



**Dmi**

Ministry of Climate and Energy

## Danish Climate Centre Report 12-05

# Porting the DMI configuration of the EC-Earth v. 2.3 GCM to the new ECMWF IBM Power 7

Peter L. Langen



# Colophon

**Serial title:**

Danish Climate Centre Report 12-05

**Title:**

Porting the DMI configuration of the EC-Earth v. 2.3 GCM to the new ECMWF IBM Power 7

**Subtitle:**

**Author(s):**

Peter Lang Langen

**Other contributors:**

**Responsible institution:**

Danish Meteorological Institute

**Language:**

English

**Keywords:**

Climate Centre Report, EC-Earth, general circulation model, ECMWF, IBM Power 7

**Url:**

[www.dmi.dk/dmi/dkc](http://www.dmi.dk/dmi/dkc)

**ISSN:**

1399-1957

**ISBN:**

978-87-7478-624-5 (Online)

**Version:**

**Website:**

[www.dmi.dk](http://www.dmi.dk)

**Copyright:**

Danish Meteorological Institute



## Content:

Abstract .....	4
Resumé.....	4
Background .....	5
Compiling .....	5
Running the coupled system .....	7
Post-processing .....	8
Speed and cost.....	9
Output and benchmark test .....	9
Conclusion .....	11
References.....	12
Previous reports.....	12



## Abstract

The EC-Earth v. 2.3 coupled atmosphere-ocean general circulation model has been ported to the European Centre for Medium-Range Weather Forecasts IBM Power7 (c2a), which finally replaces the older Power 6 in late January 2013. This port includes i) compiling and linking the executables and the necessary libraries, ii) setting up the run-environment in terms of input data and scripts to configure, execute and restart the experiments, iii) setting up the post-processing as part of the run-restart cycle, but being performed on the ecgate server rather than the c2a HPC machine. The job flow and logic of the execution and post-processing follows that used at the DMI HPC facilities, easing usability across the systems.

Using 128 cpus, 1 year of integration at T159L62/1deg resolution takes about 1.75 hours and costs 2200 billing units.

The output has been inspected following the EC-Earth benchmark test, where global averages of a number of variables are compared to other groups' implementation of the model for the first and last year of a 10-year coupled run with present day conditions. The numbers for current port compare well with those of the other groups participating in the benchmark.

## Resumé

Den koblede atmosfære-ocean generelle cirkulationsmodel EC-Earth v. 2,3 er blevet porteret til European Centre for Medium-Range Weather Forecasts IBM Power7 (c2a), som endeligt erstatter den gamle Power 6 sidst i januar 2013. Denne portering inkluderer i) kompilering og linking af de eksekverbare programmer og de nødvendige biblioteker, ii) opsætning af run-environment som inputdata og scripts til at konfigurere, eksekvere og genstarte eksperimenterne, iii) opsætning af post-processering som en del af kørsel-genstartscyklen, men udført på ecgate-serveren i stedet for på c2a supercomputeren. Jobflowet og logikken af kørsel og post-processering følger den, der bruges på DMIs supercomputerfaciliteter, og letter derved brugen på tværs af systemerne.

Kørt på 128 cpu'er, tager 1 års inegration ved opløsningen T159L62/1grad omkring 1,75 time og koster 2200 billing-enheder.

Outputtet er blevet inspiceret ifølge EC-Earth's benchmark-test, hvor globale midler af en række variable sammenlignes med andre gruppers implementering af modellen for første og sidste år en af 10-års koblet kørsel under nutidige betingelser. Tallene for denne portering er fint sammenlignelige med dem fra de andre grupper, der deltager i benchmarktesten.



## Background

Given the relatively large amount of resources allocated to the DMI at the European Centre for Medium-Range Weather Forecasts high-performance computing (HPC) system and the fact that the in-house HPC system is often at its maximum capacity, it is desirable that the EC-Earth could be run at ECMWF in the configuration used at DMI. This includes using the same scripts and environment variables for running the model and using the same procedures for post-processing of the output.

With the EC-Earth being developed on the ECMWF facilities (among others), the port is relatively straightforward. It had not been set up on the new IBM Power 7 (called c2a), however, and since the old Power 6 (c1a) will be discontinued on 28 January 2013, it was clear that it was the Power 7 that we should be setting it up on.

In the following, compilation and running of the model takes place on the HPC machine called **c2a**. The scripts that control this run perform some of the post processing, but the majority is done by jobs on the server **ecgate** that are called from within the run-scripts. The code and scripts are stored in the \$PERM-directory (limited but permanent storage, not backed up), the runs are performed in the \$TEMP (large but temporary storage, not backed up) and the post-processed output is copied to the ECFS storage system. This is done using a series of special commands developed for the ECFS (see <http://www.ecmwf.int/services/computing/docs/archives/ecfs/index.html>).

## Compiling

For compiling the model, the documentation by Stefanescu (2009), was followed very closely. The scripts and settings for the IBM Power 6 implementation were followed by copying all `ibmpower6`-scripts and making `ibmpower7` versions of them. For example, in the directory `ifs/src/build/arch` the file `Makefile.in.ibmpower6` was copied to `Makefile.in.ibmpower7` and changes were made as outlined below. Here, and in the following, all paths are relative to the top-level `EC-Earth_v2.3` directory (`/perm/ms/dk/nhy/ecearth2.3`).

The documentation by Stefanescu (2009) details the manner in which the changes are to be made. The following is an overview of files that are created and/or edited and how they differ from the corresponding `ibmpower6`-versions. To repeat the port, it is recommended to use Stefanescu (2009) as the primary guide and to refer to this document for specific values of the edited variables.

### The build directory

In the build directory, the file `settings` should be edited as outlined below. After all the changes have been made in all the component directories (see below), the model is compiled with the scripts `makeoasis3.ksh`, `makeifs.ksh` and `makenemo.ksh`. The Oasis3 must be compiled first and the ifs-compilation takes some time, so it is suggested to submit this job to the queue using `llifs.ksh`.

```
build/settings:
=====
ARCH=ibmpower7
export FORTRANCOMPILER=xlf90_r
export NETCDFDIR=/usr/local/apps/netcdf/3.6.3/LP64
export OASIS3ARCH=aix
```



```
export NEMOARCH=ibmpower7
export NEMOPROCX=4
export NEMOPROCY=8
export NEMOGRID=ORCA1
```

### Compiling the OASIS3 coupler

Compile the oasis-coupler as the first component using the script `build/makeoasis3.ksh`, after having edited the following file:

```
oasis3/prism_2-5/prism/src/mod/oasis3/util/make_dir/make.aix:
=====
MPI_INCLUDE = /opt/ibmhpc/pecurrent/ppe.poe/include/thread64
F90          = mpxlf90_r
F            = $(F90)
f90         = $(F90)
f           = $(F90)
CC          = mpcc_r
LD          = $(F90)
```

### Compiling the IFS atmosphere model

Construct/edit the following files. Wherever `ibmpower7` enters in the name, the file is constructed from the corresponding `ibmpower6` file. After completing these edits, build the IFS using the script `build/makeifs.ksh` or submit to queue using (an edited version of) the script `build/llifs.ksh`

```
ifs/src/build/arch/Makefile.in.ibmpower7:
=====
FC          = xlf_wrap
CC          = xlc_r
MPCC        = mpcc_r -qsmp=omp
MPIF_PATH=/opt/ibmhpc/pecurrent/ppe.poe/include/thread64
use the flags -qarch=pwr7 and -qtune=pwr7 (where there already were pwr6 flags).
```

```
ifs/scripts/cmpif.ksh:
=====
if [ $TARGET = ibmpower7 ];then
  if [[ $COMPILER = auto ]]; then
    COMPILER=xlf90_r
  fi
  MPIF_PATH=/opt/ibmhpc/pecurrent/ppe.poe/include/thread64/mpif.h
  narch=ibm_power7
  SFMFLAG=""
  CPP_DEFINES=""
  export arch=ibm_power7
  export OBJECT_MODE=64
fi
```

```
ifs/scripts/mkabs_fc:
=====
Set LD=mpxlf90_r inside the statement if [[ $ARCH = ibmpower* ]]
```



EMOS-libraries:  
=====

- ifs/src/emos/config.ibmpower7R64: Make from corresponding ibmpower6 file and edit the flags -qarch=pwr7 and -qtune=pwr7
- ifs/src/emos/options\_ibmpower7: Copy from corresponding ibmpower6 file
- ifs/src/emos/bufrdc/sources.ibmpower7: Copy from corresponding ibmpower6 file
- ifs/src/emos/bufrdc/tools/config.ibmpower7R64: Make from corresponding ibmpower6 file and edit the flags -qarch=pwr7 and -qtune=pwr7
- ifs/src/emos/gribex/sources.ibmpower7: Copy from corresponding ibmpower6 file
- ifs/src/emos/pbio/sources.ibmpower7: Copy from corresponding ibmpower6 file

### Compiling the NEMO ocean model

After editing/creating the following files, compile NEMO with the script build/makenemo.ksh and the settings in build/settings:

```
export NEMOARCH=ibmpower7
export NEMOPROCX=4
export NEMOPROCY=8
export NEMOGRID=ORCA1
```

nemo/build/config/ibmpower7.config:  
=====

```
CPP=/usr/ccs/lib/cpp
CPPFLAGS=-P -C -I/opt/ibmhpc/pecurrent/ppe.poe/include/thread64
FC=mpxlf90_r
LDNOMPI=xlF90_r
```

use the flags -qarch=pwr7 and -qtune=pwr7 (where there already were pwr6 flags).

After building the NEMO model, copy the script

```
/nemo/modips1/modeles/IOIPSL/tools/rebuild
to
```

```
/nemo/modips1/modeles/IOIPSL/tools/rebuild_localflio
```

In the latter, edit the line

```
#../../../../nemo_build/flio_rblld < tmp.$$
```

to

```
./flio_rblld < tmp.$$
```

## Running the coupled system

After successfully compiling the model components, go to the directory /perm/ms/dk/nhy/ecearth2.3\_dmi/testrun/scripts/run\_benchmark

This entire directory was copied from the corresponding one on the DMI hpcdev. The idea is to stay as close as possible to the way of setting up, executing and post-processing experiments on the DMI machines. In this directory, the following scripts (among others) are found:



```
run_test_v2.3.ksh
EC_Earth_T159L62_NEMO4x8_ibm7_wPostP.ksh
setupifs.ksh
setupoasis3.ksh
setupnemo.ksh
run_EC_Earth_ppcmip_ECGATE.ksh
```

They have all been modified from the original ones to accommodate for

- A different machine architecture
- A different directory structure
- A different storage structure
- A different node/processor-layout
- A different load balancing system (LoadLeveler instead of PBS)
- Performing run-execution and post-processing on two different machines (c2a and ecgate)

The flow of a job-cycle is as follows:

Edit `run_test_v2.3.ksh` for experiment name, length of run (number of re-submits of one year's duration), number of processors, nodes, etc.

1. Execute this script on the interactive node. This will set up a RUNDIR on c2a, a SAVEDIR on ecgate and a directory for storage on ECFS. To the RUNDIR, it will make local copies of the `EC_Earth_T159L62_NEMO4x8_ibm7_wPostP.ksh`, `run_EC_Earth_ppcmip_ECGATE.ksh` and setup-scripts. It also links/copies input data to the RUNDIR (they are currently stored in the \$TEMP directory and backed up on ECFS). Finally, it calls the local RUNDIR-version of `EC_Earth_T159L62_NEMO4x8_ibm7_wPostP.ksh` (local version named `EC_Earth_run.ksh`).
2. The `EC_Earth_run.ksh` script creates a LoadLeveler-submittable jobscript called `RUN_${YEAR}` and submits this to the queue.
3. The `RUN_${YEAR}` script sets the environment on the backend nodes and calls the setup-scripts which generate the namelists for the three model components. It then executes the model for one year.

After the one year run has completed, the `RUN_${YEAR}` script copies over the output from c2a to ecgate. A part of the ocean-postprocessing is done in this script, because it uses an executable that was built along with the NEMO-model. This was compiled on the c2a and must thus be executed there. The `RUN_${YEAR}` script then calls three different post-processing scripts (see next section). Finally it checks whether it should resubmit, and if so, it updates the proper environment variables and calls the `EC_Earth_run.ksh` script. The cycle then continues from step 2. above.

## Post-processing

The post-processing mainly takes place on ecgate. This has several advantages and some disadvantages. The advantages are that the ecgate is for exactly this kind of job and that ecgate-jobs are not billed to the user's account (i.e., it is free!). The disadvantage lies mainly in the difficulties that were associated with setting up the above scripts.

As mentioned above, a small part of the post-processing takes place within the run-script and is

therefore performed on the c2a. This part covers some conversions of the NEMO-output which is done with a tool compiled on the c2a. The result of this conversion is copied along with all other output to the ecgate server. Next, the log-files from the different model components are tarred together into one file and stored in the Log directory under the RUNDIR. Then three post-processing jobs are called from the c2a but queued and executing on the ecgate. These jobs use the cdo and wgrib commands, so be sure that these are available. On ecgate, I installed wgrib under ~nhj/tools/wgrib. To make cdo available, “use cdo”.

- The first post-processing job creates time series and monthly mean files variable-by-variable from the collected output coming out of the IFS. This job works on the grid-point variables. The results are copied to ECFS.
- The second post-processing job also creates time series and monthly mean files variable-by-variable from the collected output coming out of the IFS. This job, however, works on the spherical harmonics variables. The results are copied to ECFS.
- With a part of the ocean post-processing already performed on the c2a, the third script mainly copies the results to ECFS.

The first of these scripts takes much longer than the others, so at the end of this script there is a call to a final script that copies over the entire SAVEDIR to ECFS.

After these jobs have been submitted to the ecgate, the restart files that came out of the one year run are tarred together and copied to the SAVEDIR and the old ones are deleted. In this way, all restart file are kept in the SAVEDIR, while only the newest that are needed by the next cycle are kept in the RUNDIR. This copy ends before the final script copies the entire SAVEDIR to ECFS, so these restarts are also included in this final job.

## Speed and cost

On the IBM Power7 machine c2a, each node has 64 logical cpus. To avoid being billed for more processors than are used, one should thus use multiples of 64 processors for the job. Until now, the setup has been tested with 128 processors distributed with 1 for the coupler, 32 for the ocean and 95 for the atmosphere. The model is run at T159L62 resolution in the atmosphere and 1 degree resolution in the ocean, and with the above configuration, one year takes about 1.75 hours and the cost is about 2200 billing units. In 2012, DMI had in total about 28 million units allocated.

## Output and benchmark test

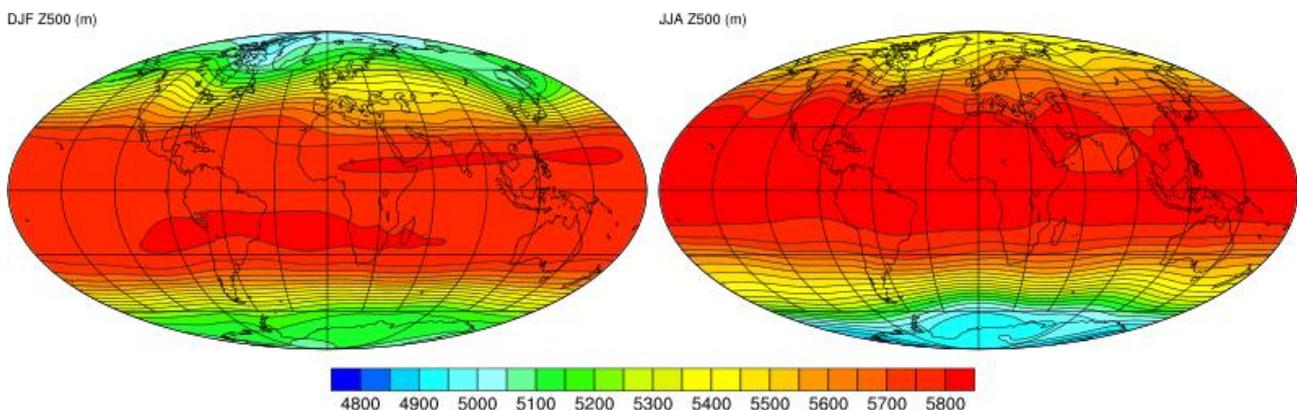


Figure 1: 10 year (1990-1999) average 500 hPa geopotential height (m) for DJF (left) and JJA (right).

A part of the port of the EC-Earth is a benchmark test, where ten years (1990 to 1999) are run in the

same configuration as the CMIP5 EC-Earth experiments. Details are found at <http://eearth.knmi.nl/index.php?n=PmWiki.Benchmark-v2-3>. As a check on whether this run produces sensible results, Figure 1 shows the 10 year DJF and JJA average over the period for the 500 hPa geopotential height. The numbers are as expected in terms of absolute values, meridional gradients, zonal asymmetries and seasonality. Averaged also over the 10 years, the March and September Arctic sea ice cover is shown in Figure 2. This field is also reasonable in its spatial pattern and seasonality.

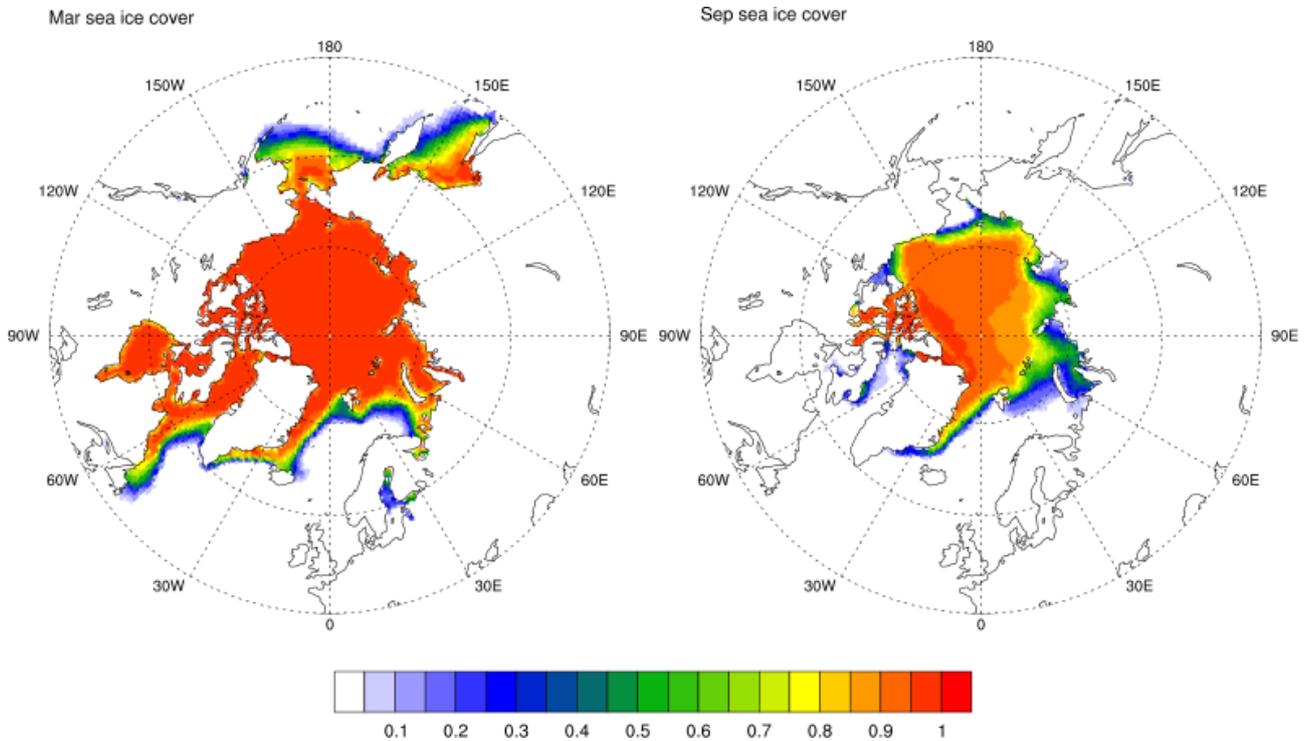


Figure 2: 10 year (1990-1999) average sea ice cover for March (left) and September (right).

In the benchmark test, the annual and global averages of a number of fields are compared between the different groups' setups for years 1990 and 1999. The following two tables repeat the other groups' values along with the values obtained in the current port (last row in red). Note that land areas are not masked when computing the means for the ocean. Land points thus enter into the means with the value 0 and therefore the reported SST, SSH and SSS are lower than expected. Moreover, the results are from a very short simulation, so they are only to be used to test if the model setup is correct.

For most of the variables, this port lies within range of values from the other groups. For the 1990 numbers, the exception is SST which is the highest of the range. Assuming randomly distributed round-off errors, it is not surprising that one value is highest or lowest of 9 setups in one out of 11 variables. For the 1999 numbers, the values of SLHF (sensible heat flux), TP (total precipitation) and SST are lower than the range. Again, with 7 models subjected to 11 tests, it is not alarming to be lowest in 3 variables. In all cases, the numbers compare well with the other groups' results, i.e., there are no concerning outliers among the numbers obtained with this port.

Year 1990	T2M (K)	TCC (fraction)	TTR (W/m <sup>2</sup> )	TSR (W/m <sup>2</sup> )	SSHF (W/m <sup>2</sup> )	SLHF (W/m <sup>2</sup> )	MSL (hPa)	TP (mm/d ay)	SST (°C)	SSS (psu)	SSH (m)
<b>MISU</b>	285.926	0.6349	-240.224	241.464	-18.465	-81.129	1010.94	2.818	12.678	24.801	0.919



<b>ISAC</b>	285.904	0.6370	-240.177	241.376	-18.461	-81.253	1010.905	2.823	12.639	24.797	0.918
<b>UL</b>	285.92	0.634	-240.30	241.57	-18.640	-80.842	1010.9	2.808	12.658	24.774	0.919
<b>ME - c1a standalone</b>	285.90	0.63815	-240.14	241.44	-18.63	-81.1	1010.9	2.82	12.693	24.661	0.9145
<b>fjka - c1a prepIFS</b>	285.971	0.6353	-240.248	241.440	-18.638	-81.055	1010.917	2.815	12.669	24.800	0.919
<b>BSC (Mare- Nostrum)</b>	285.93	0.6347	-240.25	241.58	-18.52	-81.02	1010.91	2.814	12.67	24.797	0.9197
<b>IC3</b>	285.92	0.6363	-240.19	241.51	-18.50	-81.12	1010.93	2.811	12.69	24.799	0.9193
<b>DMI</b>	285.939	0.632	-240.449	241.879	-18.760	-80.755	1010.90	2.805	12.666	24.810	0.9194
<b>DMI/ECMWF This port</b>	<b>285.944</b>	<b>0.634</b>	<b>-240.246</b>	<b>241.645</b>	<b>-18.554</b>	<b>-81.044</b>	<b>1010.92</b>	<b>2.815</b>	<b>12.702</b>	<b>24.774</b>	<b>0.9180</b>

Year 1999	T2M (K)	TCC (fraction)	TTR (W/m2)	TSR (W/m2)	SSHF (W/m2)	SLHF (W/m2)	MSL (hPa)	TP (mm/day)	SST (°C)	SSS (psu)	SSH (m)
<b>MISU</b>	286.484	0.6325	-241.055	242.026	-18.397	-82.248	1010.96	2.856	12.969	24.800	0.926
<b>ISAC</b>	286.503	0.6342	-240.941	241.767	-18.473	-82.091	1010.965	2.850	12.954	24.796	0.926
<b>UL</b>	286.36	0.636	-240.6	241.36	-18.349	-82.118	1011.0	2.851	12.897	24.771	0.923
<b>fjka - c1a prepIFS</b>	286.478	0.6336	-240.872	241.773	-18.451	-81.989	1010.946	2.847	12.955	24.805	0.927
<b>BSC/IC3 (MareNostrum)</b>	286.323	0.634	-240.652	241.501	-18.407	-81.986	1010.940	2.848	12.893	24.774	0.926
<b>DMI</b>	286.474	0.633	-240.863	241.839	-18.490	-81.922	1010.96	2.846	12.971	24.815	0.926
<b>DMI/ECMWF This port</b>	<b>286.393</b>	<b>0.633</b>	<b>-240.804</b>	<b>241.896</b>	<b>-18.465</b>	<b>-81.781</b>	<b>1010.95</b>	<b>2.841</b>	<b>12.882</b>	<b>24.772</b>	<b>0.926</b>

## Conclusion

The EC-Earth v. 2.3 has been ported to the ECMWF's IBM Power7 (c2a), which finally replaces the older Power 6 in late January 2013. This port includes i) compiling and linking the executables and the necessary libraries, ii) setting up the run-environment in terms of input data, and scripts to configure, execute and restart the experiments, iii) setting up the post-processing as part of the run-restart cycle, but being performed on the ecgate server rather than the c2a HPC machine. The flow and logic of the execution and post-processing follows that used at the DMI facilities, easing usability across the systems.

Using 128 cpus, 1 year of integration at T159L62/1deg resolution takes about 1.75 hours and costs 2200 billing units.

The output has been inspected following the EC-Earth benchmark test, where global averages of a number of variables are compared to other groups' implementation of the model for the first and last year of a 10-year coupled run with present day conditions. The numbers for the current port compare well with those of the other groups participating in the benchmark.



## References

Stefanescu, S. (2009). Standalone environment for compiling and running the EC-EARTH system - Update of Kristian Mogensen Technical Notes - June 2007. ECMWF Technical Notes, <http://eearth.knmi.nl/eearth2.pdf>

## Previous reports

Previous reports from the Danish Meteorological Institute can be found on: <http://www.dmi.dk/dmi/dmi-publikationer.htm>