Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Autosubmit Workflow Management System

A versatile tool for managing Earth system models on HPC platforms

Domingo Manubens, Joan López De La Franca

BSC Earth Sciences Department

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

What

Environmental forecasting

Why

Our strength …
- … research …
- … operations …
- … services …
- … 60 people ...

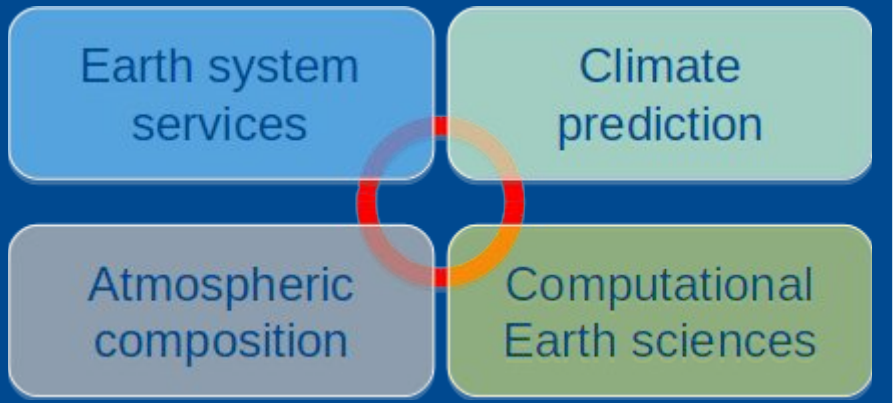
How

Develop a capability to model air quality processes from urban to global and the impacts on weather, health and ecosystems

Implement a climate prediction system for subseasonal-to-decadal climate prediction

Develop user-oriented services that favour both technology transfer and adaptation

Use cutting-edge HPC and Big Data technologies for the efficiency and user-friendliness of Earth system models



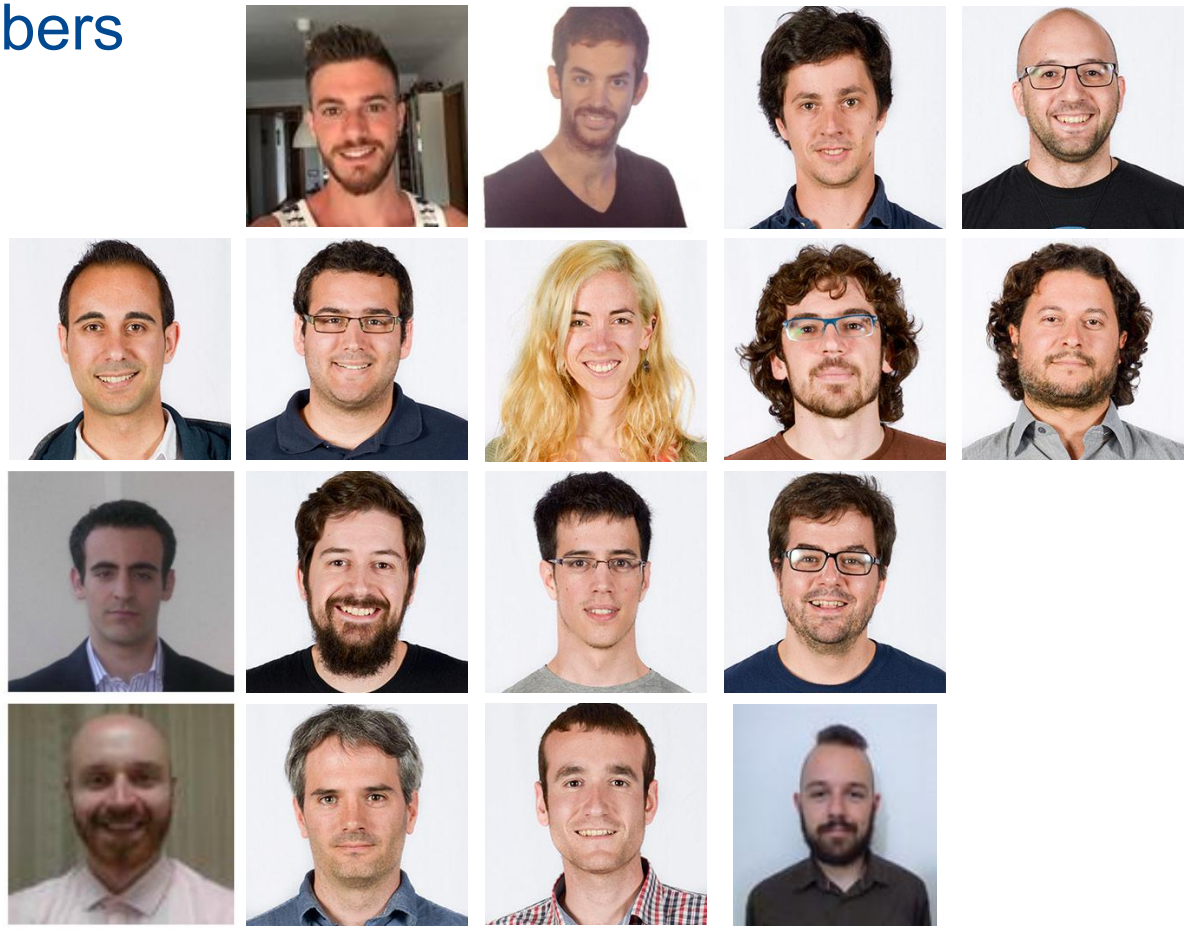
Computational Earth Sciences Group

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

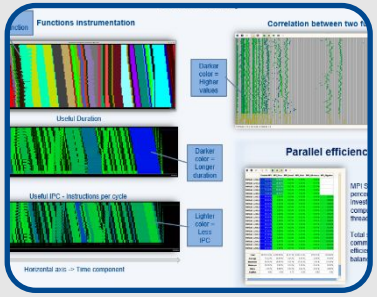
EXCELENCIA SEVERO OCHOA

- Multidisciplinary team with different IT profiles

- Currently, 17 members
  - 2 Managers
  - 10 engineers
  - 2 Postdoc
  - 1 PhD student
  - 1 Master student
  - 1 Intern



Structure



Performance Team

- Provide HPC Services such as performance analysis or optimizations for parallel computing
- Apply and develop new computational methods

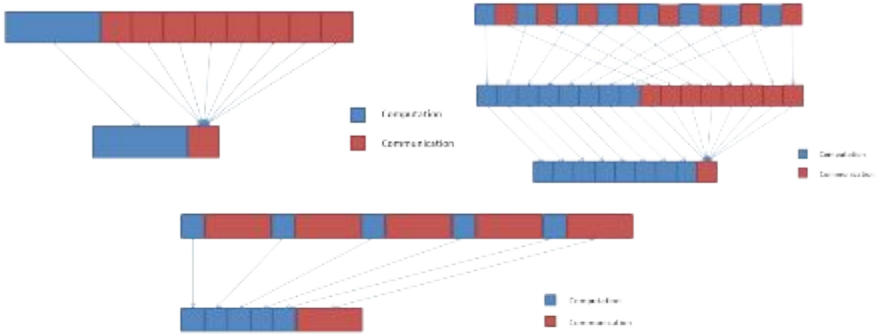
Performance activities

Collaboration with computer sciences department

BSC performance tools

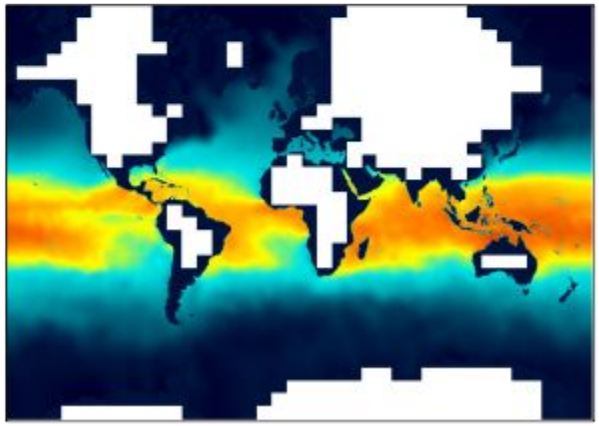
MPI communications optimizations

Reducing p2p and collective communications overhead



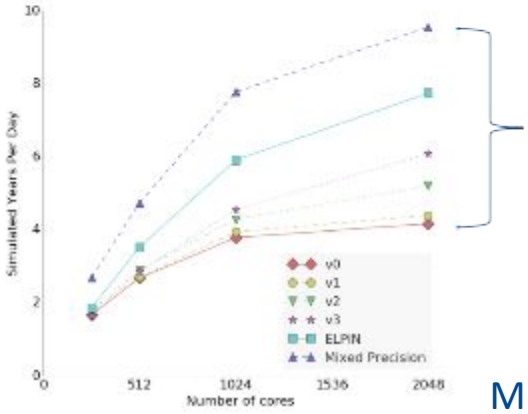
Exclude land processes in NEMO

Finding an optimal domain decomposition



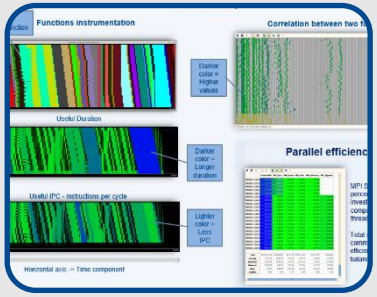
Explore mixed precision

Which precision is needed in NEMO?



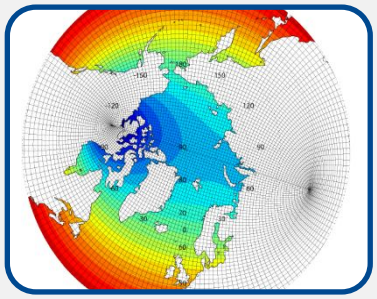
Potential overall improvement: 2x

Structure



Performance Team

- Provide HPC Services such as performance analysis or optimizations for parallel computing
- Apply and develop new computational methods

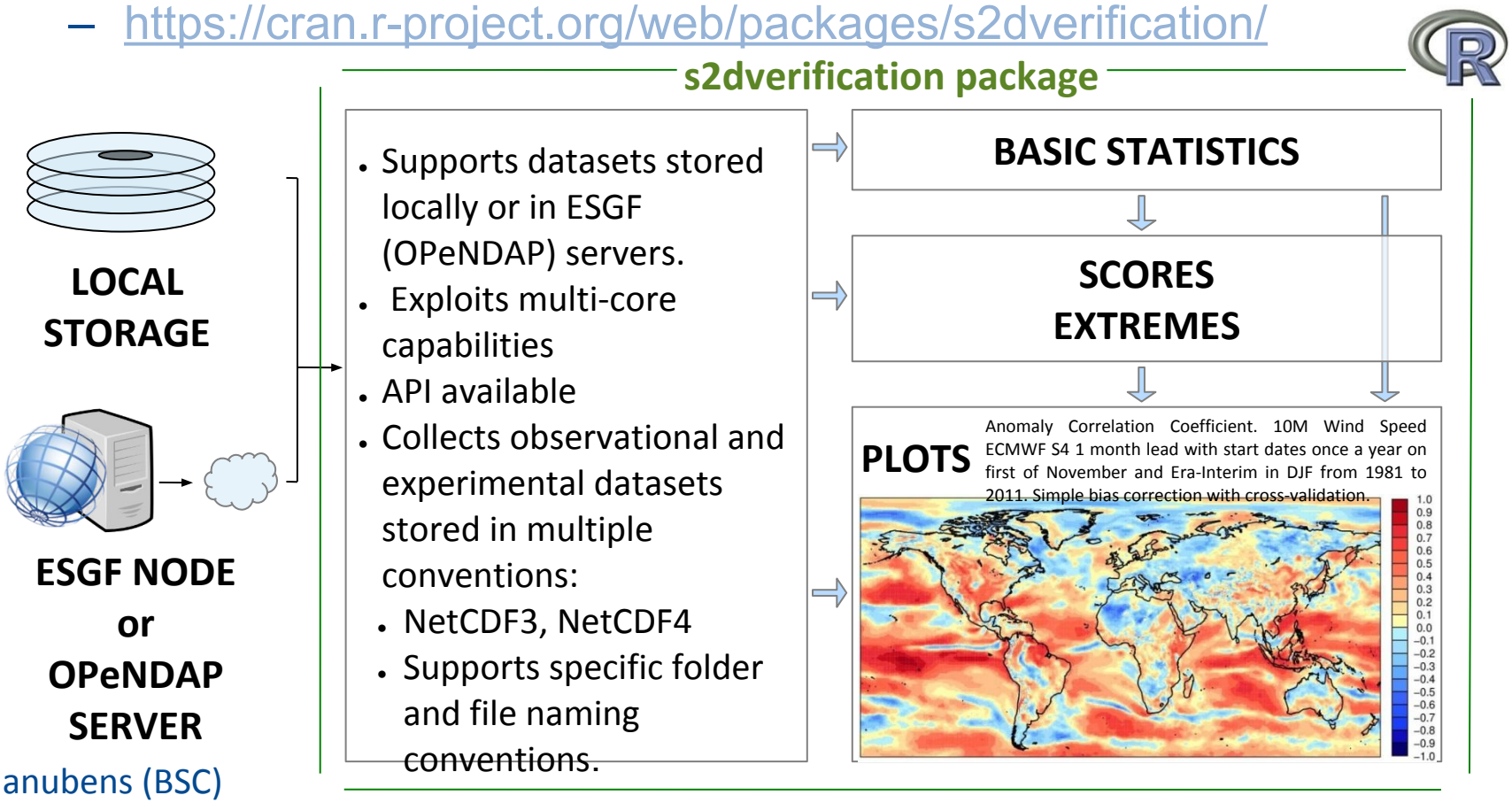


Data and Diagnostics Team

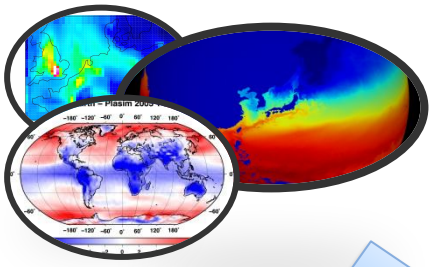
- Big Data in Earth Sciences
- Provision of data services
- Visualization

Verification and visualisation tools

- s2dverification
  - Set of tools to verify forecasts through the computation of typical prediction scores against one or more observational datasets or reanalyses
  - https://cran.r-project.org/web/packages/s2dverification/



Online diagnostics for climate models



Climate Models

Current approach: offline diagnostics

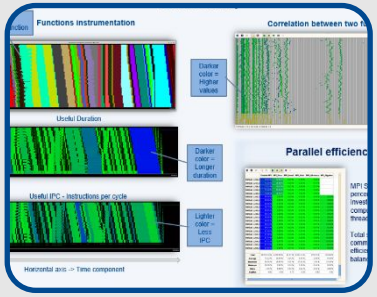
HPC → Model outputs → Serial node → **Diagnostics** ← Users need diagnostics for their analyses

Future approach: online diagnostics

HPC → Model outputs + **Diagnostics** ← Users need diagnostics for their analyses

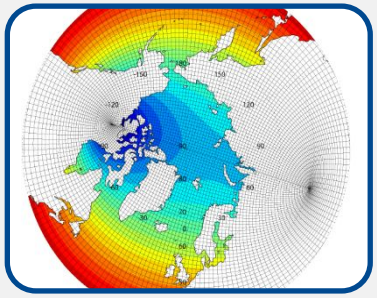
- Diagnostics computed as Analytics as a Service
  - Diagnostics online (during model run)
  - Reduced data traffic
  - Diagnostics possible on the computing nodes
  - New diagnostics (data mining of extremes) possible
  - The user gets the results faster, crucial to adapt to climate change and to develop climate services (public and private)

Structure



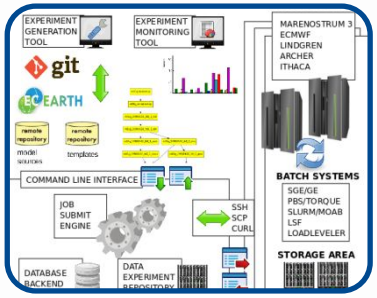
Performance Team

- Provide HPC Services such as performance analysis or optimizations for parallel computing
- Apply and develop new computational methods



Data and Diagnostics Team

- Big Data in Earth Sciences
- Provision of data services
- Visualization



Models and Workflows Team

- Development of HPC user-friendly software framework
- Support the development of atmospheric research software

Data and computing infrastructures for climate modelling

Barcelona Supercomputing Center
Centro Nacional de Supercomputación
EXCELENCIA SEVERO OCHOA

enes EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING
http://enes.org/

IS-ENES : Infrastructure for ENES
FP7 project « Integrating Activities »

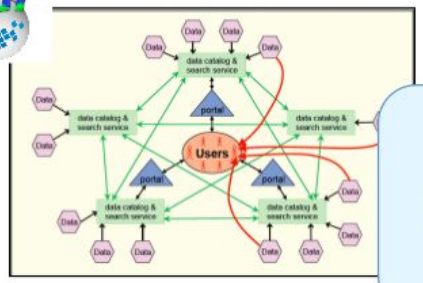
is-enes INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING
http://is.enes.org/

1rst phase: March 2009- Feb 2013, 18 partners
2nd phase: Apr 2013- March 2017, 23 partners

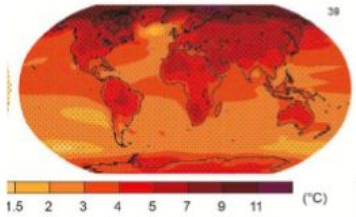
Climate models
Environment software tools
ESM ca 1000 man years

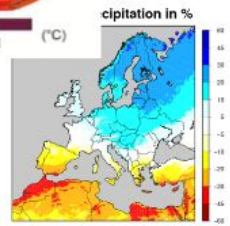
High-performance computers
& storage facilities

ESGF Earth System Grid Federation



Data & metadata
Standards
Distributed database ESGF
10 000 registered users worldwide
Open access





Climate research & Impact research
Climate services

Support WCRP international experiments
Used in IPCC Assessments Reports

WCRP World Climate Research Programme

www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

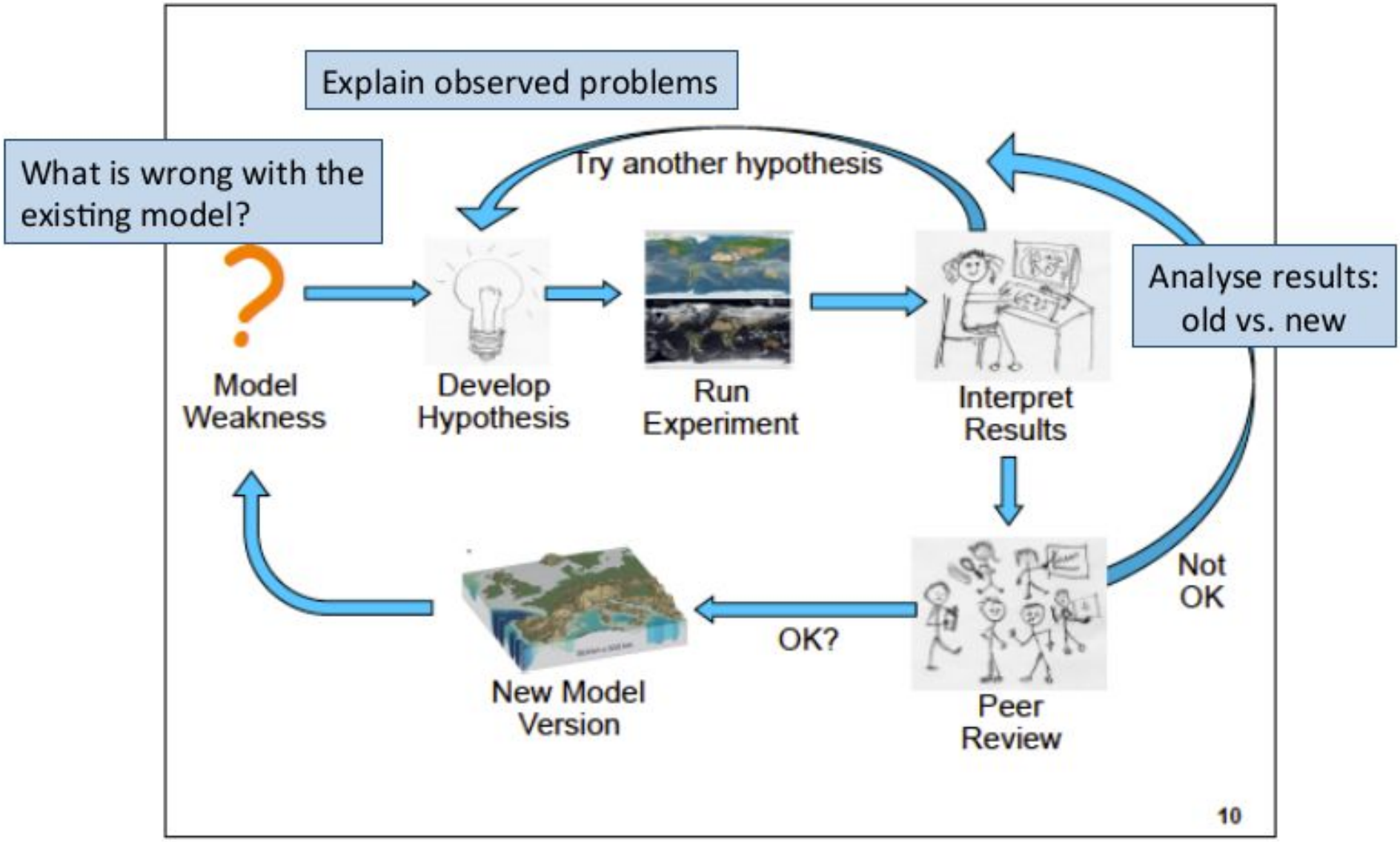
EXCELENCIA SEVERO OCHOA

Models and workflows



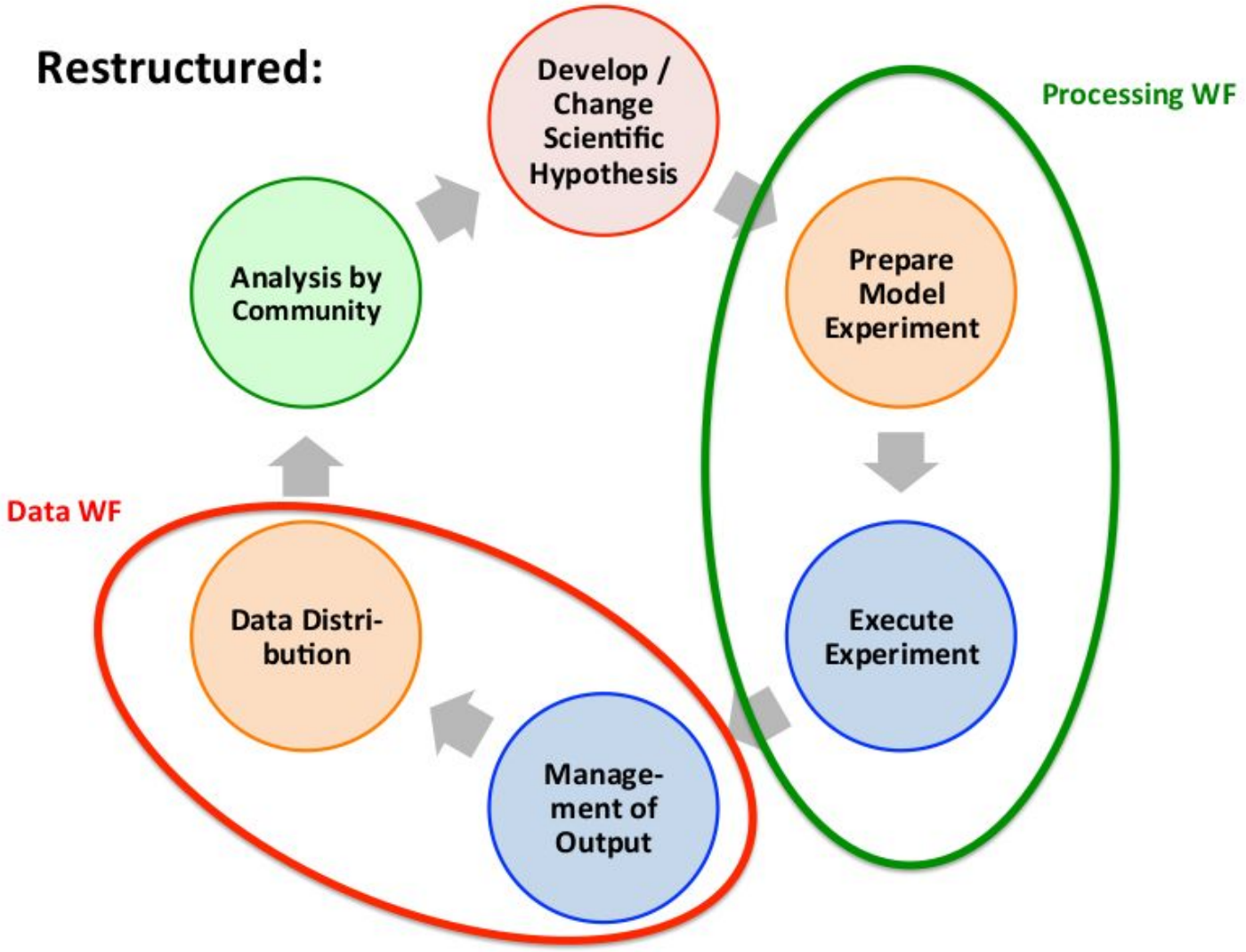

YouTube

Why are we doing research on workflows ?

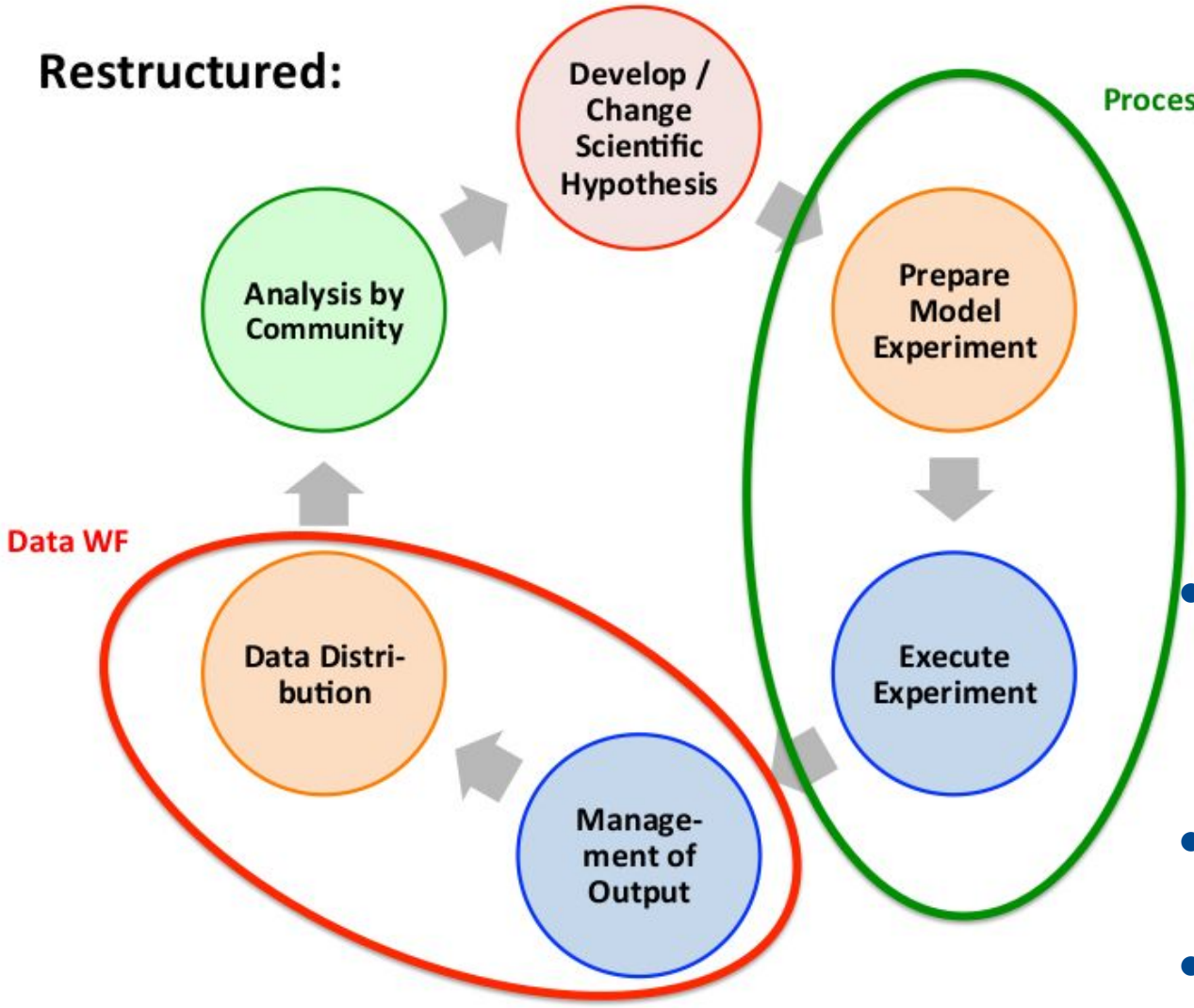


S. Easterbrook, from 1st WS on WF

Processing and data workflows



Processing and data workflows



- What files or environment it needs to run ?
- What it produces ?
- When it might need to run ?

Multiple High Performance Computing infrastructures

Computing resources funded by: National / EU / International projects

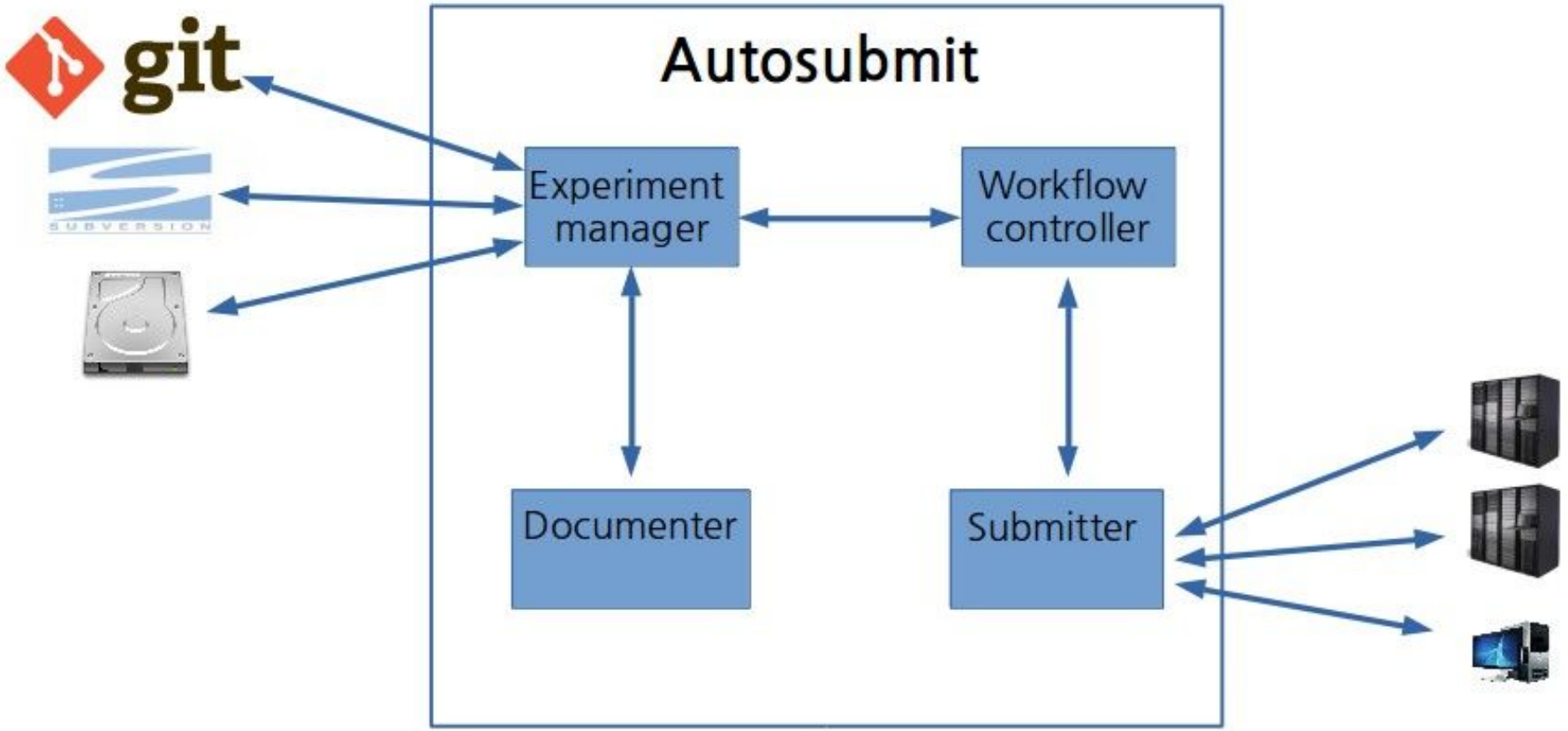
Workflow Management Systems for weather and climate

Criteria	Autosubmit	Cylc	ecFlow
Seniority	2011	2010	2011
Original authors/sponsors	IC3, BSC	NIWA, Met Office	ECMWF
License	GNU GPL v3	GNU GPL v3	Apache License v2.0

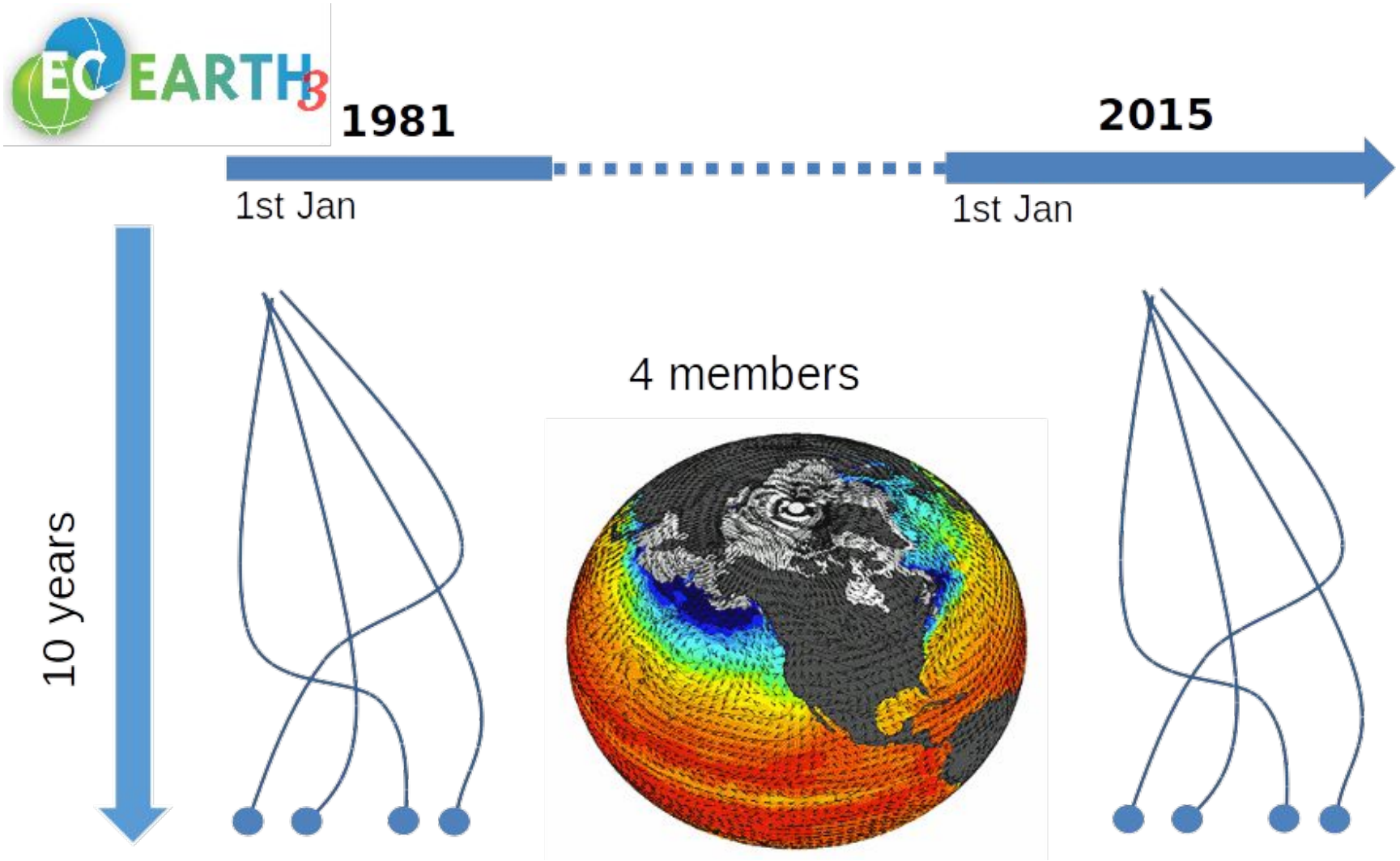
Report available here: https://goo.gl/IKqIjV

D. Manubens-Gil, J. Vegas-Regidor, C. Prodhomme, O. Mula-Valls and F. J. Doblas-Reyes, "Seamless management of ensemble climate prediction experiments on HPC platforms", 2016 International Conference on High Performance Computing & Simulation (HPCS), Innsbruck, 2016, pp. 895-900. doi: 10.1109/HPCSim.2016.7568429

Autosubmit in a nutshell ...

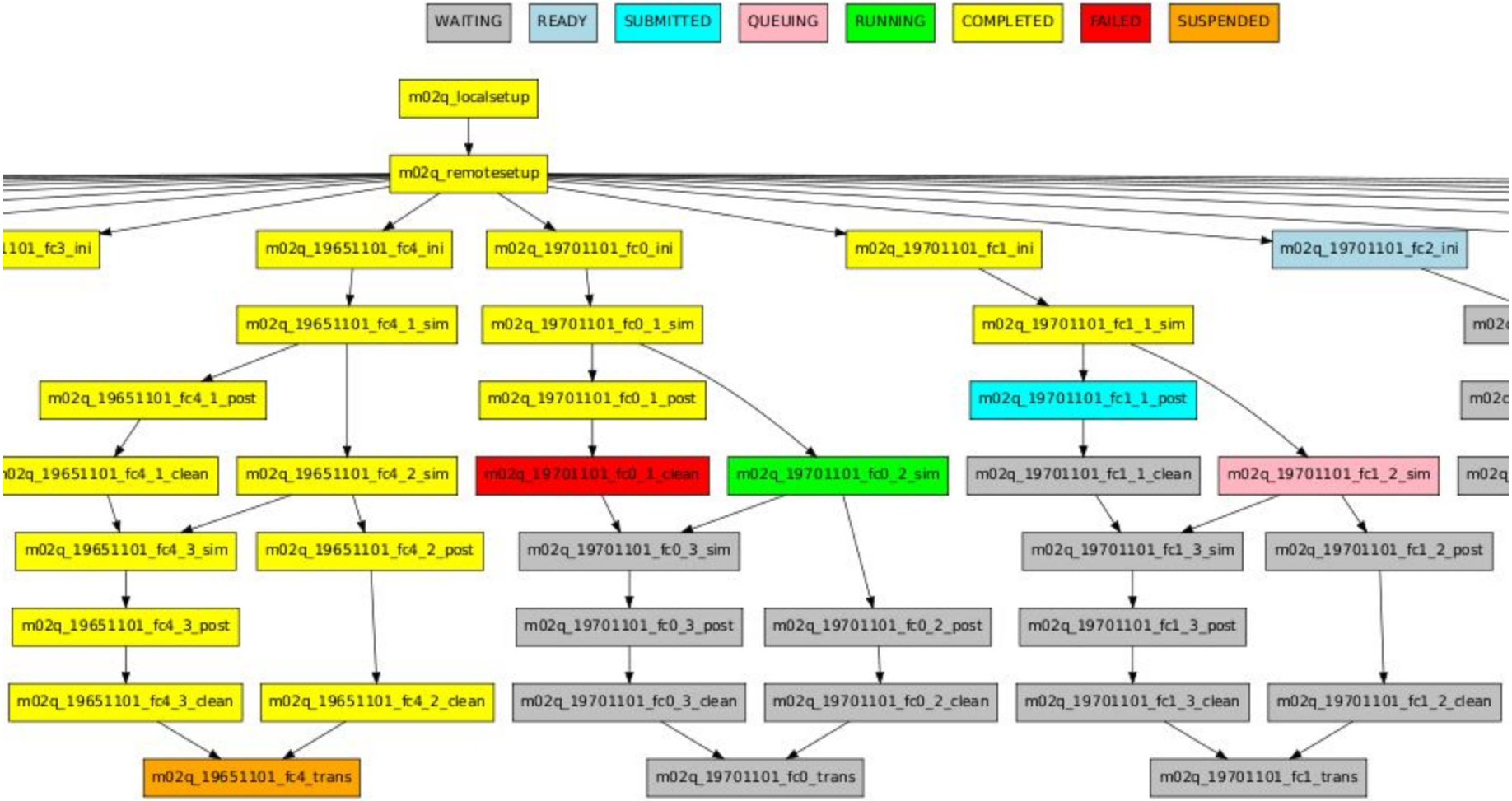


Multi-member climate experiment



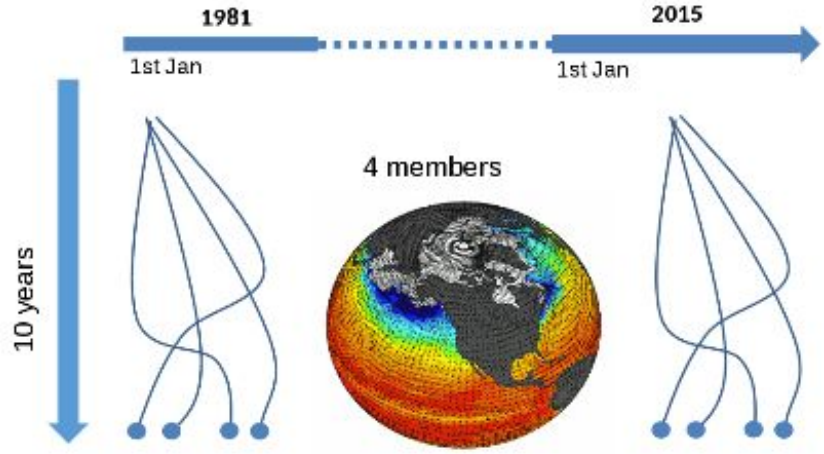
Computing cost: ~ 12000 CHPSY

Autosubmit monitoring

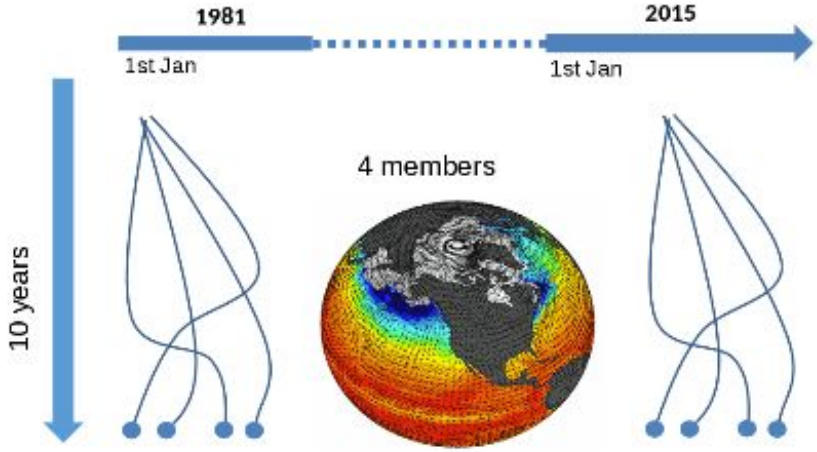


Multi-model multi-member climate experiment workflow

EC-EARTH3



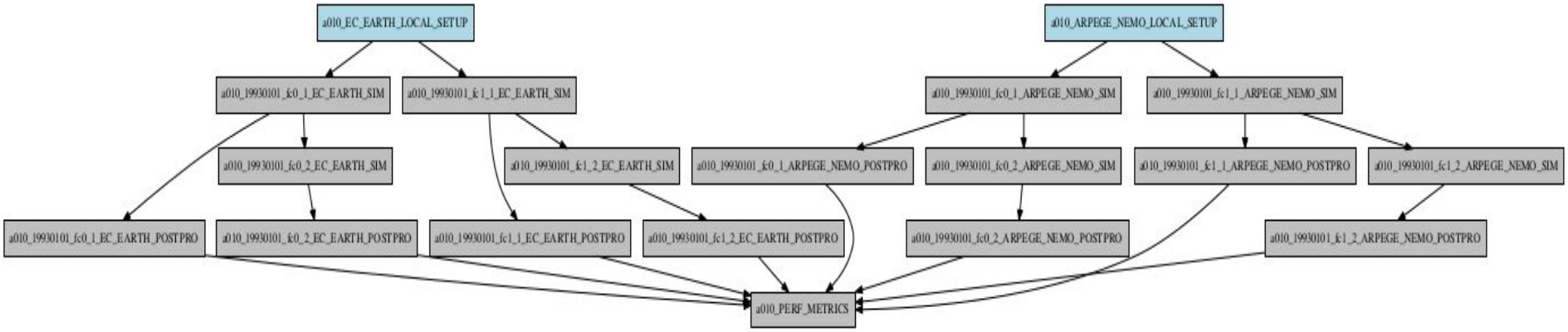
CNRM UMR3589

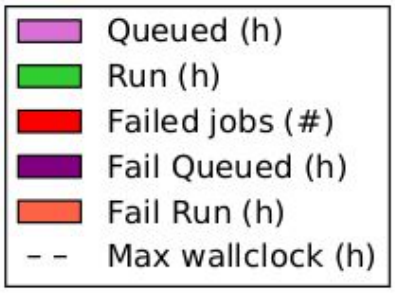


Report available here: https://goo.gl/TNv3uJ

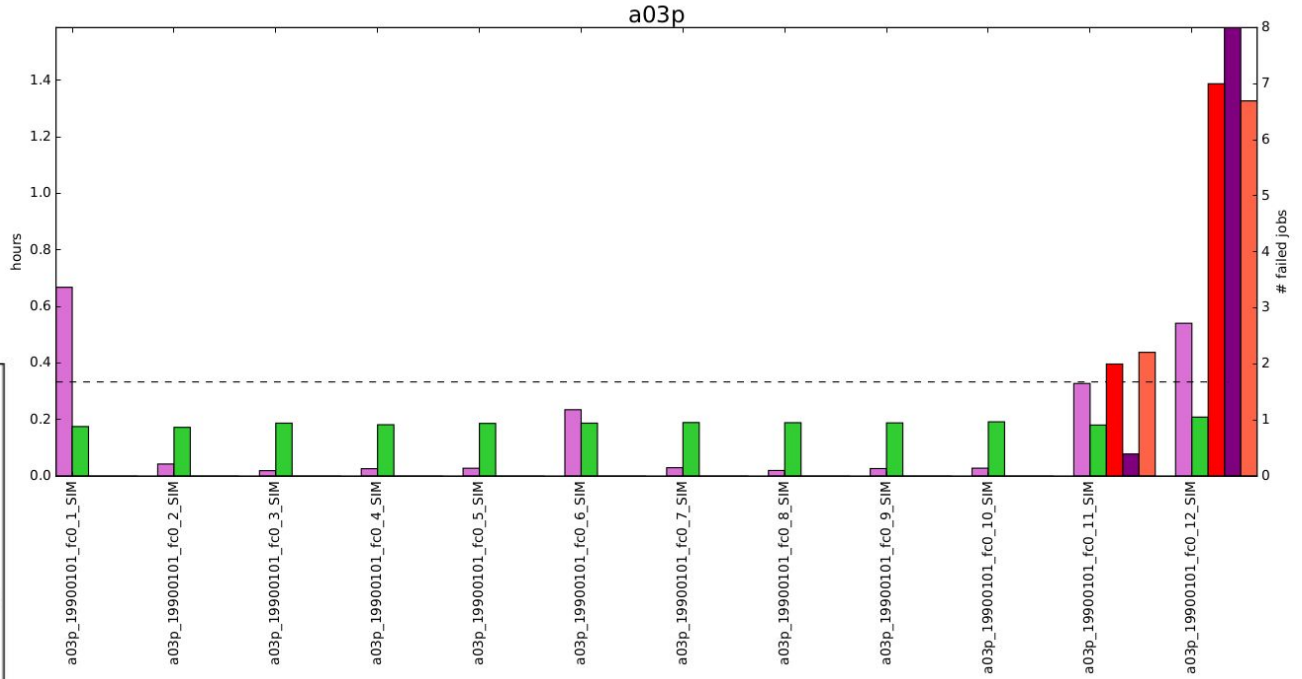
Experiment run from BSC, CERFACS, SMHI on MareNostrum 3

Autosubmit monitoring





Period: None ~ 2016-01-29 20:16:00
Submitted (#): 93
Run (#): 93
Failed (#): 26
Completed (#): 67
Expected consumption real (h): 40.0
Expected consumption CPU time (h): 24400.0
Consumption real (h): 23.13
Consumption CPU time (h): 14107.61
Consumption (%): 57.82



www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Prepare an experiment step-by-step with Autosubmit




YouTube

Example

ftp://autosubmit:sgH_123jHS$Q@bscesftp.bsc.es/Autosubmit_EC-Earth_Hands_On.ova

1- Configure Autosubmit
2- Run 2 months simulation (199311-199312, hindcast)
3- Post-process
4- Transfer results locally

Configure Autosubmit

1. **Replace** your user-ID in the **~/.ssh/config** in the following lines:

```
Host mn3
      HostName mn1.bsc.es
      User <mn3 user-ID>
      IdentityFile ~/.ssh/id_rsa
```

2. Now check you can login to MN3 without password. To do it, run the next command:

```
$ ssh mn3
```

3. **First step: create EC-Earth experiment**

```
$ autosubmit expid -y a000 -H marenostrum3 -d "Test for EC-Earth & Autosubmit training"
```

Now you can check that in **~/autosubmit** there is a new folder named **'a002'** that will be the one that will store your experiment configuration and monitoring files.

4. **Configure platform.** Y**ou have to edit** the file: **~/autosubmit/a002/conf/platforms_a002.conf** with your prefered text editor. You will have to use this configuration:

```
[marenostrum3]
TYPE = LSF
HOST = mn1.bsc.es
PROJECT = nct01
BUDGET = nct01:ESE
USER = user-here
QUEUE = training
SCRATCH_DIR = /gpfs/scratch
PROCESSORS_PER_NODE = 16
```

Configure Autosubmit

5. **Configure jobs.** Now you need to configure the workflow. To do that, edit the file: **conf/jobs_a002.conf**.

```
[LOCAL_SETUP]
FILE = runtime/autosubmit/copy-runtime.sh
PLATFORM = LOCAL

[SIM]
FILE = runtime/autosubmit/ece-ifs+nemo.sh
DEPENDENCIES = LOCAL_SETUP SIM-1
RUNNING = chunk
WALLCLOCK = 01:00
PROCESSORS = 130
CHECK = False

[POST_ATM]
FILE = runtime/autosubmit/post-ifs-cmor.sh
DEPENDENCIES = SIM
RUNNING = chunk
WALLCLOCK = 00:20
PROCESSORS = 1
```

```
[POST_OCE]
FILE = runtime/autosubmit/post-nemo-cmor.sh
DEPENDENCIES = SIM
RUNNING = chunk
WALLCLOCK = 00:10
PROCESSORS = 1

[LOCAL_TRANSFER]
FILE = runtime/autosubmit/transfer.sh
PLATFORM = LOCAL
DEPENDENCIES = POST_ATM POST_OCE
RUNNING = member
```

Create experiment

6. Autosubmit **create.** This command prepares the experiment to run.

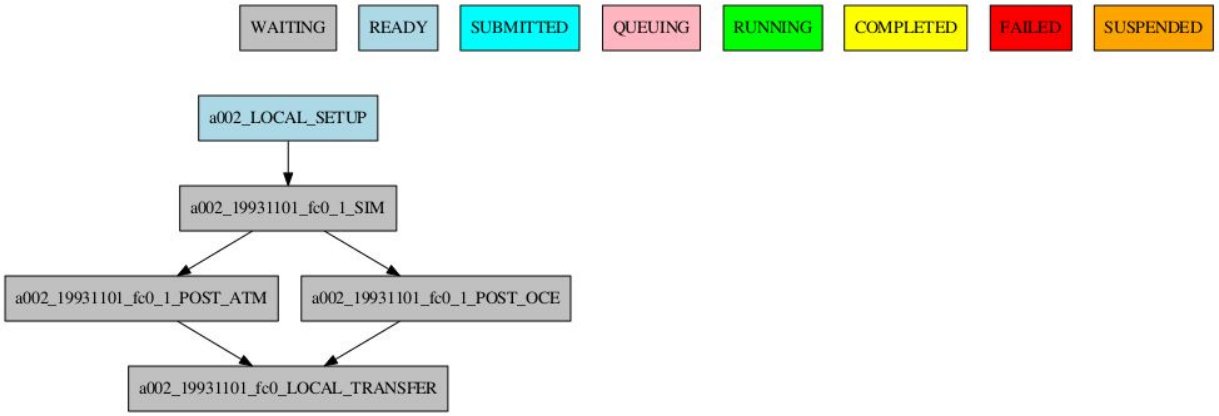
```
$ autosubmit create a002
```

If everything has worked well, you'll see something like this:

```
autosubmit@autosubmit:~/autosubmit/a000$ autosubmit create a000
Preparing .lock file to avoid multiple instances with same expid.

Checking configuration files...
autosubmit_a000.conf OK
platforms_a000.conf OK
jobs_a000.conf OK
expdef_a000.conf  OK
Configuration files OK

Loading parameters...

Creating joblist...
Creating jobs...
Adding dependencies...
Removing redundant dependencies...

Saving joblist...

Ploting joblist...
Plotting...
Plot created at /home/autosubmit/autosubmit/a000/plot/a000_20160912_0846.pdf

Job list created succesfully
Remember to MODIFY the MODEL config files!
```

Run experiment

Then you will see the tree of your experiment's jobs in PDF format:



7. **Autosubmit run.** Now the experiment will start:

```
$ autosubmit run a002
```

```
Checking configuration files...
autosubmit_a000.conf OK
platforms_a000.conf OK
jobs_a000.conf OK
expdef_a000.conf  OK
Configuration files OK

Starting job submission...
Creating jobs...
Loading JobList
Adding dependencies...
Removing redundant dependencies...
Checking scripts...
Scripts OK


2 of 2 jobs remaining (08:50)
```

www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Techniques to improve the experiment throughput with Autosubmit

Why are we doing research on this ?

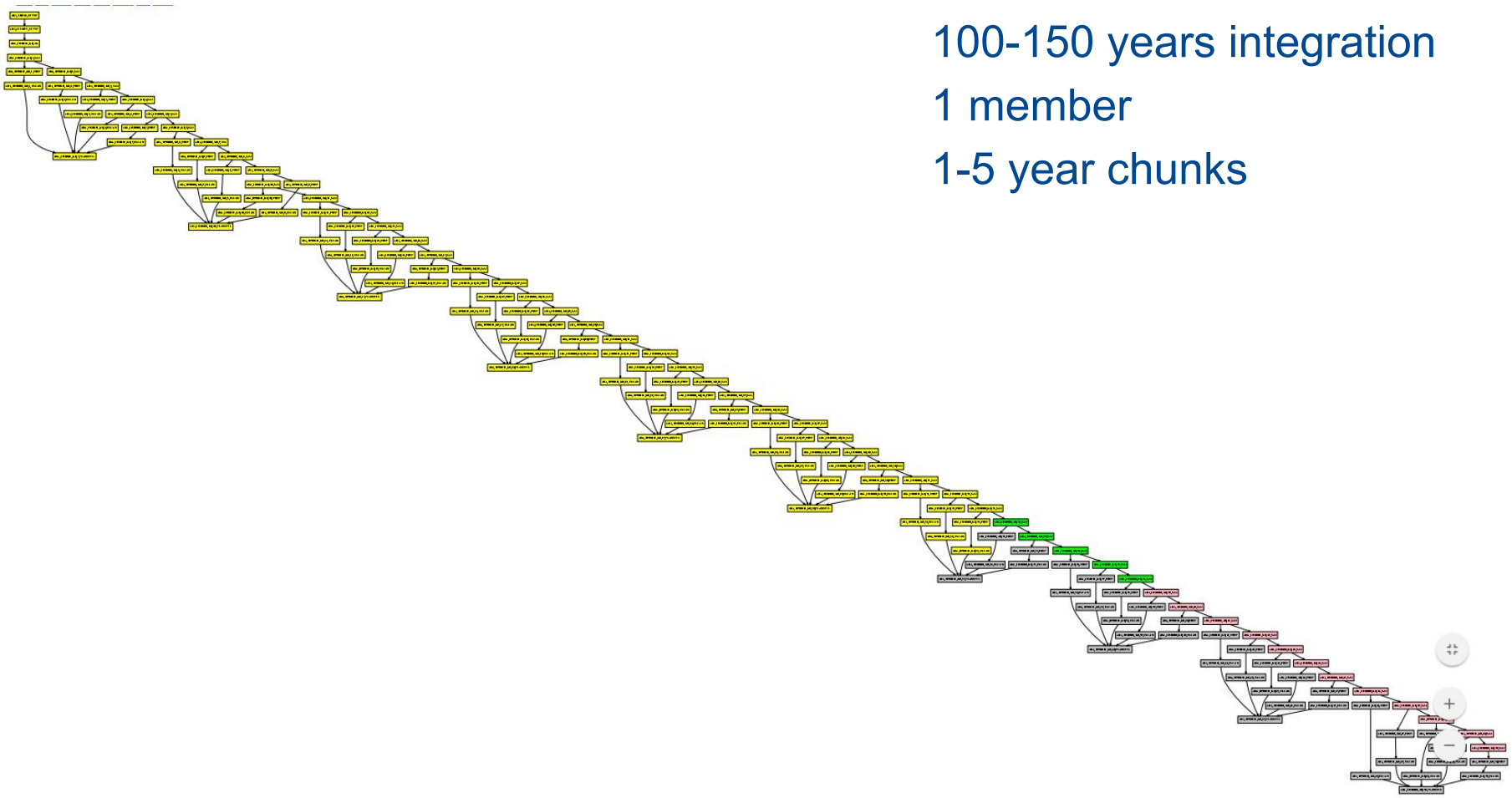
- Waiting time on queues reduces too much the throughput in some situations. We need to be faster !

Why are we doing research on this ?

- Waiting time on queues reduces too much the throughput in some situations. We need to be faster !

  A) Long integration of dependent jobs

Why are we doing research on this ?



100-150 years integration

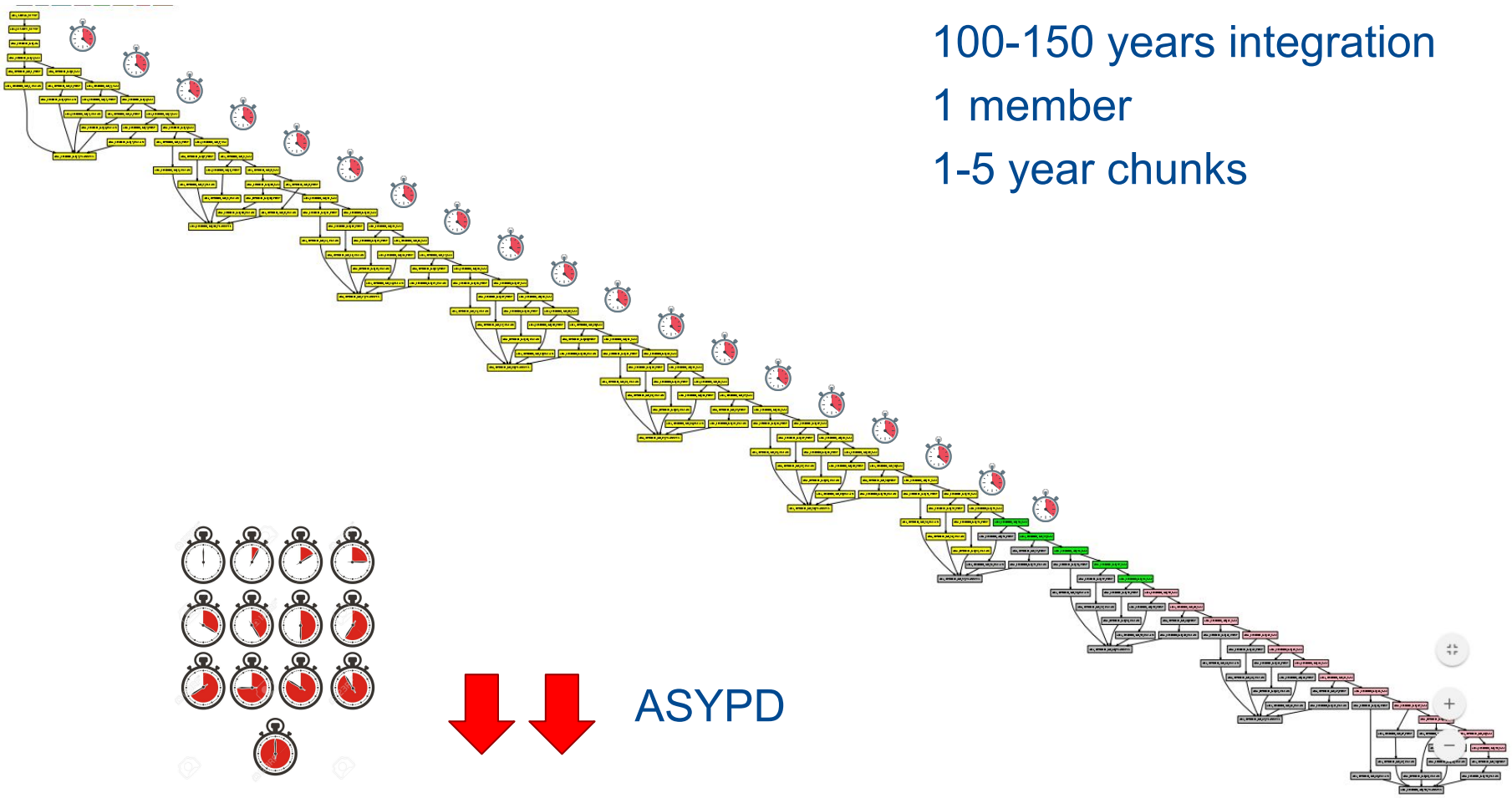
1 member

1-5 year chunks

E. Exarchou

E. Toruigny

Why are we doing research on this ?



100-150 years integration

1 member

1-5 year chunks

ASYPD

Why are we doing research on this ?

- Waiting time on queues reduces too much the throughput in some situations. We need to be faster !

  A) Long integration of dependent jobs

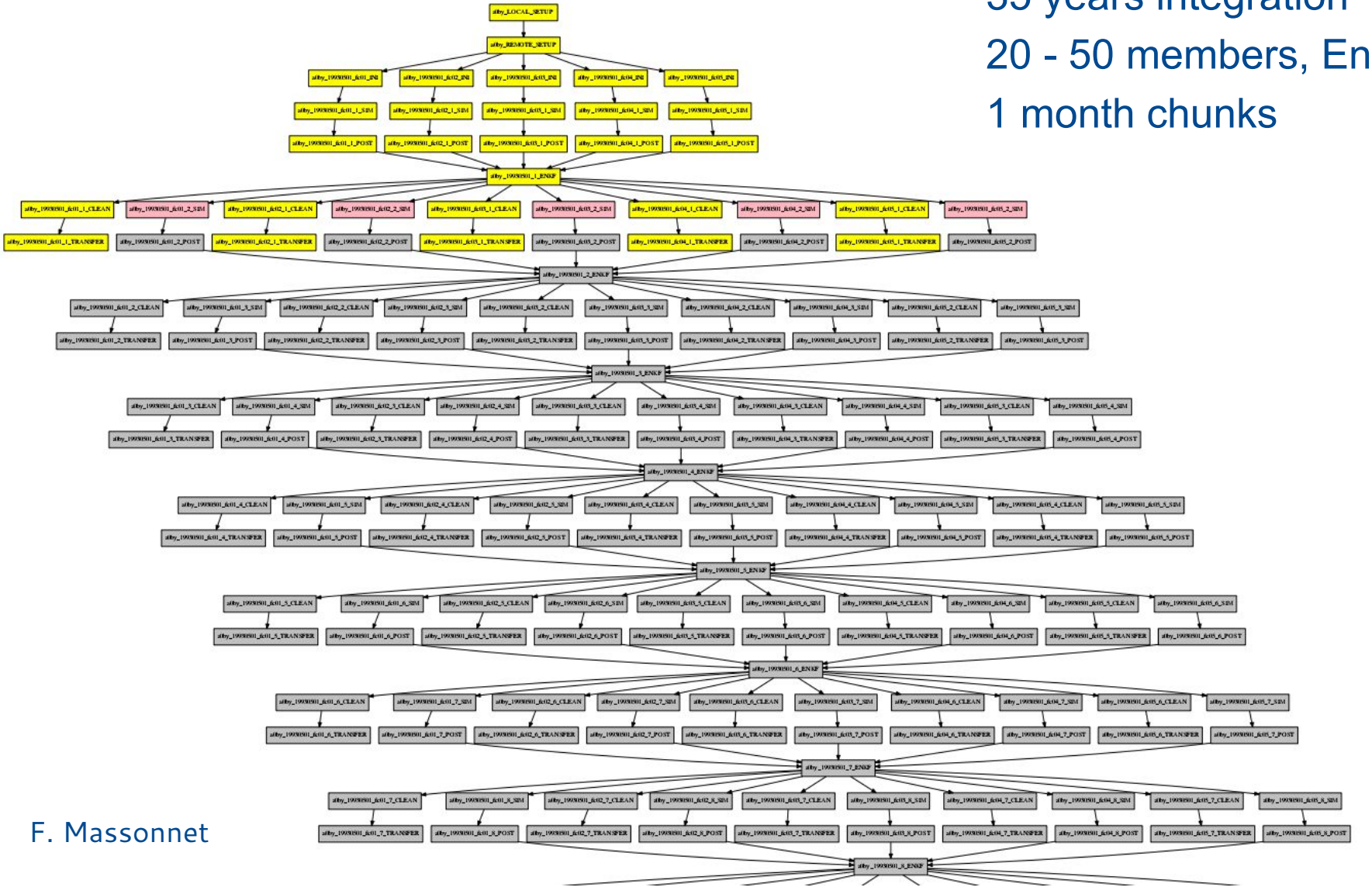
  B) Large ensemble of small independent jobs
  that need to synchronize with data assimilation job

Why are we doing research on this ?

35 years integration

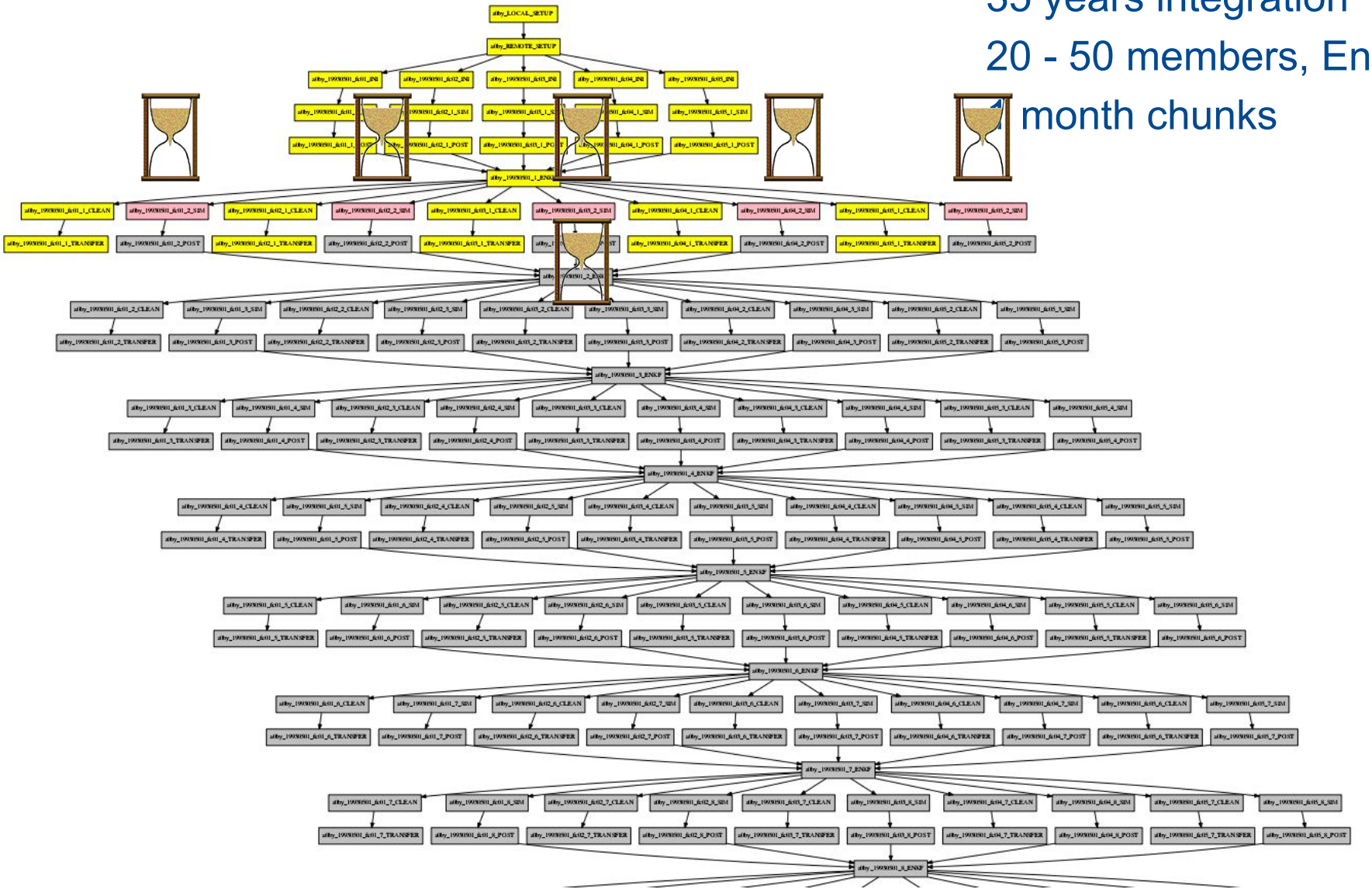
20 - 50 members, EnKF

1 month chunks



F. Massonnet

Why are we doing research on this ?



35 years integration

20 - 50 members, EnKF

month chunks

Why are we doing research on this ?

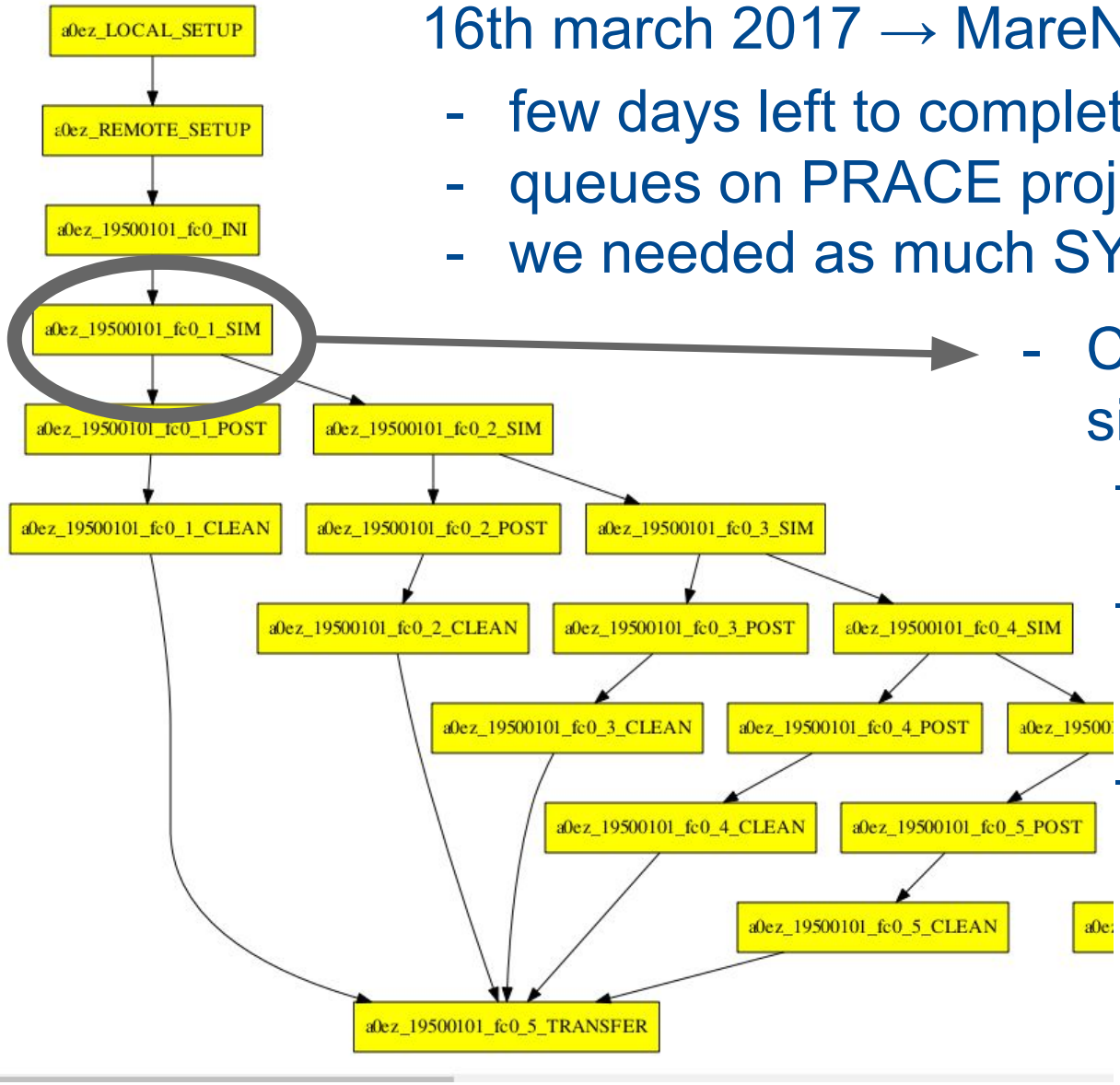
- Waiting time on queues reduces too much the throughput in some situations. We need to be faster !

  A) Long integration of dependent jobs

  B) Large ensemble of small independent jobs
  that need to synchronize with data assimilation job

- Optional functionality to improve throughput using Autosubmit (adjusting parameters is enough)

Case A) MareNostrum3 - a0ez

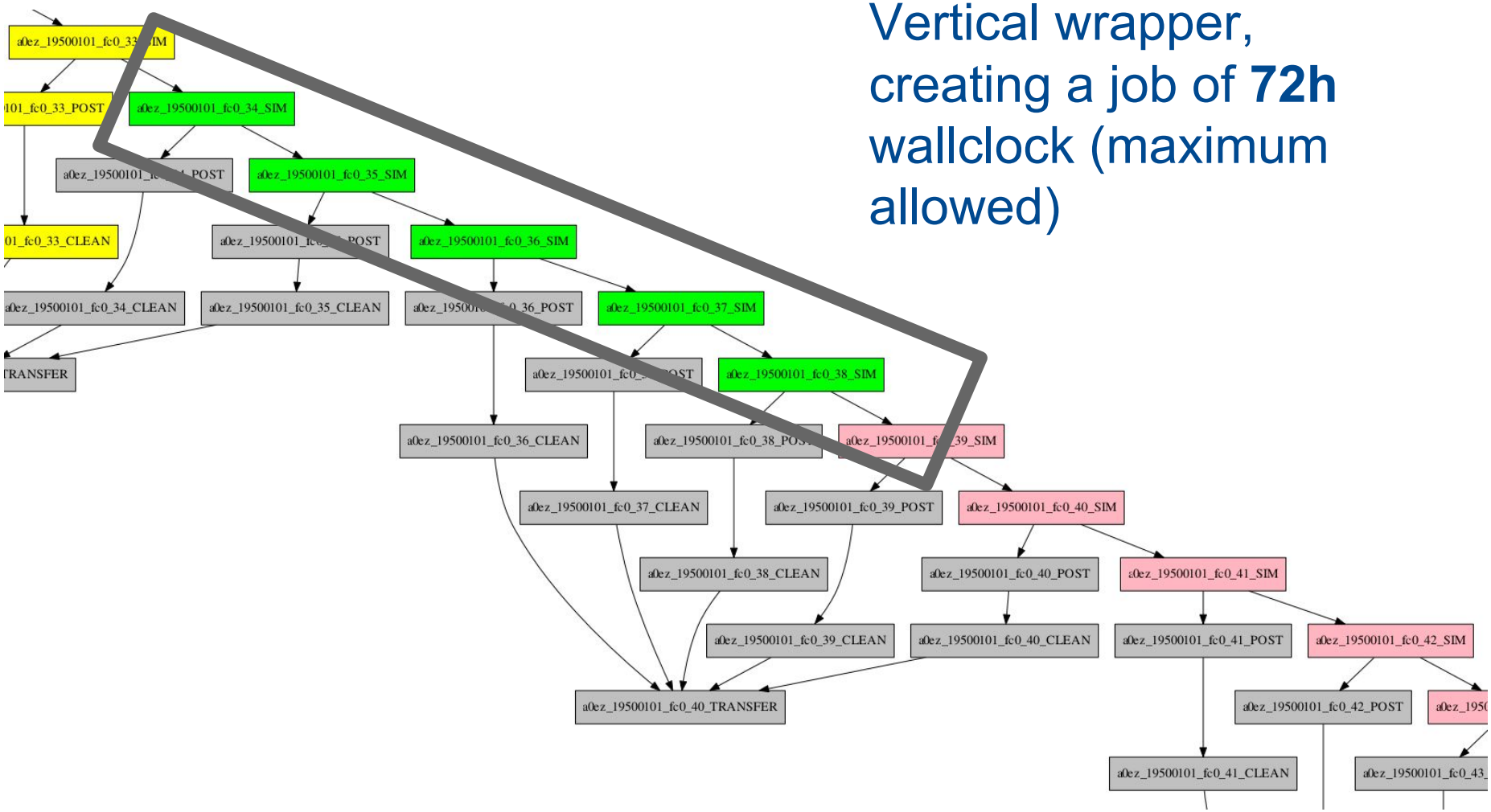


16th march 2017 → MareNostrum3 shutdown !

- few days left to complete EC-Earth spin-up HR
- queues on PRACE project full
- we needed as much SYPD as we could.

- Configuration for 1 simulation job:
  - Number of proc = **2032** (high resolution)
  - Size of chunk = **1 year** (we need restarts for each year)
  - Wallclock time = **14h** (there is room for more!)

Case A) - Vertical wrapper

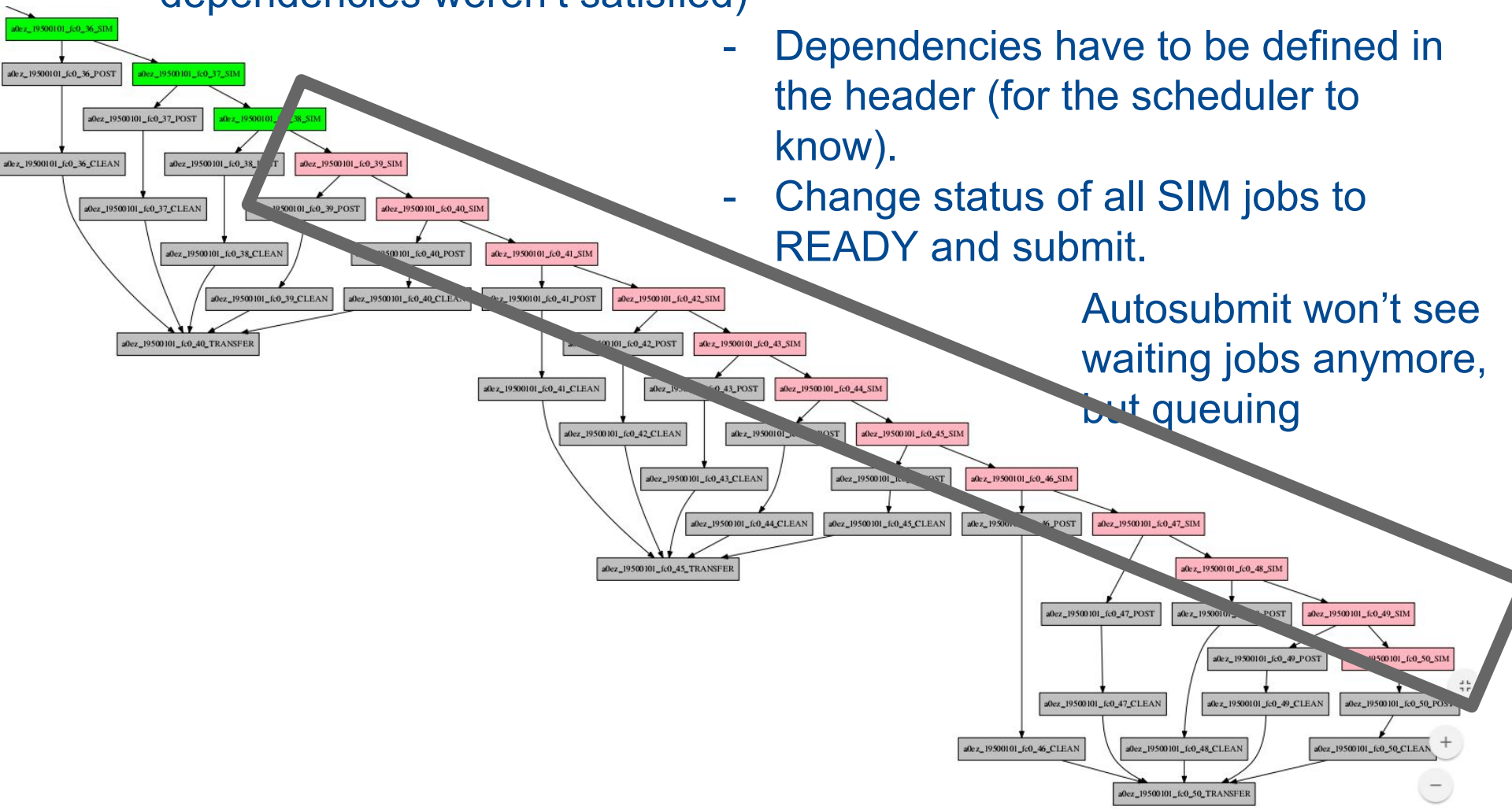


Vertical wrapper, creating a job of **72h** wallclock (maximum allowed)

Case A) - Vertical wrapper

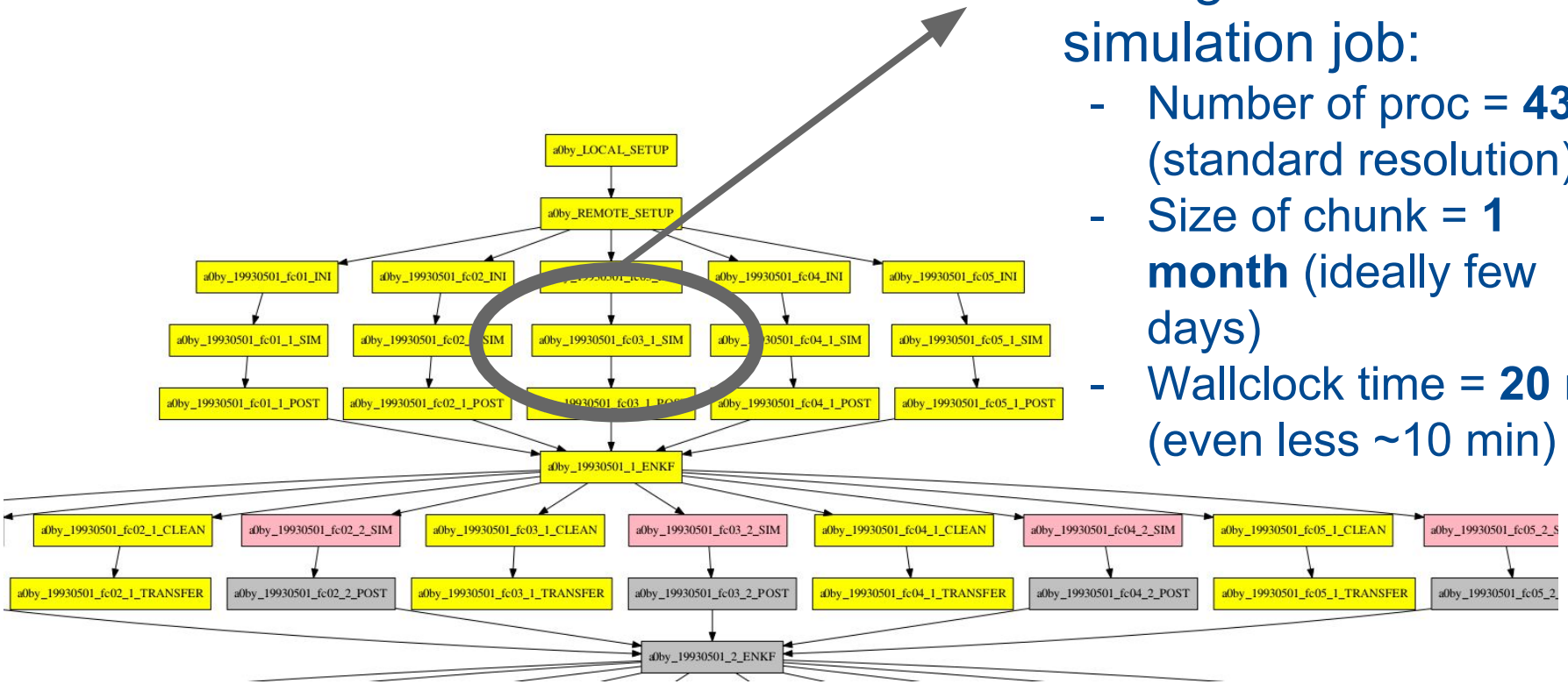
- Additionally: The sooner a job is submitted the better
- More priority for jobs having been for long time in queue (even if dependencies weren't satisfied)
  - Dependencies have to be defined in the header (for the scheduler to know).
  - Change status of all SIM jobs to READY and submit.

Autosubmit won't see waiting jobs anymore, but queuing



Case B) MareNostrum3 - a0by

24 independent members, SIM, at ORCA1, EC-Earth 3.2, it takes between 2 hours and 5 hours of physical time (depending on load on machine) to complete a full SIM-EnKF cycle. Expect at least 3 physical days to finish one year (0.3 ASYPD).

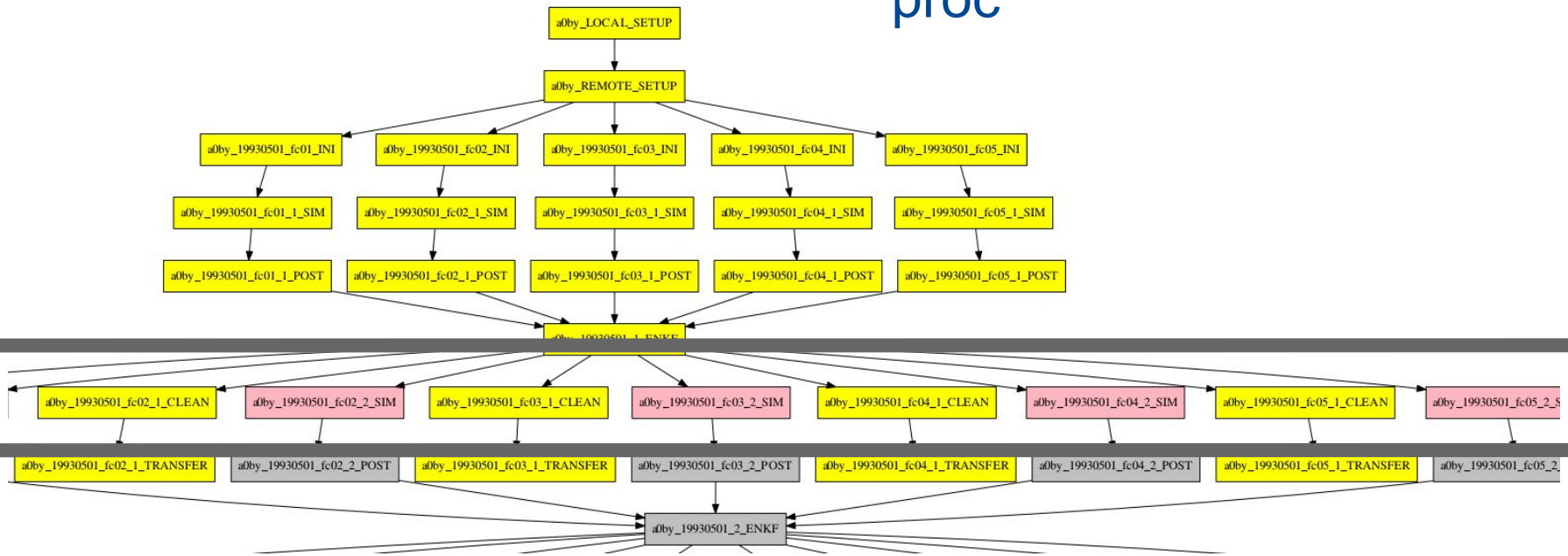


- Configuration for 1 simulation job:
  - Number of proc = **432** (standard resolution)
  - Size of chunk = **1 month** (ideally few days)
  - Wallclock time = **20 min** (even less ~10 min)

Case B) - Horizontal wrapper

Barcelona Supercomputing Center
Centro Nacional de Supercomputación
EXCELENCIA SEVERO OCHOA

Horizontal wrapper, creating a job of **2160** proc



Summary: advantages of Autosubmit

- **Automatisation:** Preparing and running the model(s), postprocessing, online CMORization and transferring the outputs managed by Autosubmit. No user intervention needed.
- **Provenance:** Assigns unique identifiers for each experiment and stores information about model version, its configuration options and computing facilities used in overall process.
- **Failure tolerance*:*** Automatic retrials and ability to rerun chunks in case of corrupted or missing data, repeating postprocessing and transfers if needed.
- **Versatility:** Currently used to run EC-Earth, NEMO and NMMB models on several platforms.
- **Throughput:** Packing simulations into a single executable to reduce queuing time. Paper in preparation.

Q & A

Get involved or contact us:	
Autosubmit GitLab:	https://earth.bsc.es/gitlab/es/autosubmit
Autosubmit Mailing List:	autosubmit@bsc.es

https://pypi.python.org/pypi/autosubmit

http://autosubmit.readthedocs.io

Download
autosubmit-3.8.1.tar.gz

www.bsc.es

Thank you!

For further information please contact

domingo.manubens@bsc.es