at different number of cores

Times here have been taken on tornado, they are not available on blizzard so far. We only have the execution time:

CPUs	nproca	nprocb	nthread	reprod.	Real Time	User Time	System Time	Job-id
4	1	4	1	(ref)	244.69 s	944 s	1 s	306838
4	2	2	1	ok	238.50 s	947 s	1 s	306846
4	4	1	1	ok	238.98 s	943 s	1 s	306868
8	1	8	1	ok	182.00 s	1416 s	5 s	306842
8	2	4	1	ok	213.40 s	1675 s	3 s	306848
8	4	2	1	ok	221.13 s	1728 s	5 s	306872
8	8	1	1	ok	123.42 s	960 s	3 s	306962
12	2	6	1	ok	94.74 s	1051 s	6 s	306849
12	3	4	1	ok	146.52 s	1737 s	5 s	306860
12	4	3	1	ok	135.41 s	1546 s	5 s	306876
12	6	2	1	ok	144.79 s	1715 s	5 s	306935
12	12	1	1	ok	93.79 s	1064 s	6 s	309384
16	2	8	1	ok	82.88 s	1190 s	9 s	306855
16	4	4	1	ok	117.03 s	1720 s	8 s	306878
16	8	2	1	ok	115.27 s	1799 s	8 s	306963
20	4	5	1	ok	67.11 s	1186 s	11 s	306880
20	5	4	1	ok	97.62 s	1892 s	12 s	306905
20	10	2	1	ok	114.97 s	2175 s	11 s	309347
24	3	8	1	ok	75.37 s	1663 s	15 s	306864
24	4	6	1	ok	62.47 s	1330 s	14 s	306887
24	6	4	1	ok	87.79 s	1901 s	15 s	306938
24	8	3	1	ok	88.54 s	1996 s	16 s	306967
24	12	2	1	ok	79.66 s	1822 s	15 s	309388
28	4	7	1	ok	56.46 s	1364 s	13 s	306890
28	7	4	1	ok	78.14 s	2132 s	11 s	306949
32	4	8	1	ok	52.82 s	1529 s	27 s	306904
32	8	4	1	ok	66.10 s	2017 s	13 s	306977
36	6	6	1	ok	52.01 s	1671 s	30 s	306941
36	9	4	1	ok	70.78 s	2265 s	18 s	309331
36	12	3	1	ok	68.68 s	2210 s	16 s	309399

40	5	8	1	ok	50.14 s	1768 s	18 s	306910
40	8	5	1	ok	55.60 s	1942 s	23 s	306997
40	10	4	1	ok	76.73 s	2839 s	25 s	309349
44	11	4	1	ok	63.70 s	2496 s	22 s	309375
48	6	8	1	ok	44.62 s	1993 s	23 s	306944
48	8	6	1	ok	49.16 s	2042 s	30 s	307000
48	12	4	1	ok	55.65 s	2493 s	33 s	309403
56	7	8	1	ok	48.20 s	2345 s	27 s	306951
56	8	7	1	ok	42.89 s	1837 s	204 s	307022
60	10	6	1	ok	46.65 s	2528 s	52 s	309365
60	12	5	1	ok	47.00 s	2316 s	33 s	309407
64	8	8	1	ok	42.88 s	2507 s	44 s	307035
72	9	8	1	ok	39.04 s	2384 s	38 s	309340
72	12	6	1	ok	44.71 s	2764 s	38 s	309430
80	10	8	1	ok	42.67 s	2928 s	43 s	309367

Recommendations	See above

2.4.2.2 NEMO

General	
Name of the model	Nucleus for European Modelling of the Ocean
Reference person	Claire Lévy, Claire.Levy@locean-ipsl.upmc.fr
Version	Reference version tag nemo_v3_1
License policy	CeCILL2, free licence
Programming Language(s)	Fortran95, partially Fortran2003
Libraries	MPI, NETCDF
Parallelization method supported	MPI
Download	Yes, available for download: contact reference person
Reference for detailed information	See: http://www.nemo-ocean.eu/About-NEMO/Reference-manuals.
Configuration	ORCA12 (global 1/12 degree resolution). Already used as benchmark at CINES, GENCI, France
Input configuration	ORCA12 : (Global ocean model with 1/12 deg resolution at the Equator) Grid size: 4322 x 3059 x 50.

	Time step is 240 sec. Bench runs 360 time steps (1 day).
Input files dimension	Bathymetry = 158 Mb Initial conditions : 5.3 Gb
Output files dimension	model output = 12 Gb 1 restart file = 78 Gb (likely not in 1 single file)

Platform	
Execution platform	MPP
Details of the execution platform	Has been run on SGI Altix ICE 8200, (Intel Quad core E5472), under Suse Linux
Libraries	Netcdf, MPI, ifort compiler
Compilation flags	-O -i4 -r8 -xS -ip -ftz -fpe3 -fno-alias -sox -I -assume byterecl -convert big_endian
General comments	Improvements on this part of the suite (if portability is taken in account) could directly benefit to the whole NEMO community. The test should be extended to include IO which may be one of the strongest bottleneck for NEMO
Scalability details	
360 time steps , no I/O	

#cores	**Wall clock (secs)**
1000	1412

Other supported platforms	• Cluster linux • NEC vector architecture • Mac OS X • IBM (Power 5, 6, Blue Gene) • Cray XT4 • Bull • PRACE platforms

Recommendations	
Expected potential for Peta/Exascaling	High, but portability remains a necessity

Expected effort to reach potential Peta/Exascaling	Major work expected on vendors side to optimize the IO
Expected potential for optimization	The NEMO reference now includes an IO Server allowing to dedicate some process to IOs. This IO server will be the contribution of the NEMO System Team to the effort
Expected effort to reach the optimization potential	Any action allowing to keep up with the portability

Table 7: NEMO Model