Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Implementation of EC-Earth 10km Global Coupled ESiWACE Demonstrator and Recent Computational Improvements

Mario Acosta, Miguel Castrillo, Kim Serradell

Computational Earth Sciences

13/06/2019

PASC19 – MS28: High-Resolution Weather and Climate Simulations

Outline

- Introduction
- ESiWACE Center of Excellence
- EC-Earth and demonstrator
- Implementation and results
- Future work and improvements

Ce … ón

SCIENCE|BUSINESS® Bringing together industry, research and policy

11 Jun 2019 | Network update from Barcelona Supercomputing Center

These updates are republished press releases and communications from members of the Science|Business Network

EuroHPC selects Barcelona Supercomputing Center as entity to host one of the largest European supercomputers

By Communication from Barcelona Supercomputing Center

The European Union would contribute around 100 million euros to MareNostrum 5, the highest EU investment in a research infrastructure in Spain.

The Spanish candidacy to host one of the pre-exascale supercomputers has been supported by Portugal, Turkey, Croatia and Ireland as EuroHPC member states

The new supercomputer will have peak performance of 200 petaflops (two hundred trillion floating-point operations per second)

The decision guarantees Barcelona a leading position in the future map of supercomputing in Europe

The European Commission (EC) has announced that EuroHPC has selected Barcelona Supercomputing Center (BSC) as one of the institutions that will host a pre-exascale supercomputer in the high-capacity supercomputer network promoted by EC. The EC announcement describes the plan to acquire 8 supercomputers: 3 pre-exascale machines (with a peak performance of at least 150 Petaflops) and 5 petascale machines (capable of executing at least 4 Petaflops)

MareNostrum 5, the future BSC's supercomputer, is one of the pre-exascale machines. It will have a peak performance of 200 petaflops (200 x 1015 floating-point operations per second) and will come into operation on 31 December 2020.

The European Union would invest in it around €100M, the highest investment done by this institution in a research infrastructure in Spain. The rest of the investment will be borne by the Ministry of Science, Innovation and Universities, the Catalan Government and the states that supported the bid: Portugal, Turkey and Croatia, with contributions that are yet to be defined.

Supercom…
to Sp…
EU re…

…gramme,
…y transfer,
…gagement

BSC-CN…
a consor…
that incl…

GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD

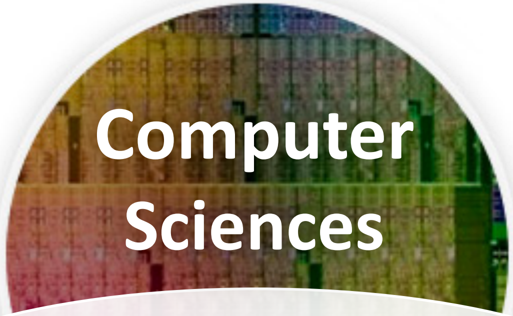
…Generalitat de Catalunya
…Departament d'Empresa i Coneixement

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

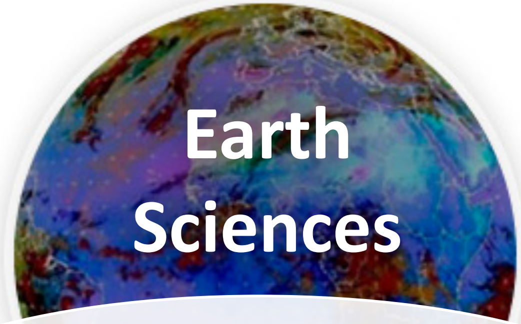
…esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Mission of BSC Scientific Departments



Computer Sciences

To influence the way machines are built, programmed and used: programming models, performance tools, Big Data, computer architecture, energy efficiency



Earth Sciences

To develop and implement global and regional state-of-the-art models for short-term air quality forecast and long-term climate applications

Life Sciences

To understand living organisms by means of theoretical and computational methods (molecular modeling, genomics, proteomics)

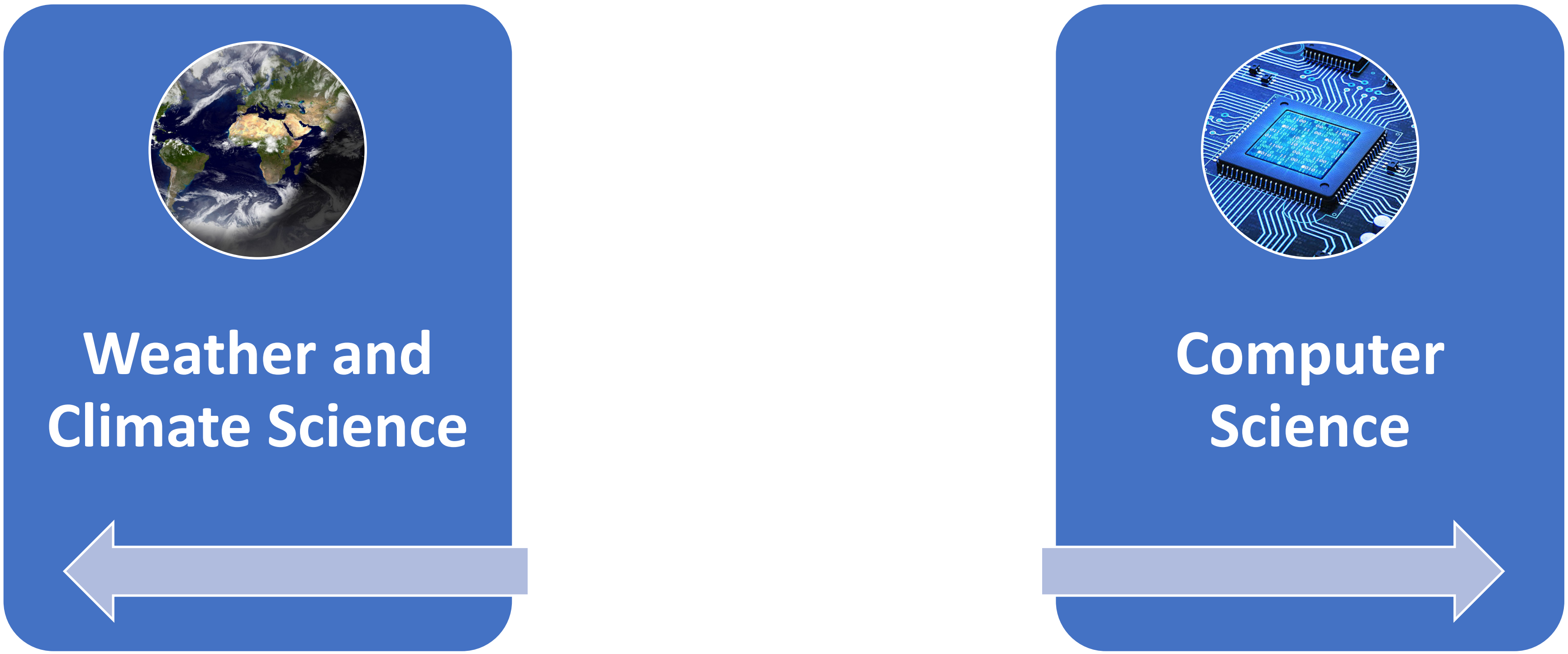
CASE

To develop scientific and engineering software to efficiently exploit super-computing capabilities (biomedical, geophysics, atmospheric, energy, social and economic simulations)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Computational Earth Sciences



- Knowledge about the mathematical and computational side of Earth System Applications
- Knowledge about the specific needs in HPC of the Earth System Applications
- Researching about HPC methods specifically used for Earth System Applications

Related talks

- In session MS43, **Computational Performance Evaluation for Hardware and Software Alternatives to Increase the HPC Efficiency of Earth System Models**
  - When: 14 June 2019
  - Where: HG D 1.1

- Mario Acosta: *New Methodologies for Computational Performance Evaluation of Climate and Weather* (10:30 - 11:00)

- Oriol Tintó: *Improving Ocean Model Computational Performance by using Mixed-Precision Approaches* (11:00 – 11:30)

ESiWACE
&
EC-Earth

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ESiWACE & ESiWACE2

- Centre of Excellence in Simulation of Weather and Climate in Europe

- Major benefit: create a framework where different institutions can work together to technically improve a community model as EC-Earth

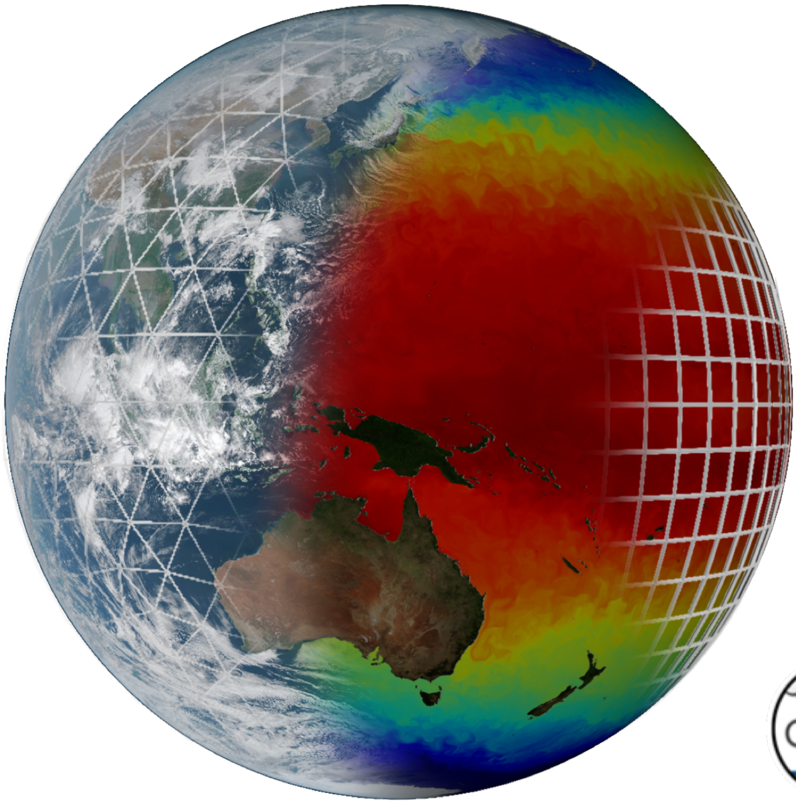
- First part of this talk is a summary from D2.11 deliverable: *“Implementation of EC-Earth 10km global coupled demonstrator and performance analysis”* - Zenodo

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

EC-Earth (I)

Atmosphere:
IFS



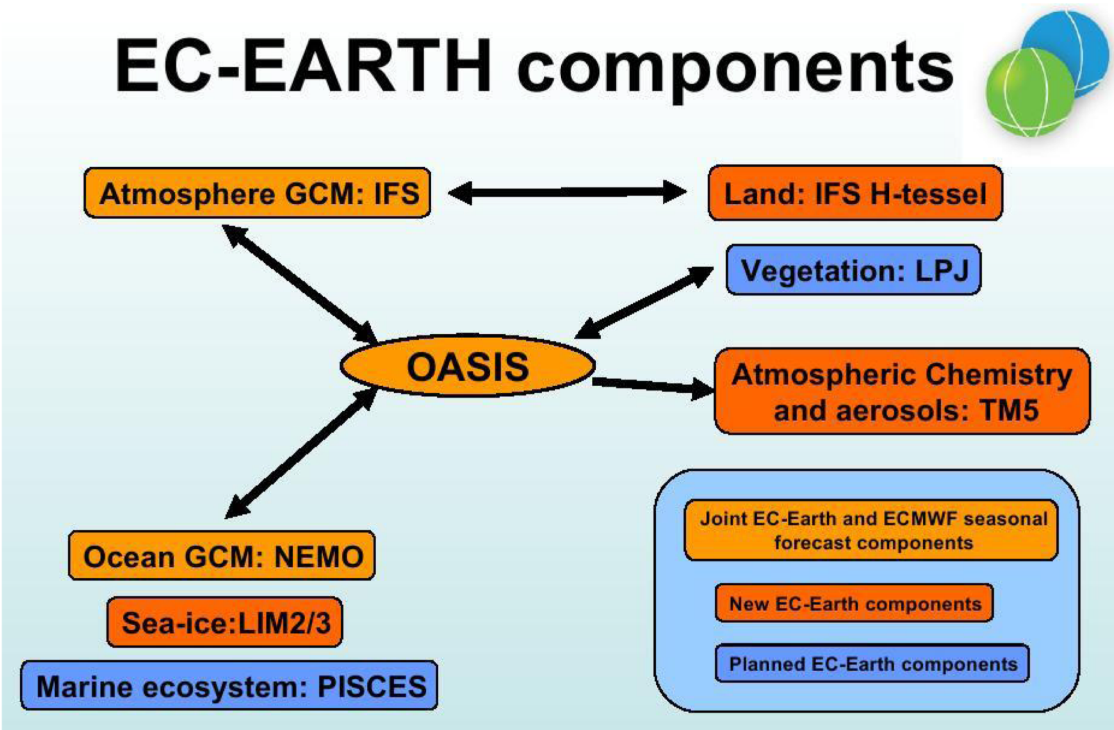
Ocean - ICE:
NEMO - LIM

Coupler:

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

EC-Earth (II)



29 partner institutes

8 core partners

KNMI, AEMET, DMI, Met Éireann, FMI, IPMA, CNR-DTA, SMHI

Workgroups

Technical
Tuning
Atmospheric Composition and Land
Ocean
Millennium scale studies
CMIP6

EC-Earth demonstrator

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Demonstrators

EC-Earth 10km coupled demonstrator

- **IFS** (atmosphere)
  - T1279L91: ~16km grid point distance, **2.1M** grid points
- **NEMO-LIM** (ocean – sea-ice)
  - ORCA12L75: ~9km grid point distance, **13.2M** grid points

Role of demonstrators

- COMPUTATIONAL CHALLENGE
  - **Demonstrate the computability** of atmosphere-only, ocean-only and coupled global models at the highest possible resolution that is configurable today. The task will **quantify how far away current capabilities** are from the long-term strategic target (global cloud resolving models).
- SCIENCE ENABLER
  - **PRIMAVERA** is a **Horizon 2020** project which aims to develop a **new generation of advanced and well-evaluated high-resolution global climate models.**
  - The **High Resolution Model Intercomparison Project (HighResMIP)** is a **CMIP6** endorsed MIP that applies, for the **first time**, a **multi-model approach** to the systematic investigation of the **impact of horizontal resolution**.

PRIMAVERA HIGHRESMIP

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

H2020: European HPC & science integration

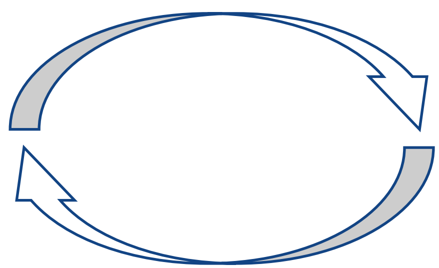
PRACE

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

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

HPC applications (CoEs)

Research infrastructure



PRIMAVERA

Climate science and HPC

Community

EC EARTH

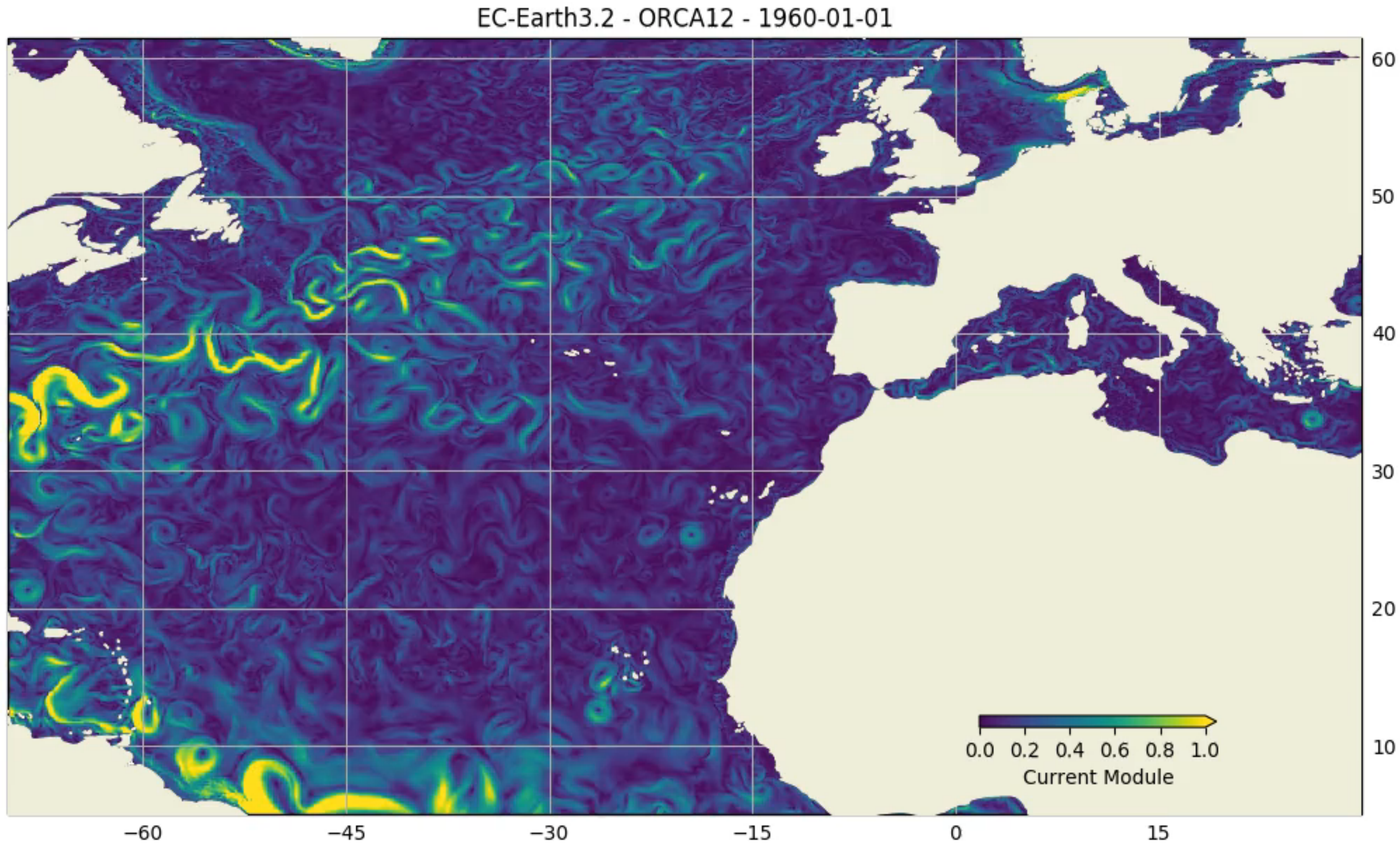
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Development of EC-Earth T1279-ORCA12

- Develop **initial data**
    - Including OASIS **interpolation weight** files
- Create **namelists** for IFS, NEMO-LIM (XIOS) and OASIS
- Adapt **source code** and existing **runscripts**
- Introduce required changes in the experiment **workflow**
- Work done in collaboration between **Research Engineers** and **Climate Scientists**
- **Scalability** tests / load **balance** studies / **profiling**

EC-Earth T1279-ORCA12: production runs



Courtesy of: Thomas Arsouze

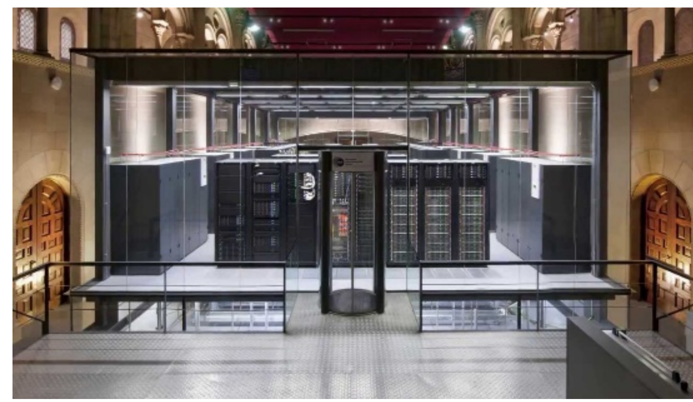
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

EC-Earth T1279-ORCA12 in MareNostrum3

First global, coupled ~10km simulations (T1279 - ORCA12):

- EC-Earth 3.2 (IFS36r4 + NEMO 3.6 + OASIS3-MCT)
- 2,035 MPI tasks - 60 SDPD
  - 1,170 NEMO
  - 848 IFS
  - 16 XIOS
  - 1 runoff mapper
- MareNostrum3 @ BSC



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

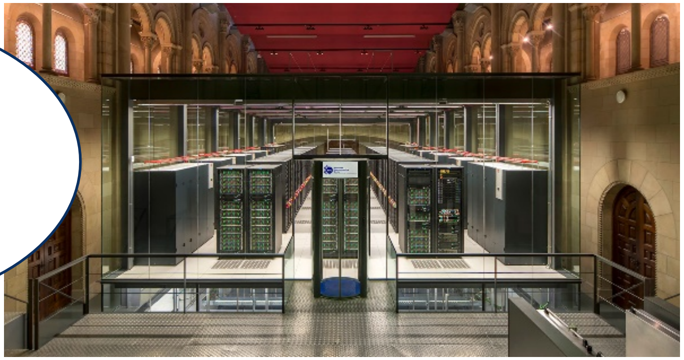
esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

EC-Earth T1279-ORCA12 in MareNostrum4

Operational global, coupled ~10 km simulations (T1279 - ORCA12):

- **EC-Earth 3.2** (IFS36r4 + NEMO 3.6 + OASIS3-MCT)
- **5,040 MPI tasks** - 0.44 SYPD, 160 SDPD
  - 3,209 NEMO
  - 1,584 IFS
  - 69 XIOS
  - 1 runoff mapper
- **MareNostrum4** @ BSC

100 year exp
~40M ch !!!



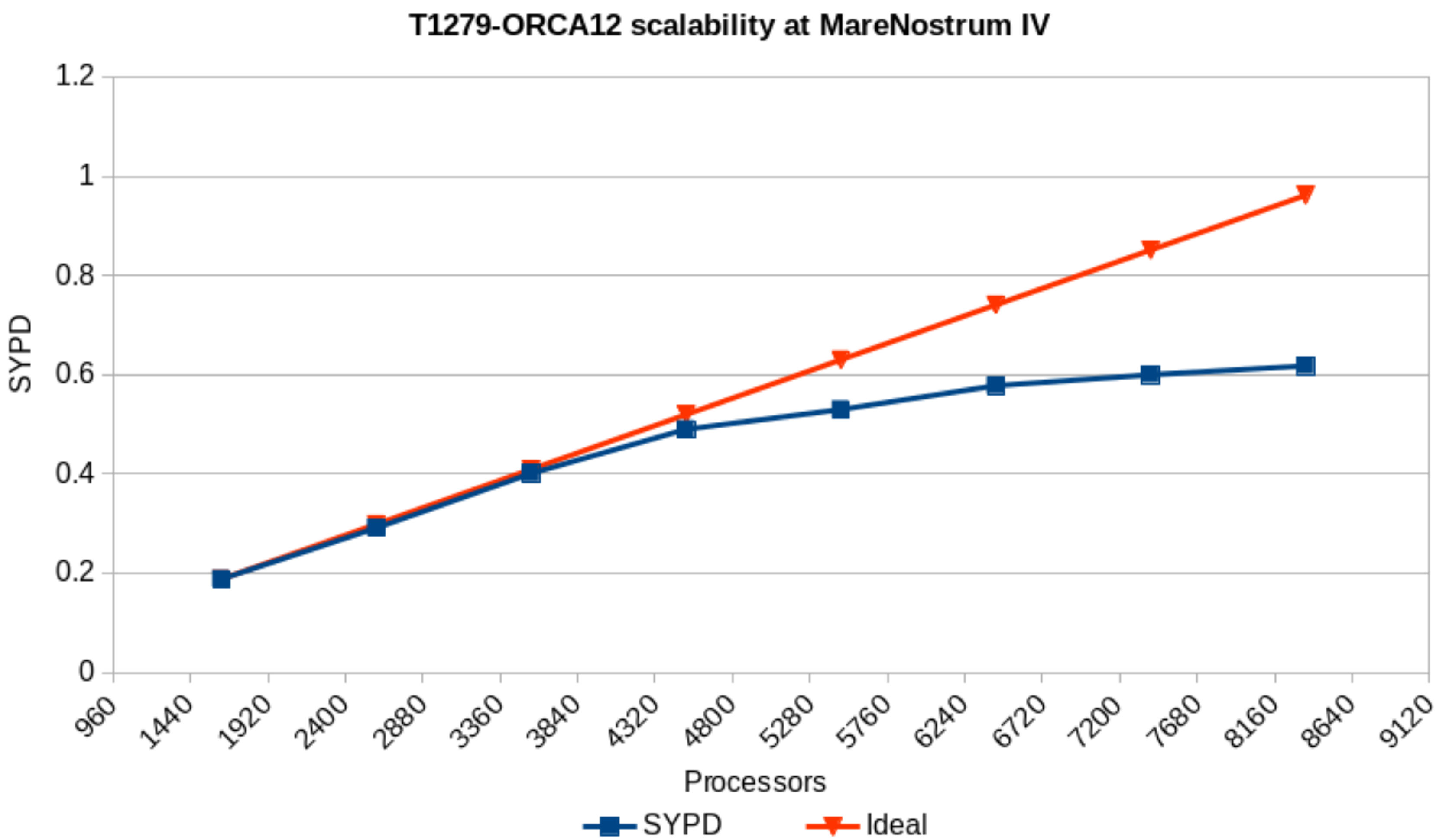
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

MareNostrum4: Tackled issues

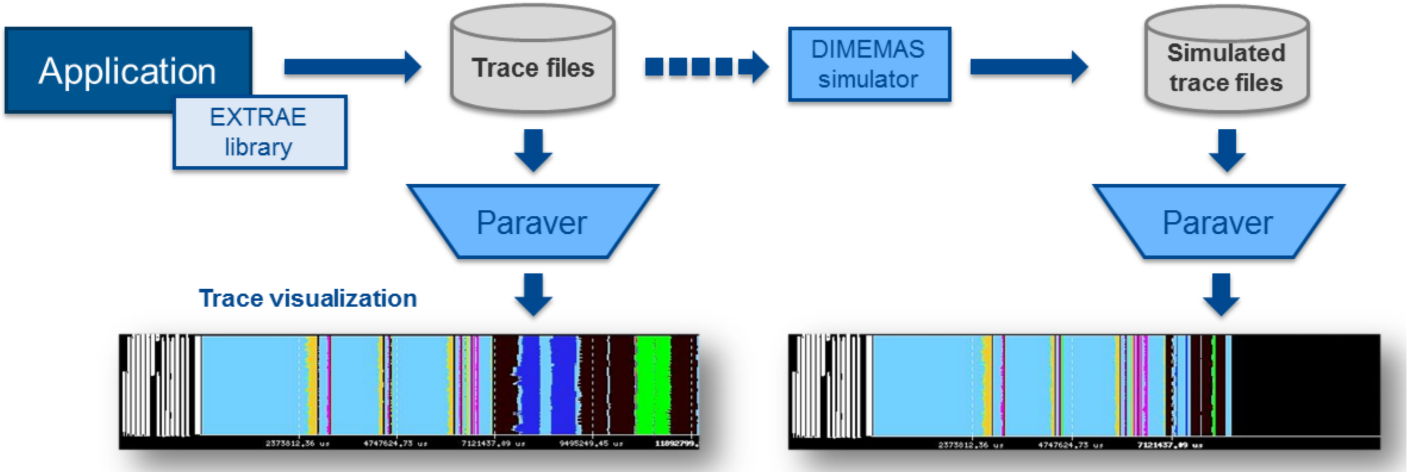
- **I/O management**
  - Use of MareNostrum4 **data-transfer nodes**
- Optimal libraries and **dependencies**
  - Never forget Usability (software stack and workflows) and user support
- Come up with a **stable environment**
  - OmniPath: numerous **tests** and **collaboration** with **operations** to find optimal **configuration** (tmi, PSM2)
  - XIOS update: decrease number of **communications**
  - Controlling process pinning: better **memory management**

Scalability

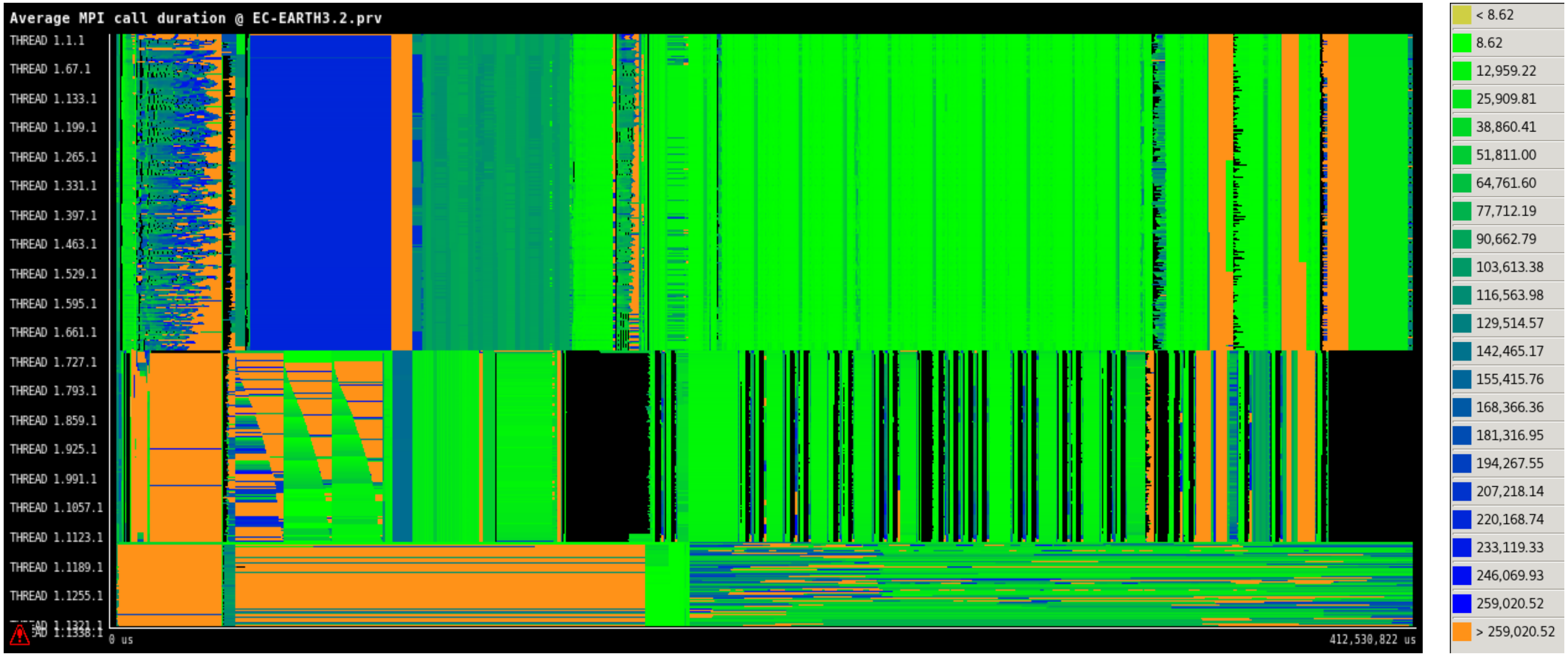


Performance analysis

- Since 1991
- Based on **traces**
- Open Source: https://tools.bsc.es
- **Extrae:** Package that generates Paraver trace-files for a post-mortem analysis
- **Paraver:** Trace visualization and analysis browser
- **Dimemas:** Message passing simulator



Tracing runs

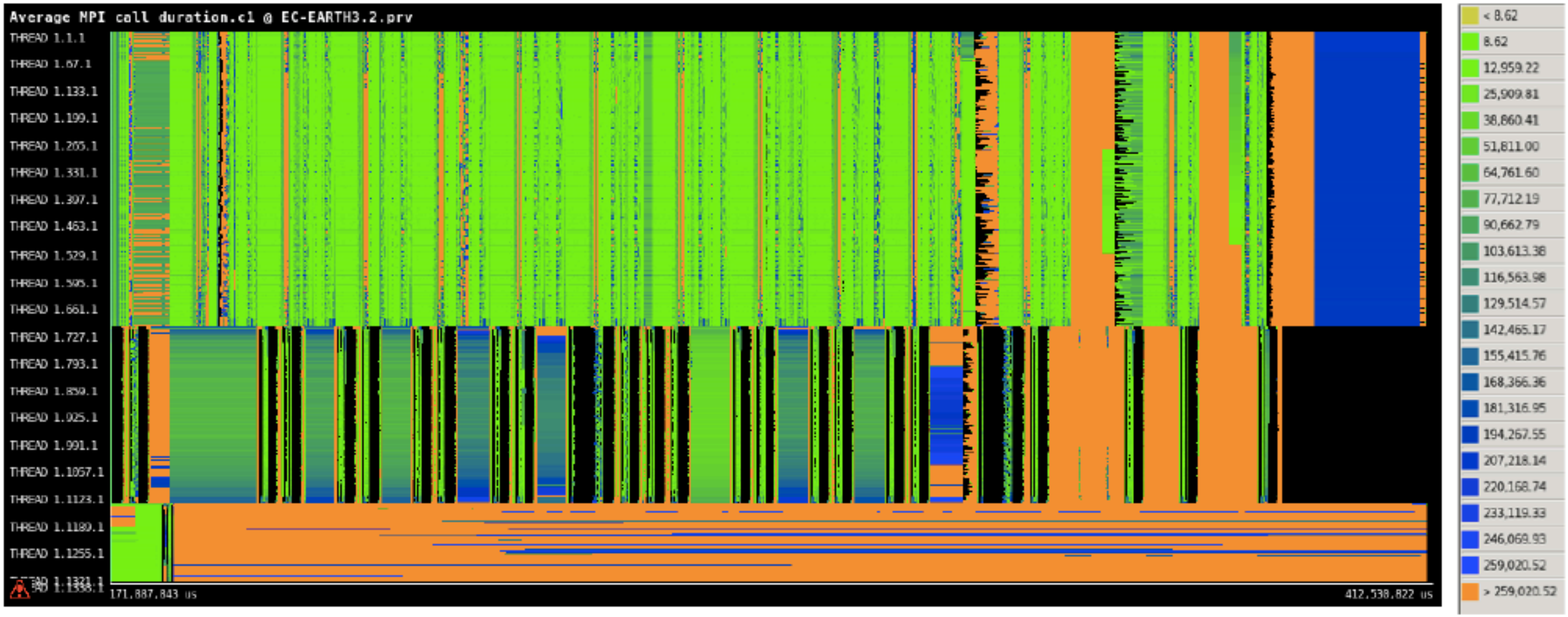


Trace of 24h T1279-ORCA12 execution representing the average MPI call duration

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Tracing runs



Trace of 24h T1279-ORCA12 execution representing the average MPI call duration with a zoom to catch more details in each time step.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Issues found

- Reduce **I/O overhead** → Interface IFS with **XIOS**
  - I/O in IFS36r4 is **sequential**
  - NEMO-LIM (ORCA12): Up to **3 SYPD** in MareNostrum4 (LIM -> 2OCE stp)
- **Sea-ice** is slow and computationally demanding
- **Update IFS** to newest cycle, using octahedral grid
- **Update NEMO** to NEMO4 (and beyond)
- Most of these improvements can be real in **EC-Earth4**

What next?

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

EC-Earth4

- In parallel with EC-Earth3.X (CMIP6), we started the development of the next version
- Still work to do in ESiWACE2 to reach *"production mode"*
- EC-Earth TWG is currently working in a first release (end 2019)
  - Objective: more regular releases
- But in this mixed approach "scientific-technical", many other things need to be discussed

EC-Earth4 TODO list

- Main components will be upgraded to last versions (OpenIFS, NEMO4, OASIS and XIOS)
- XIOS as I/O server for Atmosphere and Ocean
- Decouple sea-ice model from NEMO and use OASIS instead
- Providing a ”mixed-precision” version
- Improving workflow management, code modularity and CI/CD
- Continue deploying the reproducibility protocol among partners/machines (paper under discussion in GMD)
- Adapt post-processing tools to work with high-resolution configuration
- Increase adoption and feedback from EC-Earth developers to model developers

Exploring new I/O methods

- With this new configuration, I/O is a huge problem and should always be considered when doing projections
  - Sequential I/O in the Atmosphere
  - Issues to produce large coupler files
  - Different file formats, post-processing, big investment to have a curated and structured storage.
- Starting from the same point, we decided to tackle the problem from two perspectives
  - Based on files, more operational
  - Removing files, totally experimental

OpenIFS - XIOS

- One year ago, we started the project to integrate XIOS in IFS/OpenIFS
  - First version to validate the approach (only storing grid-point variables) has been completed
  - In April-May, thanks to ESiWACE2, a visit from BSC research engineer to NLeSC led to complete the work

- OpenIFS can now output all variables using XIOS
  - ECMWF will integrate this developments in the next OpenIFS release (expected for the next OpenIFS workshop)

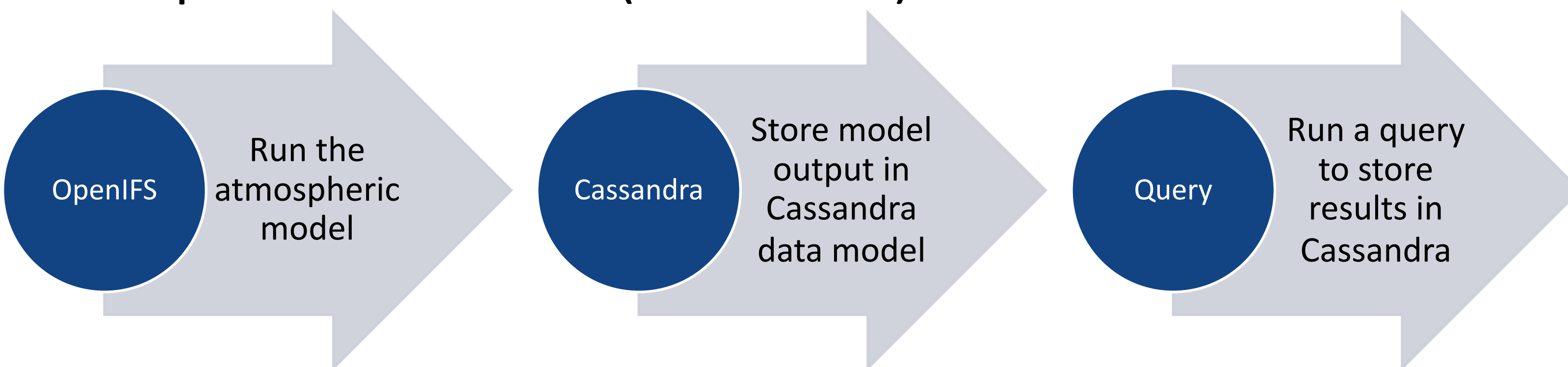
OpenIFS - XIOS

- The current operations performed in the EC-Earth post-processing task are avoidable, since XIOS has these features:
  - Output files are in netCDF format
  - Written data is CMIP-compliant (CMORized)
  - It is able to post-process data online to generate diagnostics

- The use of XIOS in EC-Earth has a twofold effect:
  - Improve the computational performance and efficiency of the model, and thus, reduce the execution time (previously mentioned)
  - Reduce the critical path of its workflow by avoiding the post-processing task

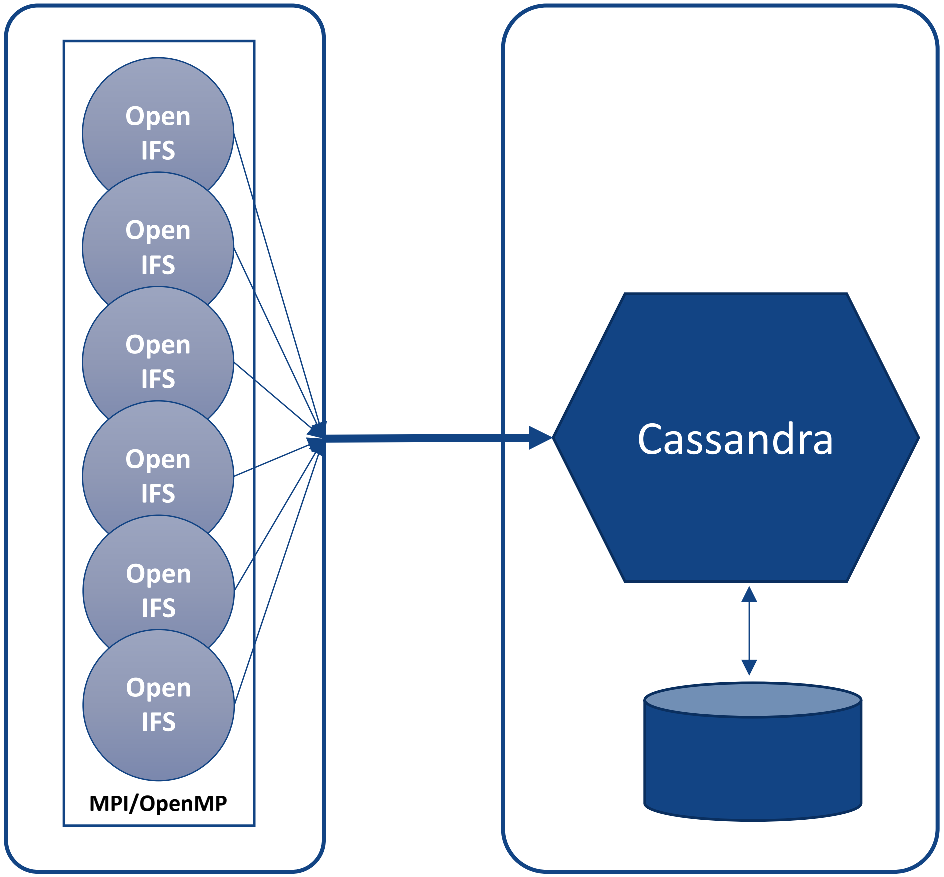
OpenIFS - Cassandra

- In collaboration with Data Driven Group in Computer Sciences, we are running an **experimental** project to couple OpenIFS with Apache Cassandra (NoSQL DB).



- In execution time, fields from OpenIFS are send to Cassandra (distributed) and stored.
- Different configurations: T255, T511 and T1279 and pushed to outputting each timestep.
- Optimized queries (using PyCOMPSs) can be developed to retrieve data and generate files (if needed).

OpenIFS - Cassandra



REMOTE

Conclusions

Now the **EC-Earth T1279-ORCA12** configuration is:

- Developed and shared among **EC-Earth consortium** partners
- **Deployed and tested** in **MareNostrum3** and **MareNostrum4** HPC systems
- Used in **production** for **H2020** projects such as PRIMAVERA using **PRACE** resources
- Used to investigate the **scalability** of **ultra-high resolution** coupled models, enabling to **push computational challenges** of the current HPC generation

And led to build a roadmap to **EC-Earth4**

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Thank you

The projects ESiWACE and ESiWACE2 have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 675191 and 823988.

kim.serradell@bsc.es