Computational issues of the CMCC Carbon Earth System Model (CMCC-CESM)

ANS Division **S. Masina**, M. Vichi, P.G. Fogli, T. Lovato*, W. McKiver
SCO Division **G. Aloisio**, S. Mocavero, I. Epicoco

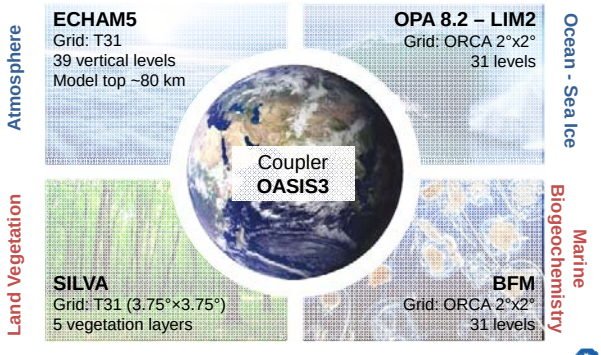
CMCC structure

The Euro-Mediterranean Centre on Climate Change (CMCC) is a national initiative of scientific research in the field of climate change.

Research Divisions

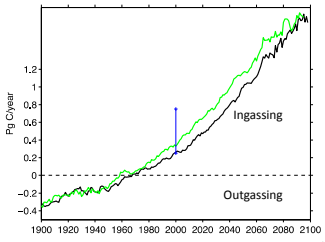
- **Numerical Applications & Scenarios**
- **Scientific Computing & Operations**
- **Climate Impacts and Policy**
- **Impacts on Forests and Agriculture**
- **Impacts on Soil and Coasts**
- **Services for Climate**

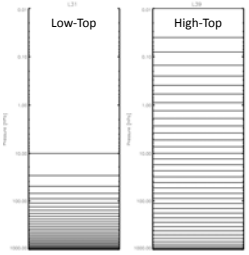
CMCC-Carbon Earth System Model



Stratosphere-ocean carbon interactions

Role of past and future stratospheric ozone changes in the SH climate and global carbon uptake, as simulated by the CMCC-CESM through troposphere-stratosphere coupling and to the CO_2 air-sea fluxes.





Atmospheric model grid: T31L39 ~3.75x3.75deg
Stratospheric component with Top at 80 km

Cagnazzo et al. (under revision in Climate Dynamics)

Key computing issues

1) The main bottleneck to the performance skills of the CMCC-CESM (CMIP5) is represented by the coupling of the NEMO and BFM models.

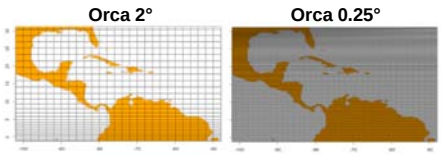
=> A specific analysis of the NEMO-BFM coupling performances is required.

2) The adoption of high resolution models is strongly promoted to achieve a better description of the system and, thus, parallelization is needed.

=> An optimal domain decomposition is indeed a crucial issue!

Impact of resolved scales on global models

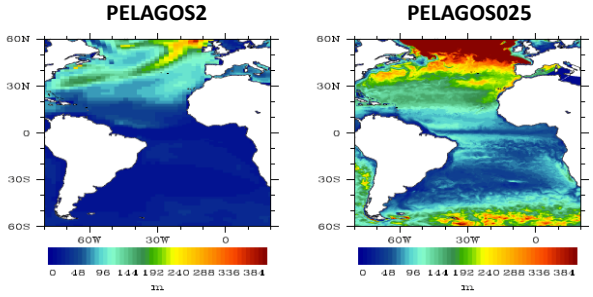
The coupling at the global scale between NEMO and BFM is referred as **PELAGOS** (after Vichi et al., 2007)



MODEL SETUP	PELAGOS2	PELAGOS025
Horizontal Grid	182 x 149	1442 x 1021
Vertical Levels	31	50
Time Step	96 min	18 min
Memory usage	~ 20 Gb	~ 850 Gb
Single output storage	~ 2 Gb	~ 18 Gb

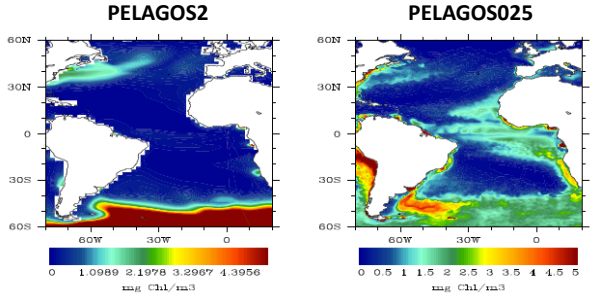
McKiver et al. @ ICES Conference 2012

Scales(1): Mixed Layer Depth



McKiver et al. @ ICES Conference 2012

Scales(2): Chlorophyll-a



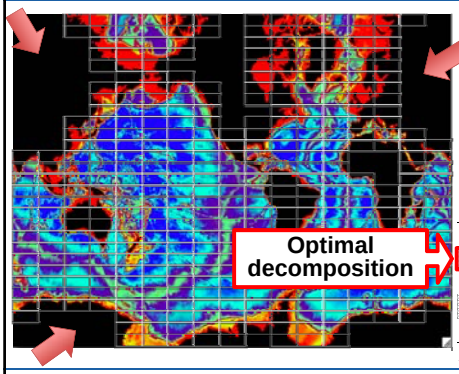
McKiver et al. @ ICES Conference 2012

PELAGOS025 computational performance

- **PELAGOS025** is actually composed by the NEMO v3.4 and BFM v4 at 1/4 of degree resolution at the global scale.
- The **Calypso** parallel computing system available at CMCC is composed by 30 computational nodes, each one mounting 16 CPU Power6 dual-core (32 physical cores per node) and 128 GB of main memory.

The **optimal domain decomposition** for each level of parallelism was identified through a two step procedure by using a modified version of the available NEMO tool (namely MPP_PREP in v3.3-later).

Optimal domain decomposition

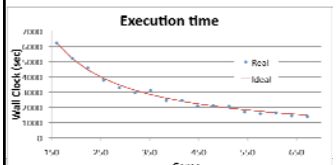


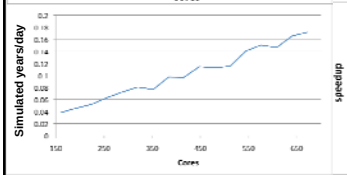
Step 1
Generate the domain decomposition by excluding land domains.

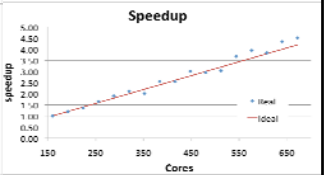
Step 2
Ranking of all the domain decomposition for each procs total by the highest # of ocean points among all the processes.

i x j (256 procs)	max (w)
30x15	25000
32x20	33500
36x16	44000
22x40	56000

Scalability analysis





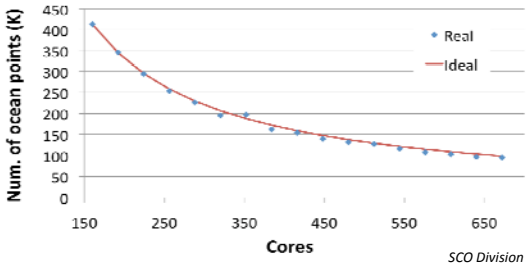


There is a good scalability of the PELAGOS025 model on Calypso, but …

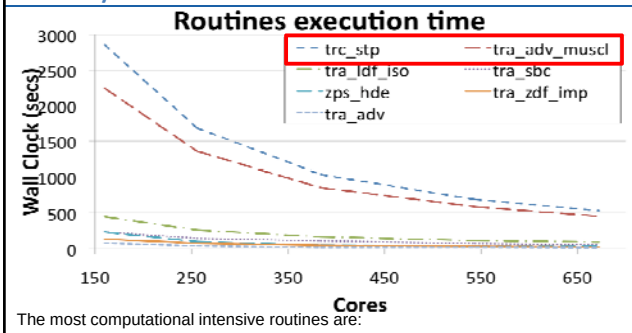
the present analysis is limited by the maximum number of available nodes!

Behaviour of the processes loads

The evolution of the most loaded process for each configuration (in number of ocean points) lead to the identification of the superlinearity effect (Gustafson, 1990). This effect is due to the different balancing level for the different configurations.



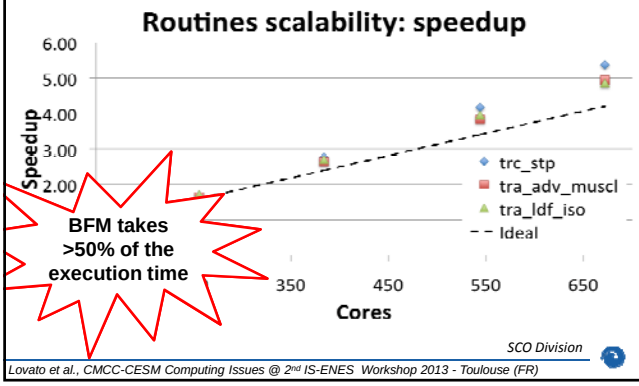
Analysis at the routine level



The most computational intensive routines are:

- **trc_stp** (BFM reaction-transport processes)
- **adv_muscl** (advection scheme usage for nearly 60 variables!!!)

Analysis at the routine level (2)



PELAGOS025 memory model

A linear memory model has been defined to establish the memory requirements for each process:

$$M(w,d) = \alpha w + \beta d + C$$

-*w*: ocean points in the subdomain;
-*d*: land points in the subdomain;
-*C*: constant memory allocated by the process.

The coefficients have been evaluated using regression analysis:

$$\alpha = 7.17KB, \beta = 1.03KB, C = 421.44MB$$

In this way it is possible to asses *a priori* the efficiency of each level of parallelism to be used within the machine.

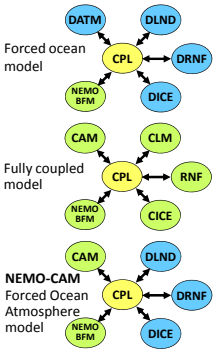
Final remarks

- An *ad hoc* decomposition of the domain allowed to improve and optimize the computational performances of PELAGOS025.
- The memory model allows to better define the use of computational resources and to avoid an inefficient deployment of the code in the machine.

Work perspectives

- The scalability of the model is significantly related to the "dimension" of the problem solved by the model (e.g., the massive use of transport subroutine).
- Actual problems are relevant also to the development of future applications of ESM, based on the same models infrastructure.

Toward the new CMCC Earth System Model



The core activity is the integration of the NEMO ocean model into the NCAR CESM infrastructure (www.cesm.ucar.edu)

CESM consists of several geophysical models plus a driver/coupler (cpl) that coordinates the models and exchanges data.
Each model component may be:

- active (green): fully prognostic, state-of-the-art model
- data model (blue): cycles input data

Two **NEMO-CAM** global setup are currently being tested:
1 deg (ORCA1 L46): CMIP5 like runs, carbon cycle, …
0.25 deg (ORCA025 L50): seasonal prediction

ANS Division, P. G. Fogli

Thanks for Your Attention

Contacts
ANS Division: *Tomas Lovato (tomas.lovato@cmcc.it)*
SCO Division: *Italo Epicoco (italo.epicoco@cmcc.it)*

Global NEMO 1/16° performances

- Configuration
  - NEMO Global at 1/16°
  - Grid points: 5762x3133x100
- Parallel Architecture
  - CPU: Intel SandyBridge - 8 cores
  - Node with 2 CPU – 16cores/node
  - Memory: 4GB/core