Case Study in Large Scale Climate Simulations: Optimizing the Speedup/Efficiency Balance in Supercomputing Environments

Muhammad Asif[1], Andrés Cencerrado[2], Oriol Mula-Valls[1], Domingo Manubens[1], Ana Cortés[2] and Francisco Doblas-Reyes[1, 3]

(1): Institut Català de Ciències del Clima (IC3), Barcelona, Spain. (2): Computer Architecture and Operating Systems Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. (3): Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. Contact: `muhammad.asif@ic3.cat`

Abstract

In this work we present the EC-Earth coupled climate model, which is a seamless Earth System Model (ESM) used to carry out climate research in 24 academic institutions and meteorological services from 11 countries in Europe. This model couples several components and it is continuously under development.

As a coupled model, EC-Earth consists of several different submodels, each of them presenting different degrees of performance because of their different features and complexities. This represents an important disadvantage at the time of consuming computing time in supercomputing environments, which usually is granted by the corresponding host institution provided that the efficient usage of the computing resources is demonstrated.

For this reason, a study to determine the best distribution of computing processors between components was carried out to assess empirically the different performance of the model depending on the amount of computing processors allocated to each component. This experimentation allowed us to optimize the speedup/efficiency balance of the model. Moreover, the obtained results highlight the important drawback caused by the low scalability of one of the EC-Earth components.

EC-Earth, the European Community Earth System Model

EC-Earth is a seamless Earth System Model (ESM) selected by IC3 to perform climate research as end users and also as part of the EC-Earth consortium. In this study EC-Earth v3.0.1 was used and here is below a brief description of EC-Earth components:

- **IFS:** is used for weather prediction at the European Center for Medium-range Weather Forecasts (ECMWF), and is also used as the atmosphere component in EC-Earth. Besides, it is used throughout Europe by many institutes and universities for weather and climate forecasts and research. IFS discretizes the 3D Navier-Stokes equations and it uses a spectral method to compute the dynamics of the atmosphere. Atmospheric physics, e.g., precipitation, is calculated on each grid-point, for which the domain is decomposed in two dimensions. Therefore, the solution needs to be converted from spectral space in order to calculate the dynamics and grid-point space so that it is possible to calculate the physics at each grid-point.
- **NEMO:** the ocean component includes several other components besides the ocean circulation, including sea-ice and biogeochemistry. It is based on a primitive equation model numerically solved on a global ocean curvilinear grid known as ORCA. It discretizes the 3D Navier-Stokes equations and uses MPI with a 2D domain decomposition approach as parallelization strategy. The ocean component is also used in several climate models that are part of the assessment reports of Intergovernmental Panel on Climate Change (IPCC).
- **OASIS3:** the coupler is a software that allows synchronized exchanges of coupling information between the IFS and NEMO components. For doing this; the coupler could use one MPI process for all coupling fields or distribute multiple coupling fields among different MPI processes (this is called *pseudo-parallelism*). OASIS3 is currently used by approximately 30 climate modeling and operational weather forecasting groups in Europe, USA, Canada, Australia, India and China.

EC-Earth works with different configurations of IFS and NEMO and for the sake of experimental analyses of this work the following two configurations were chosen:

a) **Standard configuration (T255-ORCA1)**: IFS is run with configuration T255 which works with a grid-spacing of approximately 80 km and NEMO is run with ORCA1 which translates to about 110 km. In this configuration since IFS incorporates more physical processes; it requires several times more computing resources than NEMO. The input data for T255 are about 1.5 GB in size and for ORCA1 are about 0.8 GB.

b) **High configuration (T511-ORCA025)**: in this configuration; the grid-spacing are about 40 km for T511 and 28 km for ORCA025 and the size of input data are about 8.2 GB and 11 GB respectively.

Scalability Analysis and Discussion

A well-known goal among the climate research community is to be able to simulate 10 years per wall clock day, which means to simulate approximately 1 hour per second.

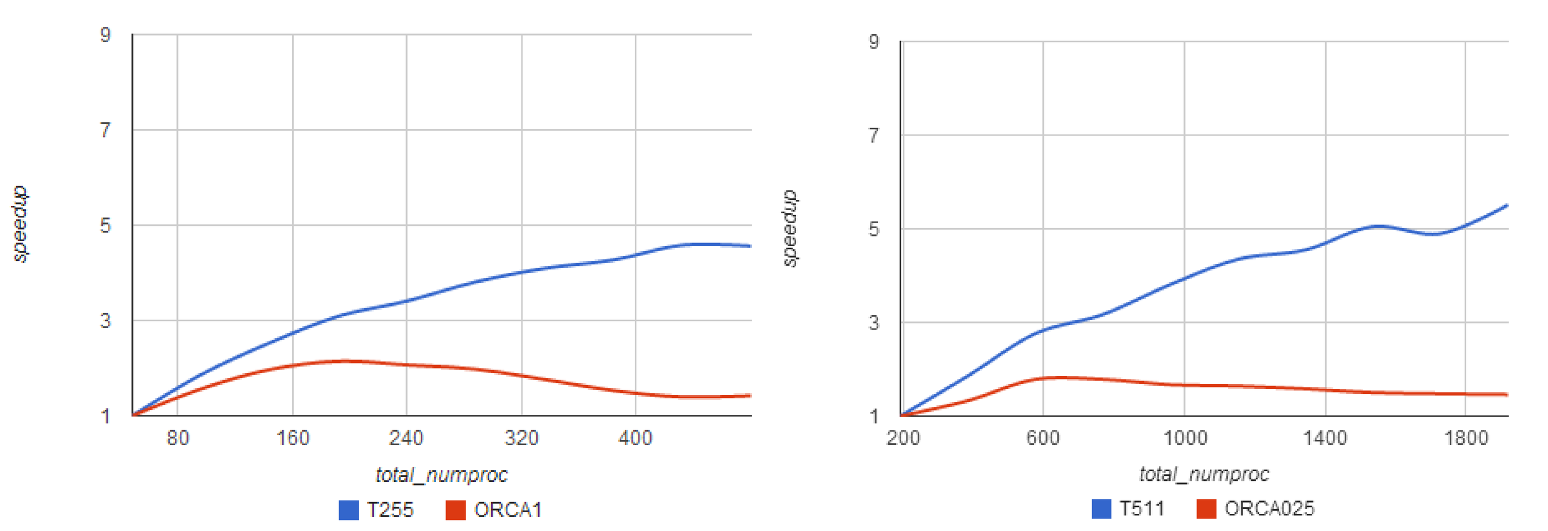

Figure 1: EC-Earth ran as standalone/uncoupled mode for 10 simulation days. Left: standard configuration. Right: high configuration.

MareNostrum III was used as computing platform for this study (2x Intel SandyBridge-EP E5-2670/1600 8-core processor at 2.6 GHz, 8x4GB DDR3-1600 DIMMS (2GB/core) and Infiniband FDR10 network) and at first, the IFS and NEMO were tested in standalone/uncoupled mode as results are shown in Figure 1.

Based on the different plots shown in Figures 2 and 3, a criteria to optimize the speedup/efficiency balance could be inferred as the model is run in coupled mode. In case of standard configuration, it could be seen that by allocating 48 processors to the NEMO component and involving approximately 300 total number of processors the most optimum speedup and efficiency results approximately 3 and 1.3 respectively could be achieved. On the othe hand, in case of high configuration the most optimum speedup value obtained is 1.85 with 1.05 efficiency and approximately 1000 total number of processors utilized where 576 processors were assigned to the NEMO component.

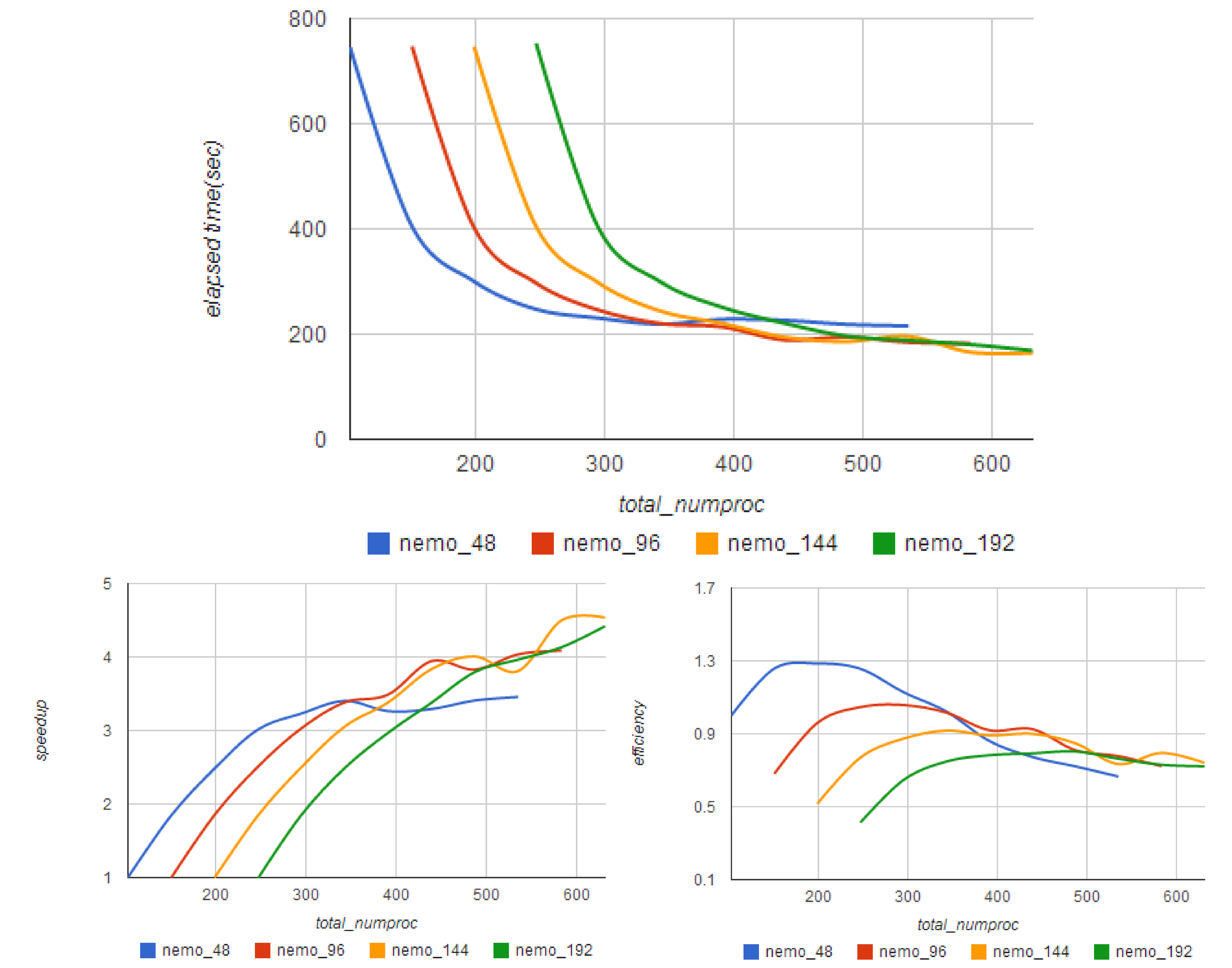

Figure 2: EC-Earth ran in coupled mode for 10 simulation days. Standard configuration (T255-ORCA1) plots (upper), (lower-left) and (lower-right) refer to elapsed time (average of 3 sets), speedup and efficiency respectively and the blue, red, orange and green curves correspond to results for tests with NEMO using 48, 96, 144 and 192 processors while speedup and efficiency values are relative to the reference test with 103 processors. In each case, the number of processors for OASIS3 is 7 and the number of processors for IFS is the remainder of subtracting the number of processors for NEMO and the number of processors for OASIS3 from the value at the horizontal axis.

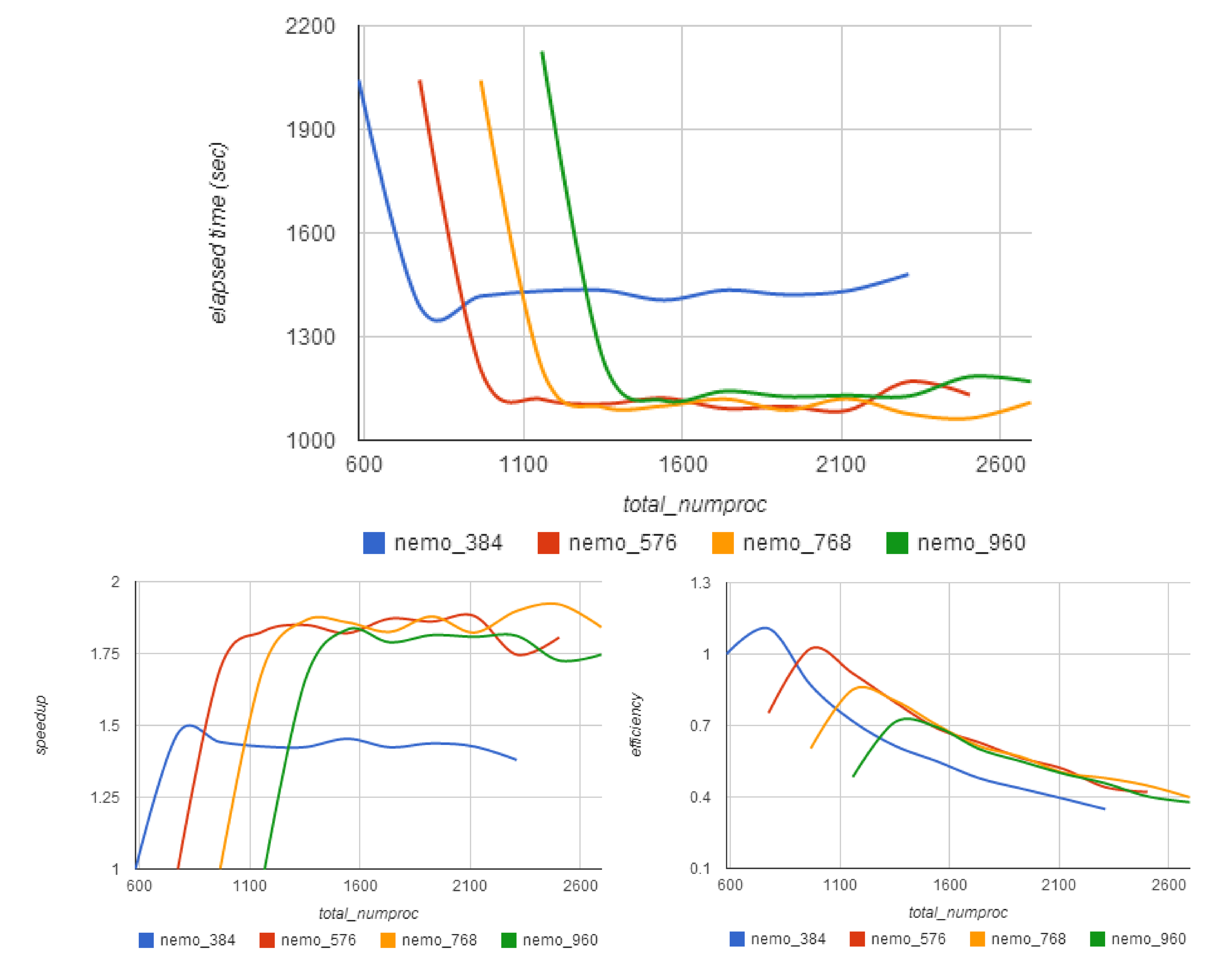

Figure 3: EC-Earth ran in coupled mode for 10 simulation days. High configuration (T511-ORCA025) plots (upper), (lower-left) and (lower-right) refer to elapsed time (average of 3 sets), speedup and efficiency respectively and the blue, red, orange and green curves correspond to results for tests with NEMO using 384, 576, 768 and 960 processors while speedup and efficiency values are relative to the reference test with 583 processors. In each case, the number of processors for OASIS3 is 7 and the number of processors for IFS is the remainder of subtracting the number of processors for NEMO and the number of processors for OASIS3 from the value at the horizontal axis.

Conclusions

- Obtained results demonstrate a good scalability of EC-Earth up to approximately a few hundred processors for standard configuration and a few thousand processors for high configuration while the scalability limit constitutes a typical problem among the climate models
- Analysis of the results revealed that the NEMO component seriously limits the ability to exploit supercomputing environments
- This work allowed us to be able to set the best number of processor allocations among different components of the model (depending on the desired configuration) when running production simulations as a standalone/uncoupled mode or coupled mode for studying the climate sciences.

Acknowledgements

This study was supported by the European Union's FP7-funded projects SPECS (308378) and IS-ENES2 (312979), the MINECO-Spain's RUCSS (CGL2010-20657), the MICINN-Spain under contract (TIN2011-28689-C02-01) and the Catalan Government. The authors thankfully acknowledge the use of computing resources via PRACE's HiResClim on Marenostrum III at Barcelona Supercomputing Center (BSC) in Spain.