Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

High Performance Computing for Earth System Models: Optimization & Profiling

Pablo Ortega, Miguel Castrillo and Mario Acosta

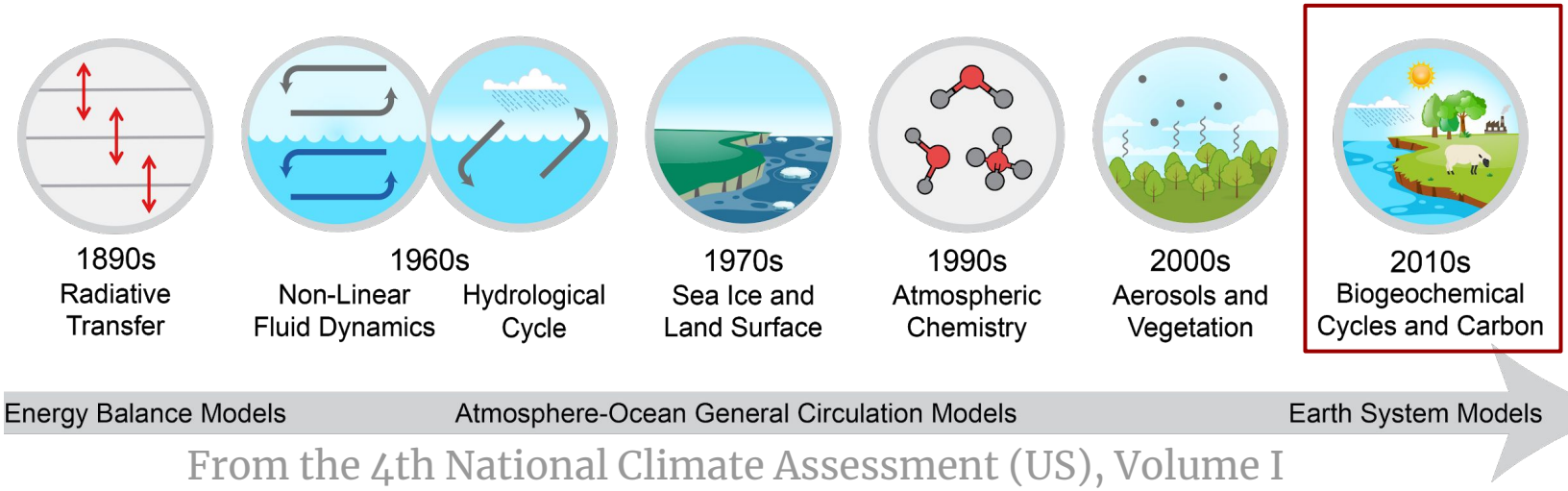
Earth Sciences Department

High Performance Computing in Earth Sciences

- Earth System Models (ESMs) are sophisticated tools with continuously increasing complexity:
  - More components of Earth System are included
  - Finer Spatial and Temporal resolutions
- This increase in complexity has only been possible thanks to the important parallel advances in HPC

A climate modeling Timeline

Inclusion of new components



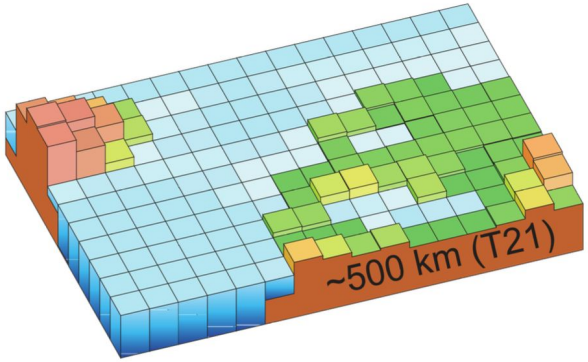
From the 4th National Climate Assessment (US), Volume I

- Allowed the representation of new climate and biogeochemical processes
- Improved the ESMs ability to represent the real world
- Provides a new framework to investigate the interactions between the different components

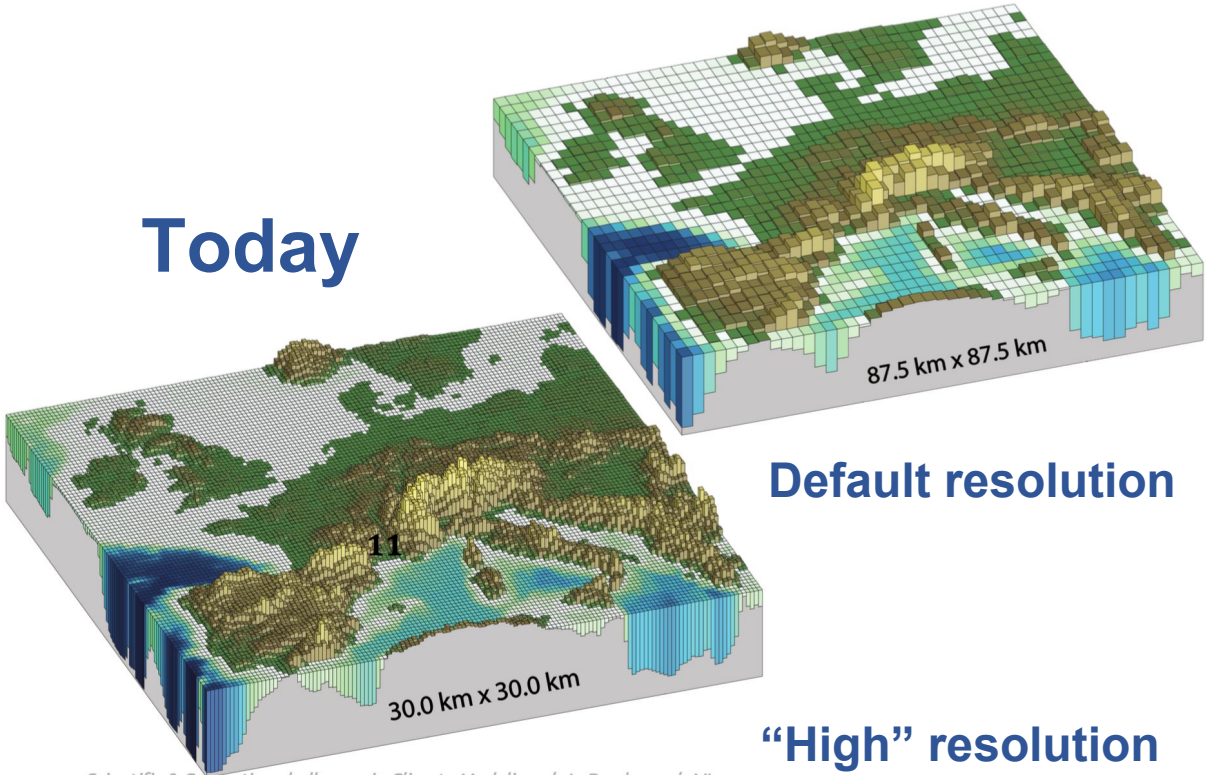
A climate modeling Timeline

Increase in spatial resolution: Atmosphere

Typical climate model in 90s



Today



Default resolution

“High” resolution

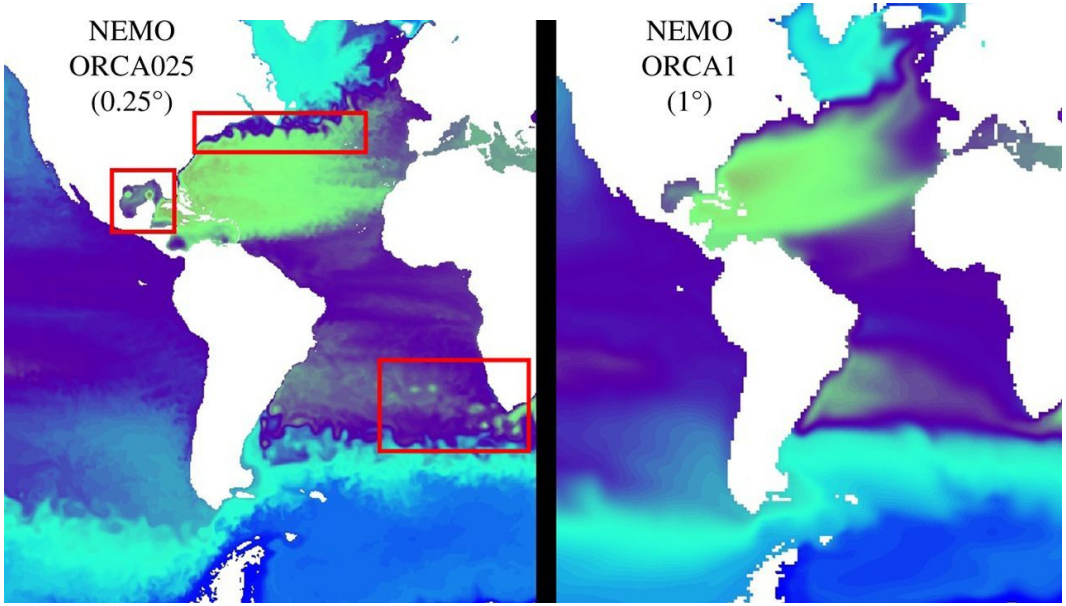
Achieving higher resolutions is essential to better represent orography, and its effect on climate (i.e. in precipitation)

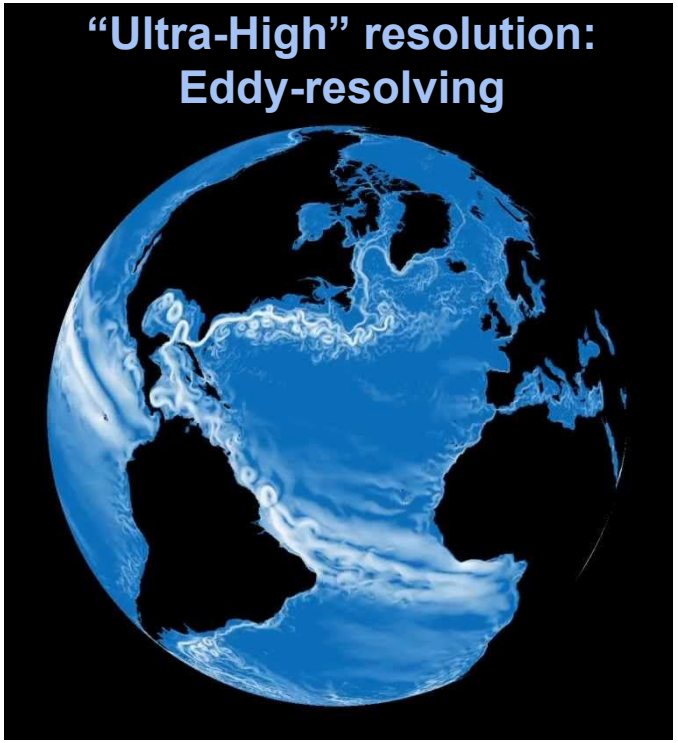
A climate modeling Timeline

Increase in spatial resolution: Ocean

High Resolution (0.25°) Eddy-permitting

Standard Resolution (1°)



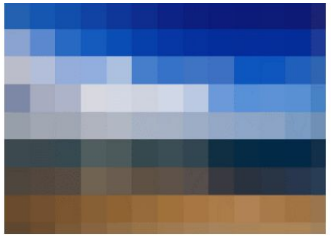


The improvements in ocean resolution translate in a better representation of eddies and ocean currents, which are key to describe realistically decadal variability in the ocean

A climate modeling Timeline

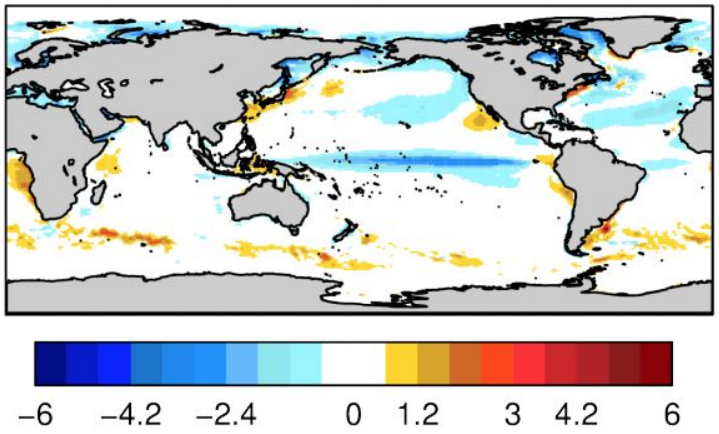
Impact of resolution on model biases

SR: Oce 1°/ Atm 70km



From Prodhomme et al (2016)

BIAS in SST [SR minus OBS]

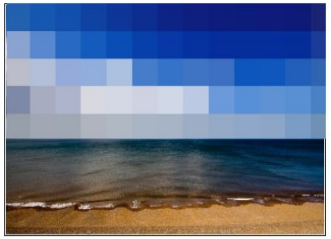


A climate modeling Timeline

Impact of resolution on model biases

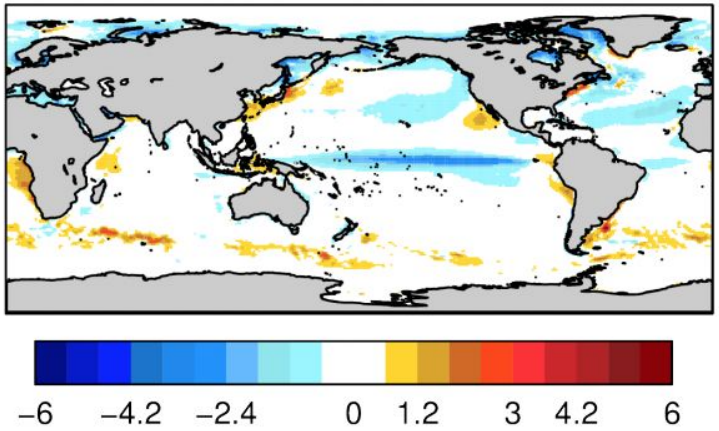
SR: Oce 1°/ Atm 70km

IR: Oce 0.25°/ Atm 70km

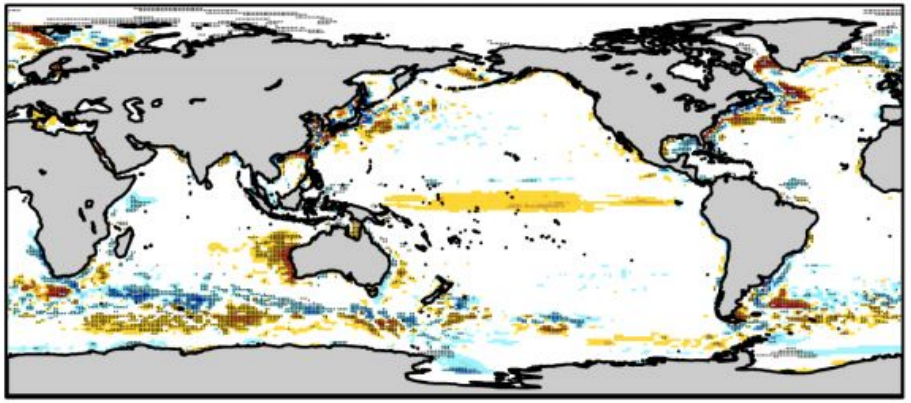


From Prodhomme et al (2016)

BIAS in SST [SR minus OBS]



Diff in SST [IR minus SR]

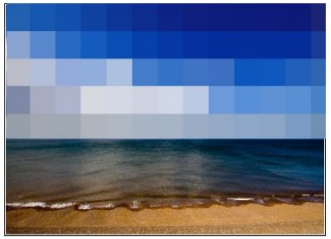


A climate modeling Timeline

Impact of resolution on model biases

SR: Oce 1°/ Atm 70km

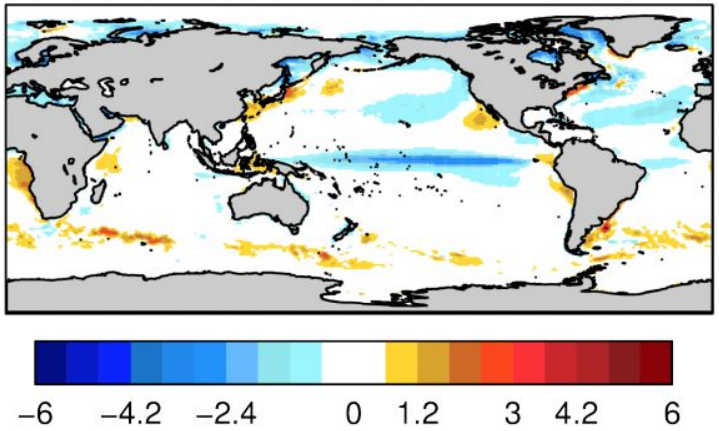
IR: Oce 0.25°/ Atm 70km



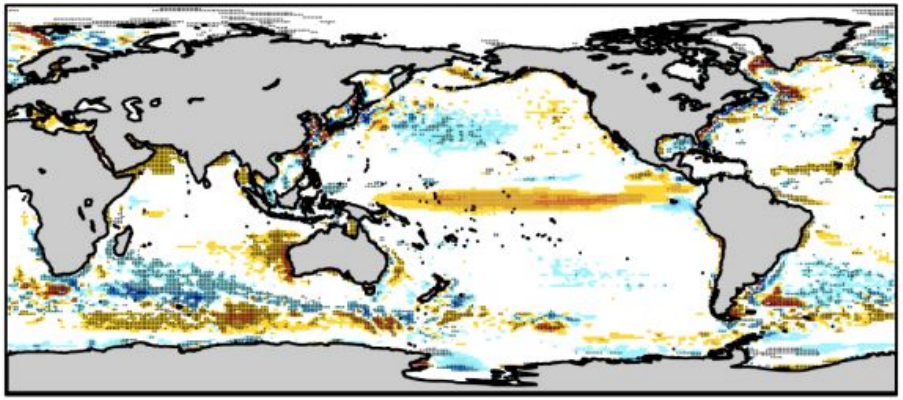
HR: Oce 0.25°/ Atm 40km

From Prodhomme et al (2016)

BIAS in SST [SR minus OBS]



Diff in SST [HR minus SR]



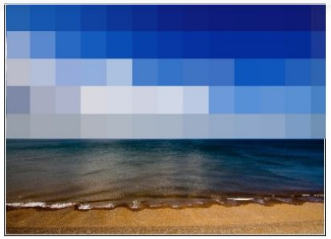
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

A climate modeling Timeline

Impact of resolution on model biases

SR: Oce 1°/ Atm 70km

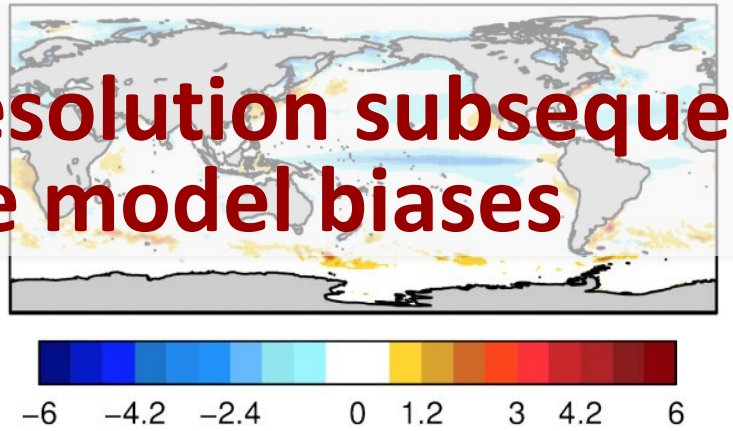
IR: Oce 0.25°/ Atm 70km



HR: Oce 0.25°/ Atm 40km

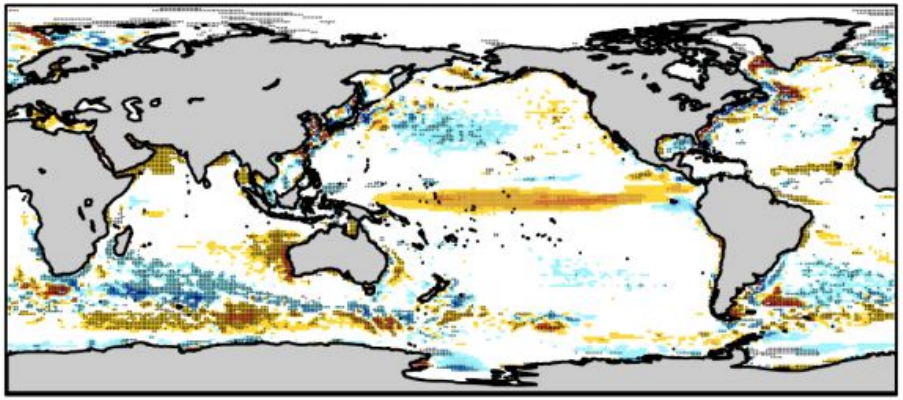
From Prodhomme et al (2016)

BIAS in SST [SR minus OBS]



The increases in resolution subsequently reduce the model biases

Diff in SST [HR minus SR]

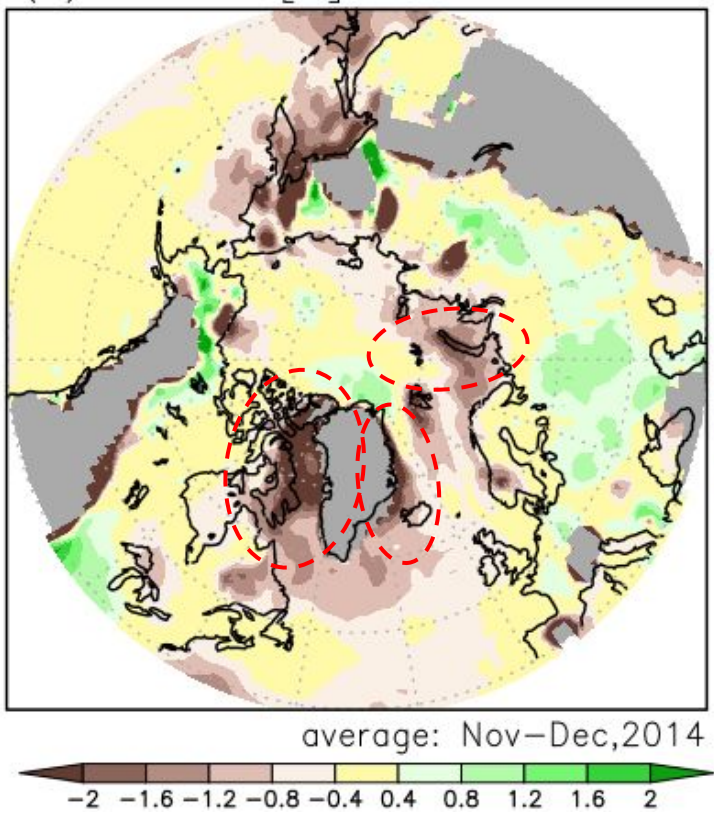


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

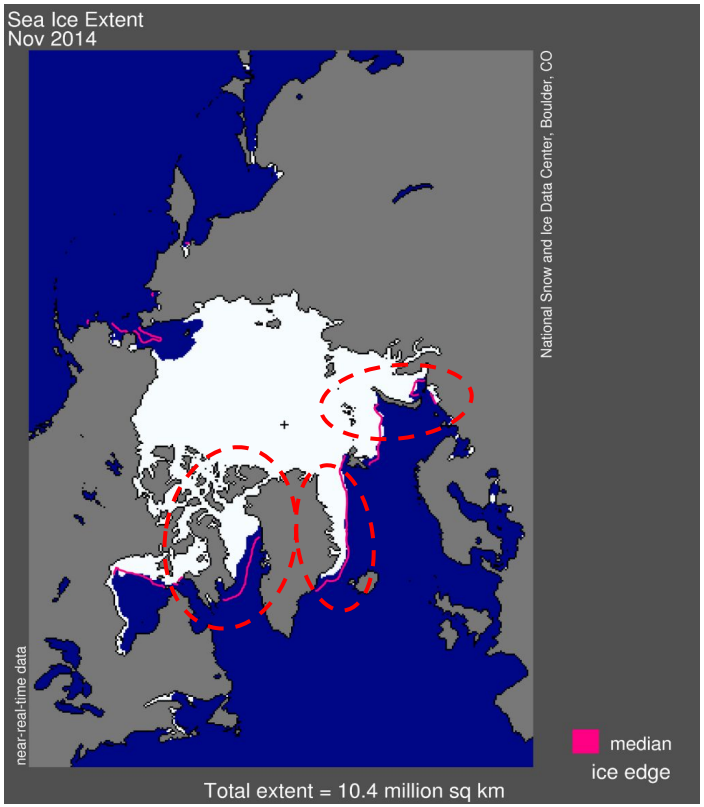
Goal of the collaboration

Reducing the model biases in the NICAM-LETKF model using HR sea ice reconstruction with EC-Earth

RIKEN R-CCS



Barcelona Supercomputing Center
Centro Nacional de Supercomputación



Porting the model to other architectures

MPMD run: different models, different needs

Model	I/O server	Output lib
IFS (atmosphere)	Not used in EC-Earth	GRIB
NEMO (ocean)	XIOS	NetCDF/HDF5

Different systems, different features

Facility	Endianness	RAM/Core	Pre-compiled libs linking
Marenostrum IV	Little endian	2 GB	Both static & dynamic
Mira	Big endian	1 GB	Static
K	Big endian	2 GB	Static

Porting the model to other architectures

Our experience in ALCF (Mira). Similar changes for RIKKEN (K).

	MareNostrum IV	MareNostrum IV
GribAPI	./configure \ --disable-jpeg	./configure \ **--build=ppc64-redhat-linux ** **--host=powerpc64-bgq-linux ** **CC=mpicc ** **CFLAGS="-O3" ** **FC=mpif90 ** **FCFLAGS="-O3" ** **F77=mpif77 ** **FFLAGS="-O3" ** --disable-jpeg \ **--enable-shared=no**
GribEX		
IFS linking		
IFS headers		
IFS compiling		

Porting the model to other architectures

Our experience in ALCF (Mira). Similar changes for RIKKEN (K).

	MareNostrum IV	MareNostrum IV
GribAPI		
GribEX		
IFS linking		
IFS headers		
IFS compiling		

```
#ifdef LITTLE_ENDIAN
  return ((*centre*1000000) +
(*subcentre*1000) + (*number &
0xff));
#else
  return ((*centre*1000000) +
(*subcentre*1000) +
((*number>>shift) & 0xff));
#endif
```

```
#undef LITTLE_ENDIAN
#ifdef LITTLE_ENDIAN
  return ((*centre*1000000) +
(*subcentre*1000) + (*number &
0xff));
#else
  return ((*centre*1000000) +
(*subcentre*1000) +
((*number>>shift) & 0xff));
#endif
```

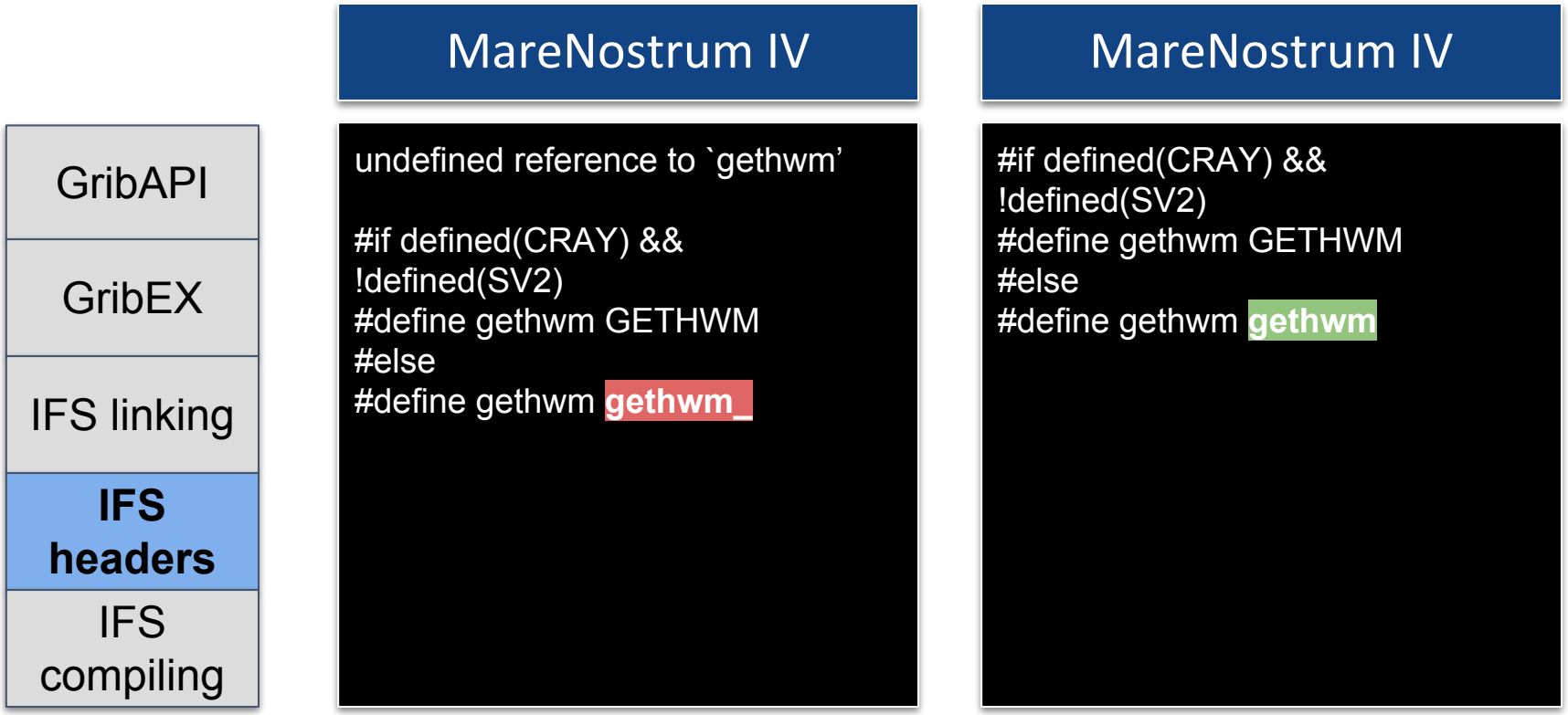
Porting the model to other architectures

Our experience in ALCF (Mira). Similar changes for RIKKEN (K).

	MareNostrum IV	MareNostrum IV
GribAPI	-L/gpfs/mira-fs1/projects/EXCEL/models/ecearth/v3.1/sources/sources/ifs-36r4/lib **-lifs -ltrans -lsurf -lifsalgor -lifsaux -lifs -ldummy**	-L/gpfs/mira-fs1/projects/EXCEL/models/ecearth/v3.1/sources/sources/ifs-36r4/lib **-lifs -lifsaux -lifsalgor -ltrans -lifsalgor -lsurf -lifsaux -lifs -ldummy**
GribEX	-L/gpfs/mira-fs1/projects/EXCEL/models/ecearth/v3.1/sources/sources/oasis3/ecconf/lib -lpsmile.MPI1 -lmpp_io -lclim.MPI1	-L/gpfs/mira-fs1/projects/EXCEL/models/ecearth/v3.1/sources/sources/oasis3/ecconf/lib -lpsmile.MPI1 -lmpp_io -lclim.MPI1
IFS linking	-L/soft/libraries/netcdf/current/cnk-gcc/current/lib -lnetcdff -lnetcdf -lpnetcdf -lhdf5_hl -lhdf5 -ldl -lz -L/projects/EXCEL/opt/grib_api-1.14.0/cnk-gcc/lib -lgrib_api_f90 -lgrib_api -L/projects/EXCEL/opt/emos_000392/cnk-gcc/ -lemosR64 -L/soft/libraries/alcf/current/gcc/LAPACK/lib -llapack	-L/soft/libraries/netcdf/current/cnk-gcc/current/lib -lnetcdff -lnetcdf -L/soft/libraries/pnetcdf/current/cnk-gcc/current/lib -lpnetcdf -L/soft/libraries/hdf5/current/cnk-gcc/current/lib -lhdf5_hl -lhdf5 -ldl -L/soft/libraries/alcf/current/gcc/ZLIB/lib -lz -lm -L/projects/EXCEL/opt/grib_api-1.14.0/cnk-gcc/lib -lgrib_api_f90 -lgrib_api -L/projects/EXCEL/opt/emos_000392/cnk-gcc/ -lemosR64 -L/soft/libraries/alcf/current/gcc/LAPACK/lib -llapack -L/soft/libraries/alcf/current/gcc/BLAS/lib -lblas
IFS headers		
IFS compiling		

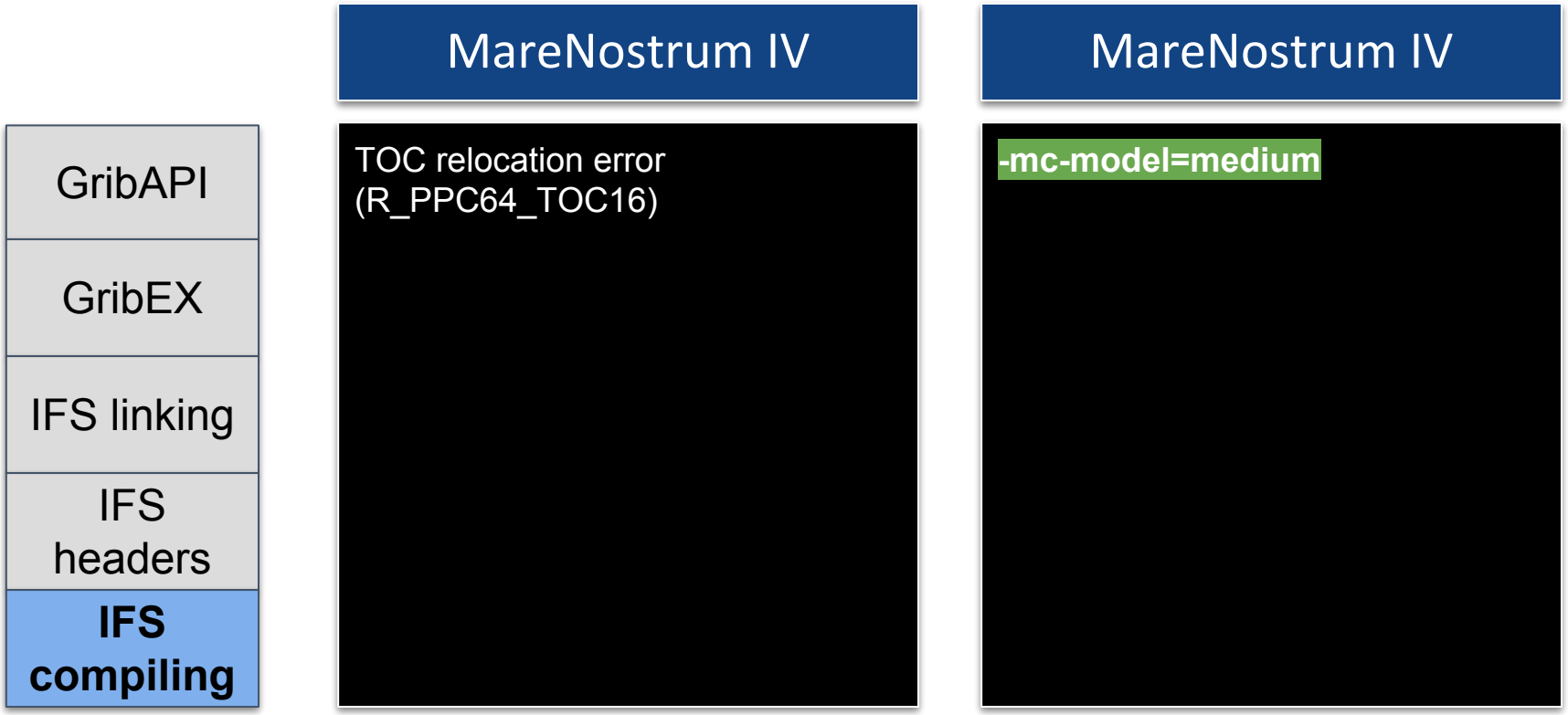
Porting the model to other architectures

Our experience in ALCF (Mira). Similar changes for RIKKEN (K).



Porting the model to other architectures

Our experience in ALCF (Mira). Similar changes for RIKKEN (K).



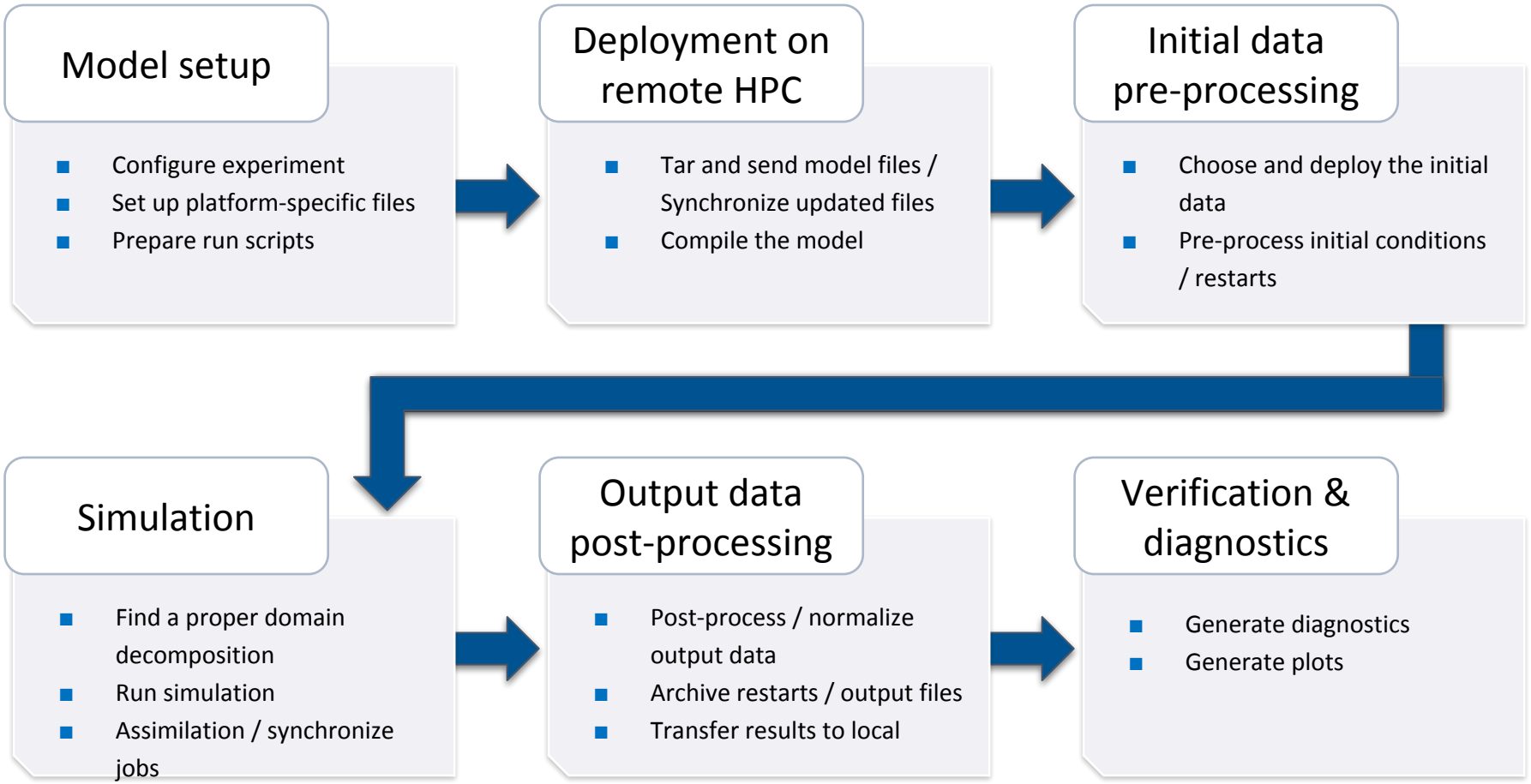
Workflow management

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

High Performance Computing in Earth Sciences

- Earth System Models (ESMs) are sophisticated tools with continuously increasing complexity:
  - More components of Earth System are included
  - Finer Spatial and Temporal resolutions
- This increase in complexity could be developed thanks to the important parallel advances in HPC

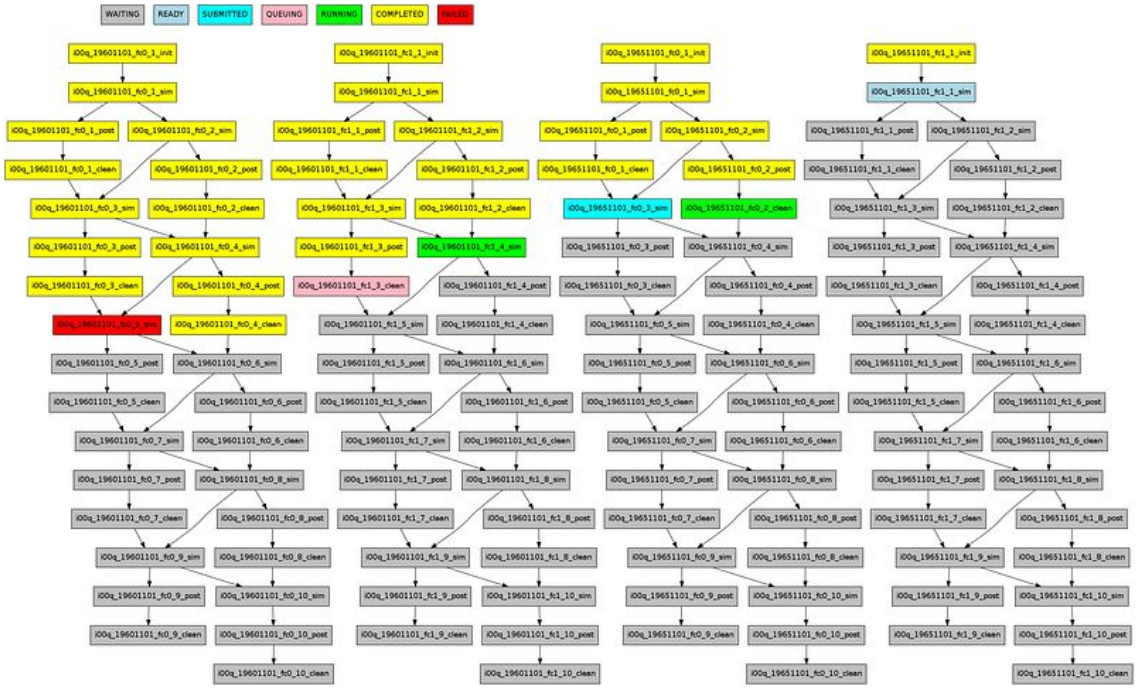
A workflow for Earth System models



Workflow managers: motivation

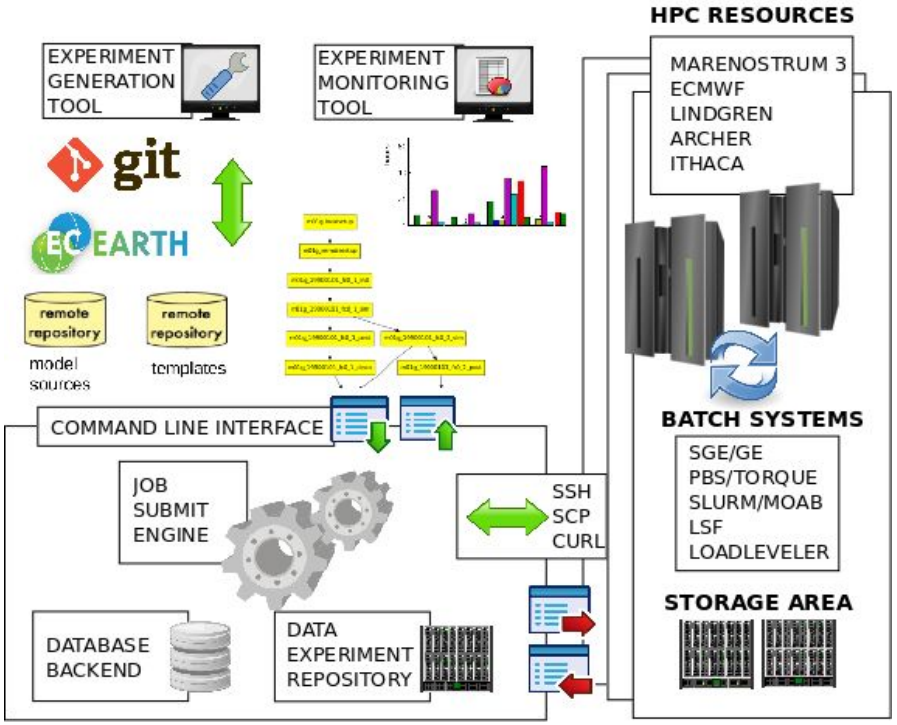
Workflow managers are **essential** to carry out production experiments in an **efficient** way

- Deal with workflow **complexity**
- Ensure **robustness** & **portability**
- **Usability** → Scientists more productive



Autosubmit

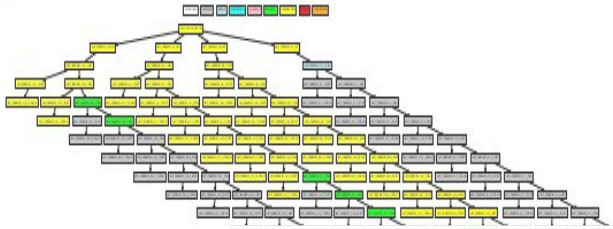
A **versatile** tool to manage Weather and Climate Experiments in diverse Supercomputing Environments:
https://pypi.python.org/pypi/autosubmit



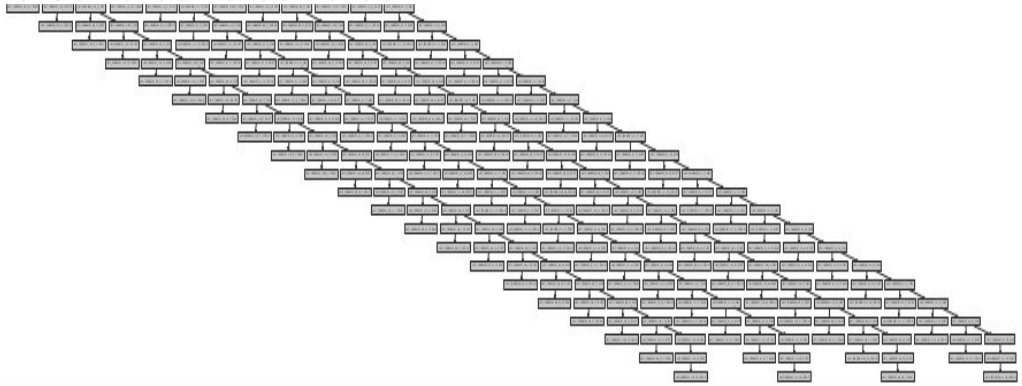
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

AUTOSUBMIT

Autosubmit: lessons learned

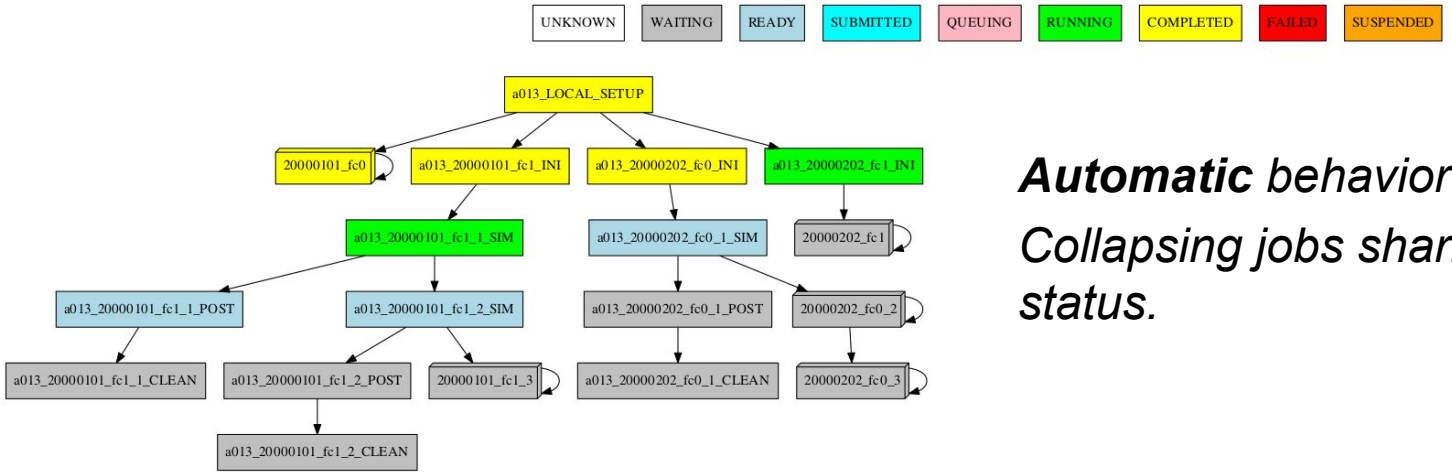


- Workflows are getting more and more **complex**
- Workflow **managers** are required to **improve** in order to **deal** with this **complexity**



Autosubmit v3.10: Visualization

Improvements in the graph **visualization** of the workflow, by **grouping** jobs by date, member, chunk, split; or automatically.

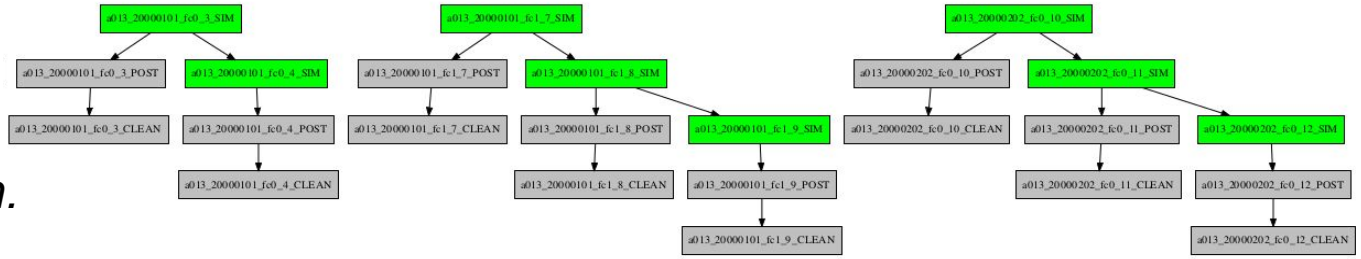


***Automatic** behavior:*

Collapsing jobs sharing status.

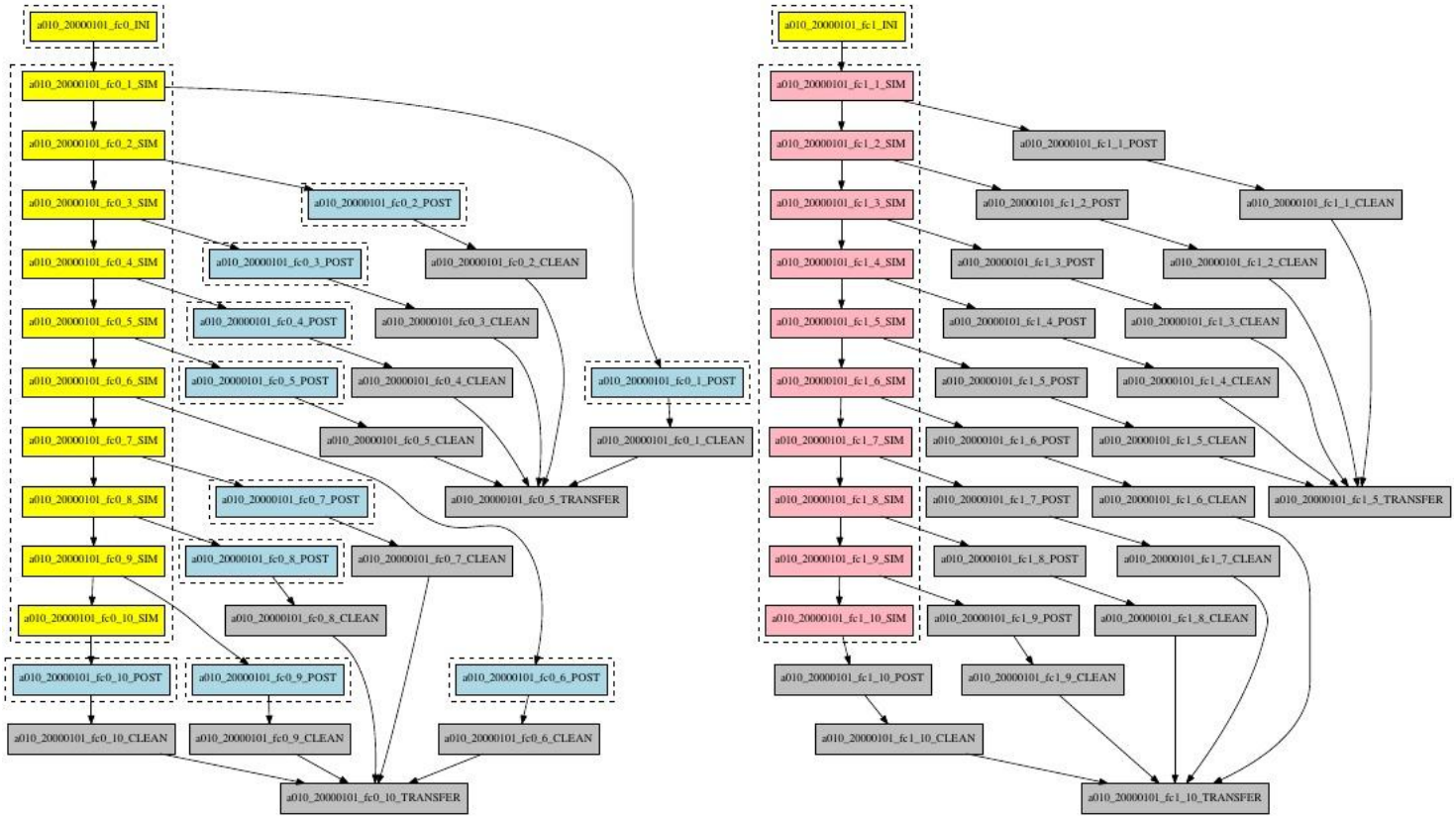
Hide groups:

*Showing the more **relevant** information.*



Autosubmit: Wrapper

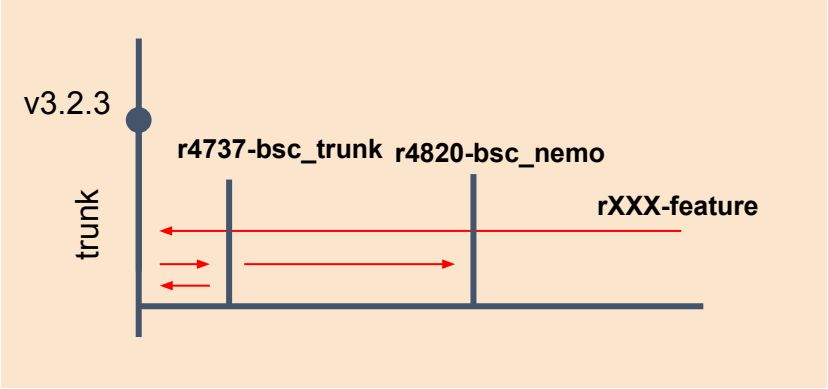
Motivation: to **improve** throughput by **reducing** queueing **time** through wrapping different jobs together.



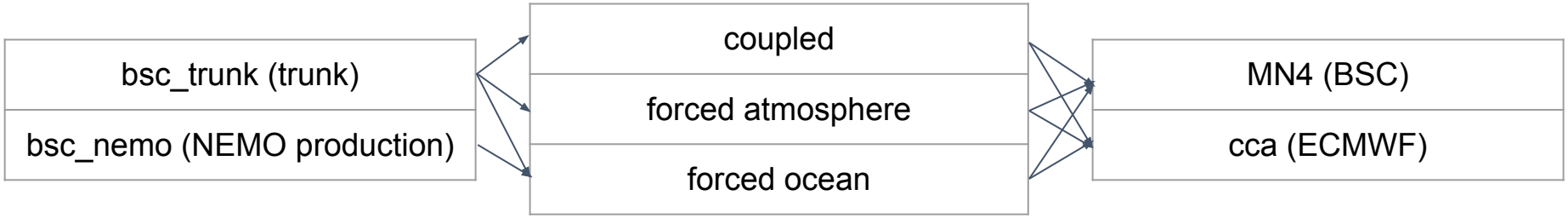
Auto-EC-Earth testing: continuous integration

Goal: To be able to **run** the last EC-Earth version, **use** new features,

merge latest developments smoothly



Every week: run a set of LR tests



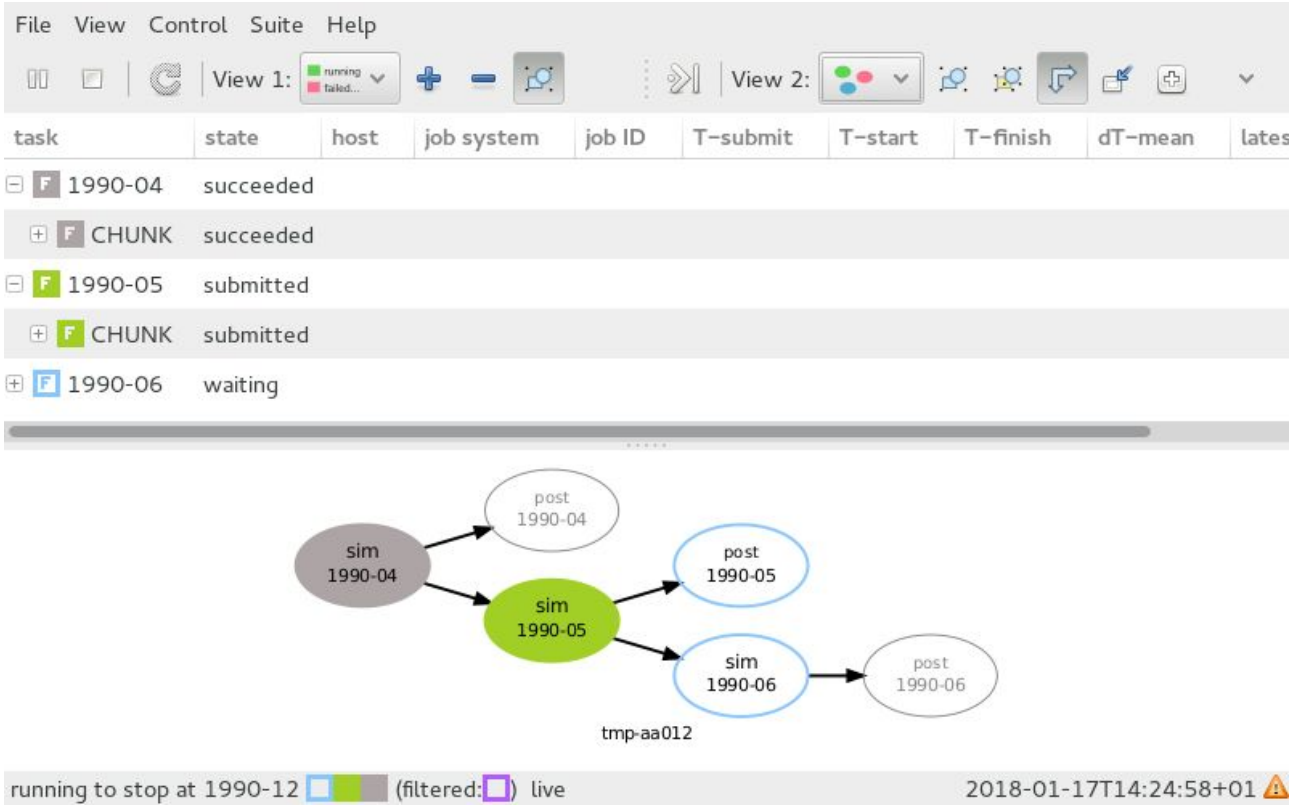
Auto-EC-Earth testing: release tests

For every verison release: run a complete set of tests

nord3	CCA	MN4	resolution	type	details
t02c	t00u	t00q	T255L91-ORCA1L75-LIM3	coupled	start from restart
		t00v	T255L91-ORCA1L75-LIM3	coupled	atmos. nudging
		t011	T255L91-ORCA1L75-LIM3	coupled	sppt
	t00s	t00o	T255L91-ORCA1L75-LIM3	coupled	cold start
	t01d	t00z	ORCA1L75-LIM3	nemo	cold start
		t01j	ORCA1L75-LIM3	nemo	cold start ocean nudging
	t01e	t00r	T511L91-ORCA025L75-LIM3	coupled	start from restart
		t01o	ORCA025L75-LIM3	nemo	cold start
	t01b	t00y	T511L91	ifs	cold start
	t00t	t00p	T511L91-ORCA025L75-LIM3	coupled	cold start

Cylc

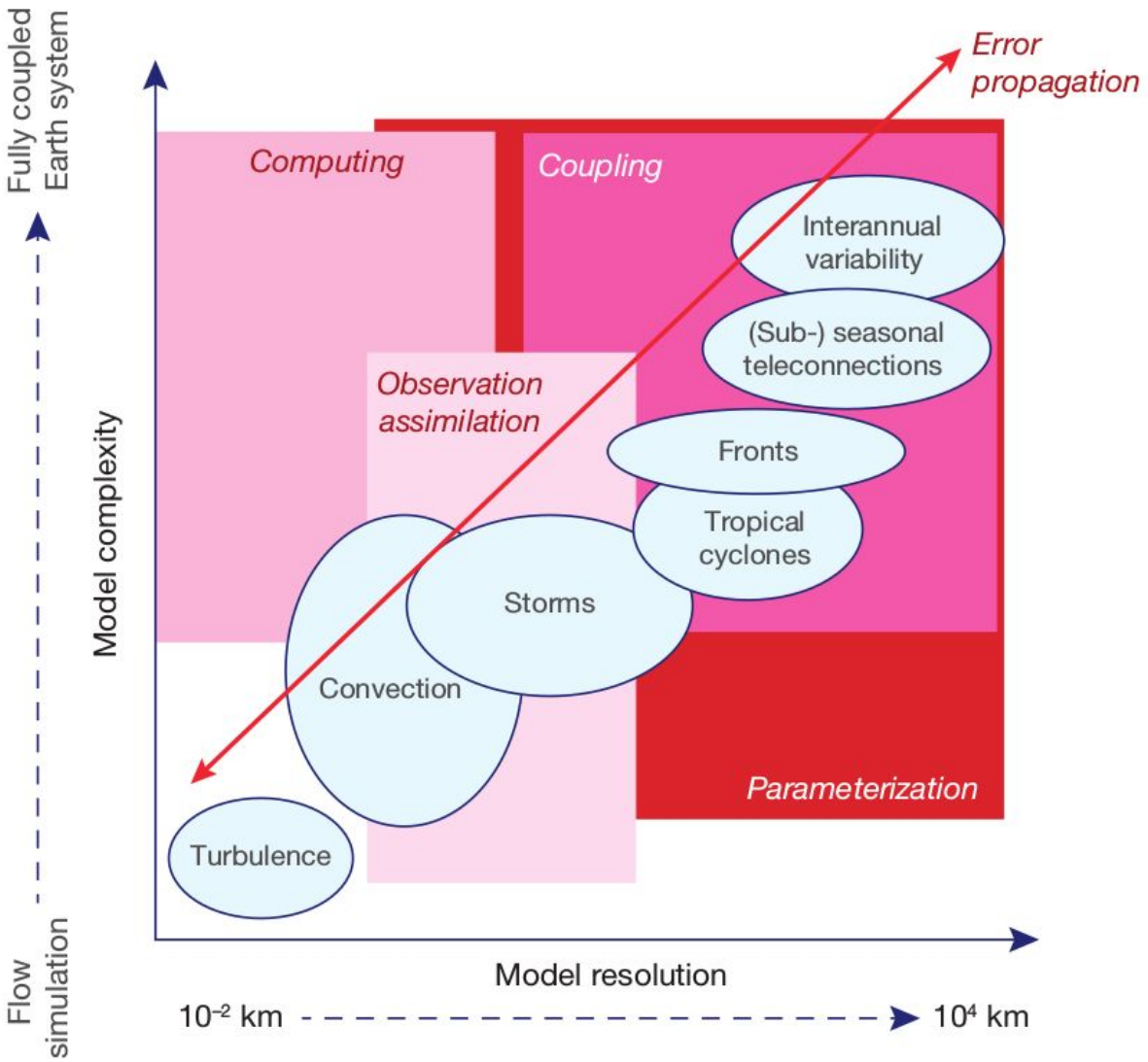
- First **proof of concept** → NEMO standalone integrations
- Testing different **interfaces** and **configurations: Rosie go, Rose Config Edit & Rose Stem**



Earth System Model performance

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HPC Challenges



Marenostrum IV

- MareNostrum IV in operation since July 2017
- One of the first HPCs featuring new Intel Scalable Processors

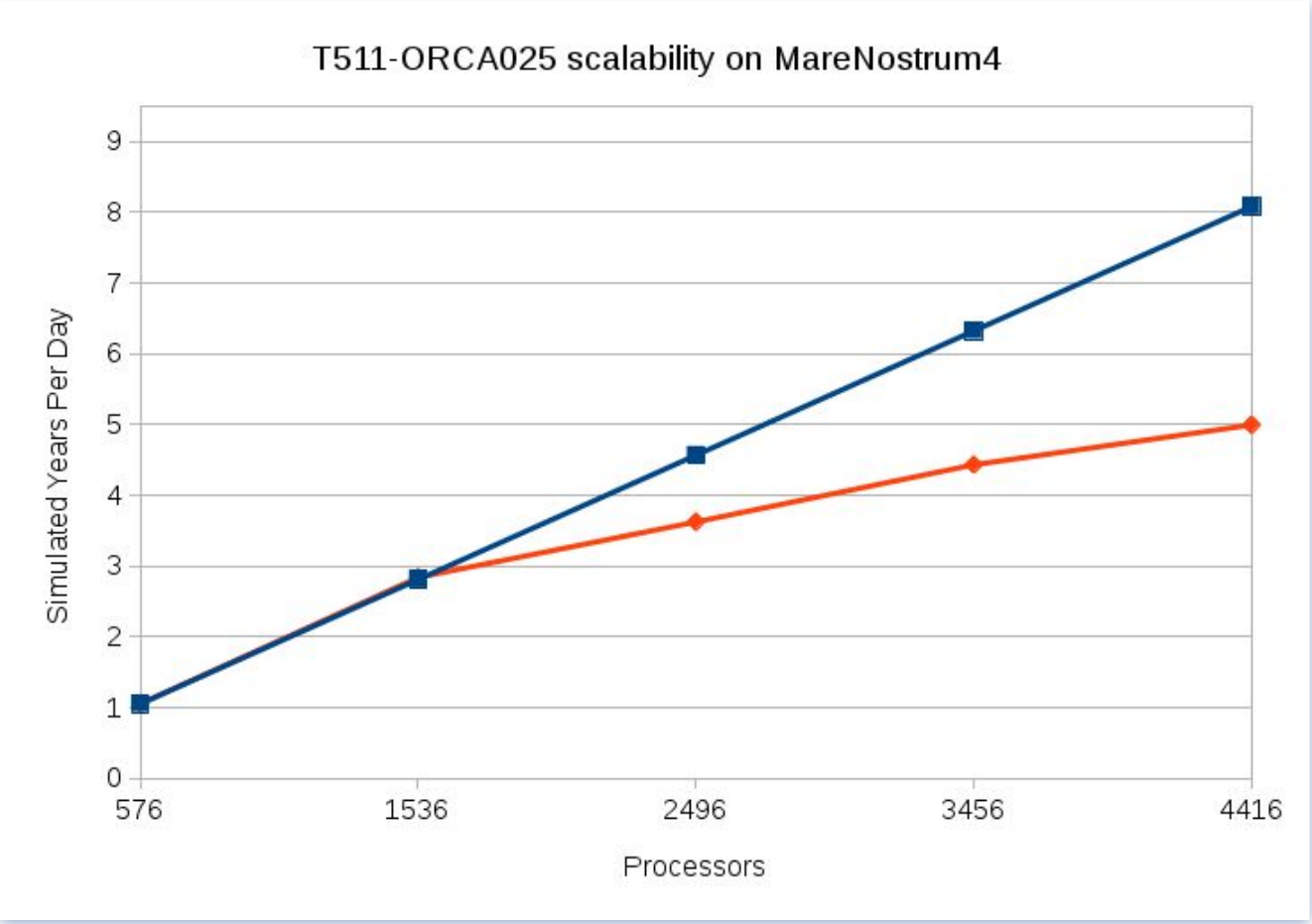
	MareNostrum III	MareNostrum IV
Processor	Intel Xeon E5-2670 2.6 GHz	Intel Xeon Platinum 8160 2.1 GHz
#Cores per socket	8	24
#Sockets	2	2
Memory	32Gb DDR3-1600 2 GB/core	96Gb DDR4-2667 2 GB/core
Interconnection	Infiniband FDR10 10Gb	Intel Omni-Path 100Gb

MareNostrum III - 11,15 petaFLOPS

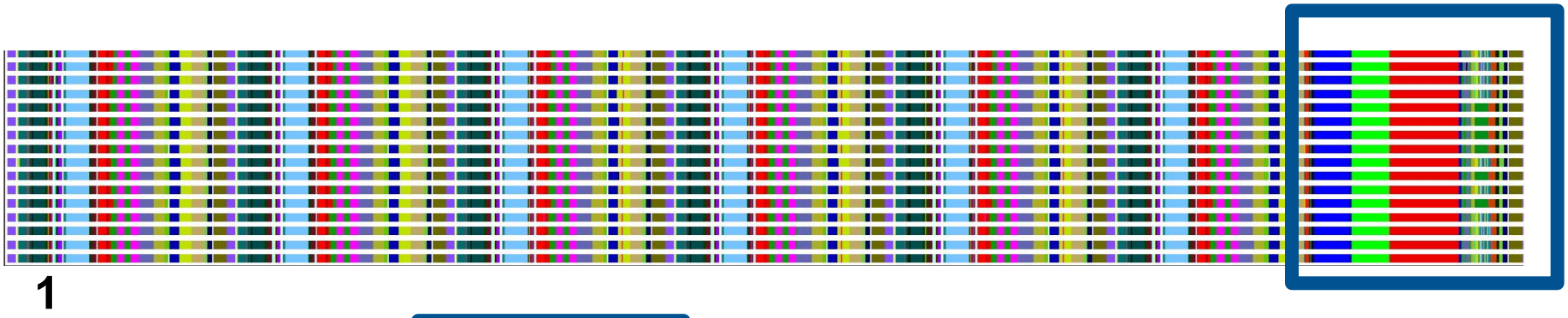
MareNostrum IV - 11,15 petaFLOPS

EC-Earth scalability on MN4

Scalability for EC-Earth trunk with default output configuration



Performance Analysis

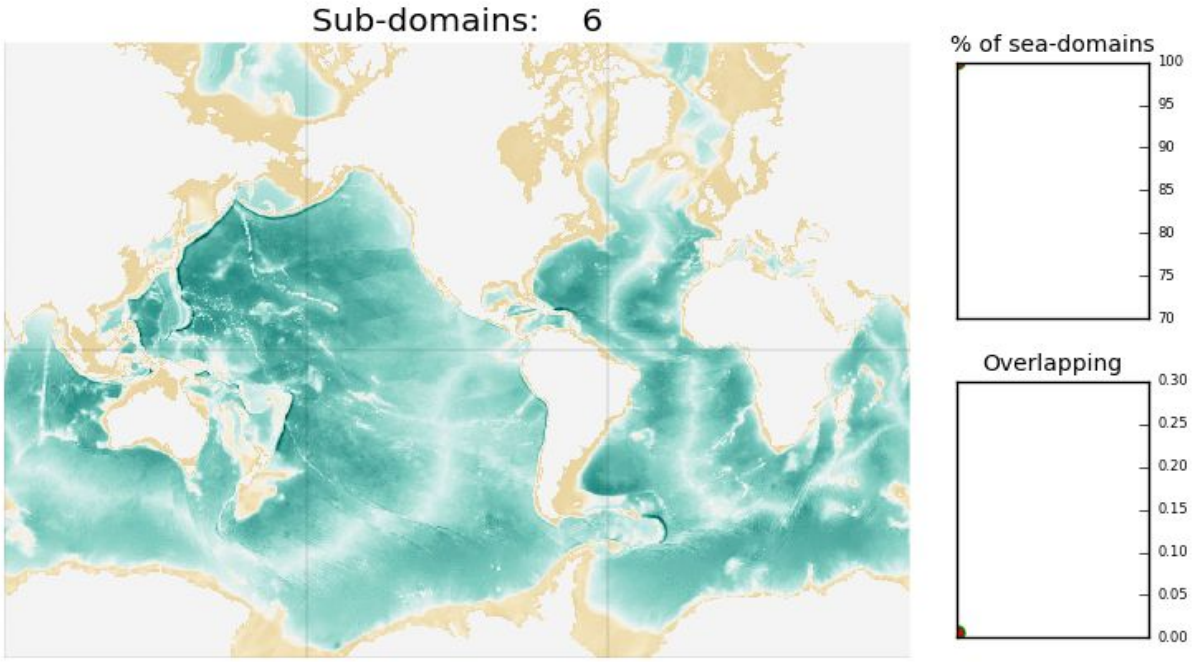


Optimizations for NEMO

- Diagnostic for NEMO:
  - Scalability is constrained by:
    - 1) Algorithms with too much communication
    - 2) Sub-optimal implementation
- Actions taken
  - Improve communication implementation to reduce number of point-to-point messages
  - Reduce number of collectives

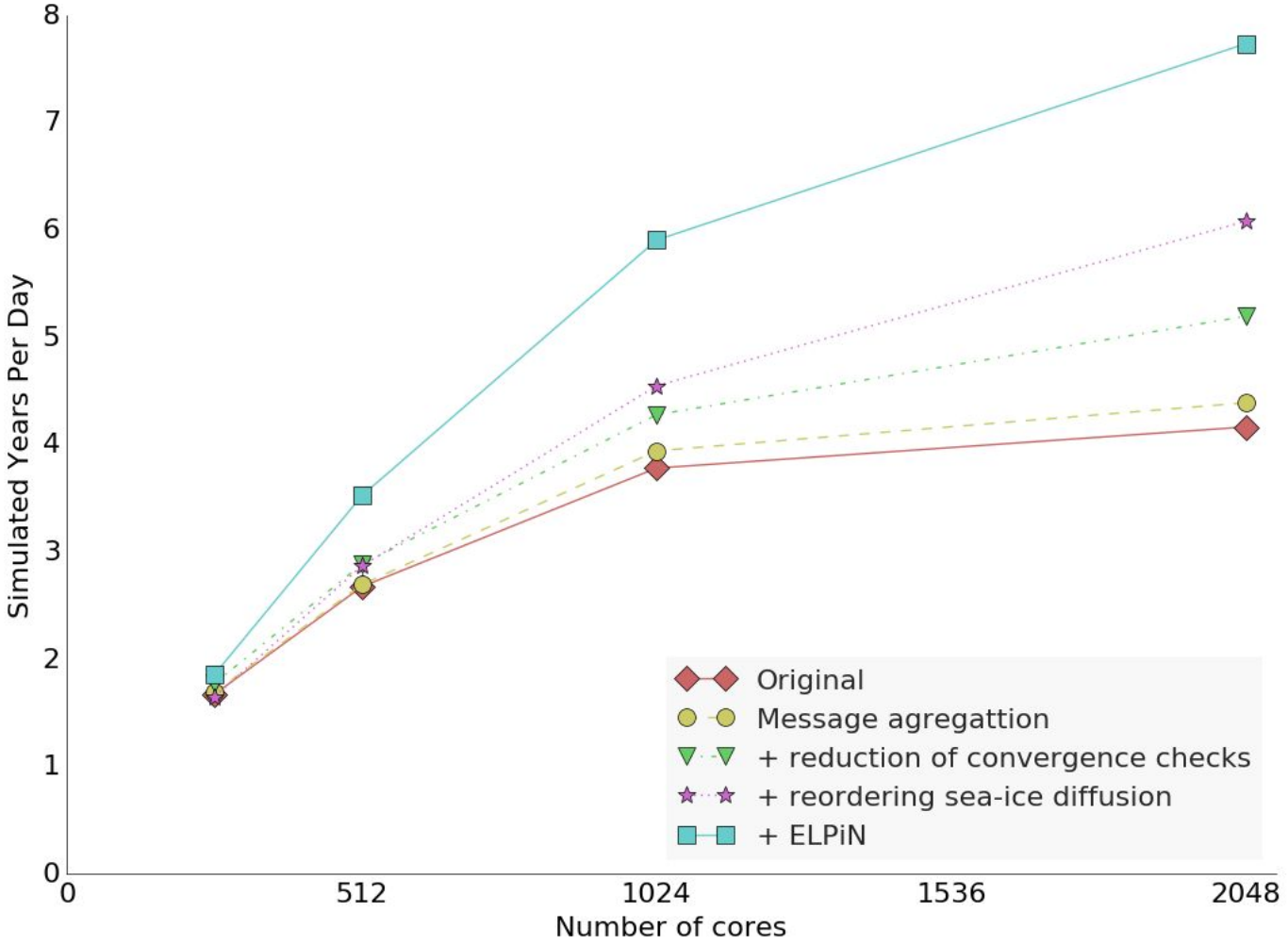
Optimizations for NEMO

- ELPiN allows to find proper namelist parameters to exclude land-only processes in NEMO simulations

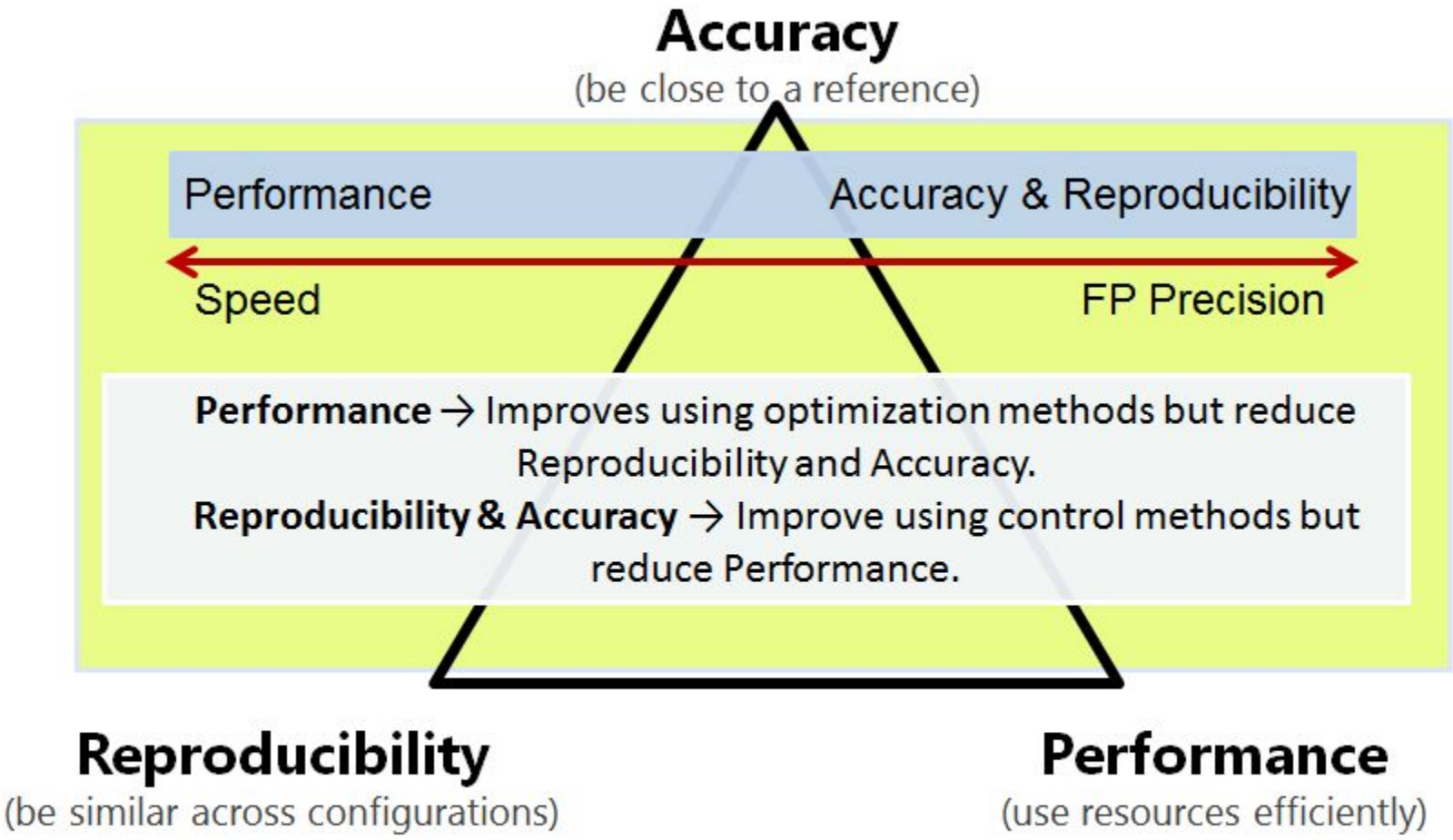


Optimizations for NEMO

- Impact of proposed optimizations on ORCA1/4º + LIM3 simulations (27km global)



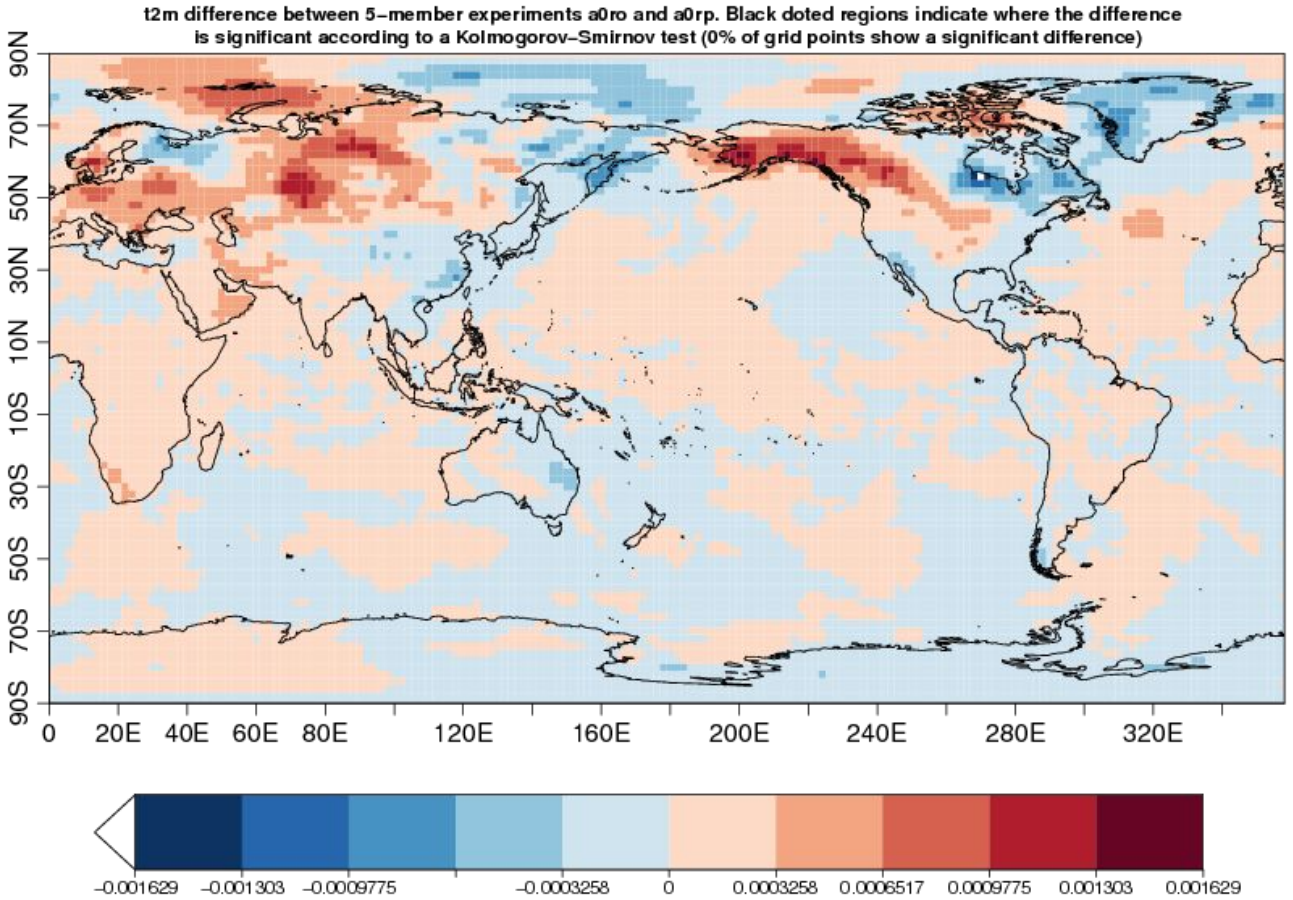
Results Verification



Find options to control the tradeoffs among accuracy, reproducibility and performance.

Reproducibility

- Compare to CMIP5 results to evaluate the accuracy and between two experiments to evaluate the reproducibility



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you

YourEmail@bsc.es