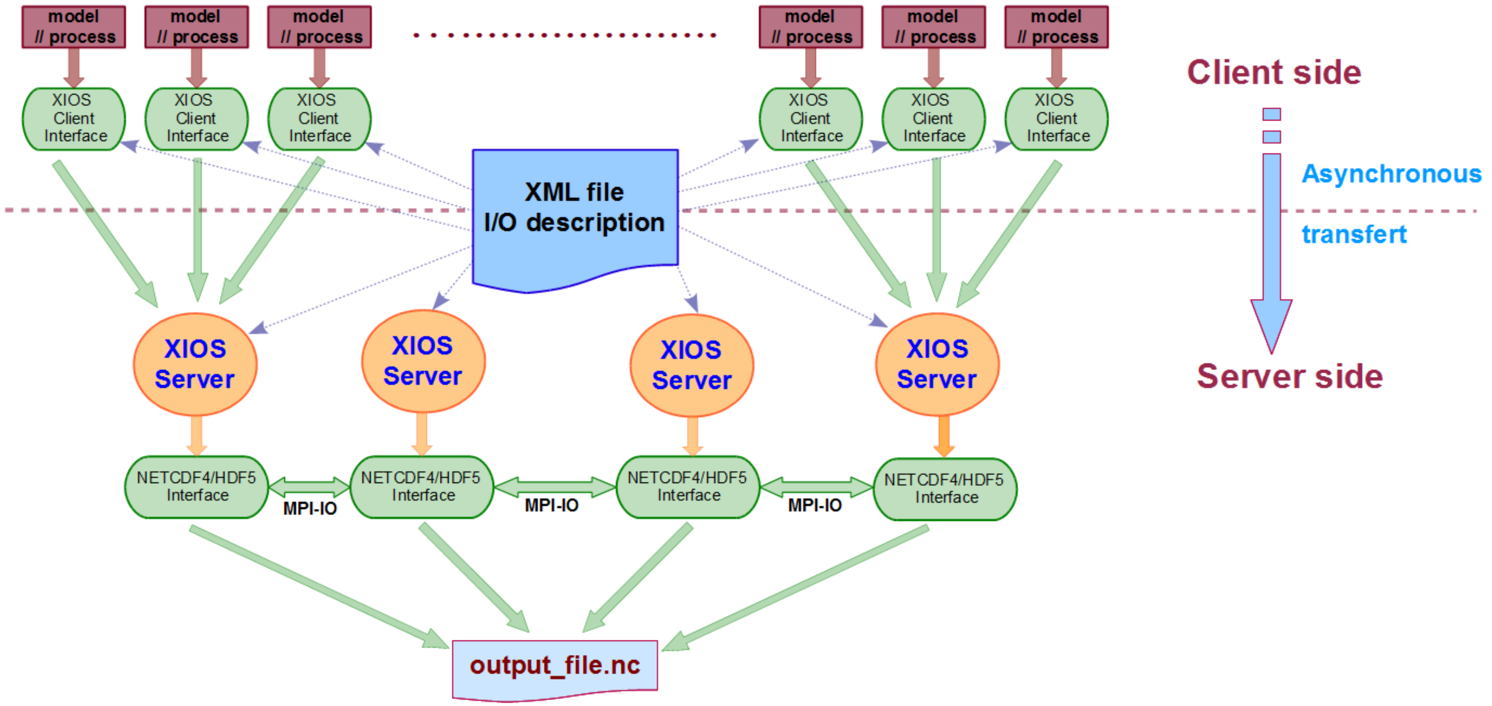
XIOS
where are we and example of use

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

XIOS – 1 : XML / IO / SERVER



Flexible data output description through an external XML file.

XIOS servers : asynchronous processes exclusively dedicated to output

- **Overlap computation and I/O**
- **Rearrange data for better output efficiency**
- **Use parallel I/O for better efficiency**
  - Aggregate I/O bandwidth of parallel file system
  - One piece files, no need to rebuild

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

XIOS - 2 : Adding Workflow Functionalities

XIOS-2 embed an internal parallel workflow/dataflow

The XML files describe a parallel task graph

- Incoming data are representing data flux, assigned to a timestamp
  - Each flux can be connected to one or more filters
- Filters are connected to one or more input flux and generate a new flux on output
  - All Filters can be chained to achieve complex treatment
  - All filters are parallel and scalable

Arithmetic filters

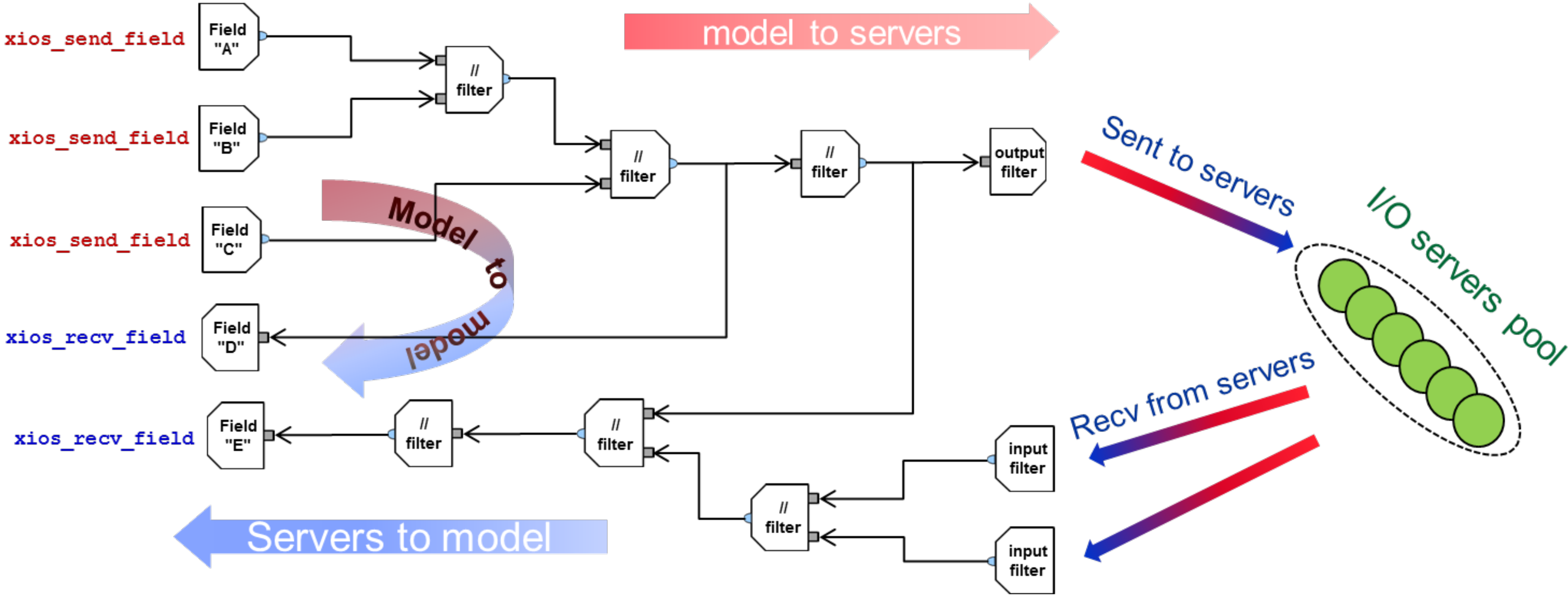
- Combining fields together
- Apply arithmetic operator or function

Temporal filters

- Temporal integration of input flux
- instant, average, maximum, minimum, accumulate

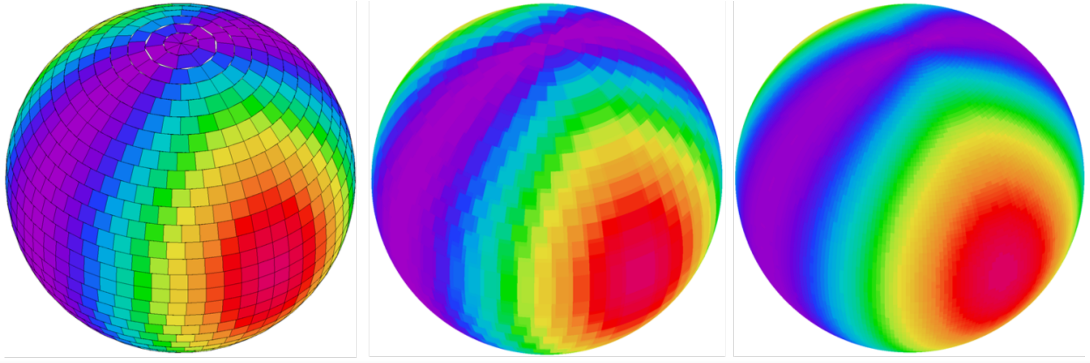
Spatial filters

- Geometrical shape of the input flux is modified
- Defined by a grid transformation from a source grid to a target grid

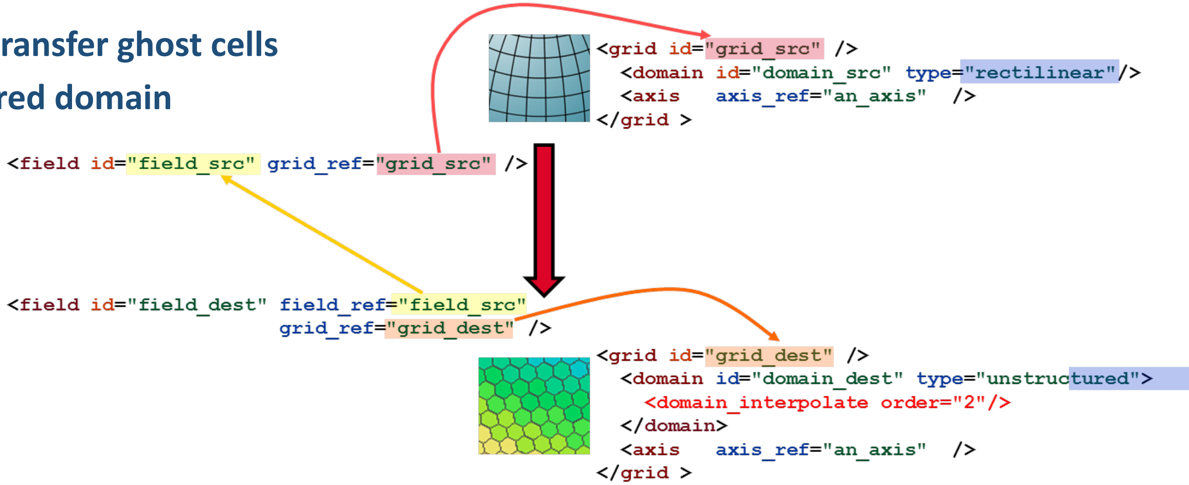


XIOS-2 : AVAILABLE SPATIAL FILTERS

- (domain -> domain): <zoom_domain /> : extract area of interest
- (axis -> axis): <zoom_axis/> : extract part of an axis
- (axis->scalar): <extract_axis_to_scalar/> : axis slice extraction
- (domain->axis): <extract_domain_to_axis/> : latitude or longitude extraction
- (axis->axis): <inverse axis/> : invert axis
- (axis->axis): <interpolate_axis> : axis interpolation, possibly on pressure level
- (domain->domain) : <interpolate_domain/> : horizontal conservative remapping
- (domain): <generate_rectilinear_domain/> : create and/or distribute a rectilinear mesh
- (domain->scalar): <reduce_domain_to_scalar/> : global domain reduction (sum, average, max, min,…)
- (domain->axis): <reduce_domain_to_axis/> : partial domain reduction along i or j direction
- (axis->scalar): <reduce_axis_to_scalar/> : axis reduction (sum, average, max, min, …)
- (scalar->axis): <temporal_splitting/> : diurnal cycle
- (scalar->axis): <duplicate_scalar_to_axis> : duplicate data along a new axis
- (domain->domain) : <reorder_domain/> : reorder indexes of horizontal domain
- (domain->domain): <expand_domain/> : expand local domain at first neighbor and transfer ghost cells
- (domain) : <compute_connectivity/> : find the connectivity of an unstructured domain



Source mesh   Remapping order 1   Remapping order 2



```
<grid id="grid_src" />
  <domain id="domain_src" type="rectilinear"/>
  <axis   axis_ref="an_axis"   />
</grid >

<field id="field_src" grid_ref="grid_src" />

<field id="field_dest" field_ref="field_src"
                       grid_ref="grid_dest" />

<grid id="grid_dest" />
  <domain id="domain_dest" type="unstructured">
    <domain_interpolate order="2"/>
  </domain>
  <axis   axis_ref="an_axis"   />
</grid >
```

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

XIOS-2.5 : The CMIP6 workflow compliant version

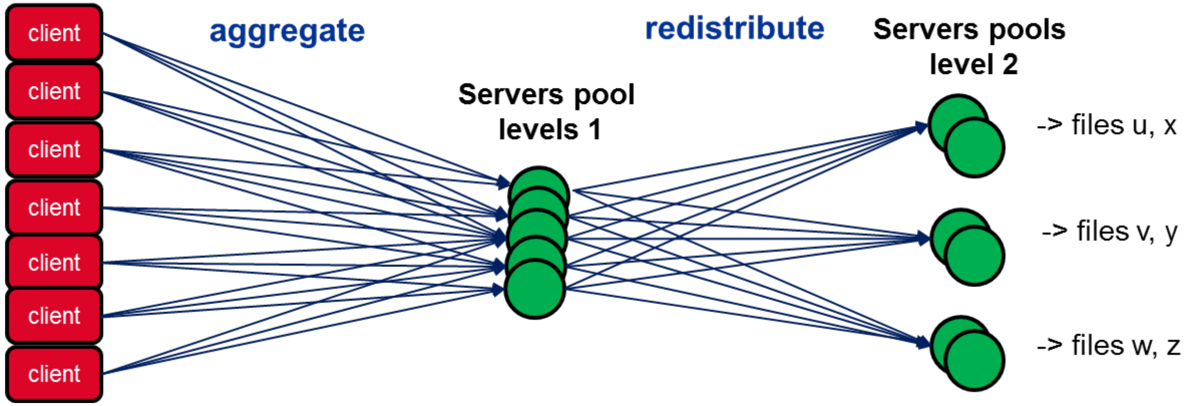
Adding missing functionalities required by CMIP6 workflow

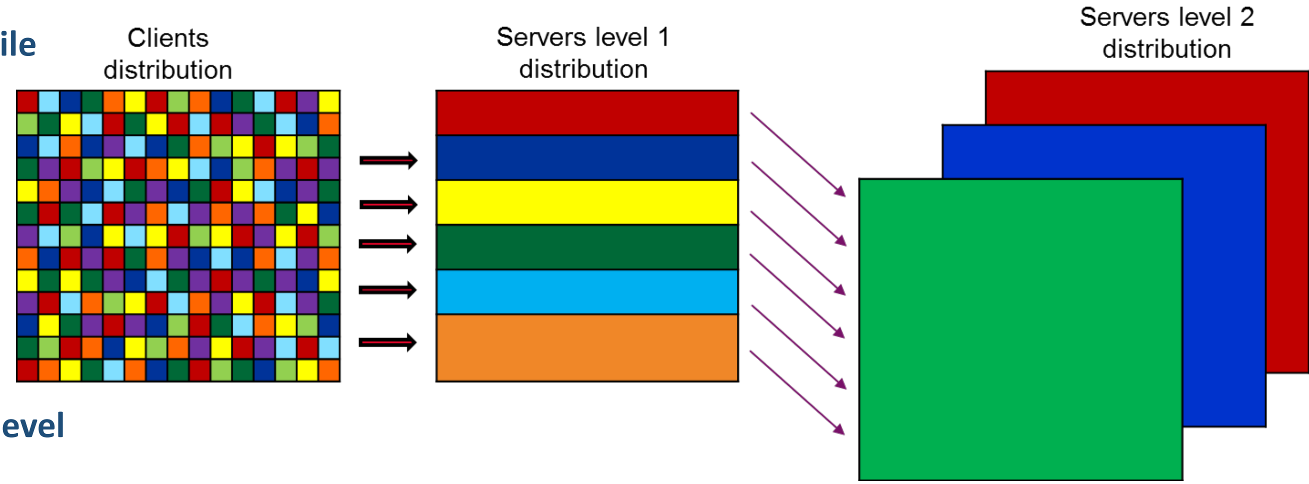
XIOS output fully CF 1.7 compliant

- Axis & coordinates
- Variables and associated metadata
- Time axis management

Automatic time series management

- One file by variable
- Automatic generation of UUID (tracking_id)
- Automatic chunk splitting at a given frequency specifically to an output file
  - Constant size for chunk of file variable
  - Automatic file name suffix corresponding to the period of chunk
- An output file can be reopen and appended by the next run





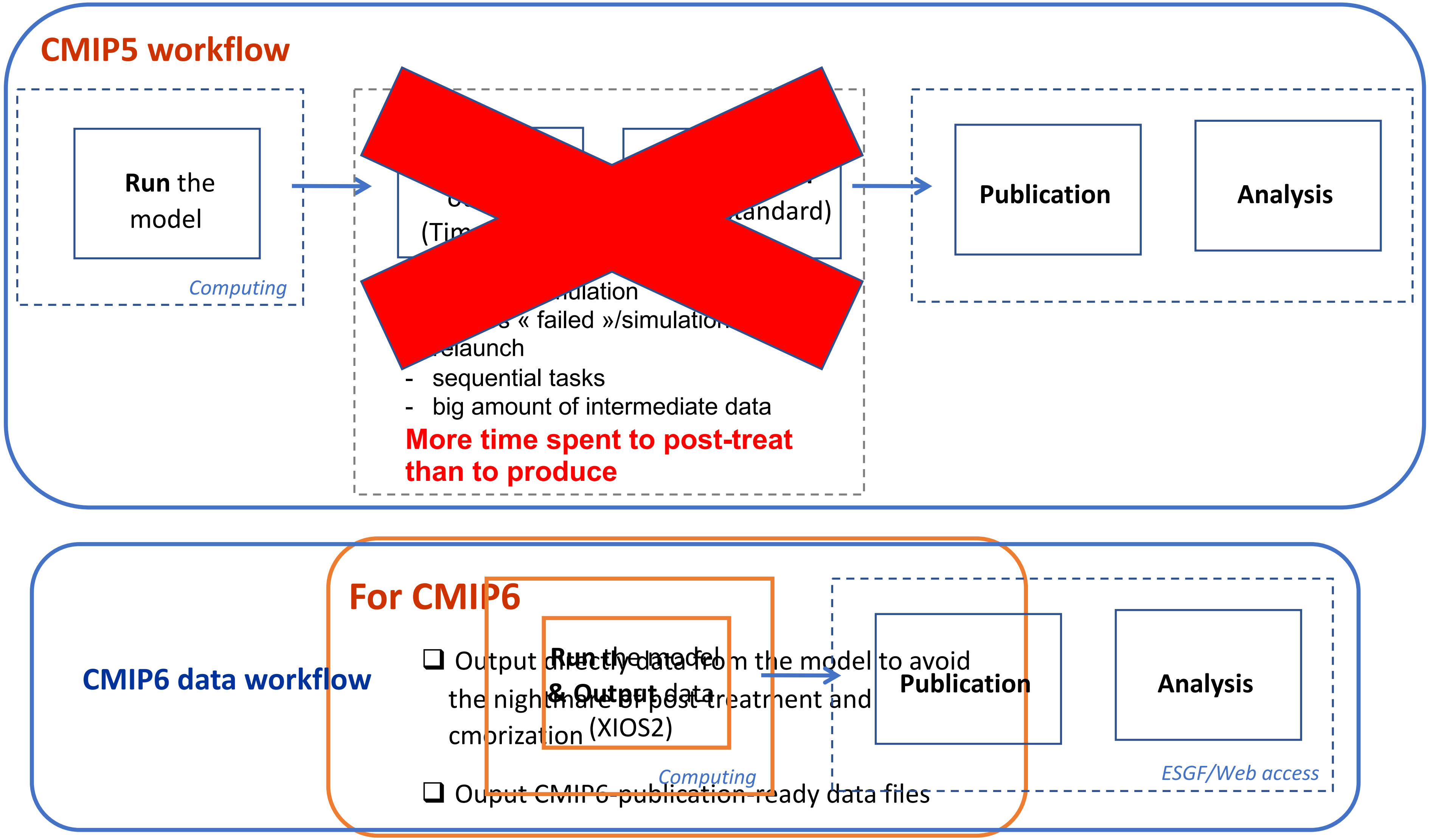
Adding a second level of servers

- First level will aggregate fields from client and redistribute its to second level
- Second level received field on global mesh and make sequential write
- I/O parallelism is achieved by write sequential files concurrently
  - Increase I/O parallelism performance
  - Enabling compression on the fly using HDF5 filters

The IS-ENES3 project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement No 824084

An example of use : the CMIP6 XIOS Workflow

FROM CMIP5 to CMIP6



ulation
s « failed »/simulation
relaunch
- sequential tasks
- big amount of intermediate data

More time spent to post-treat than to produce

For CMIP6

- ❑ Output directly data from the model to avoid the nightmare of post-treatment and cmorization
- ❑ Ouput CMIP6-publication-ready data files

The IS-ENES3 project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824084

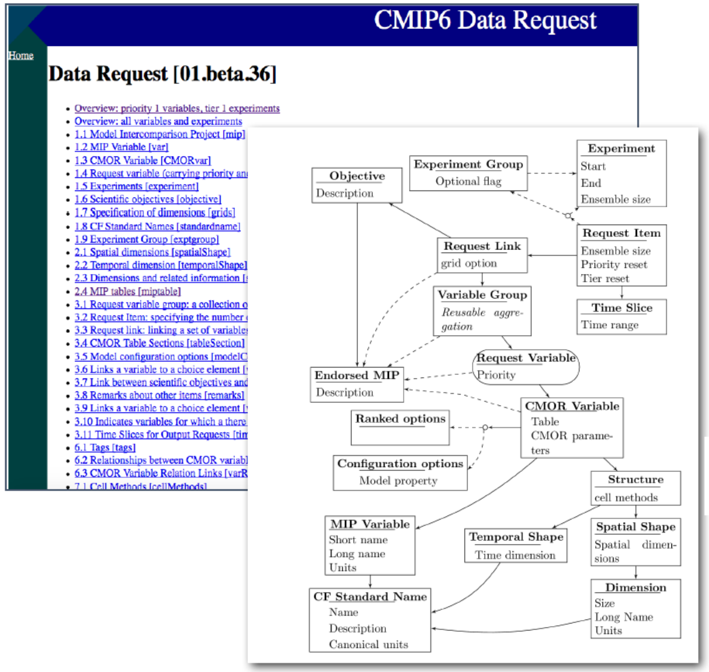
CMIP6 workflow strategy

CNRM and IPSL are sharing the same CMIP6 workflow based on XIOS-2.5

Strategy : automatic translation of CMIP6 data request into XIOS XML configuration files

- **CNRM develop DR2XML tool (S. Sénesi, M.P. Moine)**
  - Translates CMIP6 Data Request to XIOS configuration files (Python script)
- **IPSL implement XIOS missing functionalities**

Data Request python API + XML files
(Martin Juckes)



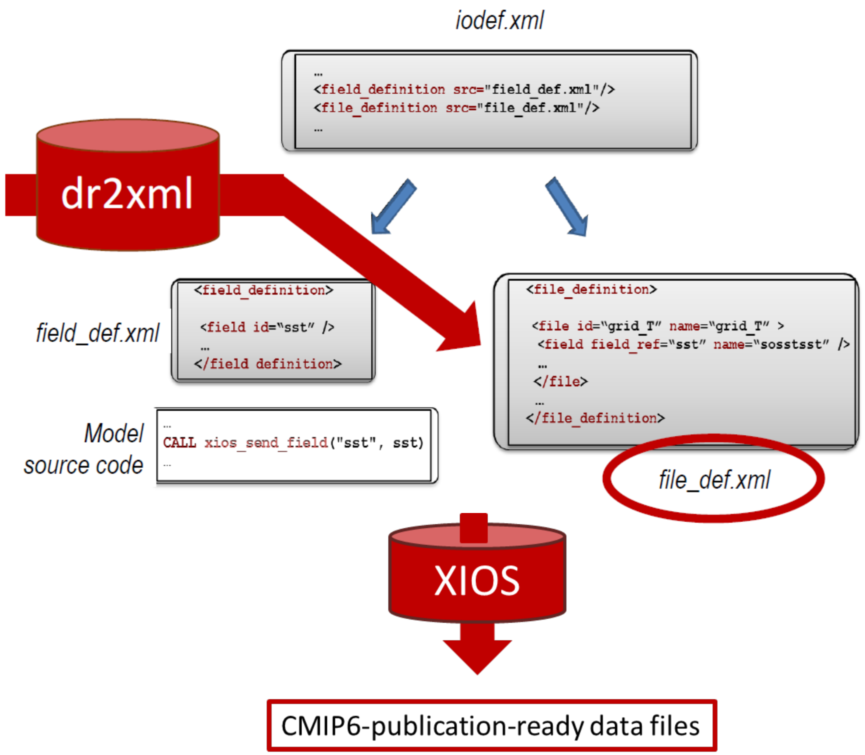
CMIP6 Data Request

Home

Data Request [01.beta.36]

- Overview: priority 1 variables, tier 1 experiments
- Overview: all variables and experiments
- 1.1 Model Intercomparison Project [mip]
- 1.2 MIP Variable [var]
- 1.3 CMOR Variable [CMORvar]
- 1.4 Request variable (carrying priority an
- 1.5 Experiments [experiment]
- 1.6 Scientific objectives [objective]
- 1.7 Specification of dimensions [grids]
- 1.8 CF Standard Names [standardname]
- 1.9 Experiment Group [exptgroup]
- 2.1 Spatial dimensions [spatialShape]
- 2.2 Temporal dimension [temporalShape]
- 2.3 Dimensions and related information [
- 2.4 MIP tables [miptable]
- 3.1 Request variable group: a collection o
- 3.2 Request Item: specifying the number
- 3.3 Request link: linking a set of variables
- 3.4 CMOR Table Sections [tableSection]
- 3.5 Model configuration options [modelC
- 3.6 Links a variable to a choice element [
- 3.7 Link between scientific objectives and
- 3.8 Remarks about other items [remarks]
- 3.9 Links a variable to a choice element [
- 3.10 Indicates variables for which a there
- 3.11 Time Slices for Output Requests [tim
- 6.1 Tags [tags]
- 6.2 Relationships between CMOR variabl
- 6.3 CMOR Variable Relation Links [varR
- 7.1 Cell Methods [cellMethods]

XIOS-enable model
(e.g. NEMO)



iodef.xml

```
…
<field_definition src="field_def.xml"/>
<file_definition src="file_def.xml"/>
…
```

field_def.xml

```
<field_definition>

 <field id="sst" />
 …
</field definition>
```

Model source code

```
…
CALL xios_send_field("sst", sst)
…
```

file_def.xml

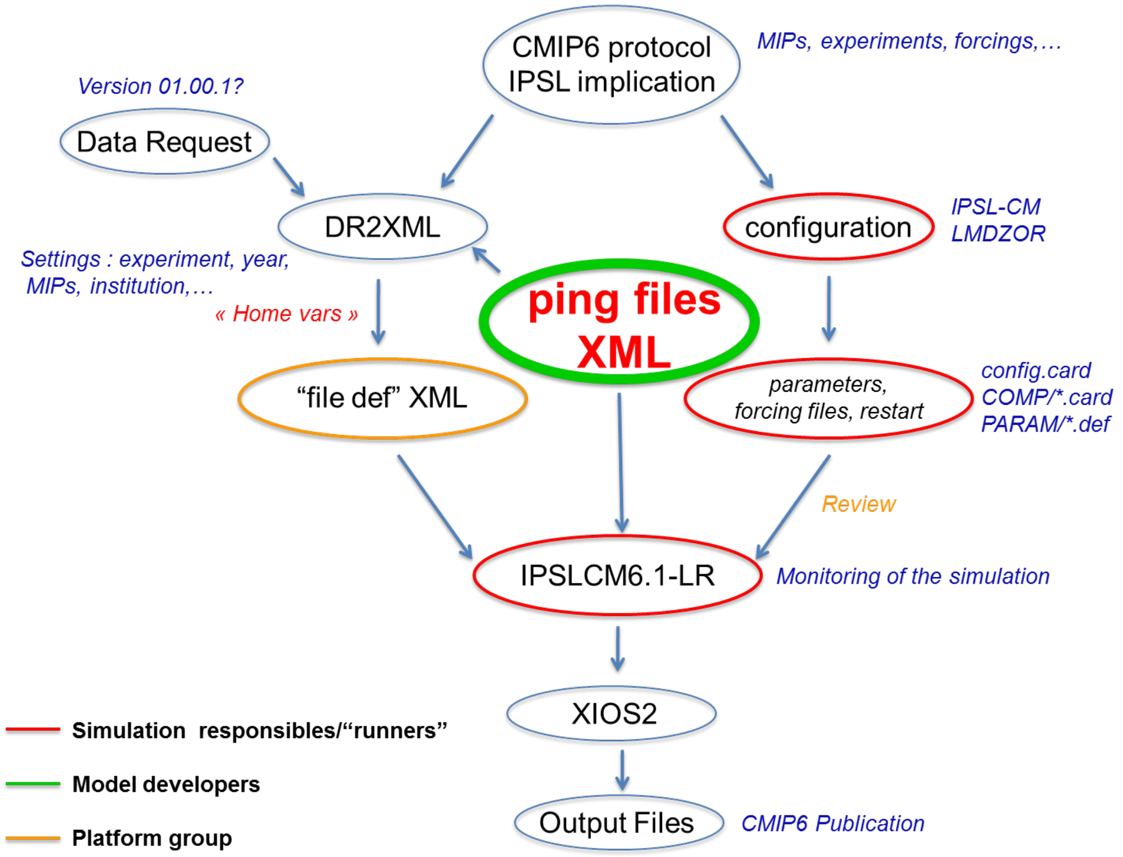
```
<file_definition>

 <file id="grid_T" name="grid_T" >
  <field field_ref="sst" name="sosstsst" />
  …
 </file>
 …
</file_definition>
```

CMIP6-publication-ready data files

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

CMIP6 "in situ diagnostics"

- Unit rescaling
- Normalization by area or level height
- Time integration (averaging, minimum, maximum)
- Vertical interpolation in pressure levels
- Extraction on specific pressure levels
- Vertical or global summation
- Horizontal remapping
- Zonal mean
- Diurnal cycle, seasonal means
- Cfsites (points station extraction)
- Transects (flux across ocean straight)
- Many more complex diagnostics (ex : Eliassen Palm flux)



CMIP6 protocol
IPSL implication

MIPs, experiments, forcings,…

Version 01.00.1?

Data Request

DR2XML

Settings : experiment, year, MIPs, institution,…

« Home vars »

configuration

IPSL-CM
LMDZOR

ping files
XML

"file def" XML

parameters, forcing files, restart

config.card
COMP/*.card
PARAM/*.def

Review

IPSLCM6.1-LR

Monitoring of the simulation

XIOS2

Output Files

CMIP6 Publication

🟥 Simulation responsibles/"runners"

🟩 Model developers

🟧 Platform group

~ 90 000 XML code line automatically generated by DR2XML for each experiment

A key for workflow portability : the ping file

- Map the model variable id to data request variable id
- Same workflow has been shared by IPSL and CNRM

IPSL CMIP6 Production

- 28 MIPs, 228 experiments, 850 simulations
- 55 000 years of simulation to perform
- 20 logins for the production campaign
- 200 logins for the climate model development and CMIP6 analysis
- in average 20 000 cores, with peak at 80 000 cores (one week at the end of Curie)
- 300 millions computing hours (development + production)
  - on TGCC (Curie and Irene) and few millions on IDRIS (development)
- 4 Pb of data produced
- 207 000 datasets published = 800 TB

Among the first group (with CNRM) to publish in ESGF

- Most of the experiments done with IPSL-CM6A-LR ~ 200km atm, 100km oce
  - About 16 SYPD on ~ 1000 cores (Intel sandy bridge)

Impact of the CMIP6 XIOS Workflow

- Between 10% and 20% depending of the experiment

Coupled model configuration

- IPSL-CM6A-LR
  - LMDZ –ORCHIDEE 144x143x79
  - NEMO (ORCA1-LIM3-PISCES)
- IPSL-CM6A-ATM-HR
  - LMDZ-ORCHIDEE 512x360x79
- IPSL-CM6A-AER-LR
  - LMDZ –ORCHIDEE 144x143x79
  - NEMO (ORCA1-LIM3-PISCES)
  - INCA (Aerosols)
- IPSL-CM5A2-CHM-VLR
  - LMDZ –ORCHIDEE 96x95x39
  - NEMO (ORCA2-LIM2-PISCES)
  - INCA (NMHC-AER-S)
- IPSL-CM7A-ATM-HR
- DYNAMICO-LMDZ-ORCHIDEE 50km (and 25 km)

The IS-ENES3 project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824084

Lessons learned from CMIP6 production workflow

Success :

- Handle complex workflow with ~ thousand of fields with numerous diagnostics
- Efficient management of time series and compression "on the flight" with ~ thousand of files
- Acceptable impact on models performance (~10% -> 20%)
- Remove the post-treatment phases

But :

- Big time overhead at starting time : some minutes depending the configuration
- Huge memory consumption, that increases with resolution
  - On clients and servers
  - need to depopulate nodes for servers
- Difficulties for parallel I/O to scale at large number of servers
- Workflow doesn't scale well at large number of cores (> 10 000)

HiResMIP production test

- IPSL : CMIP6 workflow works for 50 km and 25 km resolution, by reducing the level of output and depopulate
- CNRM : more difficulties to handle their ESM at such resolution

In the state of the art, a 10 km resolution full ESM with such outputs will be probably difficult to reach

All these bottlenecks should be solved in future to address the "exascale" era

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

IS-ENES3 : XIOS roadmap

Improve robustness and reliability

Improve XIOS error diagnostics

- Full stack is now output by the exception manager
- Full information (attribute) of the concerned object (field, file, etc…) is output all along the stack

✓ *Done in 2019*

Performance profiling logs

- detailed information to understand performance bottleneck

✓ *Done in 2019*

Output and visualize XIOS workflow graph

- Graphical view of spatial and temporal chained graph composing XIOS workflow
- Graphs generated at the end of execution
- Visualization within a standard web navigator

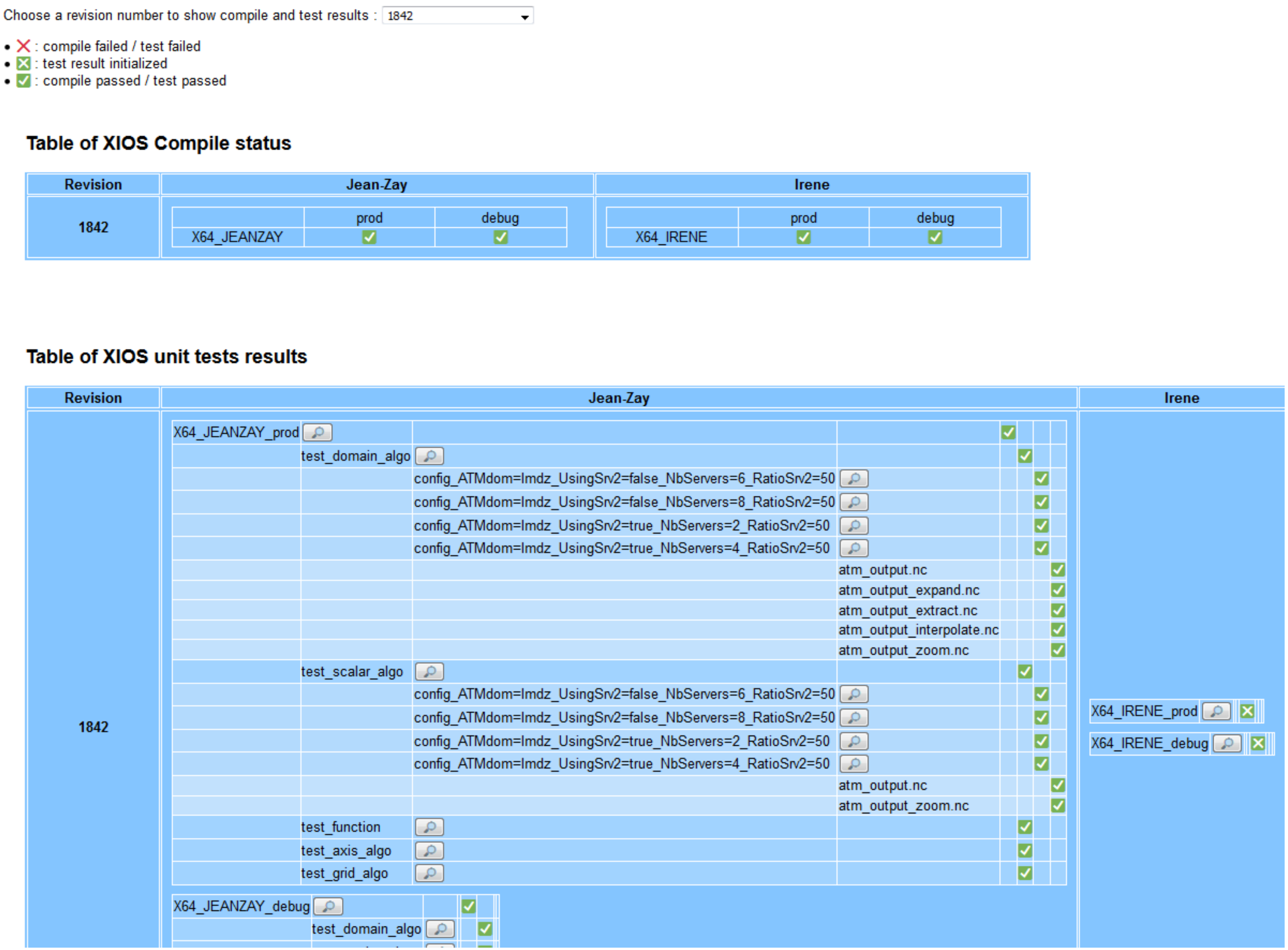
✓ *Done in 2019*

Development of a test case suite for contiguous integration

- Build a generic test case (binary) that can handle all XIOS functionalities:
  - Test all kind of mesh, including mesh indexation and mask
  - Test for fields on scalar, 1-D, 2-D, 3D or 4-D grid
- Run is defined by a set of parameters list
  - Nb models, nb proc for client, nb proc for servers, selected mesh
- All test case suite will be declined in unitary test and automated after each commit on different supercomputers
  - Compilation is also tested
- Results and regressions are exposed through a navigator

✓ *Done in 2019*
(IS-ENES3 milestone)

The IS-ENES3 project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824084



Choose a revision number to show compile and test results : 1842

- ✗ : compile failed / test failed
- ☒ : test result initialized
- ✓ : compile passed / test passed

Table of XIOS Compile status

Revision	Jean-Zay			Irene		
1842		prod	debug		prod	debug
	X64_JEANZAY	✓	✓	X64_IRENE	✓	✓

Table of XIOS unit tests results

Revision	Jean-Zay								Irene
1842	X64_JEANZAY_prod 🔍				✓				X64_IRENE_prod 🔍 ☒ X64_IRENE_debug 🔍 ☒
		test_domain_algo 🔍				✓			
			config_ATMdom=lmdz_UsingSrv2=false_NbServers=6_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=false_NbServers=8_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=true_NbServers=2_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=true_NbServers=4_RatioSrv2=50 🔍				✓		
				atm_output.nc				✓	
				atm_output_expand.nc				✓	
				atm_output_extract.nc				✓	
				atm_output_interpolate.nc				✓	
				atm_output_zoom.nc				✓	
		test_scalar_algo 🔍				✓			
			config_ATMdom=lmdz_UsingSrv2=false_NbServers=6_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=false_NbServers=8_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=true_NbServers=2_RatioSrv2=50 🔍				✓		
			config_ATMdom=lmdz_UsingSrv2=true_NbServers=4_RatioSrv2=50 🔍				✓		
				atm_output.nc				✓	
				atm_output_zoom.nc				✓	
		test_function 🔍				✓			
		test_axis_algo 🔍				✓			
		test_grid_algo 🔍				✓			
	X64_JEANZAY_debug 🔍				✓				
		test_domain_algo 🔍				✓			

IS-ENES3 : XIOS roadmap

Improvement of the internal time line management

Implementing time interpolations

- Remove current limitation : temporal filters are applied at a multiple frequency of model time step
- Time interpolation filter will uncoupled the XIOS workflow from the models time step.
- A lot of practical examples…
  - Enable models with variable time step
  - For reading, a monthly file can be interpolated daily before to be injected into model

Targeted mid 2021

Implementing XIOS restartability

- Currently XIOS is not restartable
  - Model can be stop only at a multiple of the highest frequency of the time filters (averaging)
- Will enable models and XIOS workflow to be shut down at any time and then restarted
  - Longer averaging frequency (yearly means)
  - Decadal seasonal means

Targeted end 2020

Improving spatial filters

- Implement more complex spatial filters by chaining internally already developed primary filters
  - Zonal means, grad , div and curl filters…
- Efficient station output management
- Implement still missing remapping operator
  - nearest neighbor, bilinear

Targeted end 2021

The Consortium

Coordinated by **CNRS-IPSL**, the IS-ENES3 project gathers **22 partners** in **11 countries**

cnrs with cea Sorbonne Université

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

DKRZ
DEUTSCHES KLIMARECHENZENTRUM

BSC

CERFACS

cmcc
Centro Euro-Mediterraneo sui Cambiamenti Climatici

Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

SMHI

UK Research and Innovation

UC
UNIVERSIDAD DE CANTABRIA

DLR

netherlands eScience

Norwegian Meteorological Institute

METEO FRANCE
Toujours un temps d'avance

MANCHESTER 1824
The University of Manchester

ΔΗΜΟΚΡΙΤΟΣ

WAGENINGEN
UNIVERSITY & RESEARCH

CHARLES UNIVERSITY

NORCE

LiU LINKÖPING UNIVERSITY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084

Our website
https://is.enes.org/

Follow us on Twitter !
@ISENES_RI

Contact us at
is-enes@ipsl.fr

zenodo

Join the community on ZENODO !