



EC-Earth4 - initial scalability and performance on BSC-MareNostrum5

Eric Ferrer - Gilbert Montané

Roc Salvador - Marcel Cases

18/09/2024

Deploying EC-Earth

Steps to use a new system:

1. Deploy and run
2. Reproducibility
3. Performance, coupled and by component
(scalability)
4. Optimize coupled performance (balancing)

MareNostrum 5 System overview

MareNostrum 5 GPP - P-22 Top500 June24

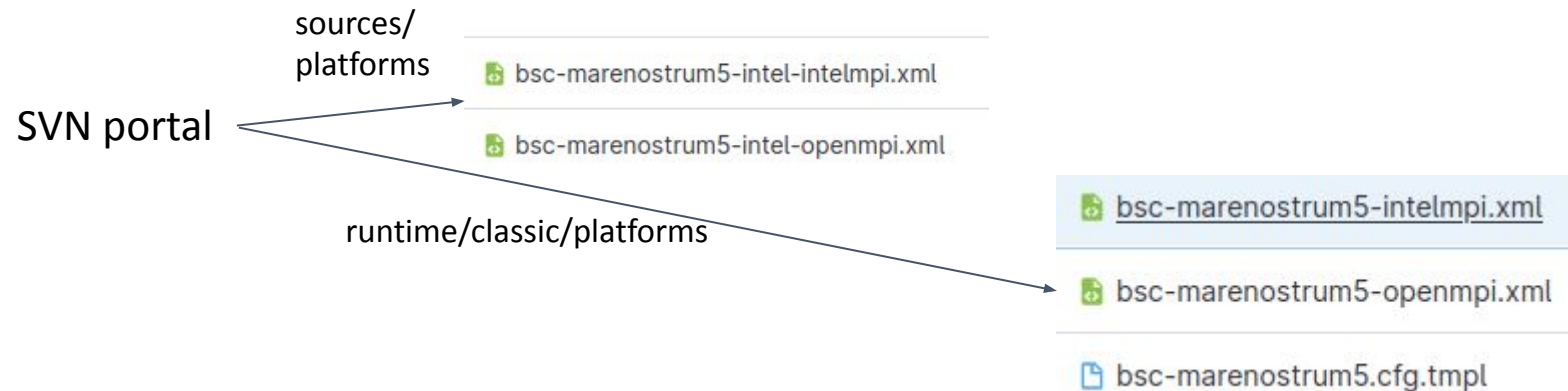
6408 nodes - Intel Sapphire Rapids
2x 8480+ at 2Ghz & 56c (112 cores x node)

MareNostrum 5 ACC P-8 Top500 June24

1120 nodes - Intel Sapphire Rapids and Nvidia
Hopper GPUs.
2x 8460Y+ at 2.3Ghz and 32c (64 cores x node)
4x GPUs with 64 HBM2 memory

Deployment of EC-Earth

- Create all the compilation and configuration files needed
 - EC-Earth3 (OpenMPI and IntelMPI options):



- EC-Earth4 (only OpenMPI):
SMHI GitLab → bsc-marenostrum5-intel+openmpi.yml

EC-EARTH 3 Repro test

- **Bit to bit** reproducibility -> not achievable
- Target -> **statistical reproducibility**

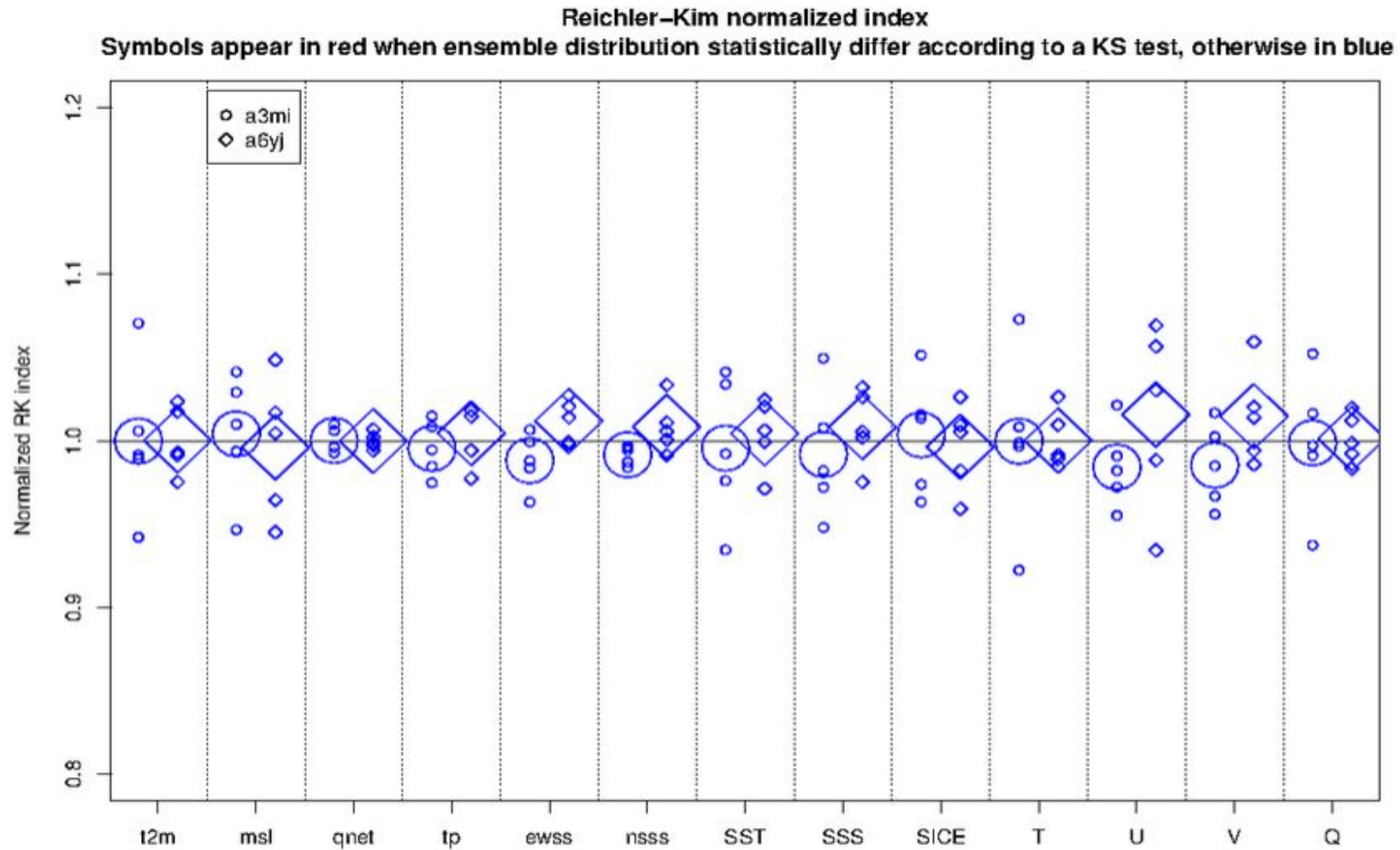
EC-EARTH 3 Repro test

- **Bit to bit** reproducibility -> not achievable
- Target -> **statistical reproducibility**

How to ensure it?

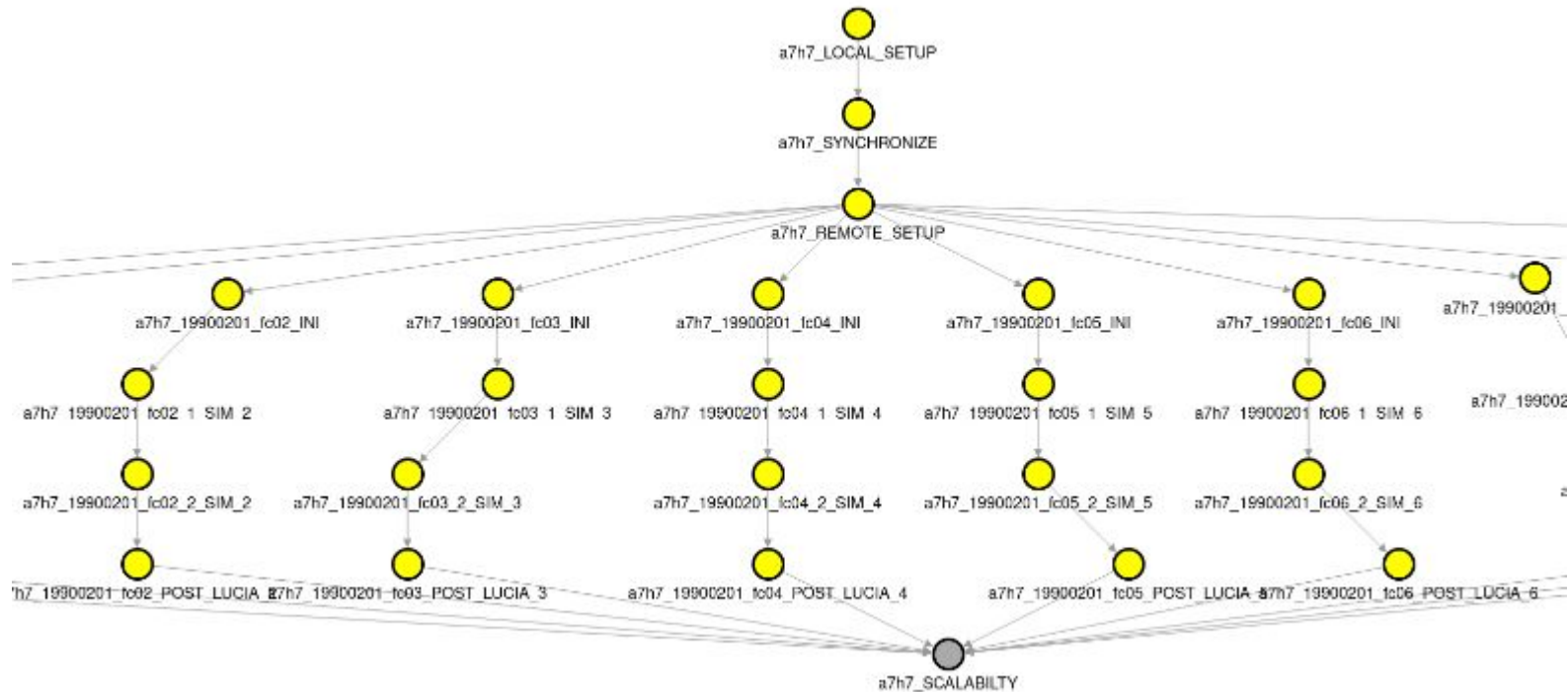
- 5 member - 20 year long ensemble
- **Compare** important variables with **baseline reference**

EC-EARTH 3 Repro test



Massonnet, F., Ménégoz, M., Acosta, M., Yepes-Arbós, X., Exarchou, E., and Doblas-Reyes, F. J.:
Replicability of the EC-Earth3 Earth system model under a change in computing environment, *Geosci. Model
Dev.*, 13, 1165–1178, <https://doi.org/10.5194/gmd-13-1165-2020>, 2020.

EC-Earth 3 Auto-Scalability

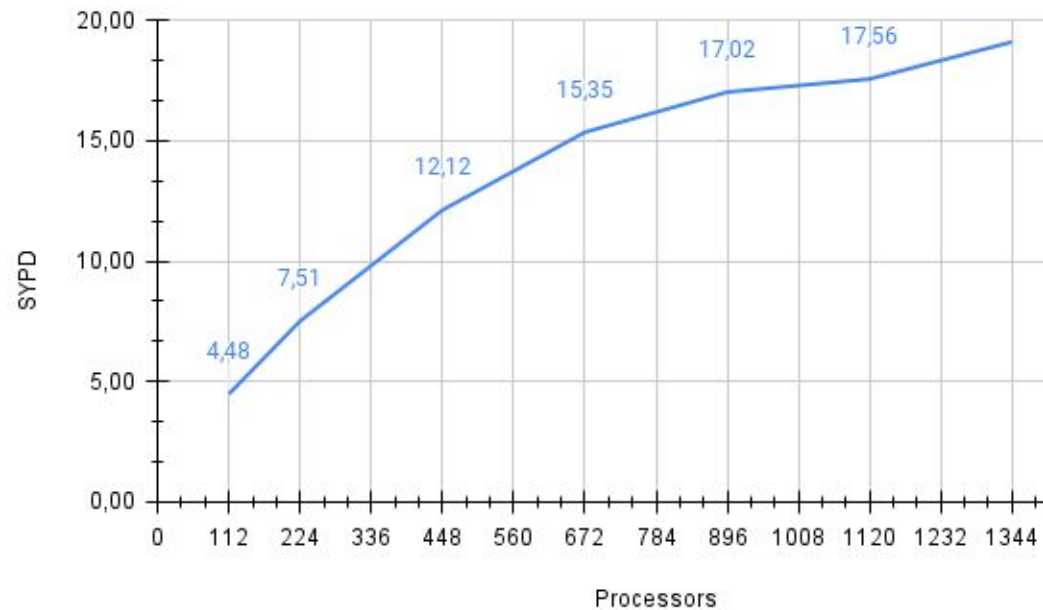


- **One** experiment does everything
- **Based** on **Lucia** tool to automatically obtain data* for:
 - NEMO & IFS
 - Individually & Coupled

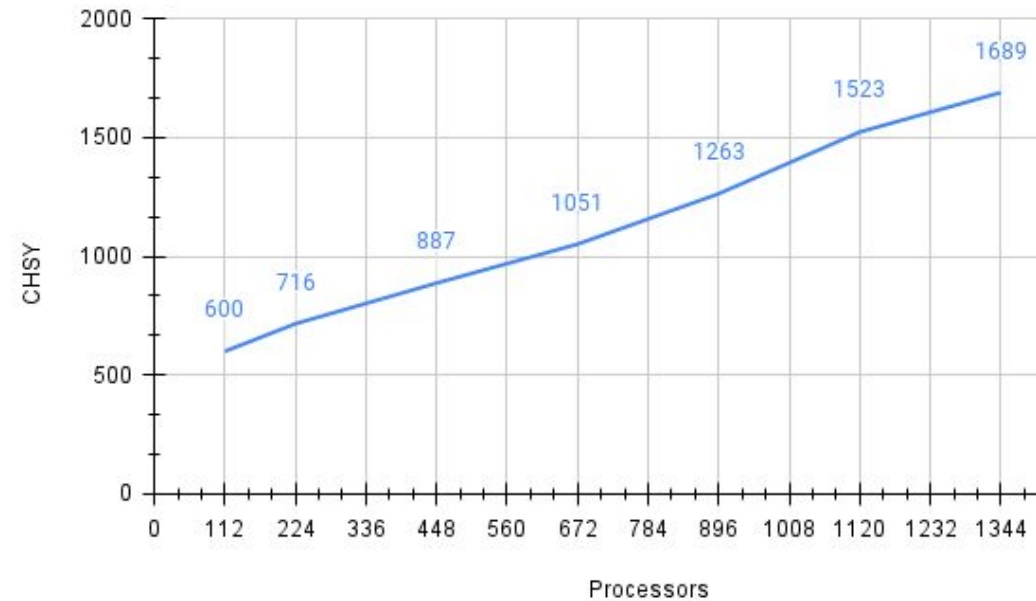
*SYDP, CHPSY, coupling cost, wait time

EC-Earth 3 (TL255L91-Orca1) Coupled Scalability

SYPD



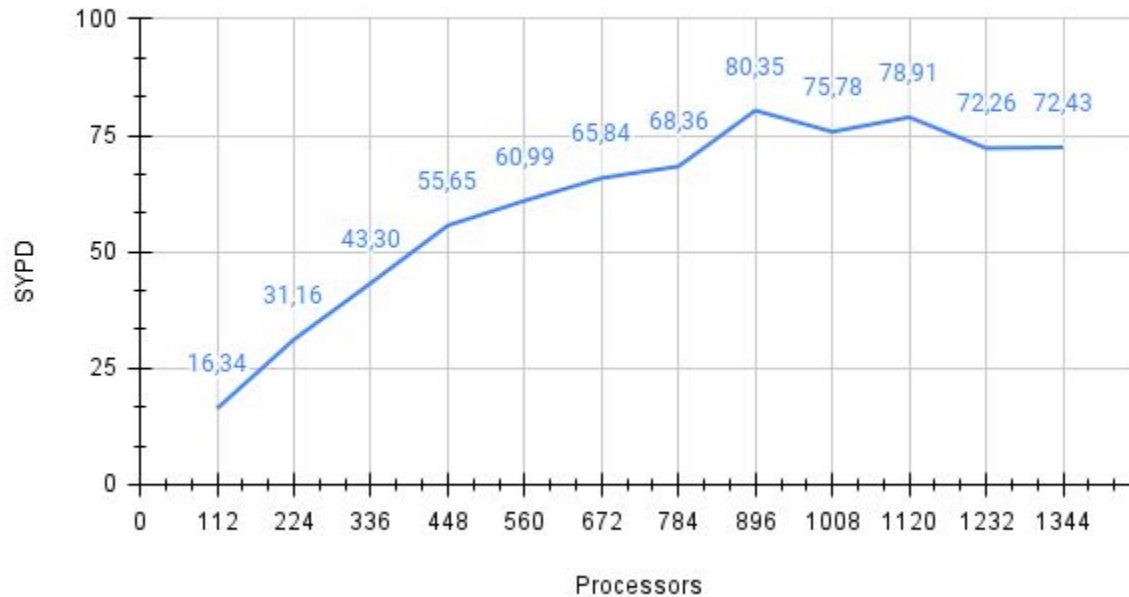
CHSY



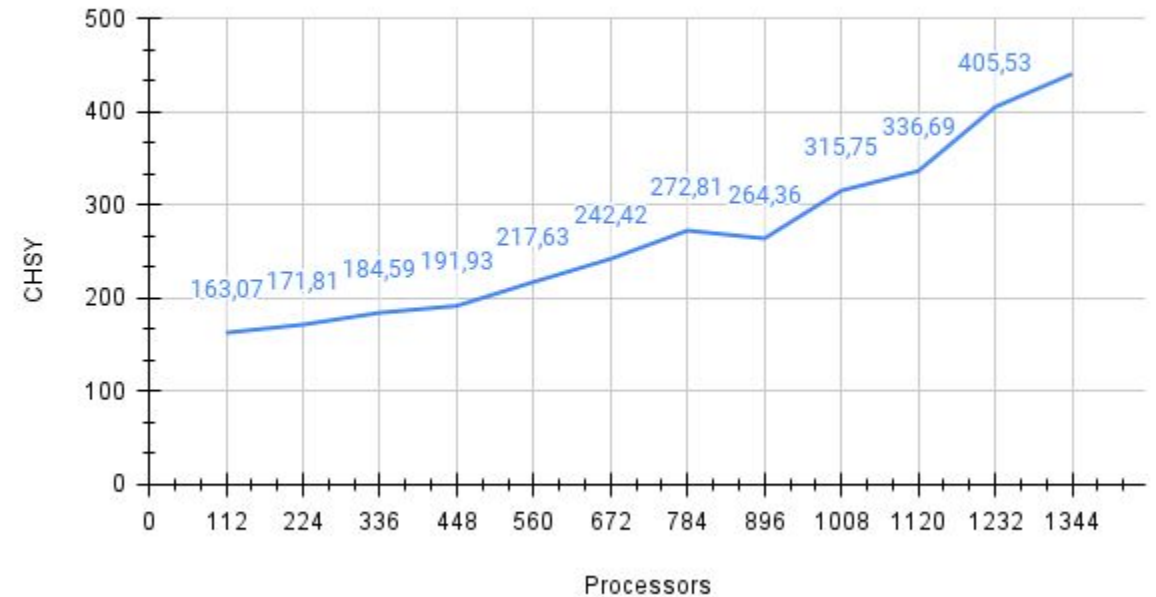
*Unbalanced, CHSY are not representative of what can be obtained when scaling up with balanced configurations

EC-Earth 4 (TL159L91-eOrca1) Coupled Scalability

SYPD



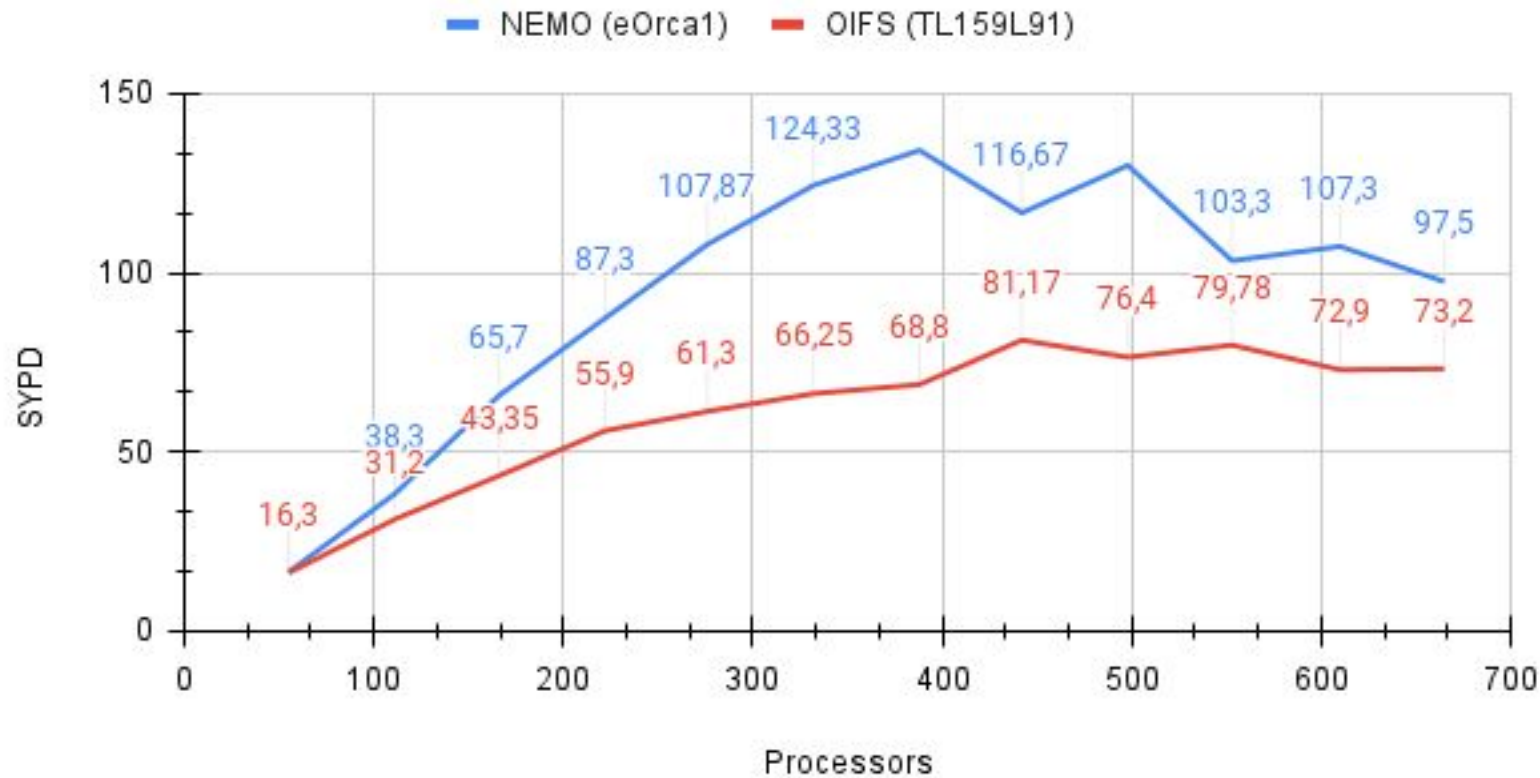
CHSY



*Unbalanced, CHSY are not representative of what can be obtained when scaling up with balanced configurations

EC-Earth 4 (TL159L91-eOrca1) Scalability by component

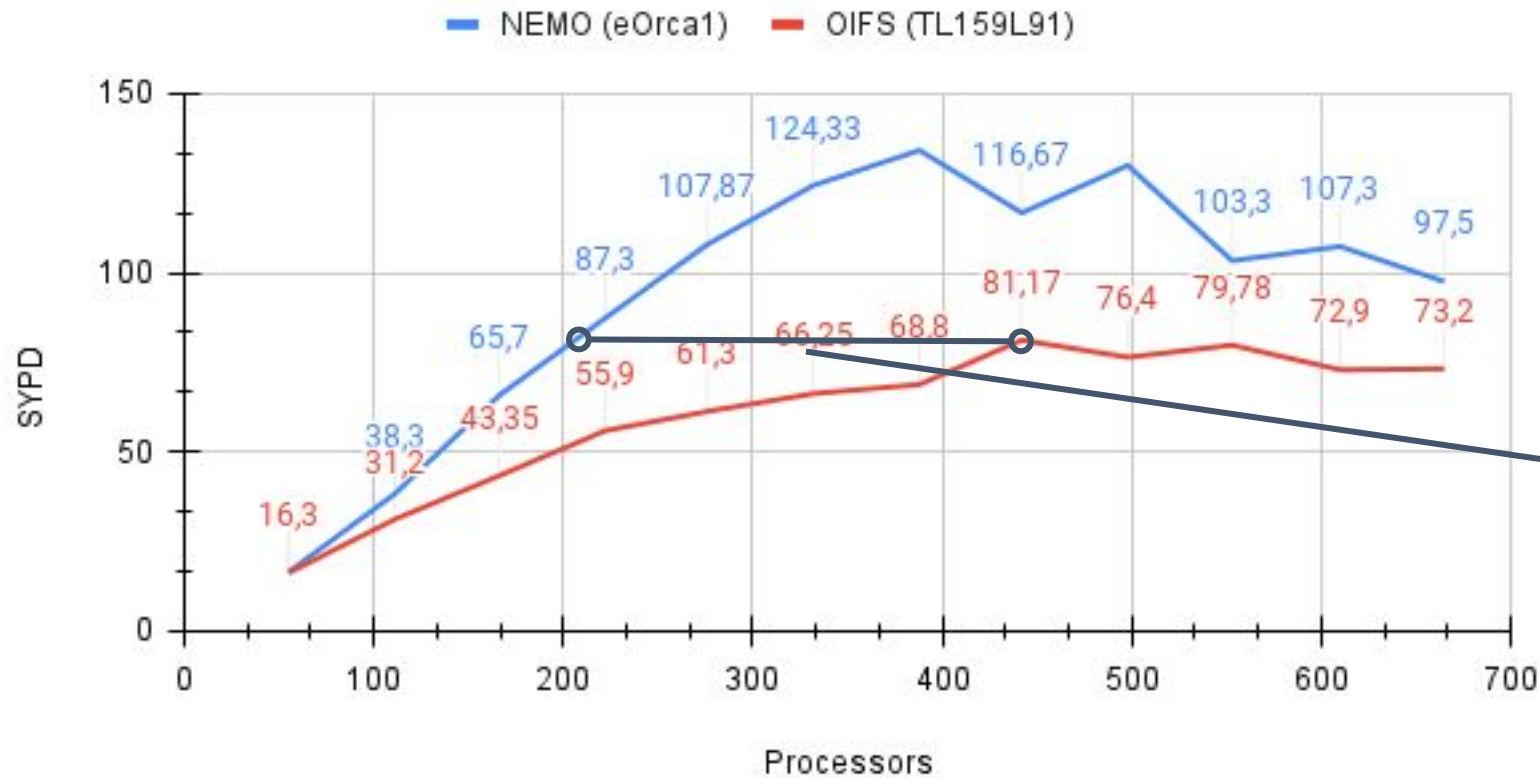
NEMO & OIFS scalability on MN5



- Lucia provides the individual scalability from a coupled run.

EC-Earth 4 (TL159L91-eOrca1) Scalability by component

NEMO & OIFS scalability on MN5



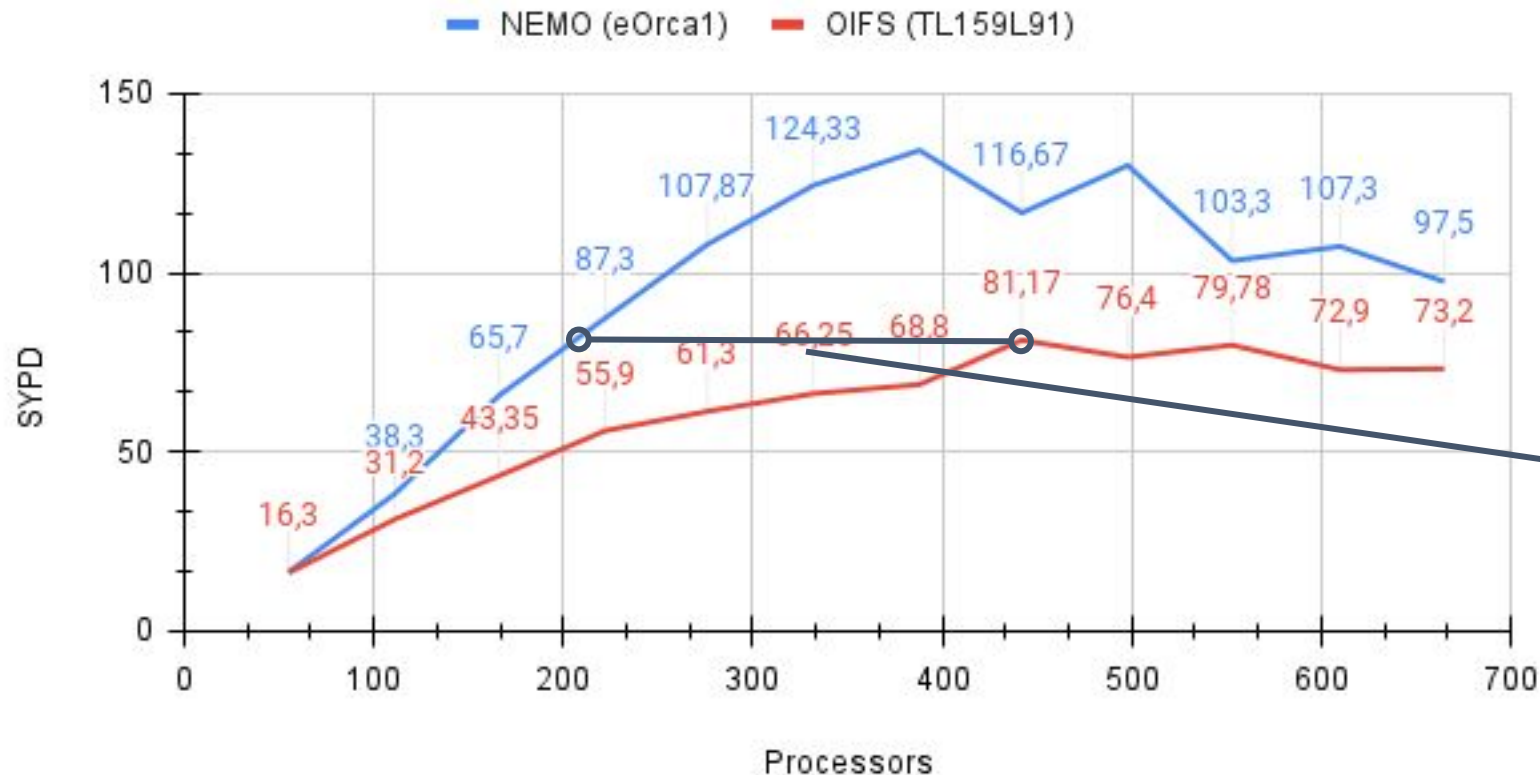
Similar OIFS & NEMO performance:

- OIFS - 444 processors ~ 81 SYPD
- NEMO - 222 processors ~ 87 SYPD

Expected Coupled SYPD: ~ 81 SYPD

EC-Earth 4 (TL159L91-eOrca1) Scalability by component

NEMO & OIFS scalability on MN5



Similar OIFS & NEMO performance:

- OIFS - 444 processors ~ 81 SYPD
- NEMO - 222 processors ~ 87 SYPD

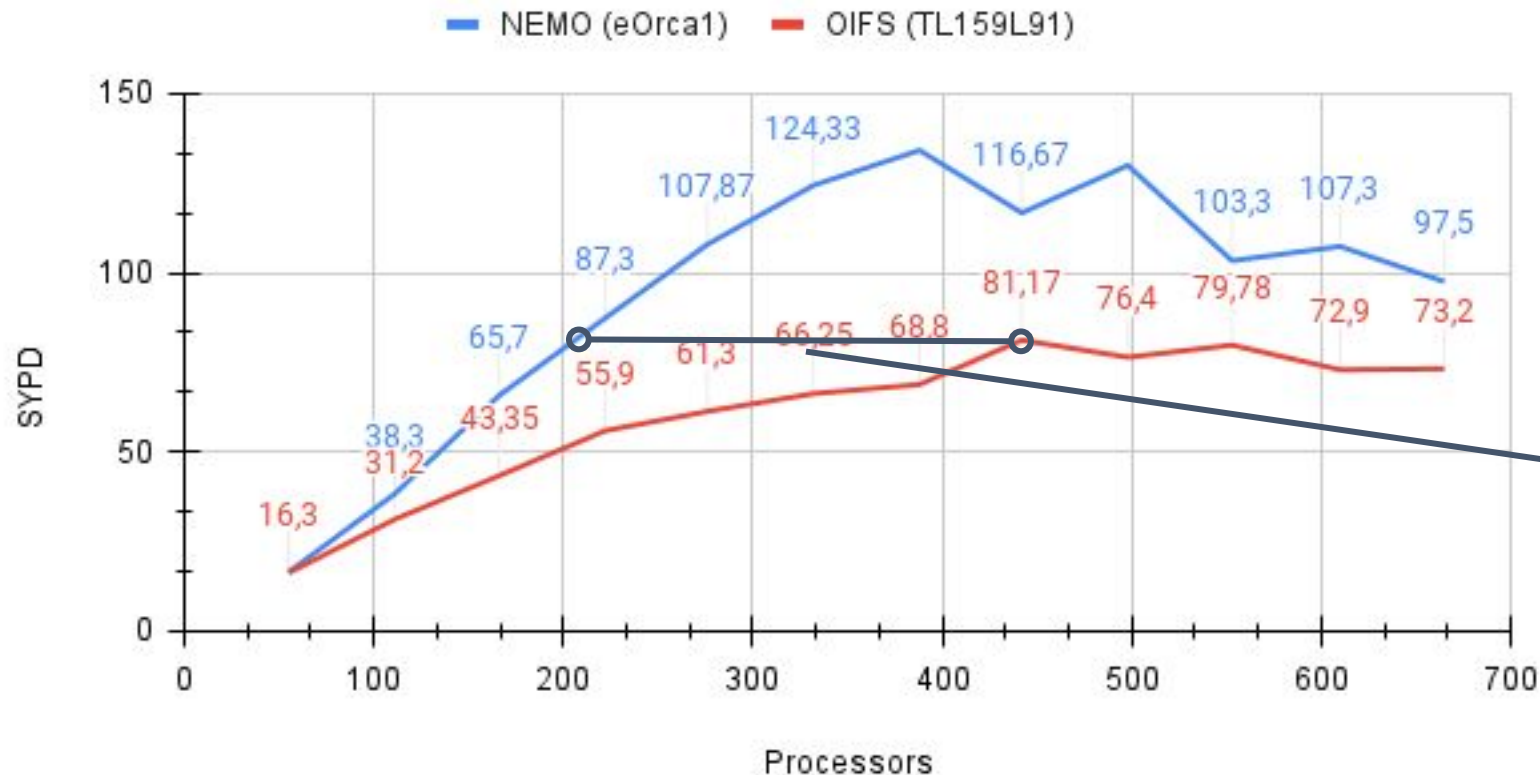
Expected Coupled SYPD: ~ 81 SYPD

MN5, results ~2 month later with same configuration:

NEMO - 88 SYPD & OIFS 73 SYPD
Obtained Coupled SYPD: ~ 71.9 SYPD

EC-Earth 4 (TL159L91-eOrca1) Scalability by component

NEMO & OIFS scalability on MN5



Similar OIFS & NEMO performance:

- OIFS - 444 processors ~ 81 SYPD
- NEMO - 222 processors ~ 87 SYPD

Expected Coupled SYPD: ~ 81 SYPD

MN5, results ~2 month later with same configuration:

NEMO - 88 SYPD & OIFS 73 SYPD
Obtained Coupled SYPD: ~ 71.9 SYPD

Why?

- Coupling cost
- MN5 state variability
- **Better balancing required**

Fine-tune components performance

run **efficiently** -> **balance multiple components**

Manual way -> back and forth runs analyzing results while changing processors for each component

Fine-tune components performance

run **efficiently** -> **balance multiple components**

Manual way -> back and forth runs analyzing results while changing processors for each component

Automatize it -> create a workflow that does it -> **Auto-LB**

Future Work

- Compare similar resolutions ECE3 vs ECE4 in MN5
- Compare MN4 vs MN5 with same ECE3 configuration
- Adapt auto-scalability and auto-lb to use them with ECE4
- Adapt reproducibility scripts to use them with ECE4



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



Thank you

Eric Ferrer - Gilbert Montané

Roc Salvador - Marcel Cases

eric.ferrer@bsc.es

18/09/2024