Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

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

Making global coupled 15km climate simulations possible

Miguel Castrillo, Thomas Arsouze, Kim Serradell, Mario Acosta
Barcelona Supercomputing Center

Glob15km – Global 15km coupled climate simulations PRACE project (ref. 2016163939)

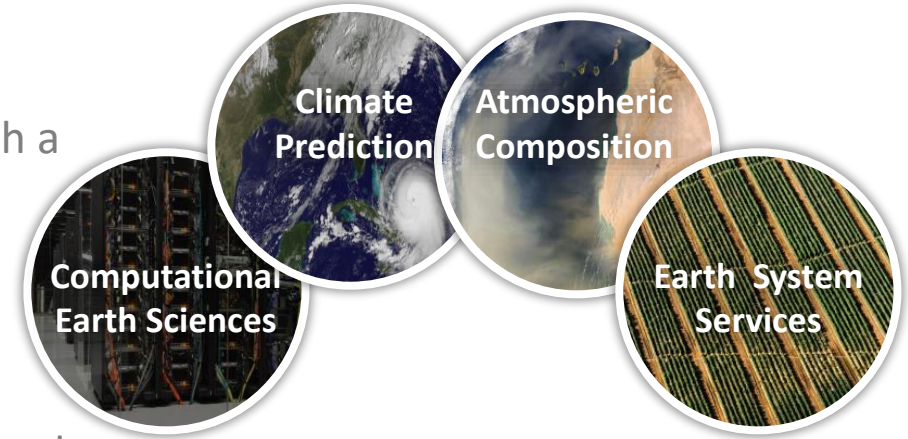
PRACE

Barcelona Supercomputing Center

- Created in 2005; more than ~~400~~ **600** employees
- Research, develop and manage information **technology**
- Facilitate **scientific** progress and its **application** in society

Earth Sciences department

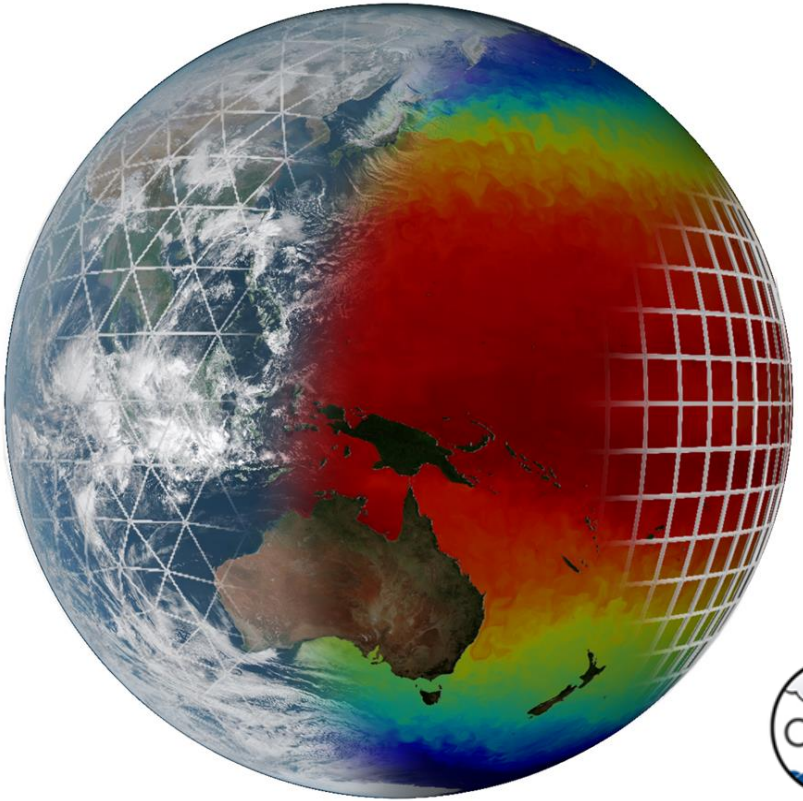
- Environmental modelling and forecasting, with a particular focus on **weather, climate and air quality**
- **>100** people
- Funding from H2020, COPERNICUS, private contracts, ESA, Spanish and regional governments



EC EARTH

Atmosphere:
IFS

ECMWF

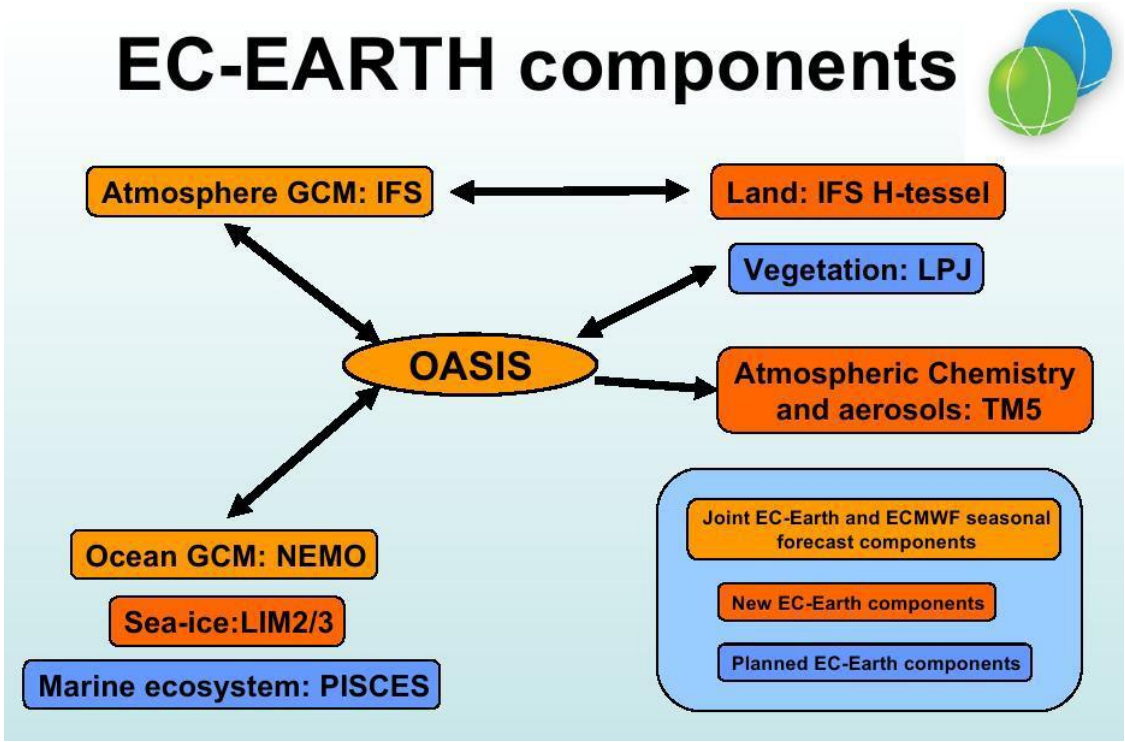


Ocean - ICE:
NEMO - LIM

NEMO

Coupler:

Oasis



29 partner institutes

8 core partners

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

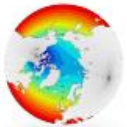
Workgroups

Technical
Tuning
Atmospheric Composition and Land
Ocean
Millennium scale studies
CMIP6

- Especially **critical** in Earth science models
- Simulations use a huge amount of computational **resources**
  - **I/O** !
- Future simulations will need much more resources

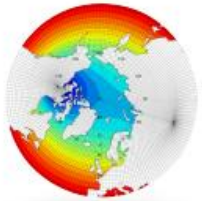
ORCA 2
550 MB of memory
8 CPU hours
10 Gigabytes of output (daily)

0.84M points



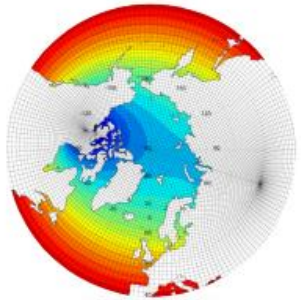
ORCA ¼
47 Gigabytes of memory
3500 CPU hours
120 Gigabytes of output (daily)

67.72M points



ORCA 1/12
414 Gigabytes of memory
90 000 CPU hours
1 Terabyte of output (daily)

991M points



ORCA 1/36
> 1 Terabytes of memory
~4 000 000 CPU hours
> 5 Terabytes of output (daily)

~2000M points

EC-Earth 10km coupled demonstrator for ESiWACE H2020 project

- **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
- Total 3D space: **1,024kM vertices**

- 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**
- **Scalability** tests / load **balance** studies / **profiling**

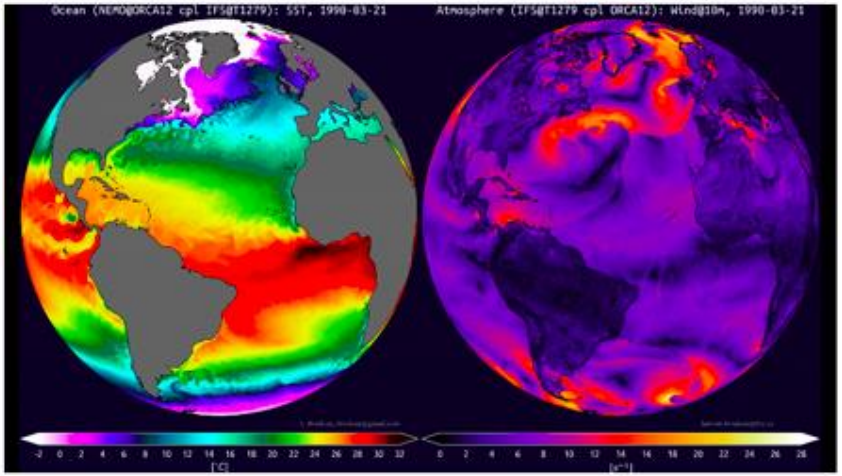
First global, coupled ~10km simulations (T1279 - ORCA12):

- **EC-Earth 3.2** (IFS36r4 + NEMO 3.6 + OASIS3-MCT)

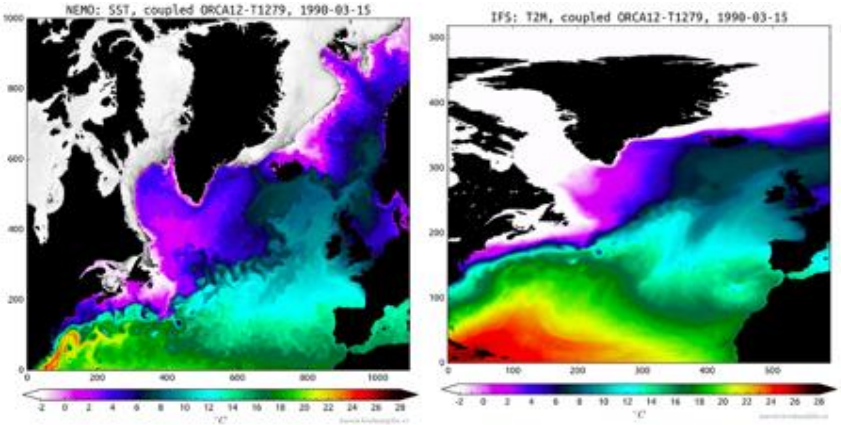
- 2,035 MPI tasks - 60 SDPD
  - 1,170 NEMO
  - 848 IFS
  - 16 XIOS
  - 1 runoff mapper

- **MareNostrum3** @ BSC

First global, coupled ~16km simulations (T1279 – ORCA12):



Left, Global Sea Surface Temperature of the ocean component NEMO. Right, Global Speed Wind at 10m of atmosphere component IFS.



Left, regional crop Sea Surface Temperature of the ocean component NEMO. Right, regional crop Temperature at 2m of the atmosphere component IFS.

MareNostrumIV in operation since **July 2017**

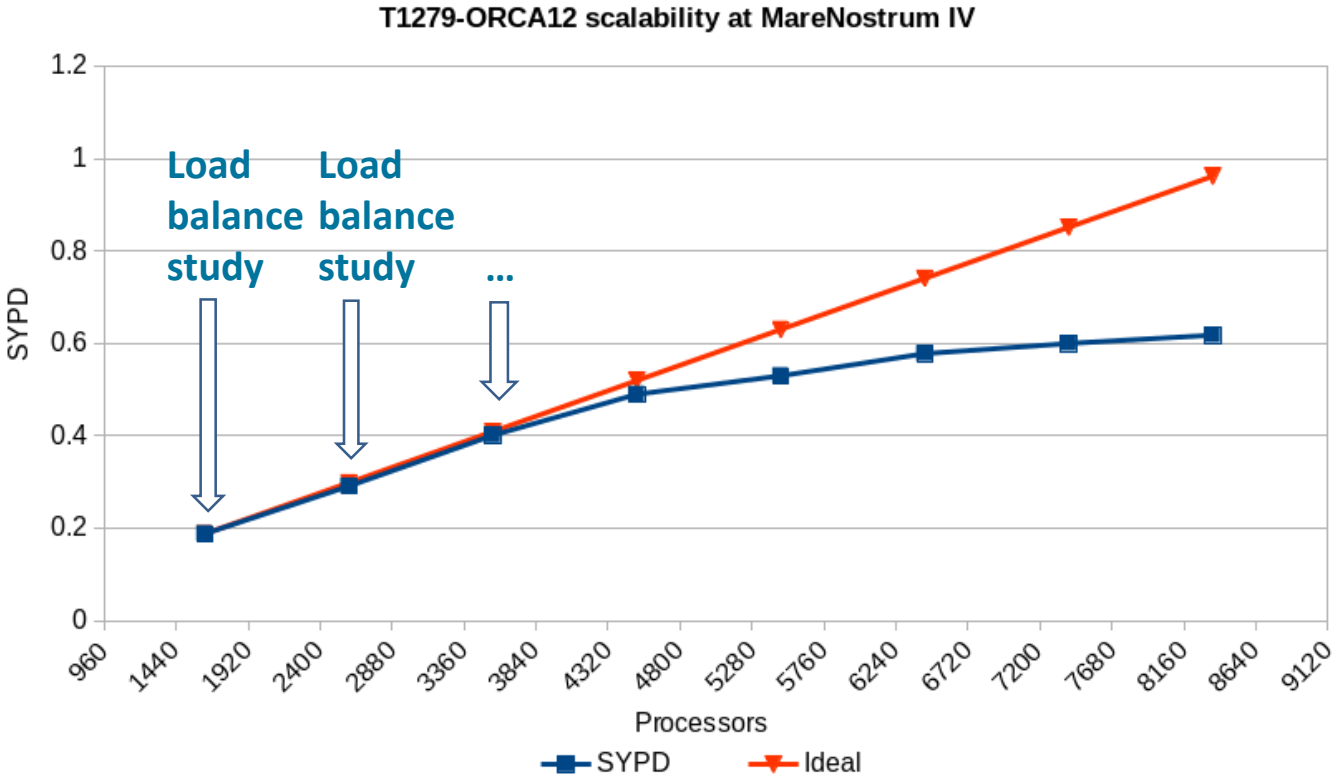
	MareNostrum III	MareNostrum IV
Processor	Intel Xeon E5-2670 2.6 GHz	Intel Xeon Platinum 8160 2.1 GHz
#Cores per socket / #Sockets	8 / 2	**24** / 2
Memory	32Gb DDR3-1600	96Gb DDR4-2667
Interconnection	Infiniband FDR10 10Gb	Intel Omni-Path **100Gb**

MareNostrum III – 1.1 petaFLOPS

MareNostrum IV – 11.15 petaFLOPS

- **I/O management**
  - Use of MareNostrum4 **data-transfer nodes**
- Optimal libraries and **dependencies**
- 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**

EC-Earth T1279-ORCA12 in MareNostrum4

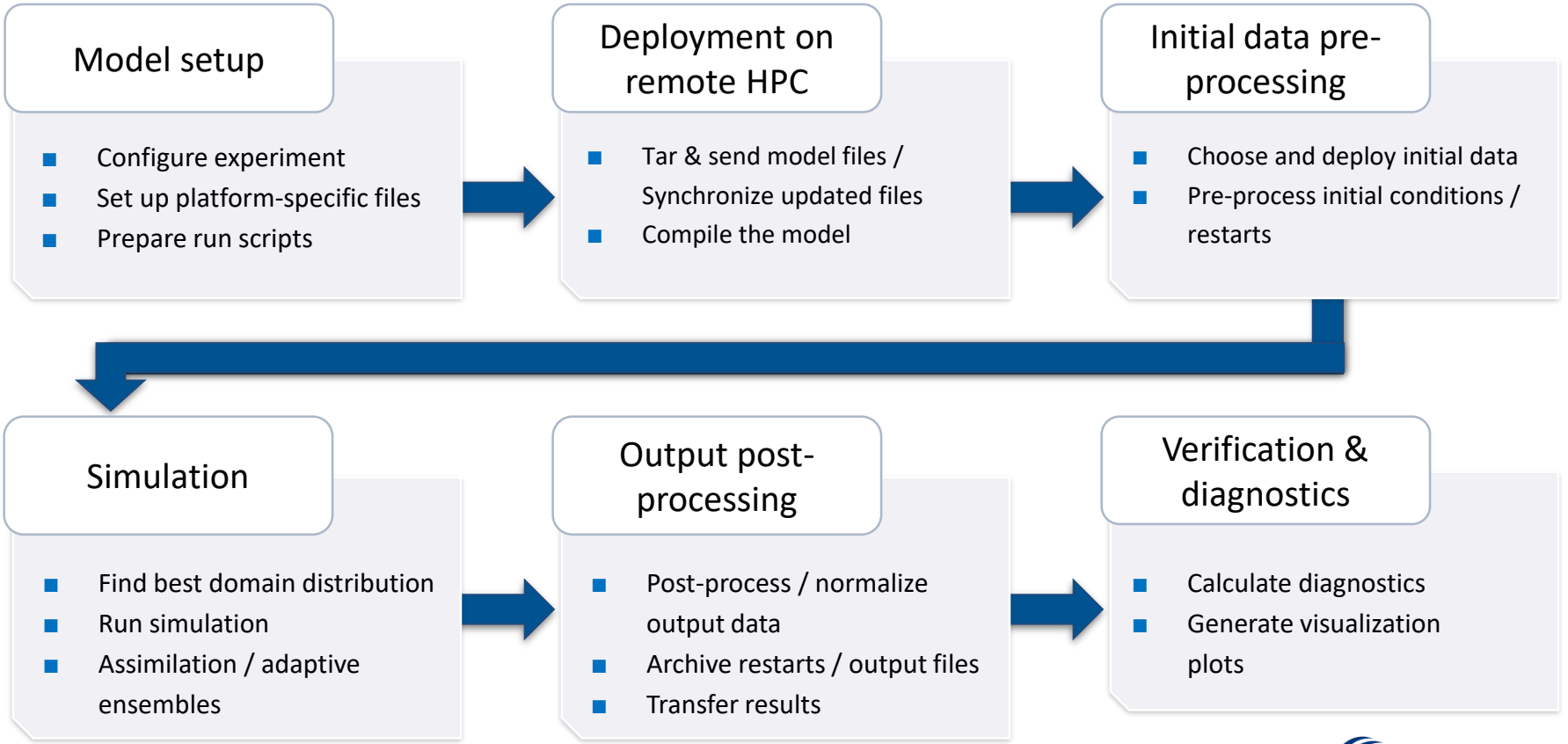


Operational global, coupled ~10 km simulations (T1279 - ORCA12):

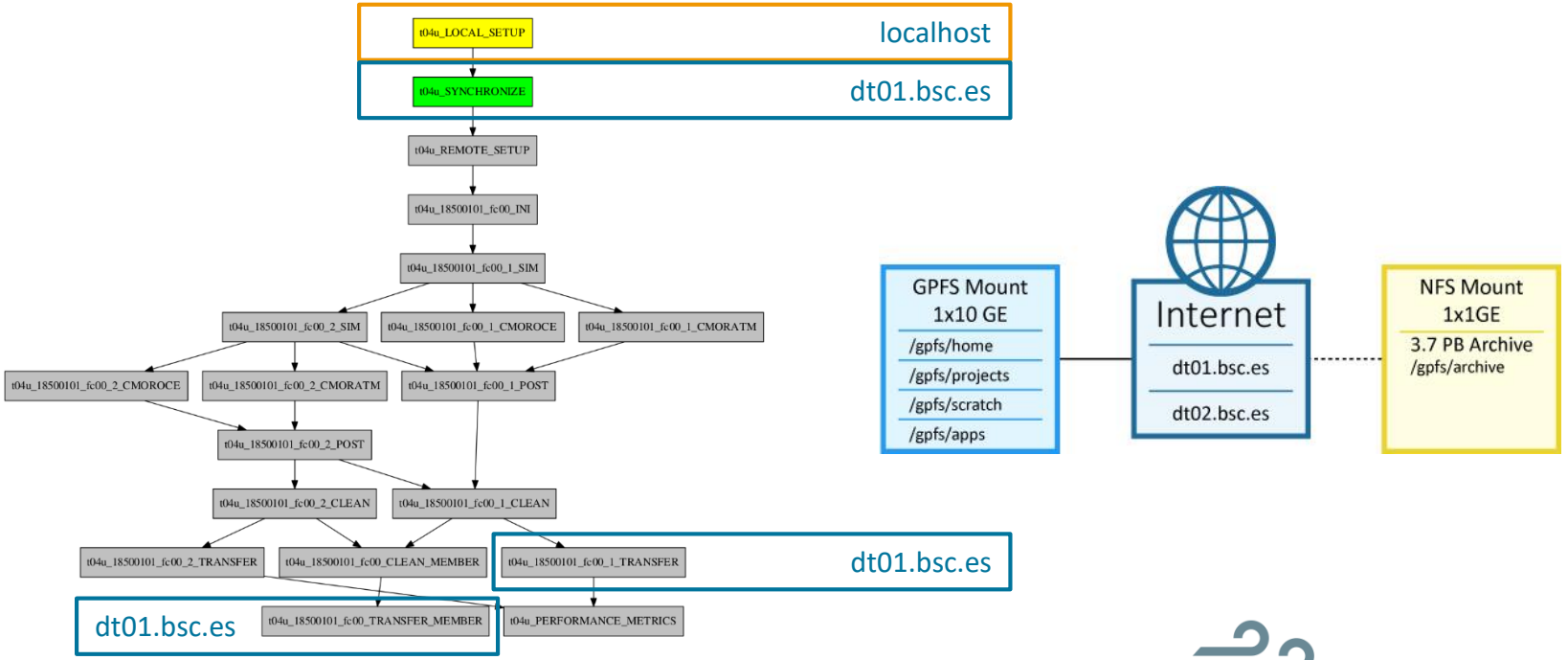
- **EC-Earth 3.2** (IFS36r4 + NEMO 3.6 + OASIS3-MCT)
- **5,040 MPI tasks** - 0.44 SYPD, 160 SDPD
  - 3,209 NEMO
  - 1,584 IFS
  - 69 XIOS
  - 1 runoff mapper
- **MareNostrum4** @ BSC

100 year exp
~40M ch !!!

An operational EC-Earth workflow in BSC-ES



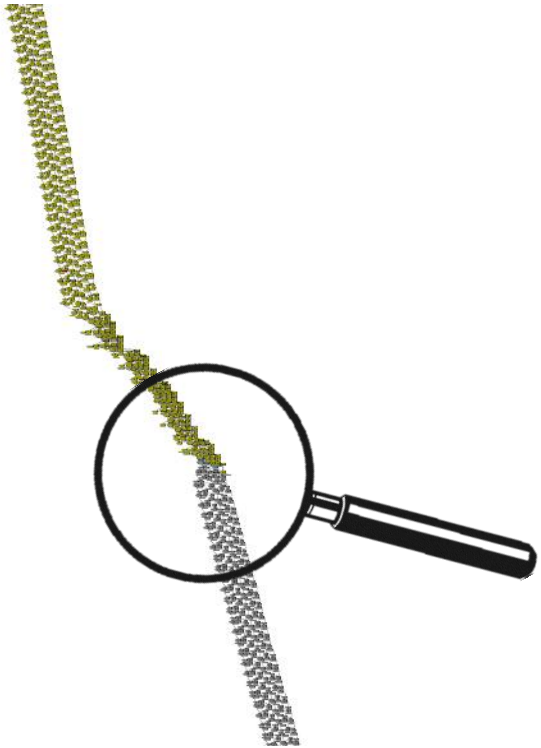
Adapting **workflow** for production: **data transfer nodes**

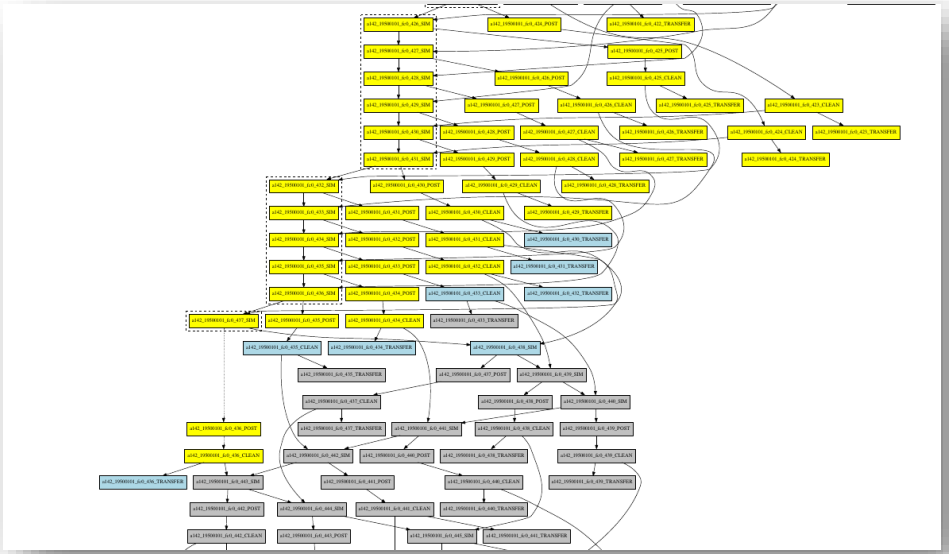


AUTOSUBMIT

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Production runs: **Managing the workflow**





AUTOSUBMIT

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

- **PRIMAVERA** is a **Horizon 2020** project which aims to develop a **new generation of advanced and well-evaluated high-resolution global climate models**, capable of simulating and predicting regional climate with **unprecedented fidelity**, for the **benefit** of governments, business and society in general.

- 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**.

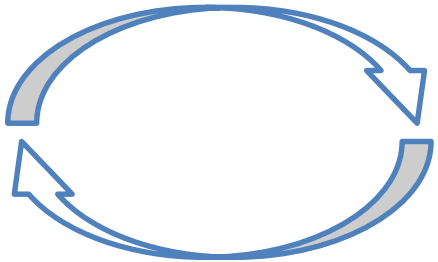
PRACE

Research infrastructure

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

PRIMAVERA

HPC applications (CoEs)

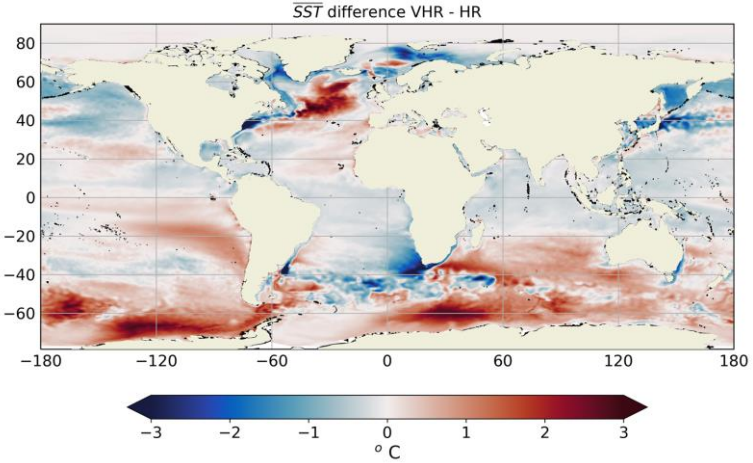


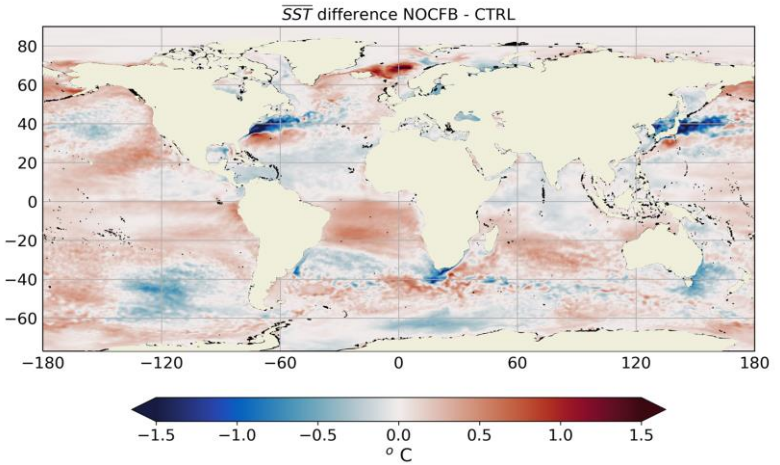
Climate science and HPC

Community

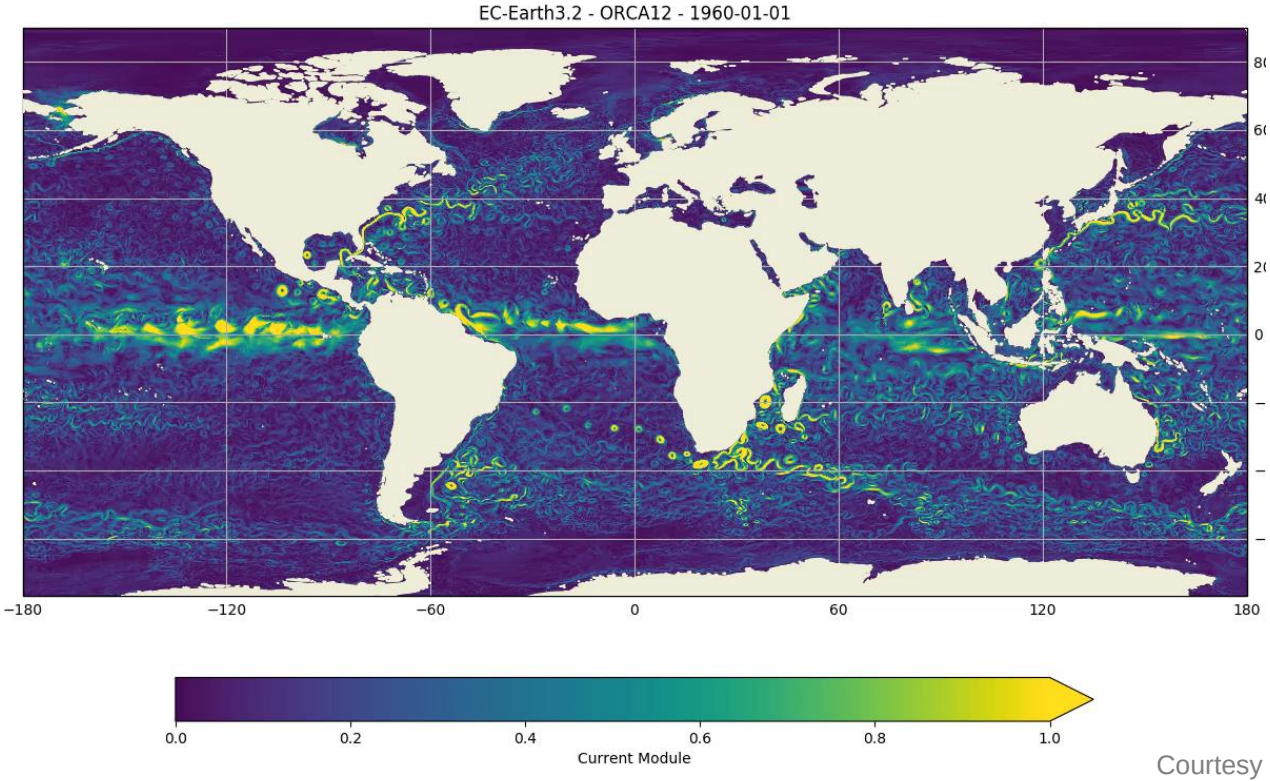
Scientific objectives

- Develop and prepare a **new generation** of **global high-resolution** climate models
- Evaluating global high-resolution climate models at a **process level**
- Focus on **air-sea interactions** at oceanic mesoscale:
  - Thermal feedback
  - Evaluate the role of the mechanical interactions between oceanic surface currents and atmospheric winds ("**current-feedback**")



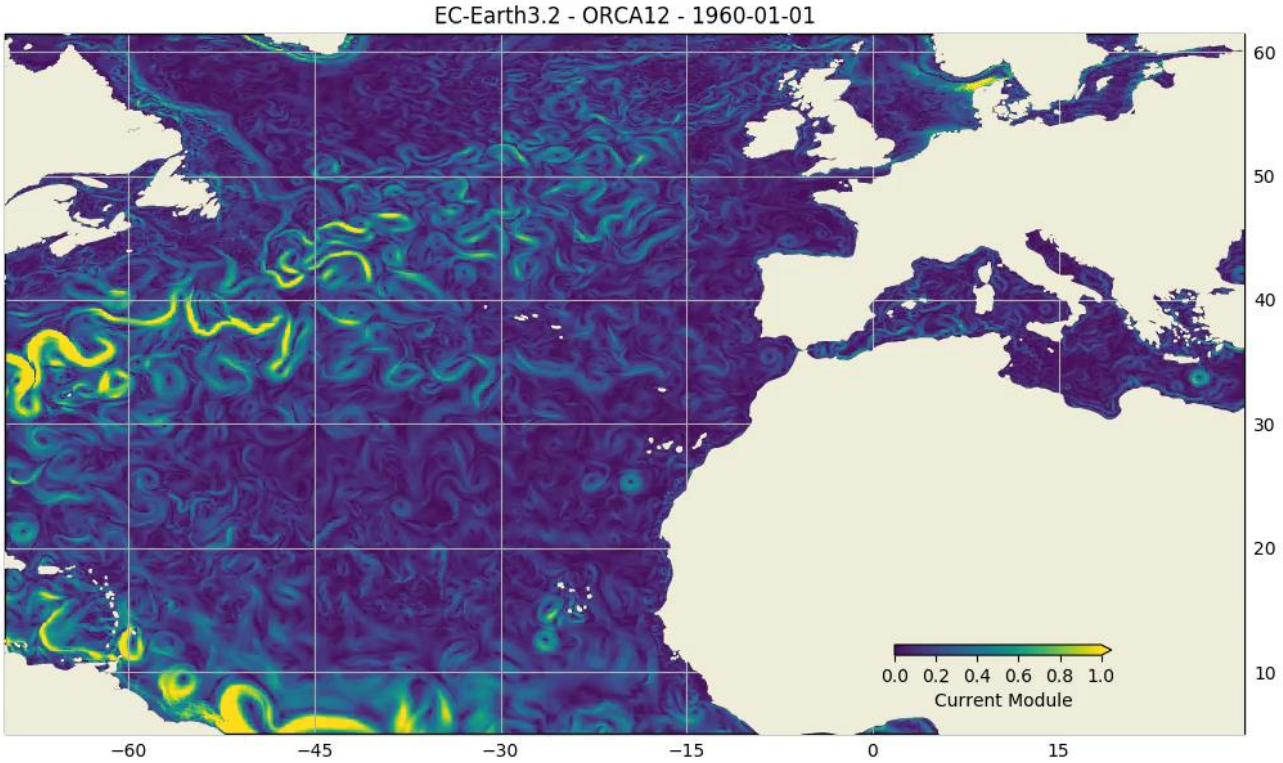


EC-Earth T1279-ORCA12: production runs



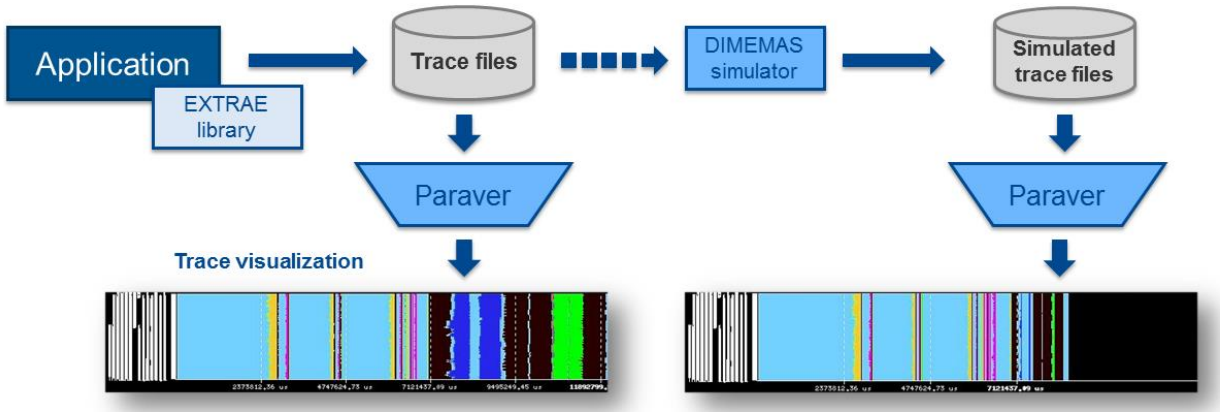
Courtesy of: Thomas Arsouze

EC-Earth T1279-ORCA12: production runs

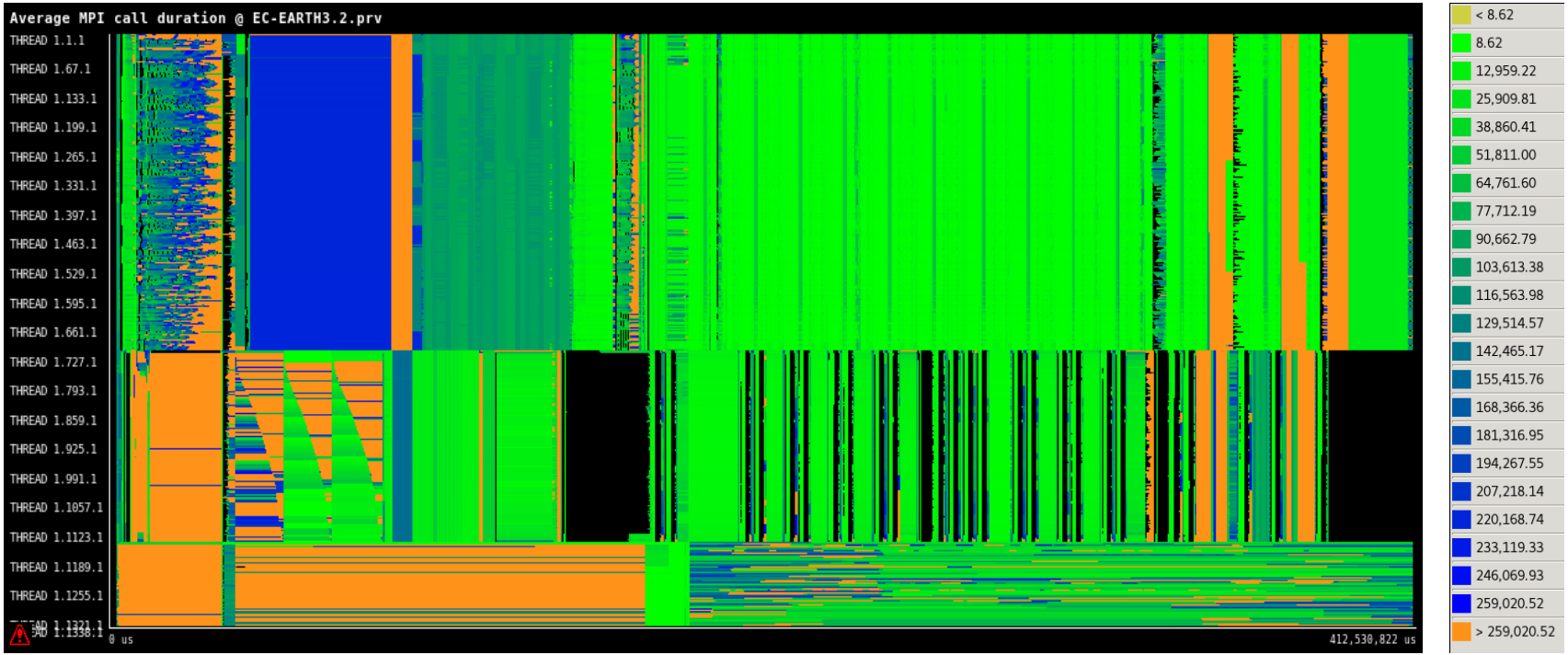


Courtesy of: Thomas Arsouze

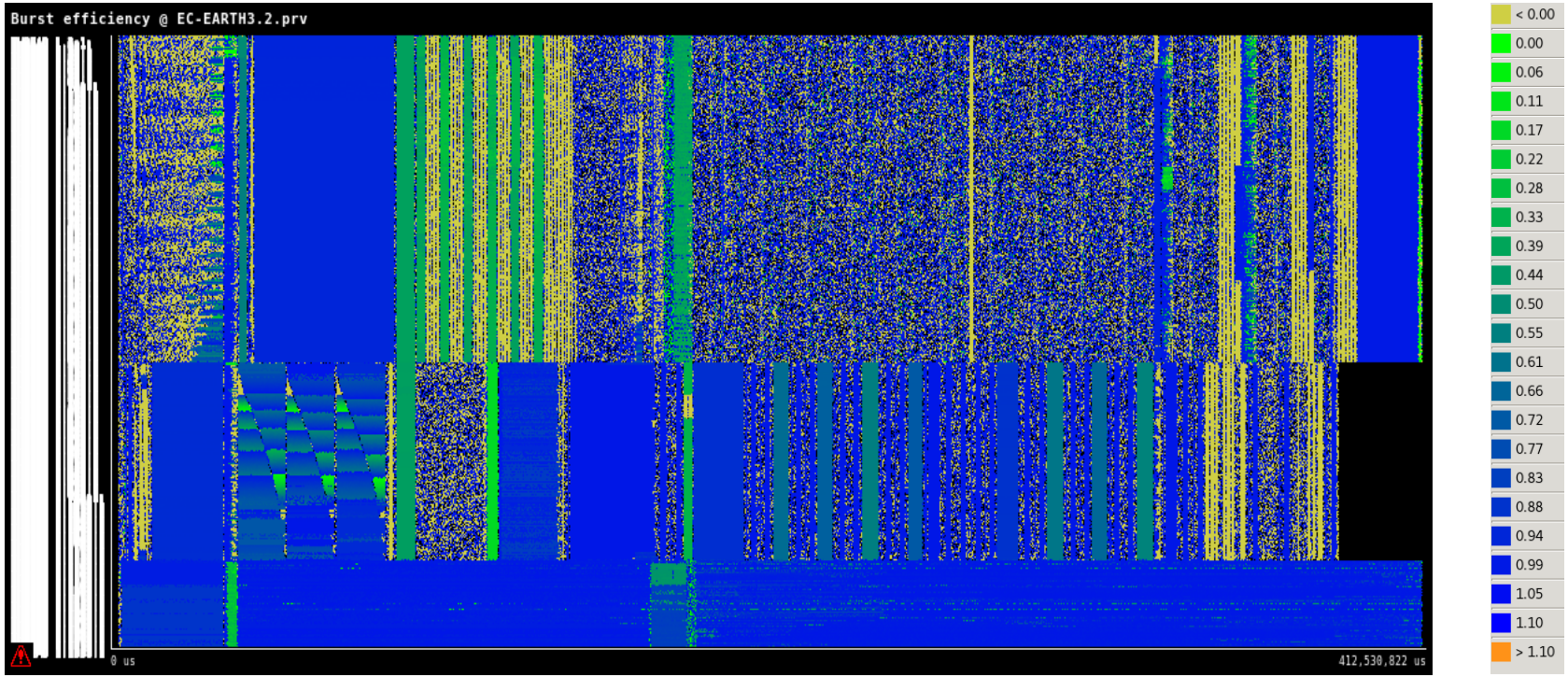
- Since 1991
- Based on **traces**
- Open Source: https://tools.bsc.es
- **Extrae**: Package that generates Paraver trace-files for a post-mortem analysis
- **Paraver**: Trace visualization and analysis browser
- **Dimemas**: Message passing simulator



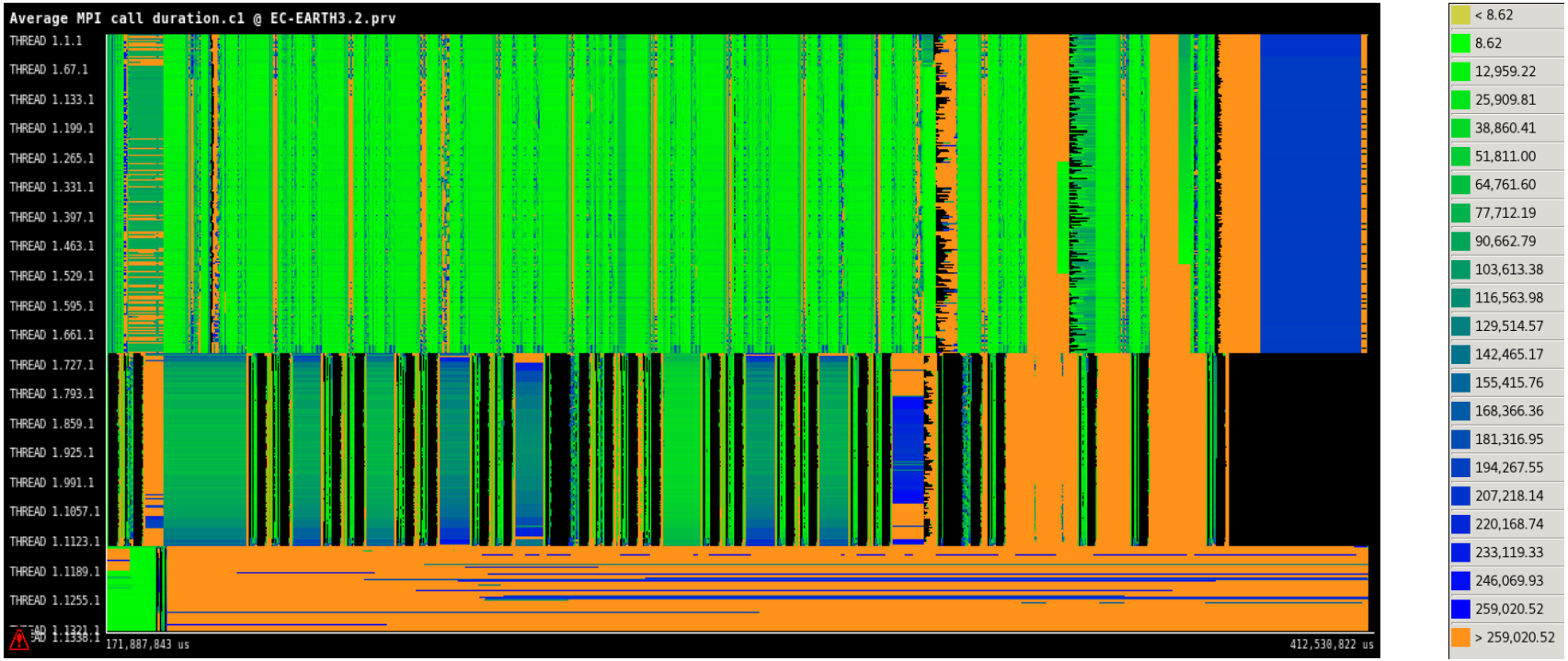
EC-Earth T1279-ORCA12: Performance analysis



EC-Earth T1279-ORCA12: Performance analysis



EC-Earth T1279-ORCA12: Performance analysis



- Reduce **I/O overhead** → Interface IFS with **XIOS**
  - NEMO-LIM (ORCA12): Up to **3 SYPD** in MareNostrum4 (LIM -> 2OCE stp)
- **Detach sea-ice** from NEMO. Couple through **OASIS**.
- **Update IFS** to newest cycle, using octahedral grid
- **Update NEMO** to NEMO4 (and beyond)
- Most of these improvements can be real in **EC-Earth4**

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

THANK YOU

miguel.castrillo@bsc.es

The Implementation Phase of PRACE receives funding from the EU's Horizon 2020 Research and Innovation Programme (2014-2020) under grant agreement 730913

The ESiWACE project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 675191