is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

**IS-ENES2 Final General Assembly**
**16-18th January 2017**
*Université Pierre-et-Marie-Curie, Paris, France*

# Coupled model benchmarks: current state, perspectives, and lessons learnt


**WP10/JRA2 Participants:**

- *Joachim Biercamp, Irina Fast (DKRZ, lead)*
- *Marie-Alice Foujols, Arnaud Caubel (IPSL, co-lead)*
- *Giovanni Aloisio, Silvia Mocavero, Fiore Sandro (CMCC)*
- *Reinhard Budich (MPG)*
- *Sophie Valcke, Eric Maisonnave (CERFACS, co-lead)*
- *Rupert Ford (STFC)*
- *Chandan Basu, Hamish Struthers (LiU)*
- *Graham Riley (University of Manchester)*
- *Paul Selwood (MetOffice)*

# Motivation

*The availability of **portable**, **easy-to-use**, and **well documented** climate model **benchmarks** can be useful for more efficient collaboration between the climate modelling community and hardware/software vendors.*

***WP10/JRA2** effort: Prepare a set of benchmarks of varying complexity based on **real production codes** used in European **climate research** to make ESMs ready for emerging architectures and extreme scale computing.*

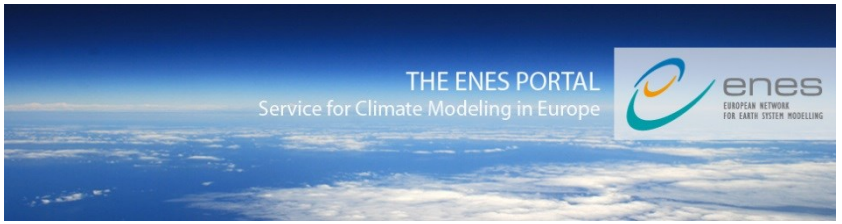
# Achievements within IS-ENES2

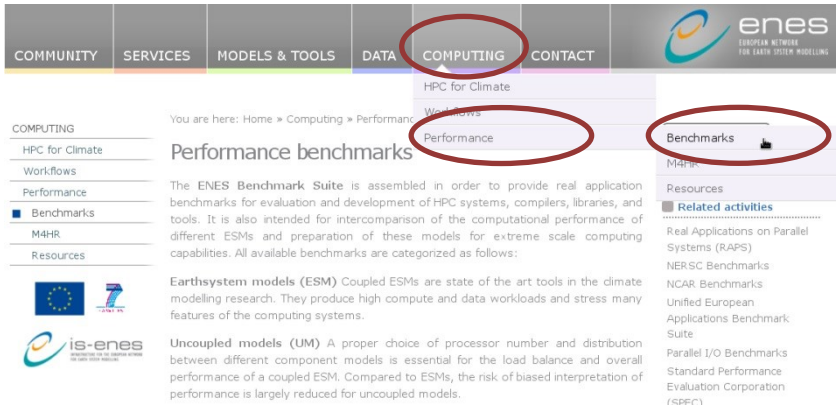
## *Available Benchmarks:*

- *Earth System Models (ESMs)*
  - *MPI-ESM1*
  - *IPSLCM*
  - *CMCC-CESM-NEMO*
  - *EC-EARTH*
- *Uncoupled Models*
  - *ICON*
  - *NEMO (?)*
  - *COSMO-CLM (in preparation)*
- *Benchmarks for evaluation of coupling strategies (→ Talk by Sophie Valcke)*
- *Kernels*
  - *NEMO tracer advection kernel*
  - *ICON decomposition and halo communication kernel (in preparation)*

## *Information on ENES Portal:*

https://verc.enes.org/computing/performance/benchmarks



# Achievements within IS-ENES2

- *Central platform for distribution of all benchmarks*
  *https://redmine.dkrz.de/projects/enes-benchmark-suite*
- *Instructions on benchmark execution*
- *Performance reference for all benchmarks*

Home Projects Help

DKRZ

ENES Benchmark Suite

Search:

Overview | Activity | News | Documents | Wiki

## ENES Benchmark Suite - Wiki

History

The **ENES** (European Network for Earth System modelling) **Benchmark Suite** is a collection of applications of different complexity ranging from simple kernels to full coupled Earth System Models (ESM).
The suite is assembled in order to provide real application benchmarks for evaluation and development of HPC systems, compilers, libraries, and tools. It is also intended for intercomparison of the computational performance of different ESMs and preparation of these models for extreme scale computing capabilities.

All available benchmarks are categorized as follows:

- Earth system models (ESM)
  - MPI-ESM1 Benchmark
  - IPSLCM Benchmark
  - CMCC-CESM-NEMO Benchmark
  - EC-EARTH Benchmark
- Uncoupled models (UM)
  - COSMO-CLM Benchmark
  - ICON Benchmark
  - NEMO Benchmark
- Coupler Benchmarks
- Application kernels derived from models
  - ICON Kernels
  - NEMO Kernels

enes
BENCHMARK SUITE

*Versions* *Size* *Downloads* *MD5 checksum*

Home Projects Help

DKRZ

ENES Benchmark Suite »

MPI-ESM1 Benchmark

Search:

Overview | Activity | Roadmap | Issues | News | Documents | Wiki | Files

## Files

| File | Date | Size | D/L | MD5 |
|---|---|---|---|---|
| MPI-ESM1 Benchmark v13.1 | | | | |
| mpiesm1_13.1_setup.tar.gz | 01/15/2016 05:58 PM | 15.5 MB | 40 | 9415d637b6db4f9eff250a7d8cd4f2b6 |
| README.txt | 01/15/2016 05:59 PM | 4.09 KB | 53 | e1b47e27db3e02ff3c6280a977cc46b3 |
| MPI-ESM1 Benchmark v13.0 | | | | |
| mpiesm1_13.0_setup.tar.gz | 12/23/2015 11:58 AM | 15.5 MB | 30 | cc491be49e8132692c5107fc2ee3b607 |
| README.txt | 12/23/2015 11:33 AM | 4.09 KB | 57 | b1bd4376c595027d84c58c3e684ebb09 |

# Open issues and lessons learnt

- ***To stay relevant, benchmarks need to be kept up-to-date.*** *Current benchmarks reflect the **state** of Earth System Models used for **CMIP5**. Any commitments to **keep** the prepared benchmark packages **up-to-date after IS-ENES2**?*

- ***Assessing of climate model correctness is hard.*** *No **evaluation** of **correctness** of benchmark runs is provided for most benchmarks. An objective, automatable methodology is needed for verification of correct benchmark execution to exclude errors due to oversights in porting, compiler bugs, too aggressive compiler optimisation etc. Can we use **ensemble-based consistency test** (Baker et al., 2015) for this?*

# Open issues and lessons learnt

- ***No automatic collection of metrics for assessment and comparison of benchmark performance.*** *Basically, only* ***execution time*** *is collected and made available. Need for a standard benchmarking environment? (CPMIP, Balaji et al. 2016?)*

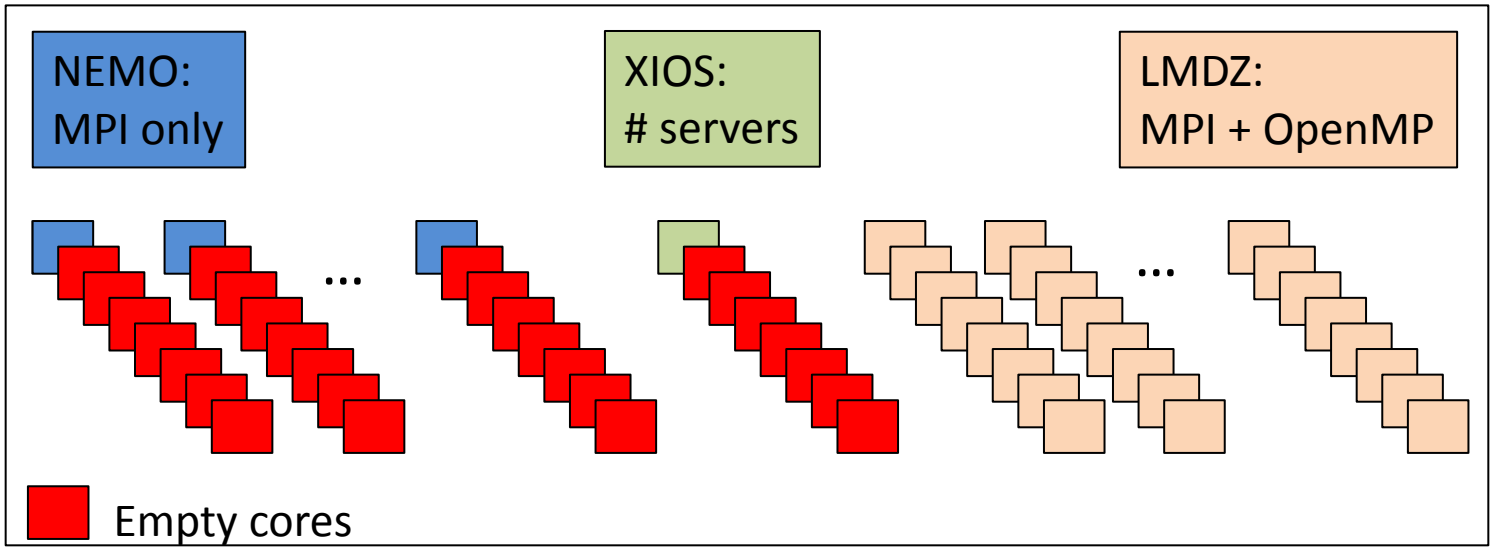
- ***The access and download policy is cumbersome.*** *Interested parties need to contact different persons to get access to different benchmarks. Can this approach be simplified? Is free download (upon registration and acceptance of licence agreements) possible?*

- ***Well-curated benchmarks can be of great help after procurements for routine monitoring of system performance.*** *Added value for continuous integration tool (Jenkins, BuildBot etc.).*

# Open issues and lessons learnt

- ***Many job scheduler/workload manager have currently only limited support for heterogeneous resources requirements of different ESM components.*** *Complex workarounds are required to avoid waste of resources due to (unintentionally) underpopulated nodes.*

Example: IPSLCM6 - coupling of MPI-only model (NEMO) to a hybrid model (LMDZ)




Courtesy Marie-Alice Foujols (IPSL)

# Open issues and lessons learnt

- ***Development and support of prepared benchmarks can be continued, extended (ESiWACE project) or stopped depending on community and vendors feedback.*** *Some vendors have already been informed about available benchmarks and asked for feedback.*

  ***NVIDIA*** *stated they were aware of ESM specific features (MPMD, hybrid parallelisation, load balancing of different components etc.) but would focus their efforts on stand-alone models and kernels. Porting and load balancing of coupled ESMs should occur in cooperation with provider of HPC system. (my account of meeting on January 12th 2017)*

# Open issues and lessons learnt

## *NEC*

- ***Generally appreciates** the early availability of application benchmarks.*
- *The value of benchmark is lowered if it is too complex, hard to execute, and imposes high requirements on human resources and available benchmarking systems.*
- *It is not sufficient to provide real applications that can be compiled and run on the development platform only. Adaptations to alternative platforms and system-specific optimizations need to be allowed if new developments and new hardware-architectures should be considered.*
- *Development of **optimal benchmarks** could be an **iterative process** of testing, learning and understanding that involves climate scientists and vendors.*