www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Seamless Management of Ensemble Climate Prediction Experiments on HPC Platforms

Domingo Manubens-Gil – Javier Vegas-Regidor – Chloé Prodhomme – Oriol Mula-Valls – Francisco J. Doblas-Reyes

HPC-WCES'16 - Innsbruck
21/7/2016

Introduction

What is the problem ?

PRIMAVERA

University of Reading

UNIVERSITY OF OXFORD

UNIVERSITY OF LEEDS

ECMWF

Science & Technology Facilities Council

Stockholms universitet

SMHI

predictia INTELLIGENT DATA SOLUTIONS

UCL Université catholique de Louvain

Max-Planck-Institut für Meteorologie

NATURAL ENVIRONMENT RESEARCH COUNCIL

Koninklijk Nederlands Meteorologisch Instituut Ministerie van Infrastructuur en Milieu

Met Office

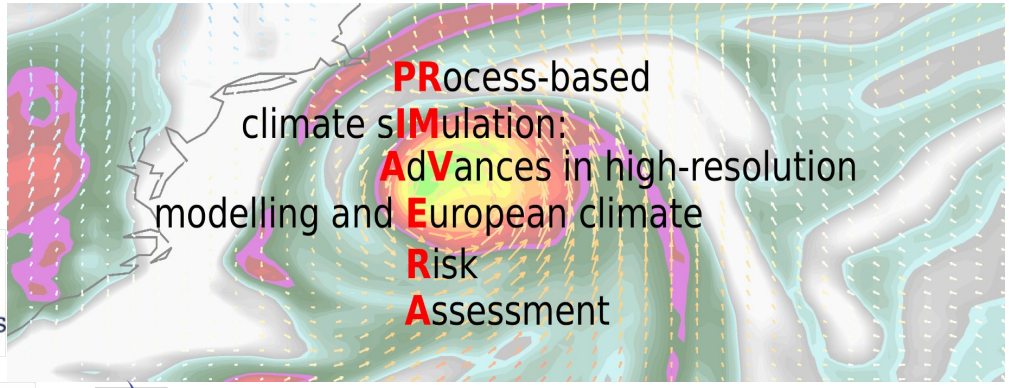
DKRZ DEUTSCHES KLIMARECHENZENTRUM

CERFACS

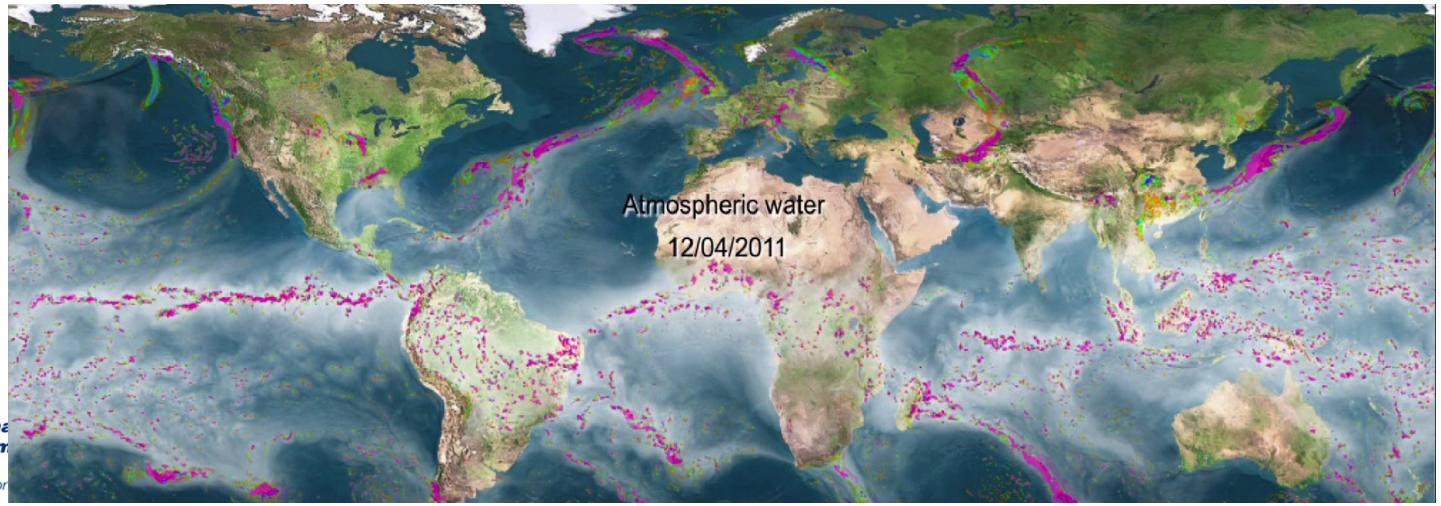
BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

AWI



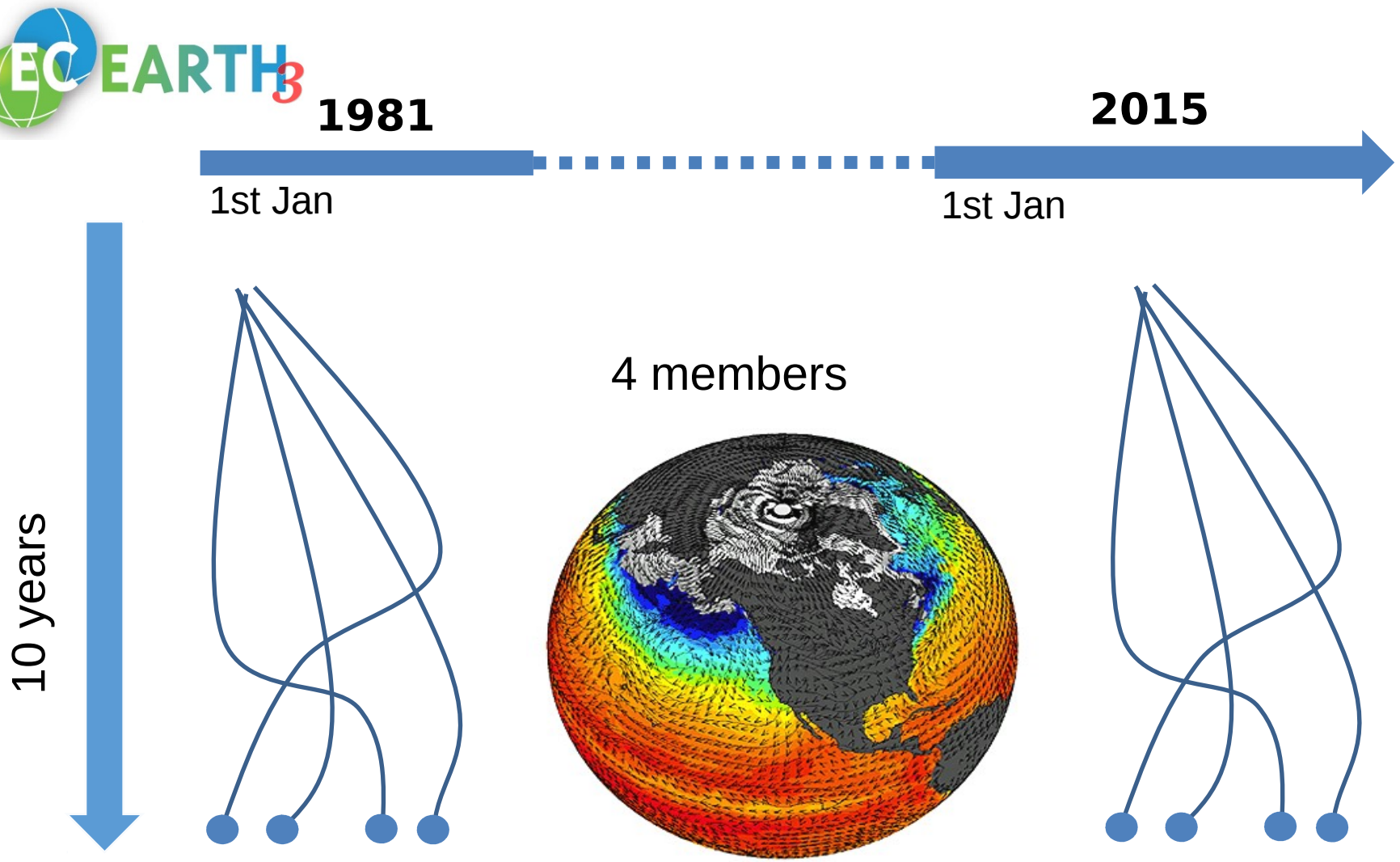
"To develop a new generation of advanced and well-evaluated high-resolution global climate models, capable of simulating and predicting regional climate with unprecedented fidelity, for the benefit of governments,



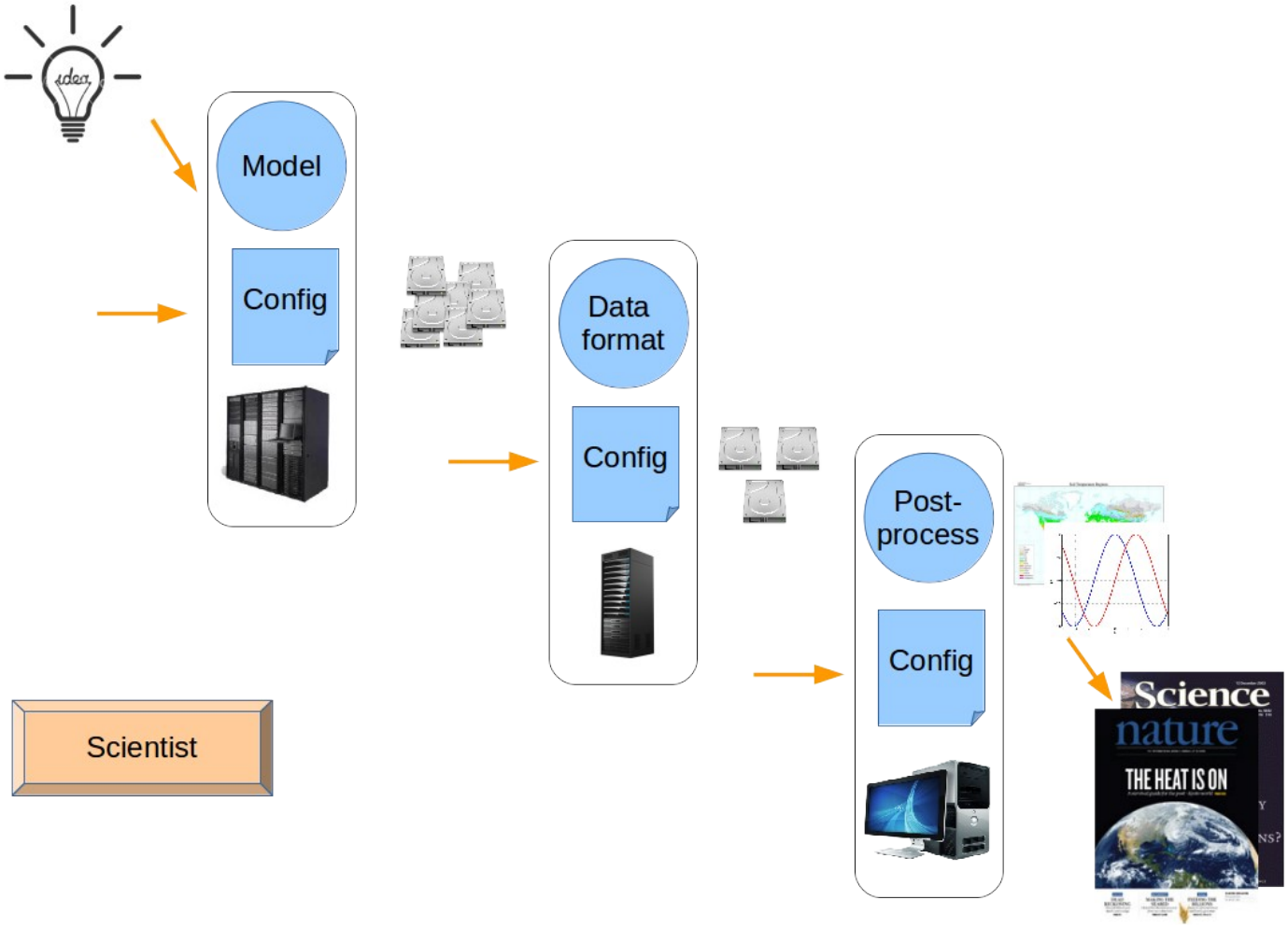
Multiple High Performance Computing infrastructures

Computing resources funded by: National / EU / International projects

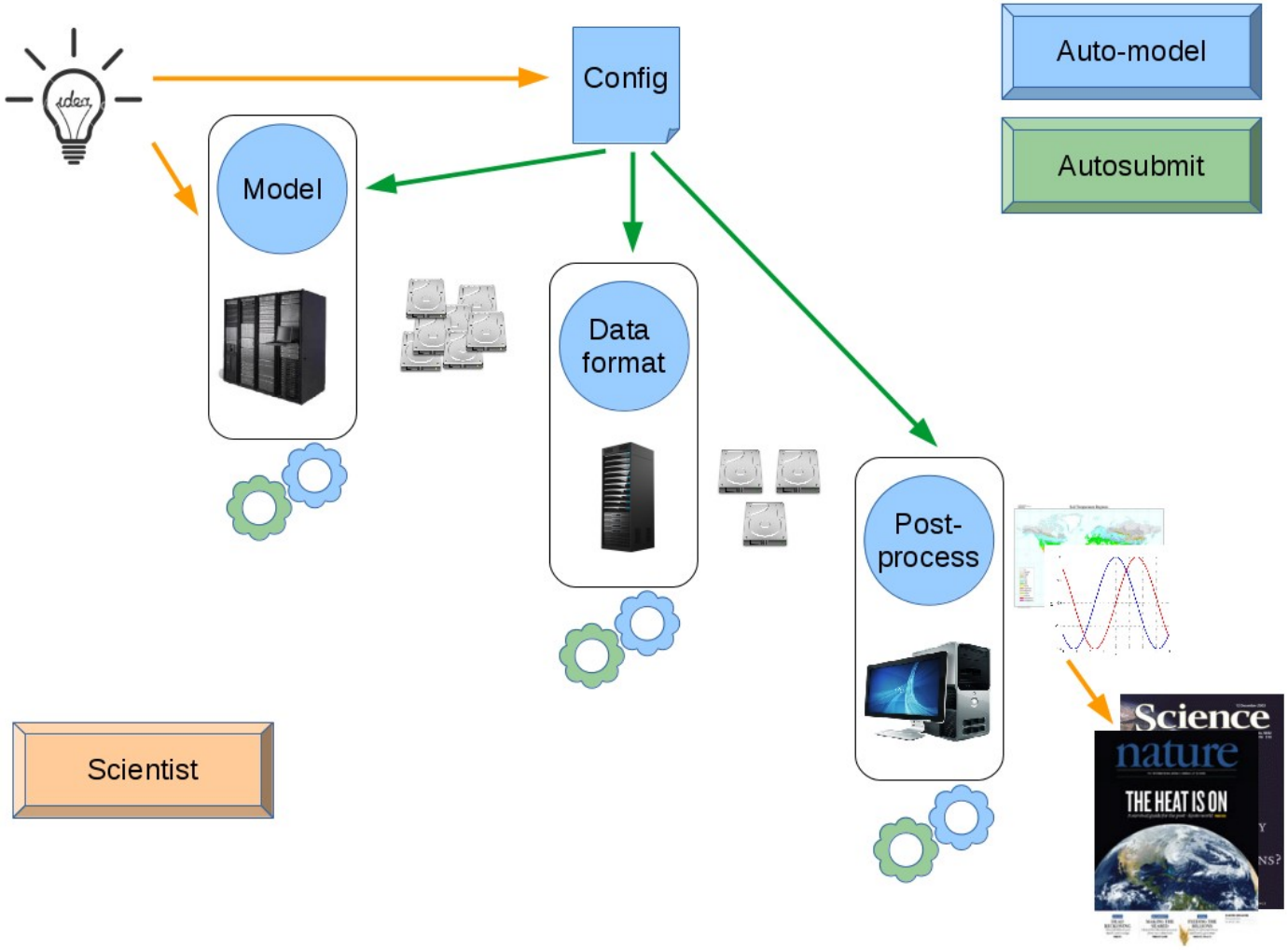
Multi-member climate experiment



Climate experiment workflow



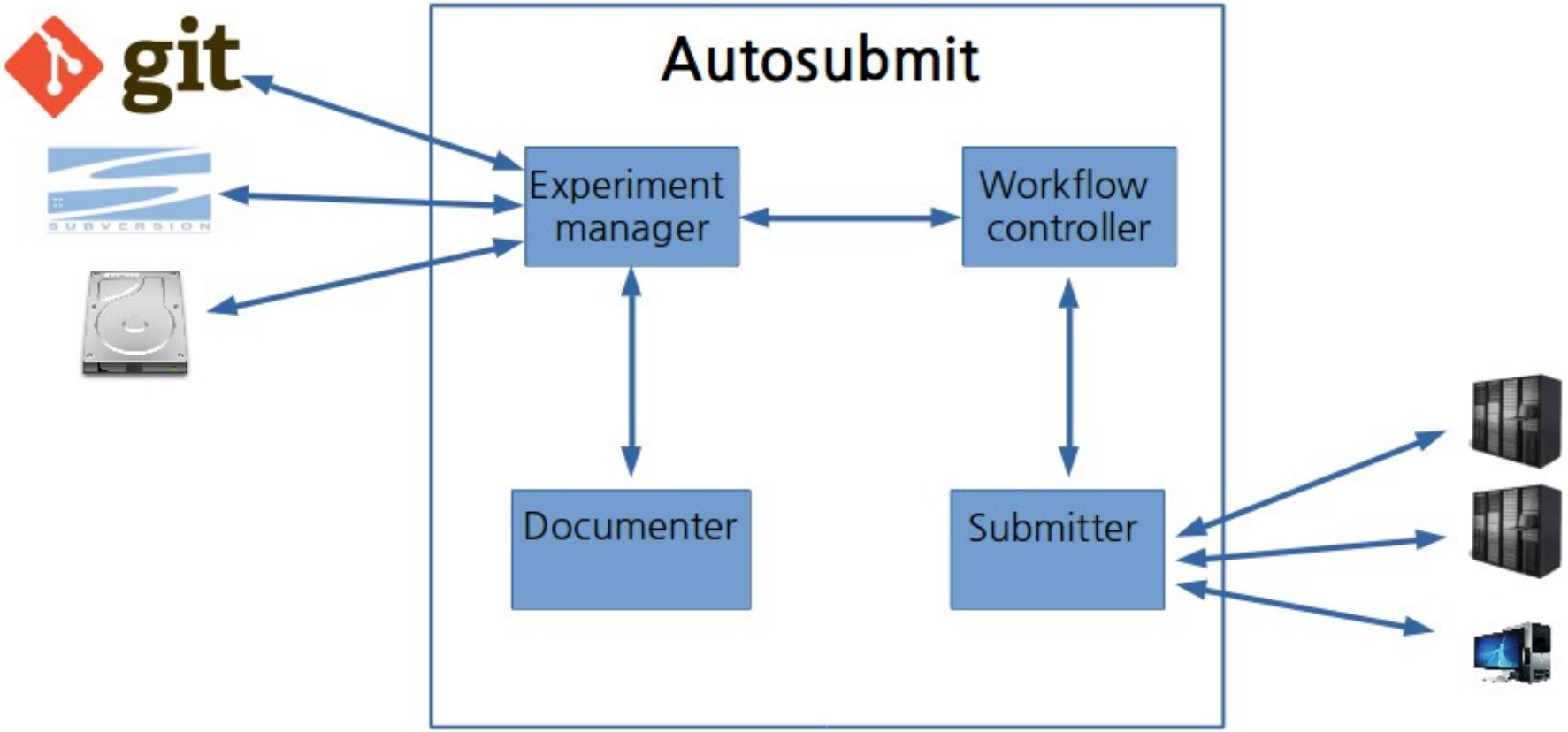
Increased automatisation



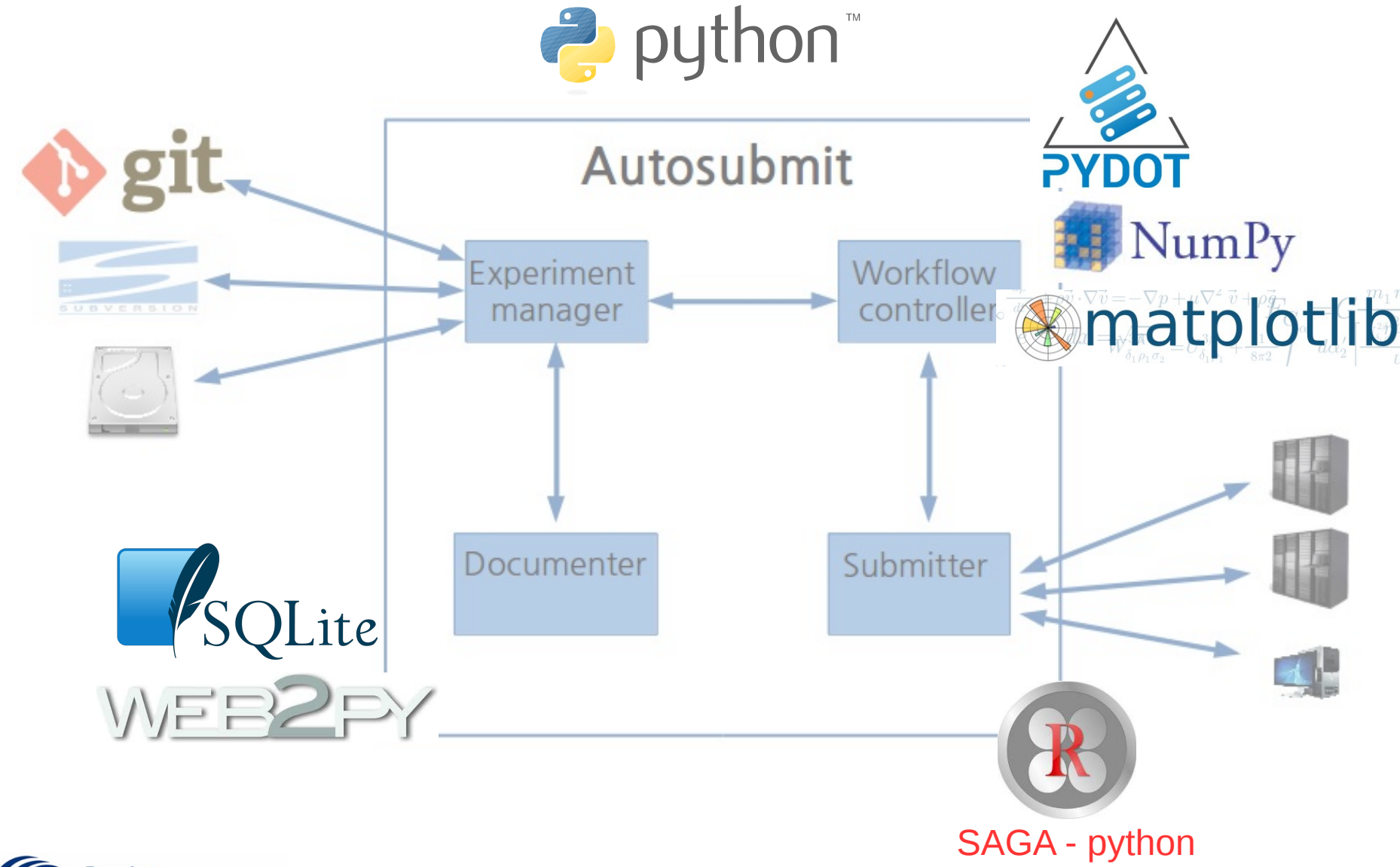

BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Autosubmit

What is Autosubmit ?



How does Autosubmit work ?

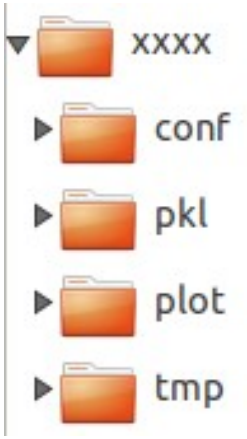



BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Example

Experiment creation

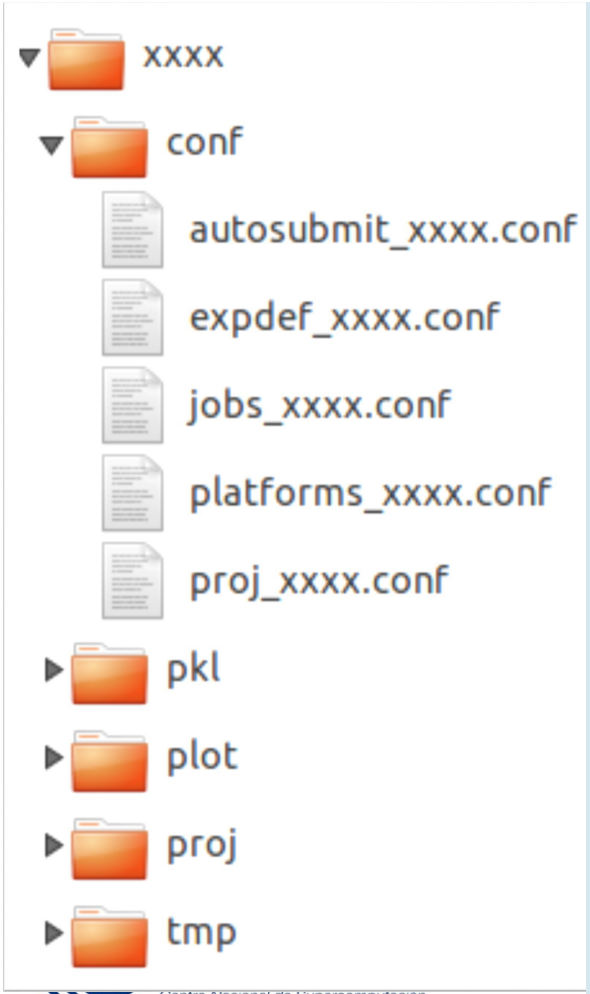
```
autosubmit expid –H HPCname
```



Experiment creation Experiment configuration

```
autosubmit expid –H HPCname
```

```
autosubmit create xxxx
```



Start dates, members and chunks (number and length).

Experiment project source: origin (version control system or path) and project configuration file path.

expdef_xxxx.conf

Workflow to be run: scripts to execute, dependencies between tasks, task requirements (processors, wallclock time...) and platform to use.

jobs_xxxx.conf

HPC, fat-nodes and supporting computers configuration.

Usually provided by technicians, users will only have to change login and accounting options for HPCs.

platforms_xxxx.conf

Project dependant experiment variables that Autosubmit will substitute in the scripts to be run.

proj_xxxx.conf

Experiment creation Experiment configuration Experiment run

```
autosubmit expid –H HPCname
```

```
autosubmit create xxxx
```

```
autosubmit run xxxx
```


Start dates, members and chunks (number and length).

Experiment project source: origin (version control system or path) and project configuration file path.

expdef_xxxx.conf

Workflow to be run: scripts to execute, dependencies between tasks, task requirements (processors, wallclock time...) and platform to use.

jobs_xxxx.conf

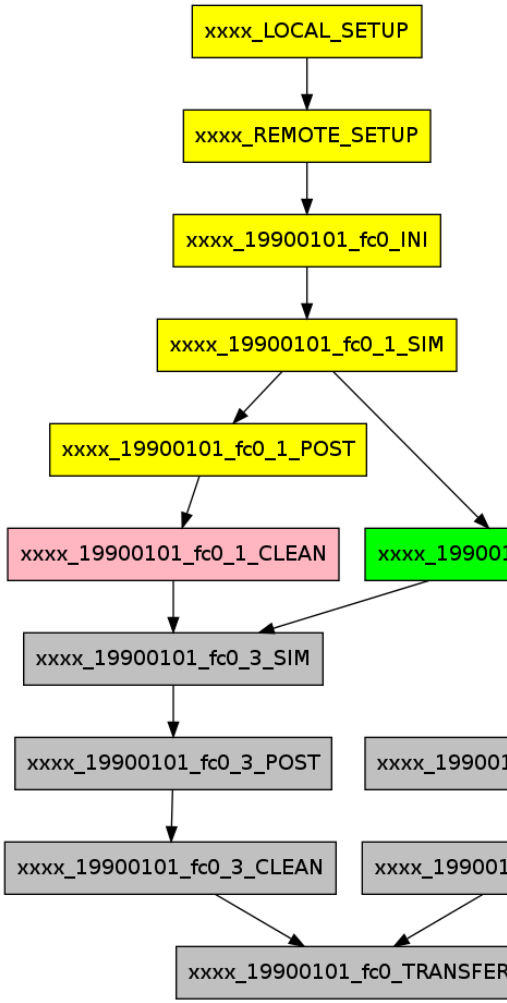
HPC, fat-nodes and supporting computers configuration.

Usually provided by technicians, users will only have to change login and accounting options for HPCs.

platforms_xxxx.conf

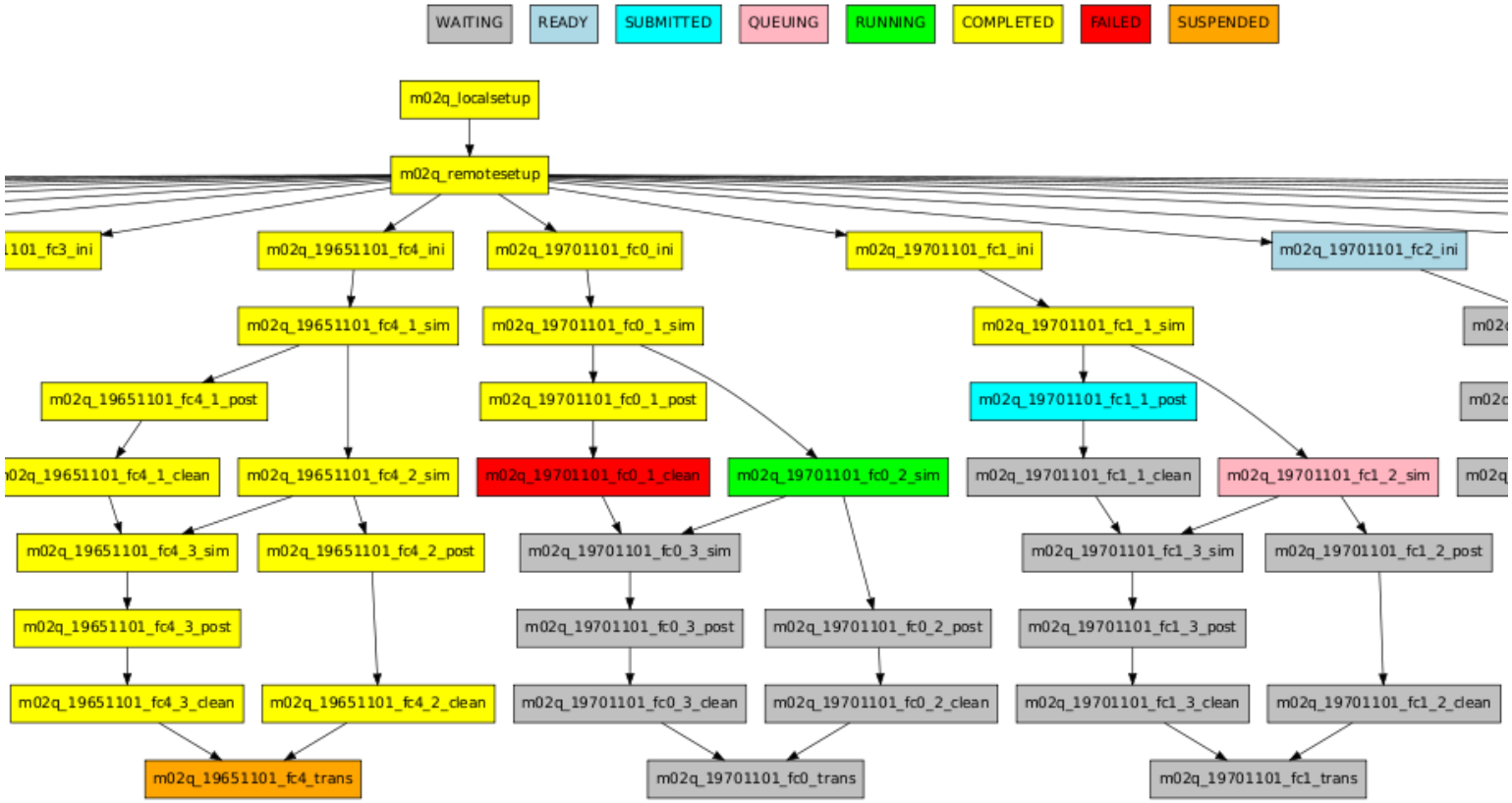
Project dependant experiment variables that Autosubmit will substitute in the scripts to be run.

proj_xxxx.conf

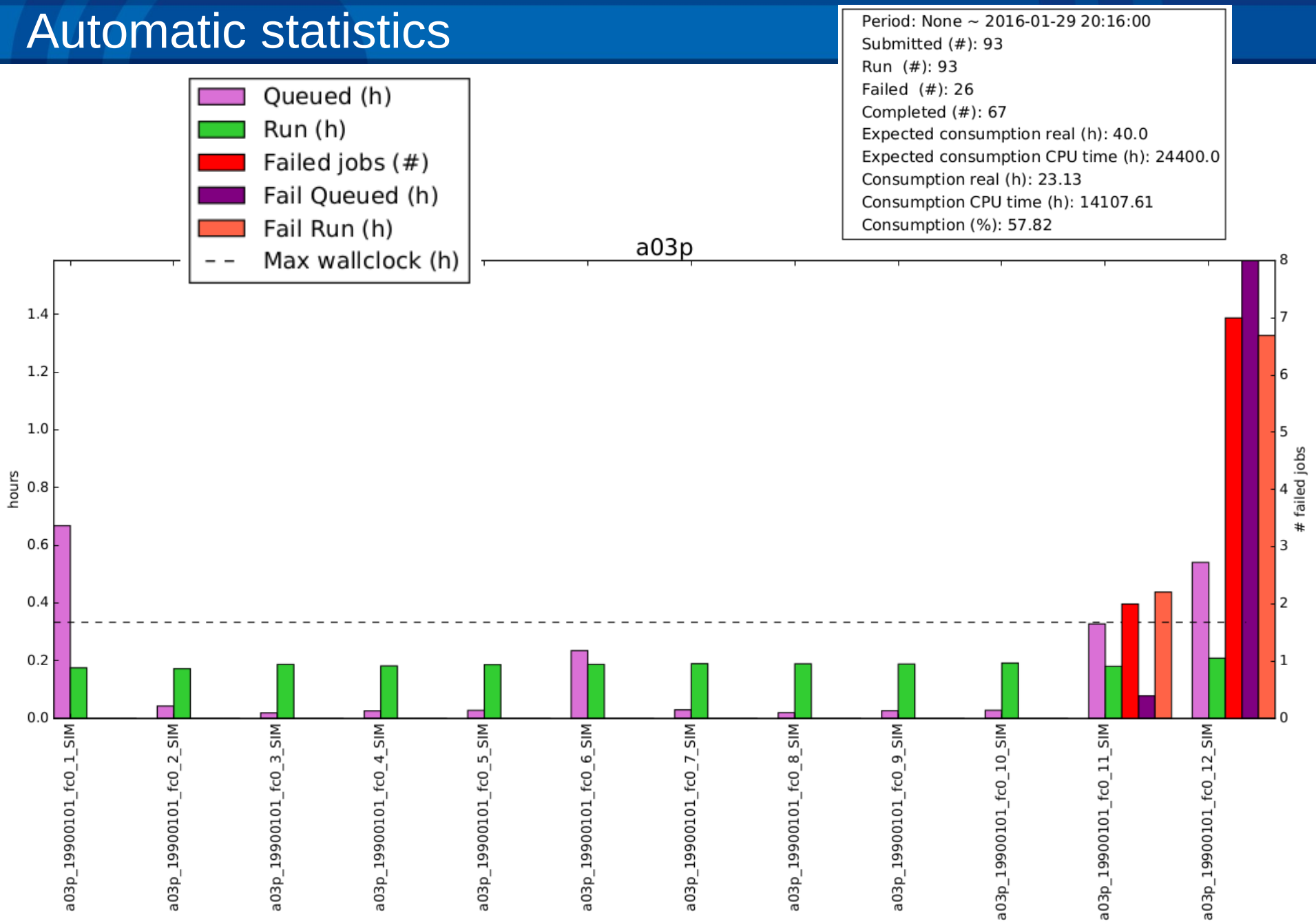


Experiment monitoring

```
autosubmit monitor xxxx
```

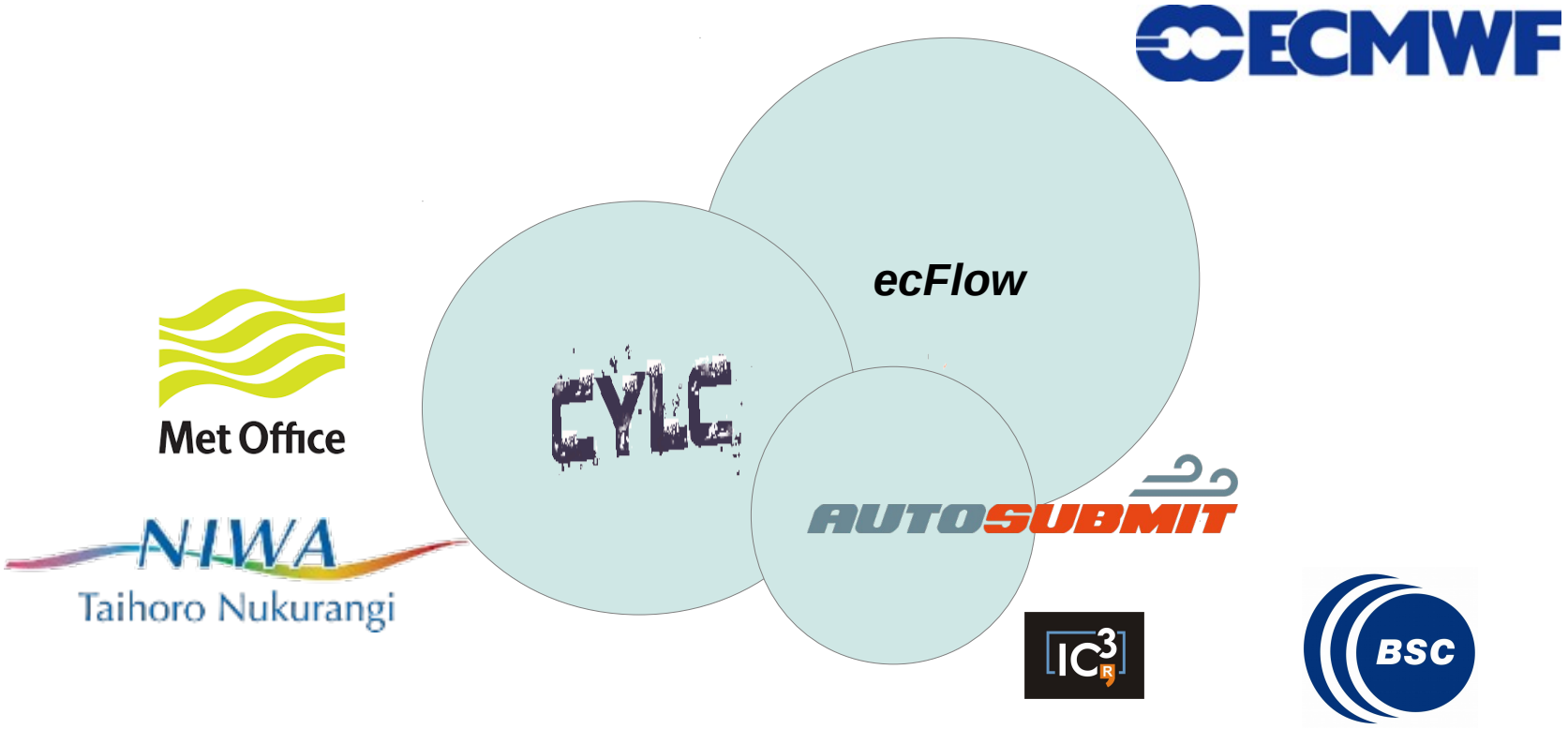


Automatic statistics



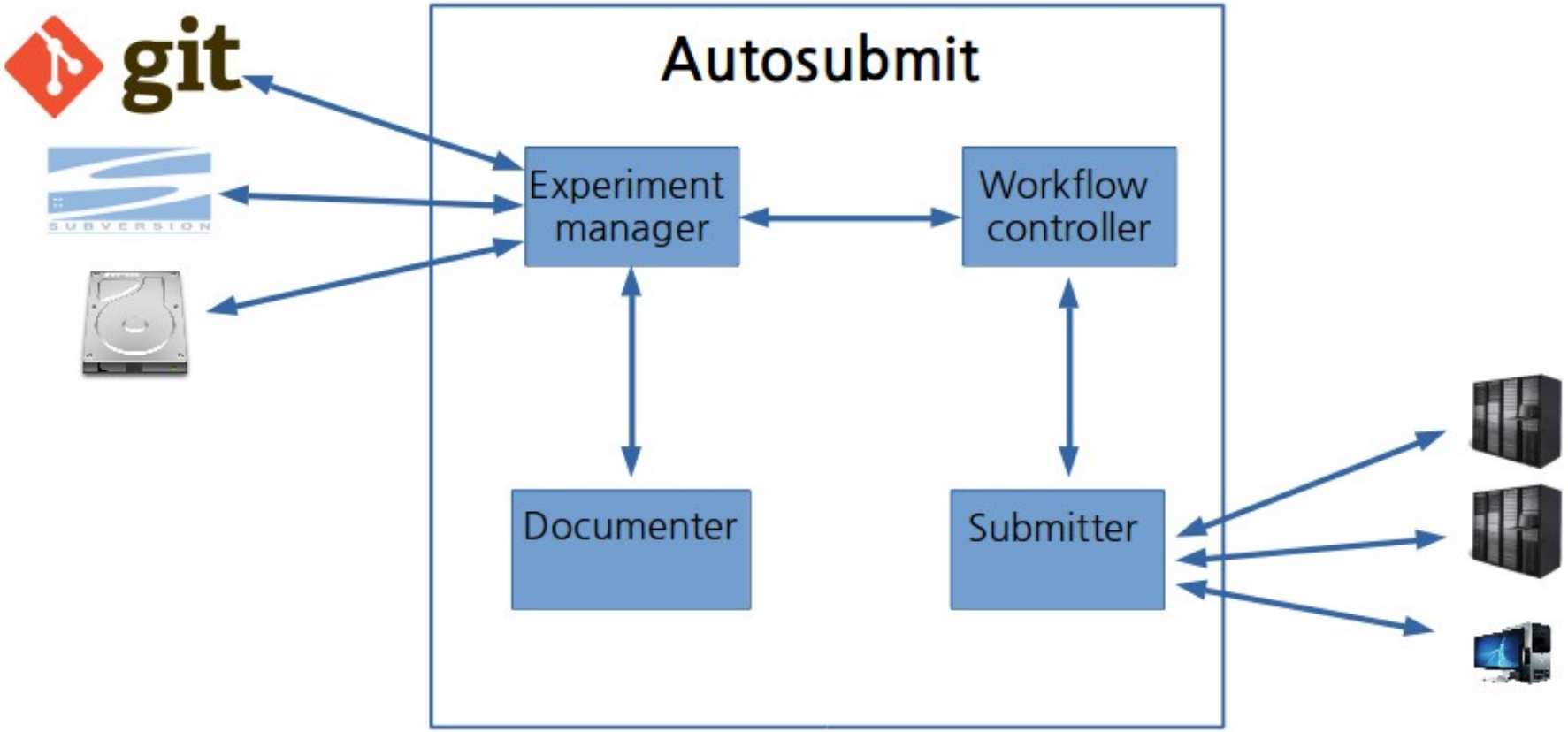
Comparison with other tools

Comparison with other workflow tools

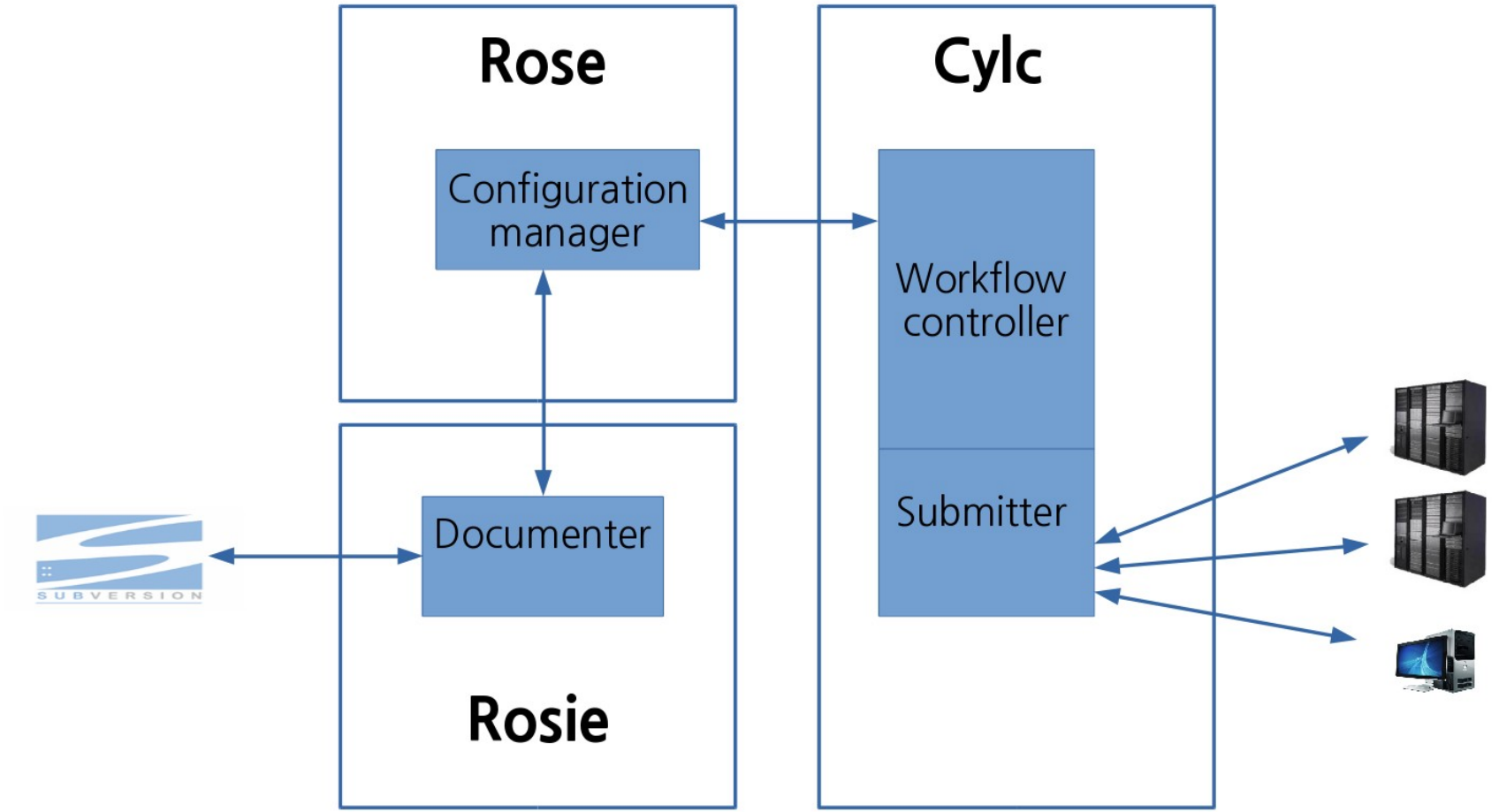


The three scheduling and submission systems have been tested and evaluated with regard to the suitability for multi-model multi-member high resolution (M4 HR) climate experiments.

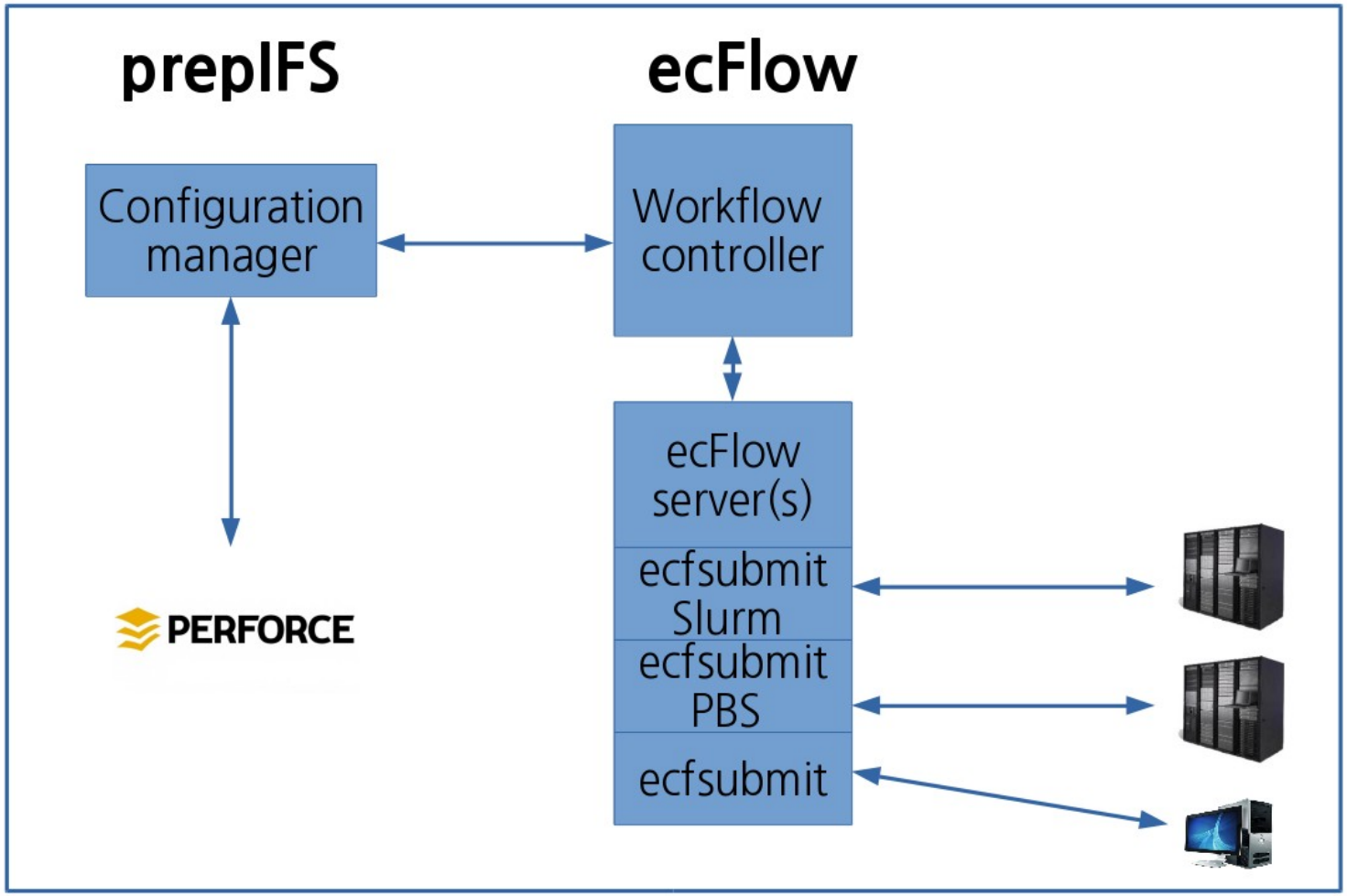
Autosubmit technical infrastructure



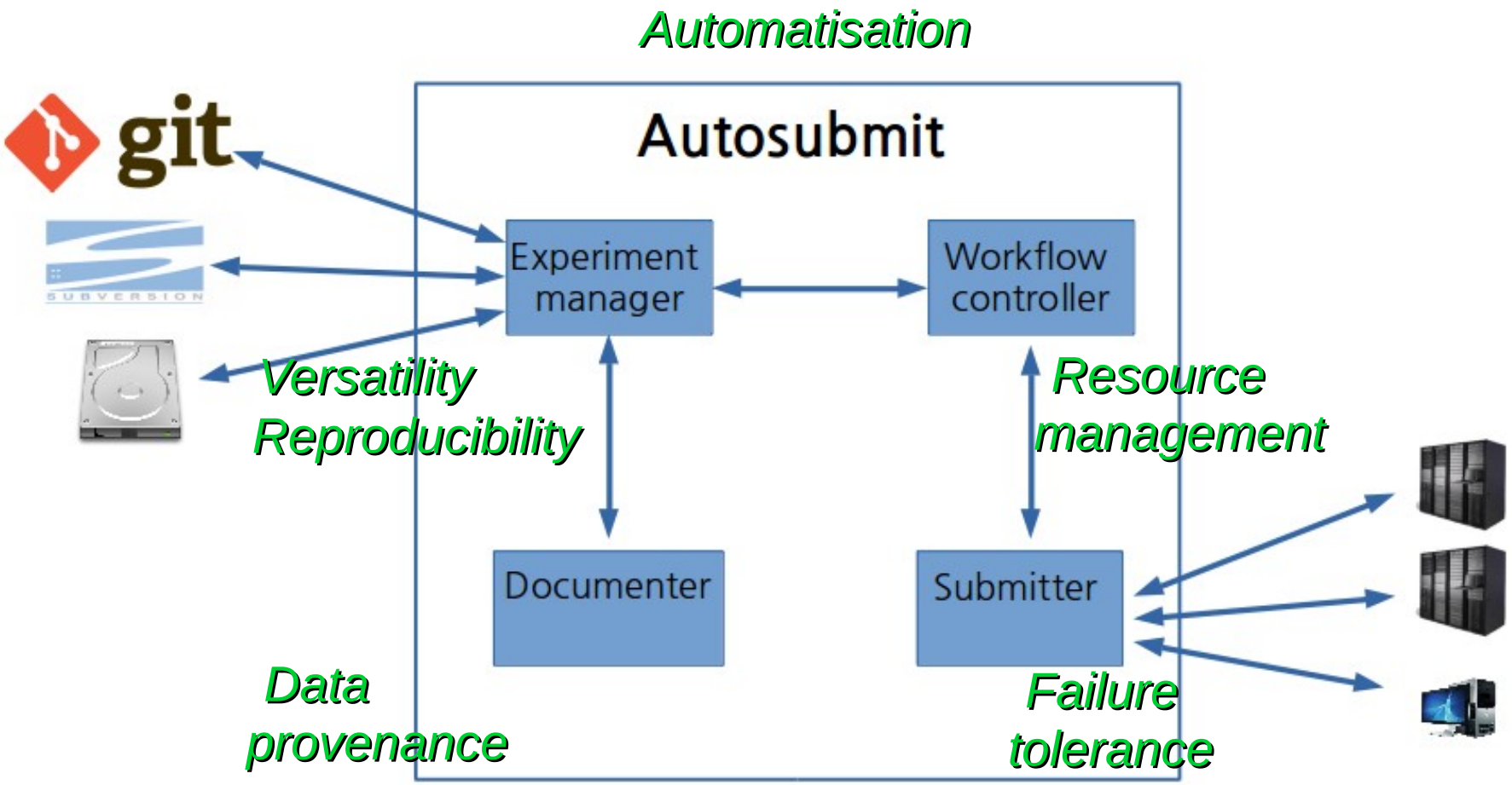
Cylc technical infrastructure



ecFlow technical infrastructure

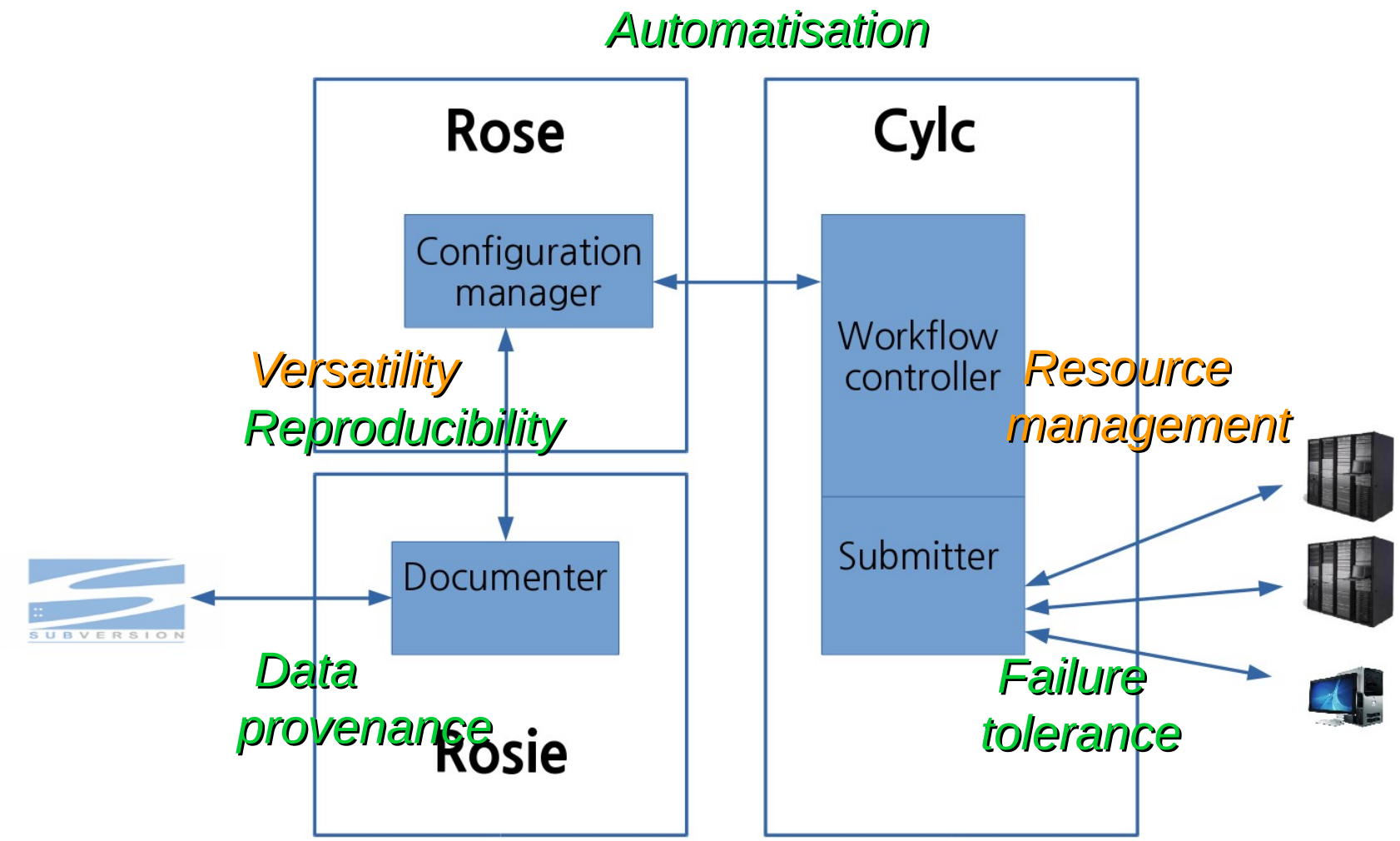


Autosubmit evaluation

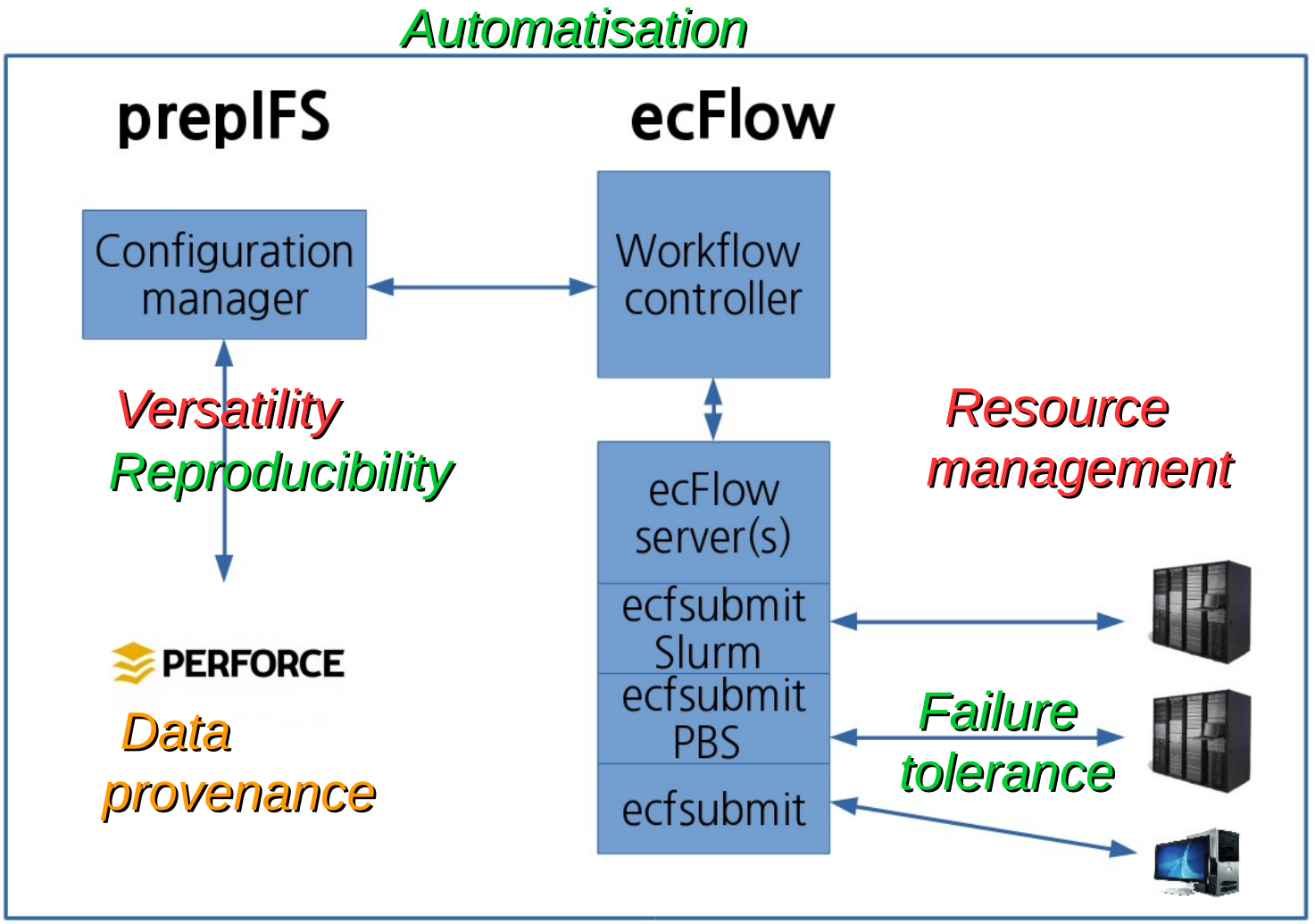


Autosubmit is not ready for operational

Comparison with Cylc



Comparison with ecFlow



ecFlow API is robust

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Case study

Case study

- m02j → ten members, four forecast months, 34 start dates → 340 independent simulations → 113 years of simulation

  Standard resolution (T255L91-ORCA1L46-LIM2)

  1,9 TB

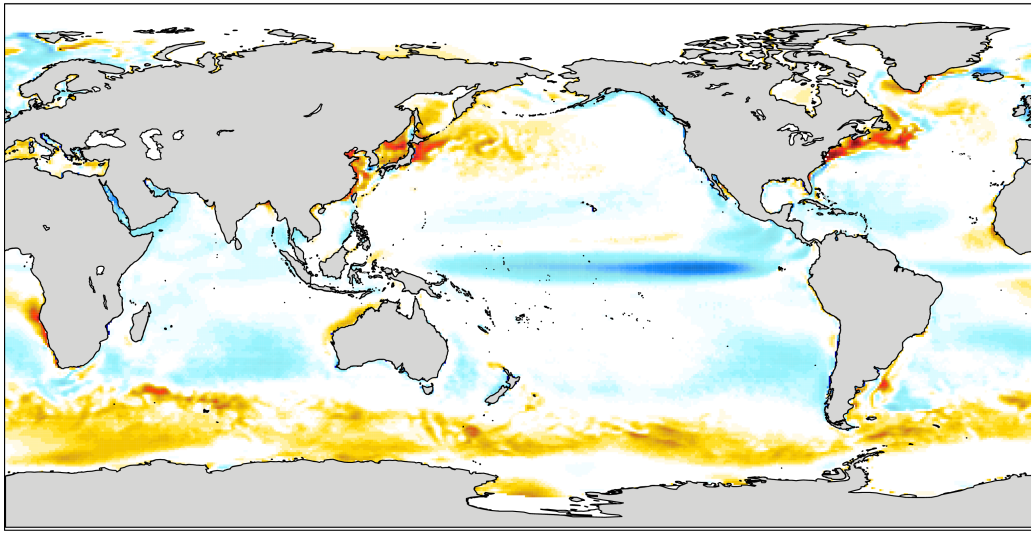
- m02s → ten members, four forecast months, 34 start dates → 340 independent simualtions → 113 years of simulation

  High resolution (T511L91-ORCA025L75-LIM2)

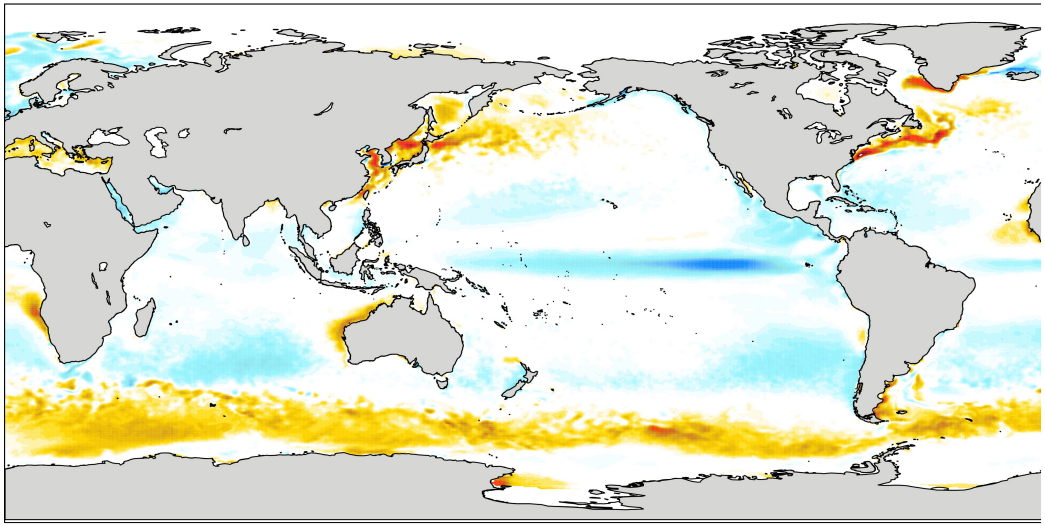
  20'5 TB

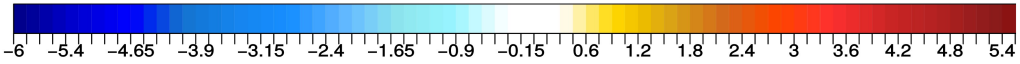
Benefits of resoultion increase in EC-Earth

- m02j



- m02s

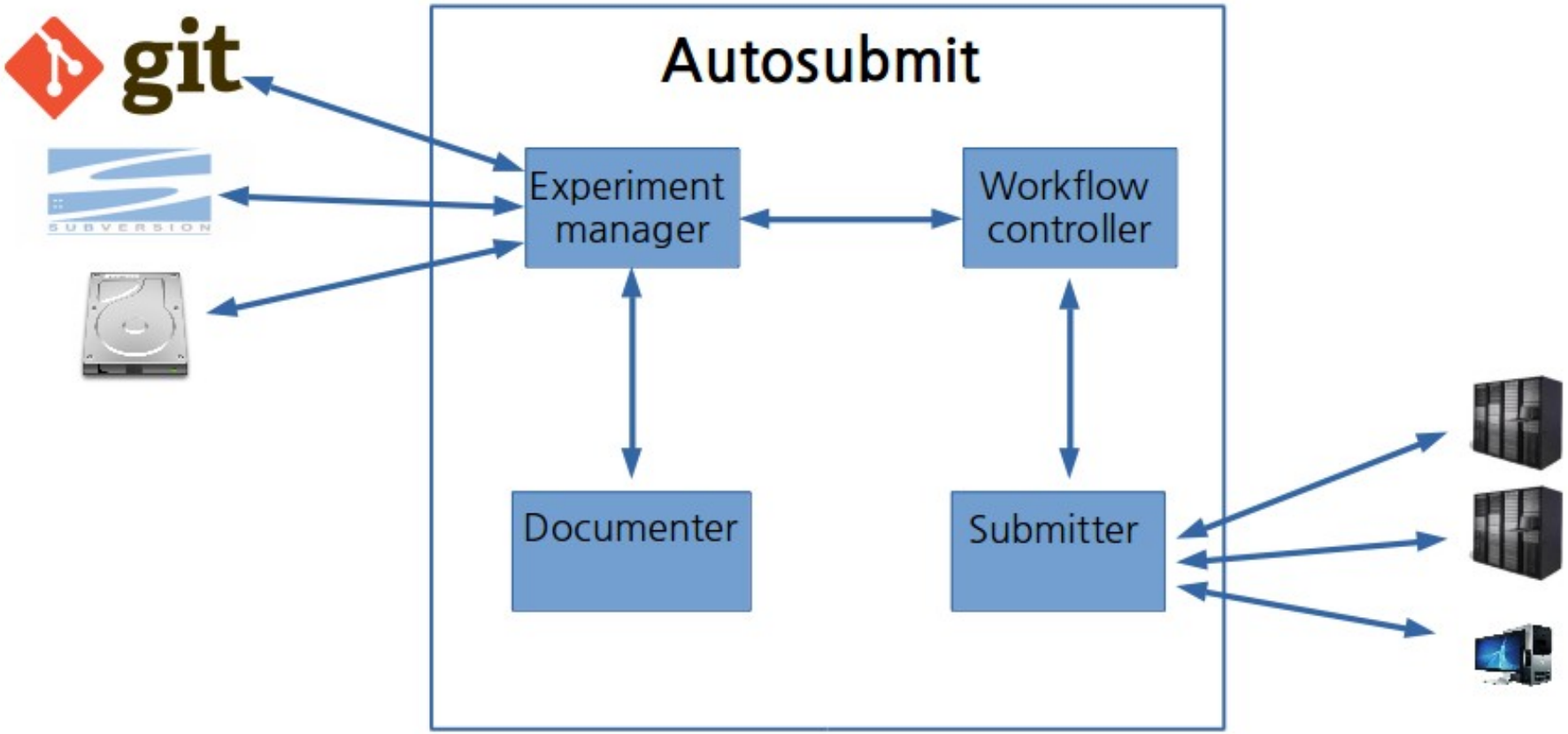




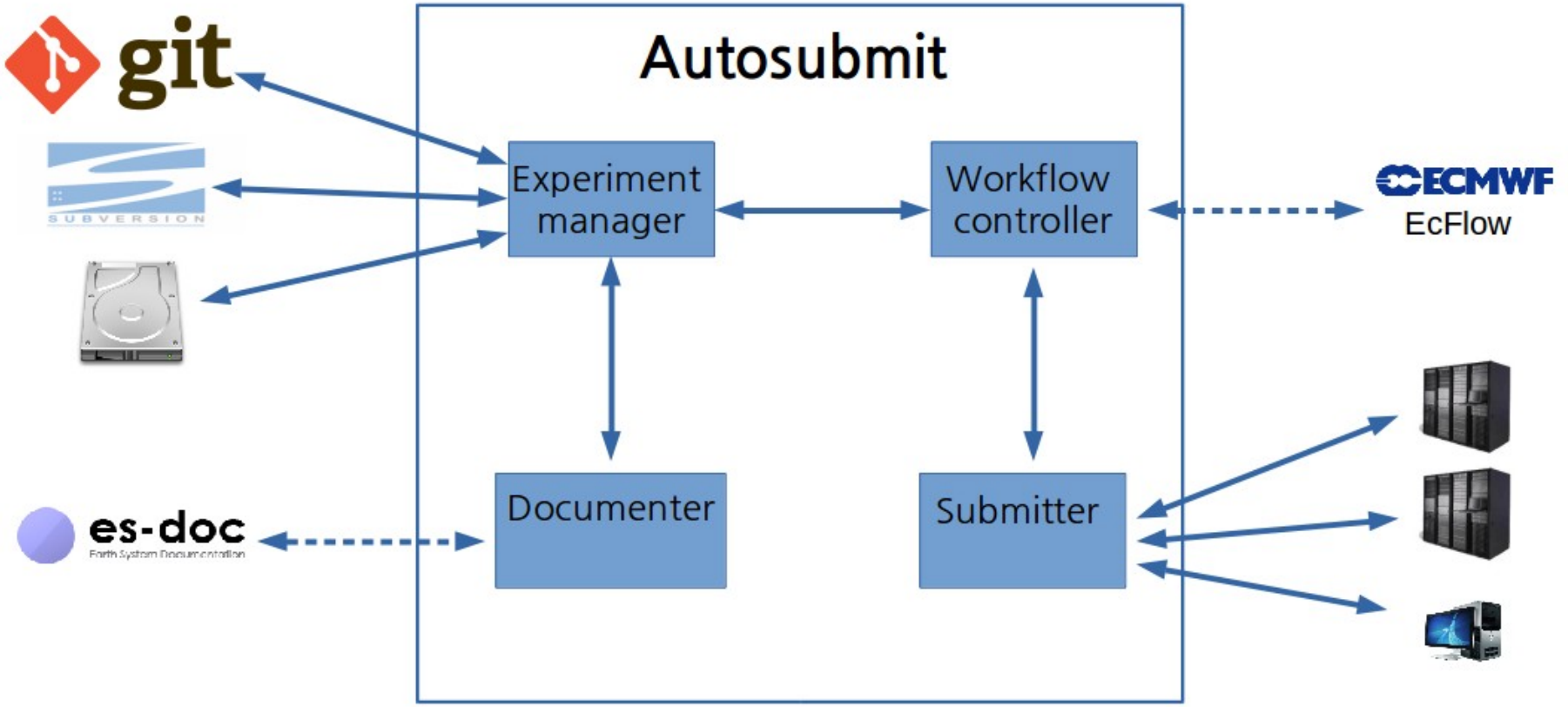

BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Future work

Future work


git
SUBVERSION
Autosubmit
Experiment manager
Workflow controller
Documenter
Submitter

Future work


git
SUBVERSION
Autosubmit
Experiment manager
Workflow controller
ECMWF
EcFlow
es-doc
Documenter
Submitter

www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thank you !

Do not hesitate to contact us
if you have any doubts or suggestions:
domingo.manubens@bsc.es
javier.vegas@bsc.es