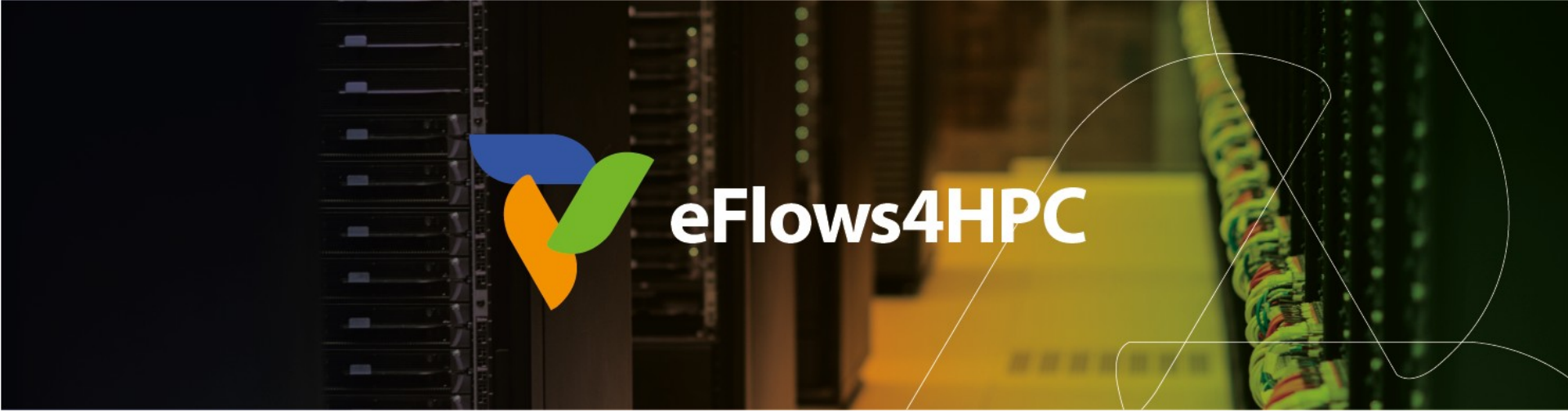
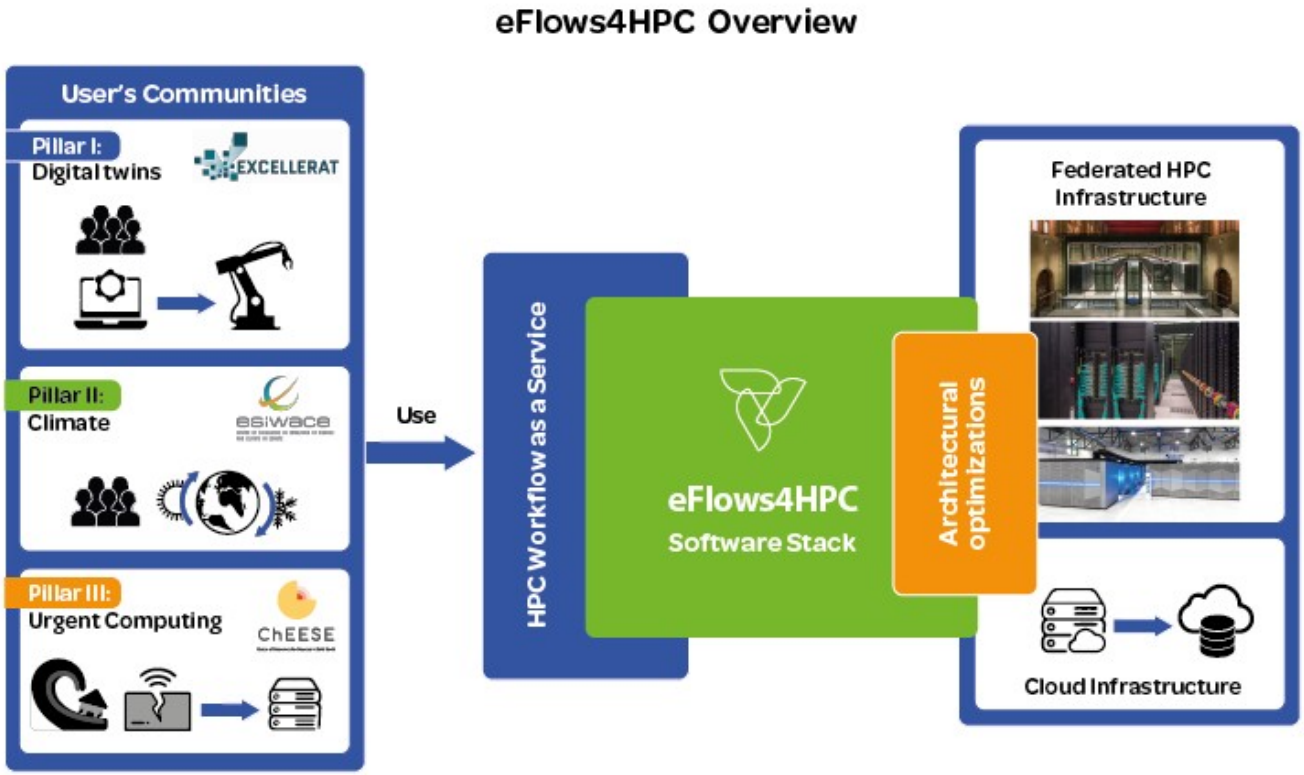


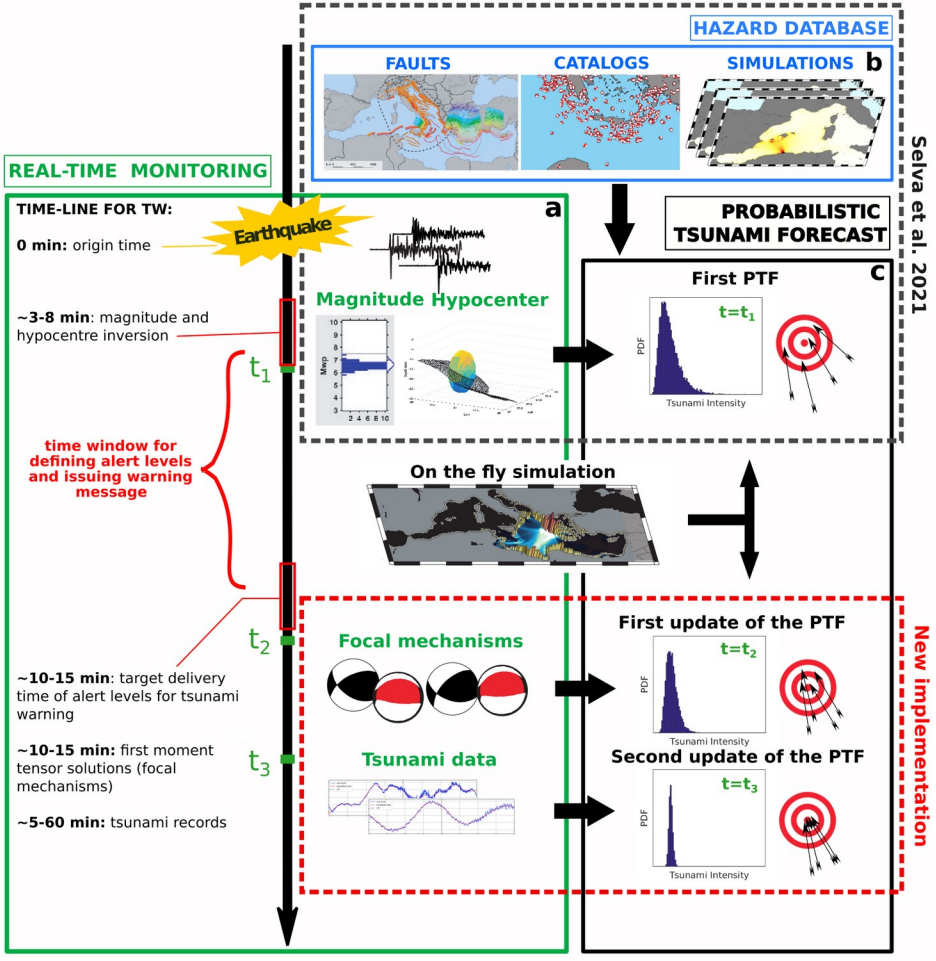
Complete workflow for tsunami simulation and hazard calculation in urgent computing using HPC services

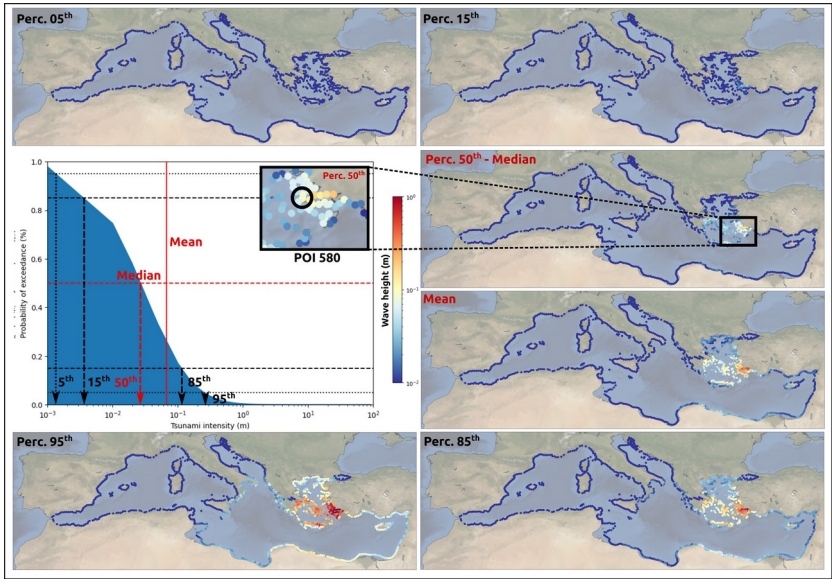
Cordrie L., Selva J., Ejarque J., Bernardi F., Tonini R., Sánchez Linares C., Macías J., Gibbons S., Løvholt F.

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955558. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Spain, Germany, France, Italy, Poland, Switzerland, Norway.

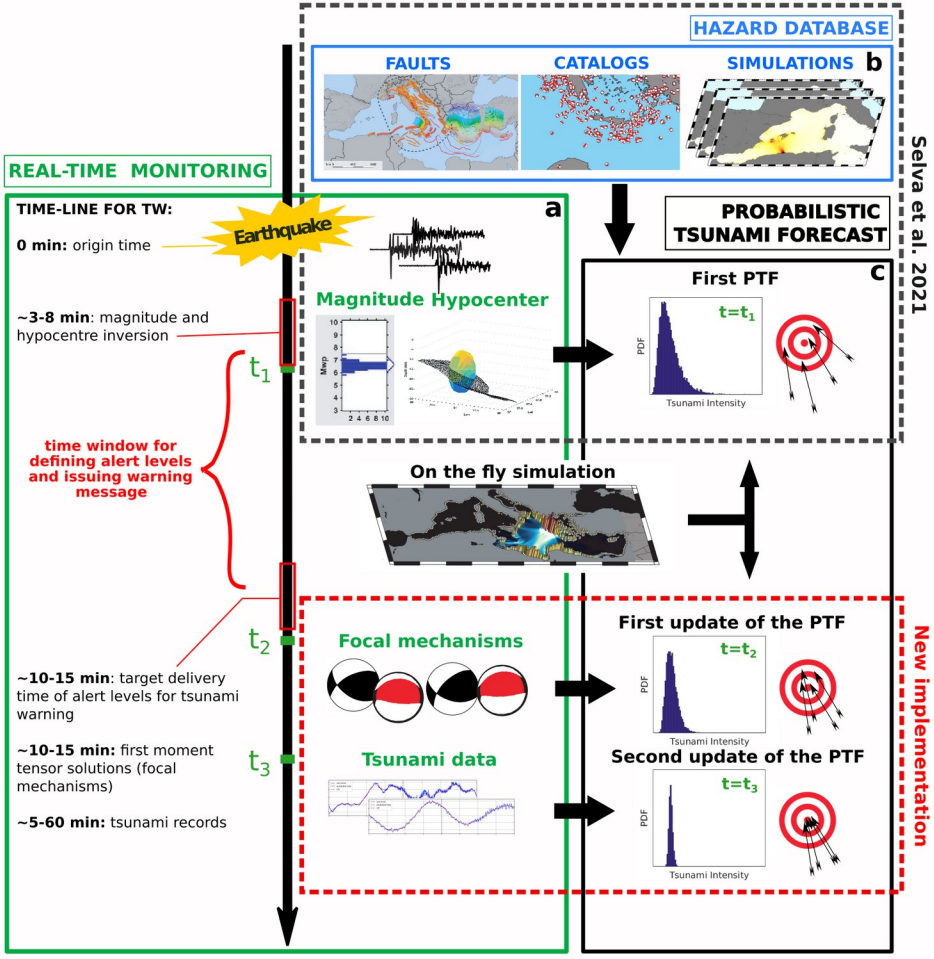

eFlows4HPC Overview
User's Communities
Pillar I:
Digital twins
EXCELLERAT
Pillar II:
Climate
esiwace
Pillar III:
Urgent Computing
ChEESE
Use
HPC Workflow as a Service
eFlows4HPC
Software Stack
Architectural optimizations
Federated HPC Infrastructure
Cloud Infrastructure

Probabilistic Tsunami Forecast (PTF) workflow





PTF workflow adjustments for HPC system



Objectives

- **Adapting the PTF workflow to the PyCOMPSS manager**
- **Reducing ensemble size** (monte-carlo sampling methods)
- **Integrating data update** (focal mechanism, tsunami data) for actualizing in real-time the forecasts
- **Delivering intermediate PTF results**
- **Integration of additional tools:** Ophidia, Datalogistic service...

Pycompss workflow manager

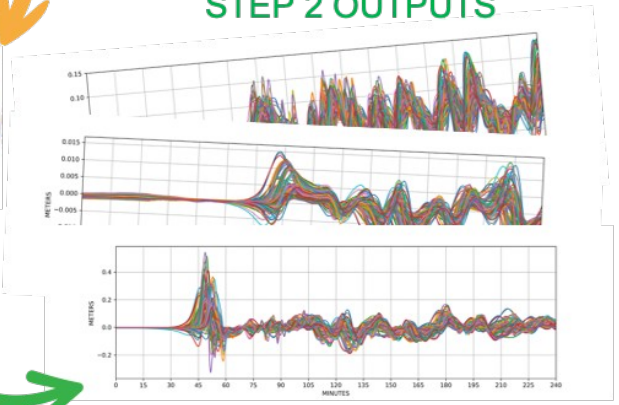
Ensemble (from step1)

SOURCE

Tsunami-HySEA
MC version

- Ophidia : - post-processing
  - data storage

STEP 2 OUTPUTS



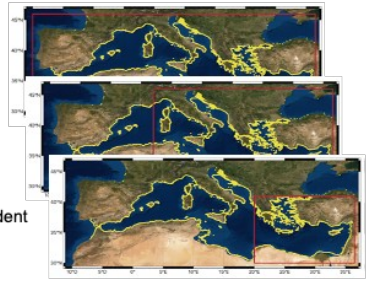
SCENARIO PARAMETRIZATION

- Computational domains
- Numerical resolutions

Pre-design files

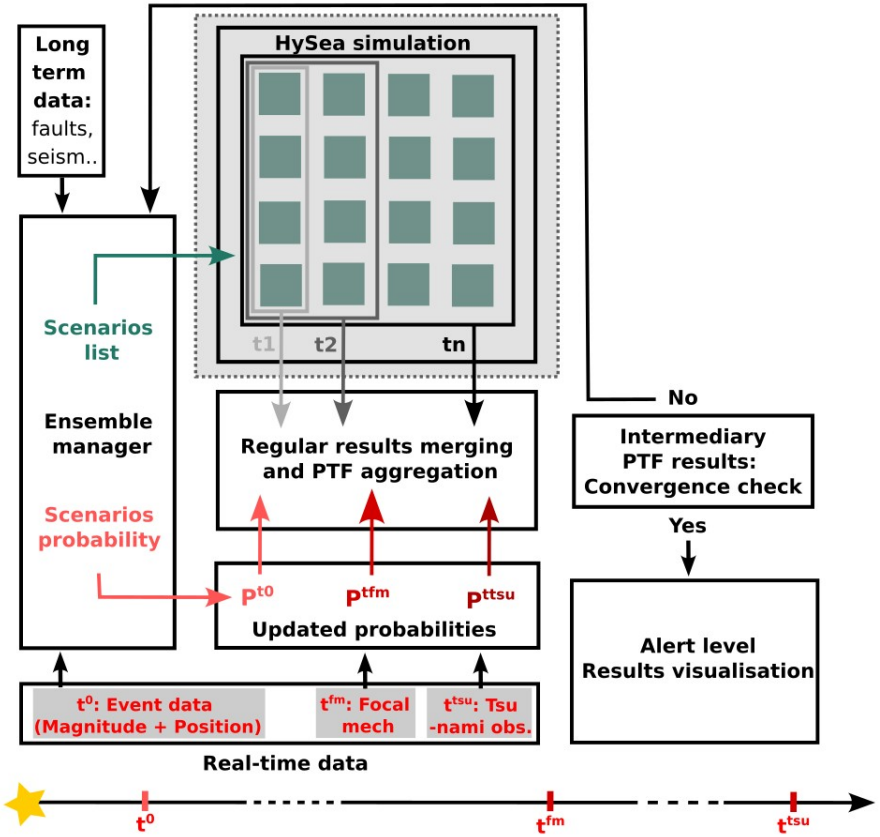
- Simulation time
- Points of interest
- Output files

Event dependent

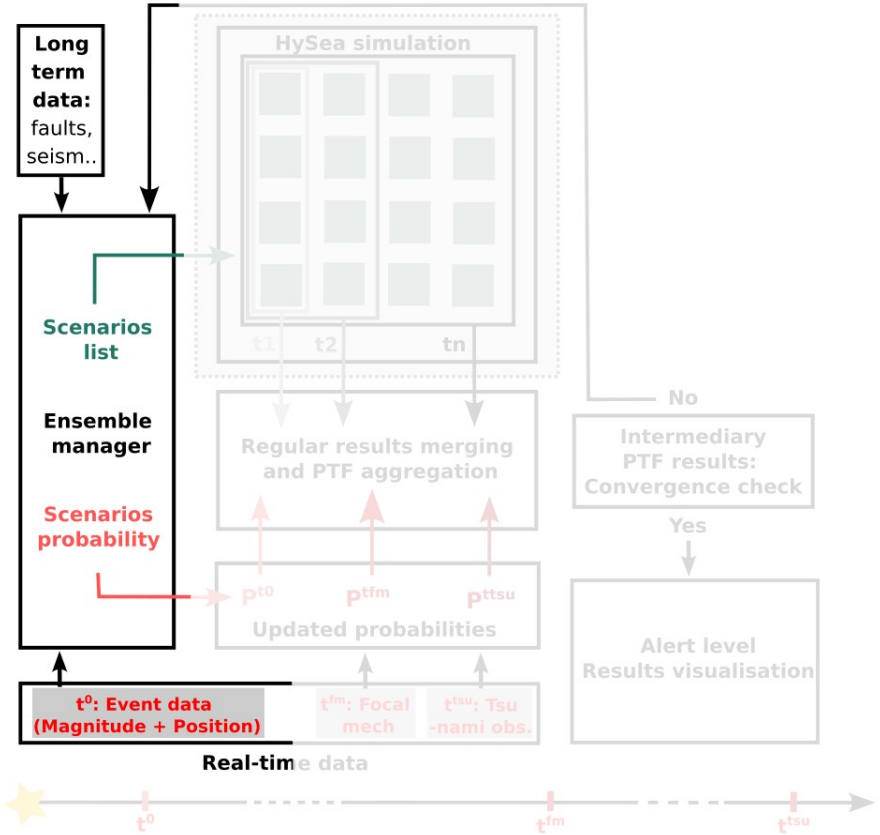


- OPTIMIZATION OF HPC RESOURCES
- AUTOMATIZATION ON DATA DEPENDENCIES

Pycompss workflow



Pycompss workflow: Step1, ensemble manager



PyCompss workflow

@task1: python script
Scenarios and probability list

@task2: binaries (.sh files)
Preparation of output folders and parameters text files for HySea simulation

Loop on the list of scenarios:

@task3: mpi task
HySea binary
NetCDF output in each scenario's folder

@task4: python script
First post processing of individual scenario's output

+ update of a **shared file**
(Commutative)

Every N scenarios added to the shared file:

@task5: python script
Second postprocessing of the N scenarios outputs

option activated by a data listener

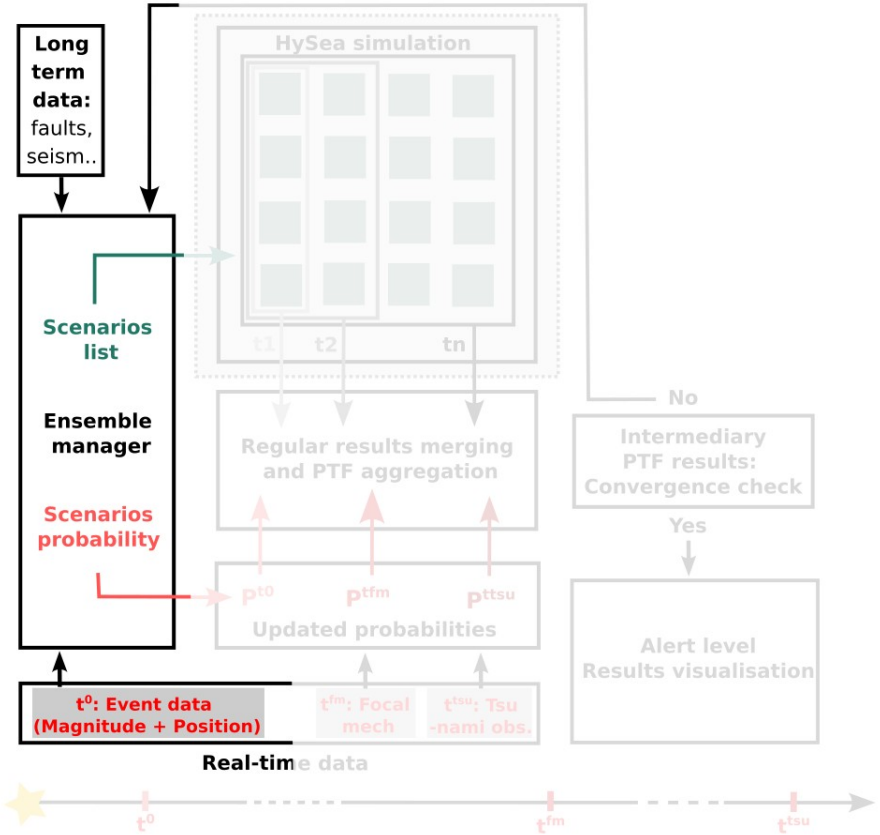
@task6: python script
Update of the probabilities (FM or Tsunami data)

@task7: python script
Aggregation of the N scenarios results and prob to create hazard curves

Arrival of data: Focal mech. or Tsunami data

t1: Int. PTF (N scenarios)
t2: Int. PTF (2*N scenarios)
t?: Int. PTF (...*N scenarios)
tf: Final PTF (All scenarios)

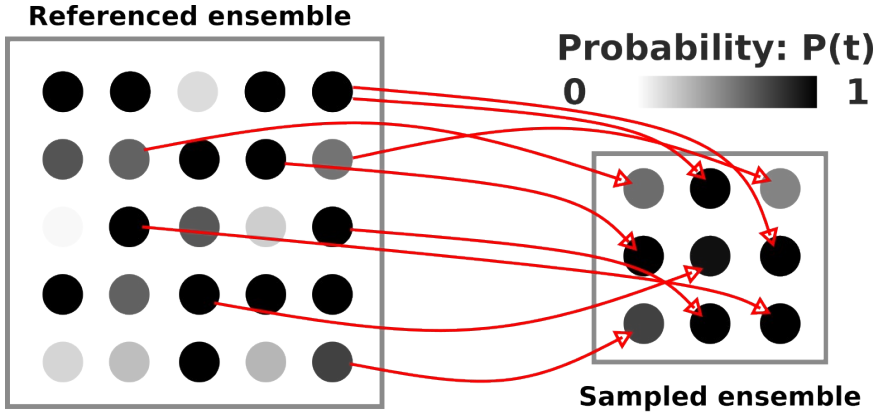
Pycompss workflow: Step1, ensemble manager



PyCompss workflow

@task1: python script
Scenarios and probability list

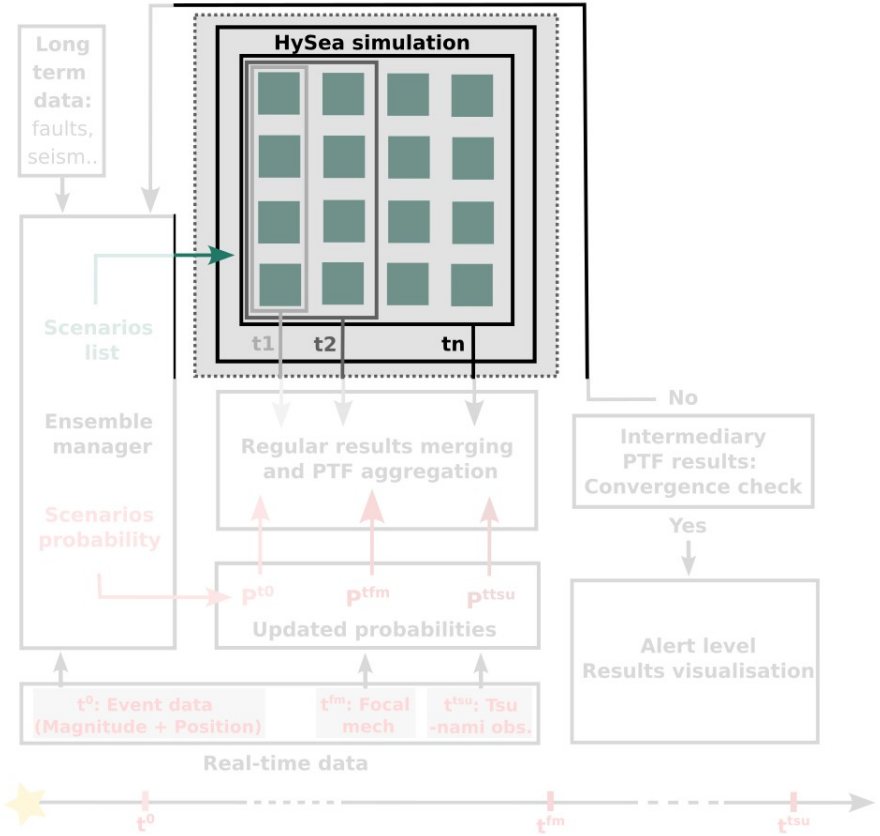
@task2: binaries (.sh files)
Preparation of output folders and parameters text files for HySea simulation

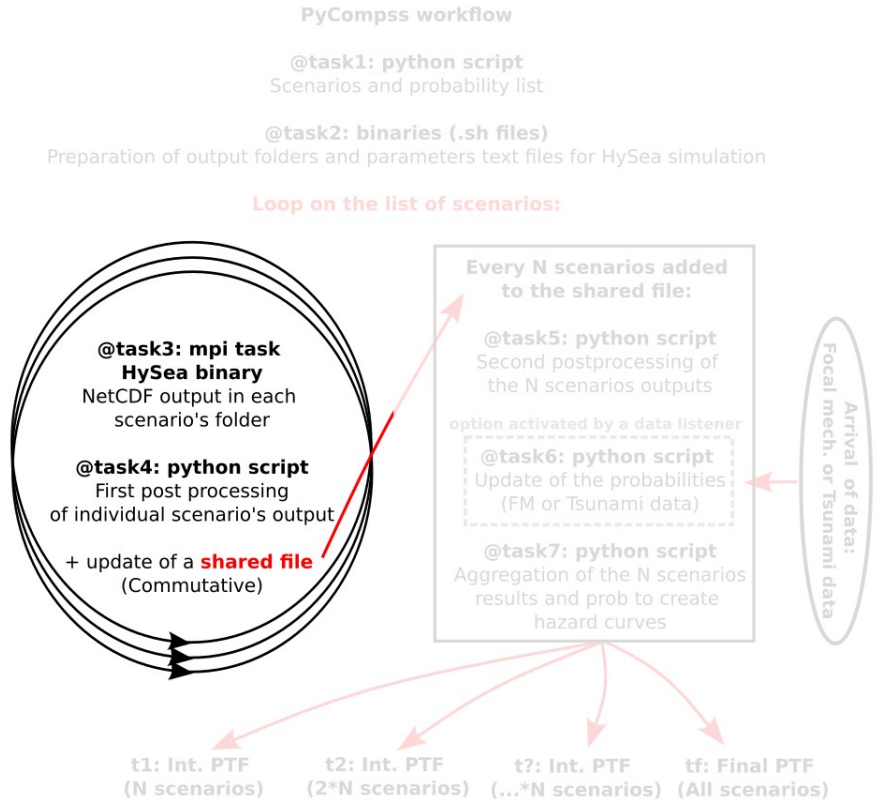


Monte-Carlo sampling method

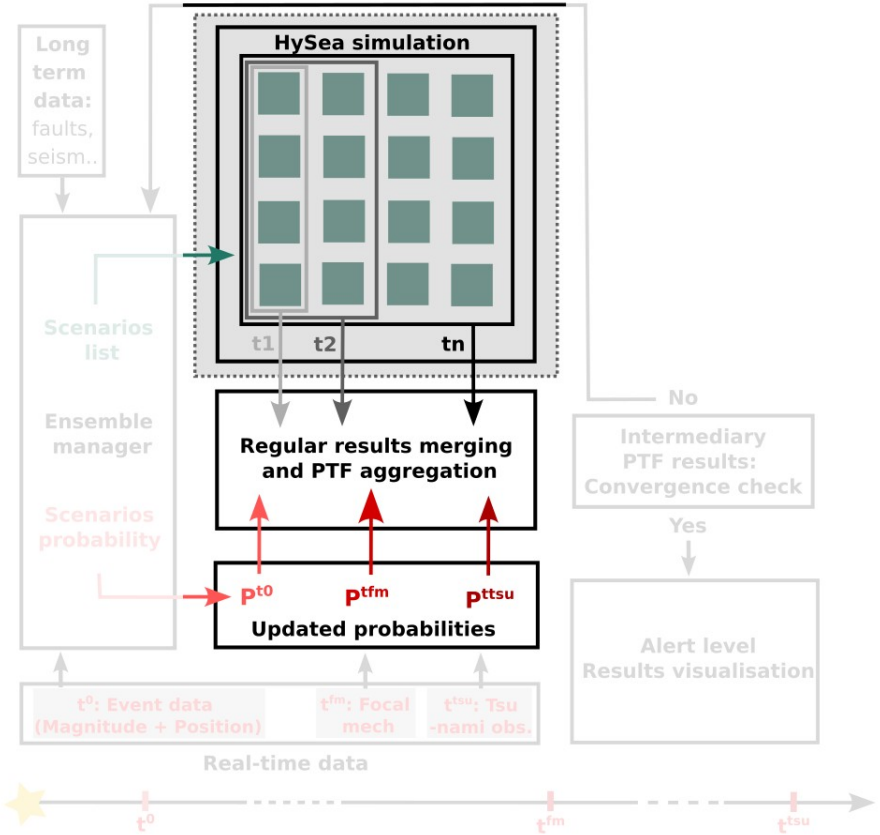
10000-50000 scenarios → 500 scenarios

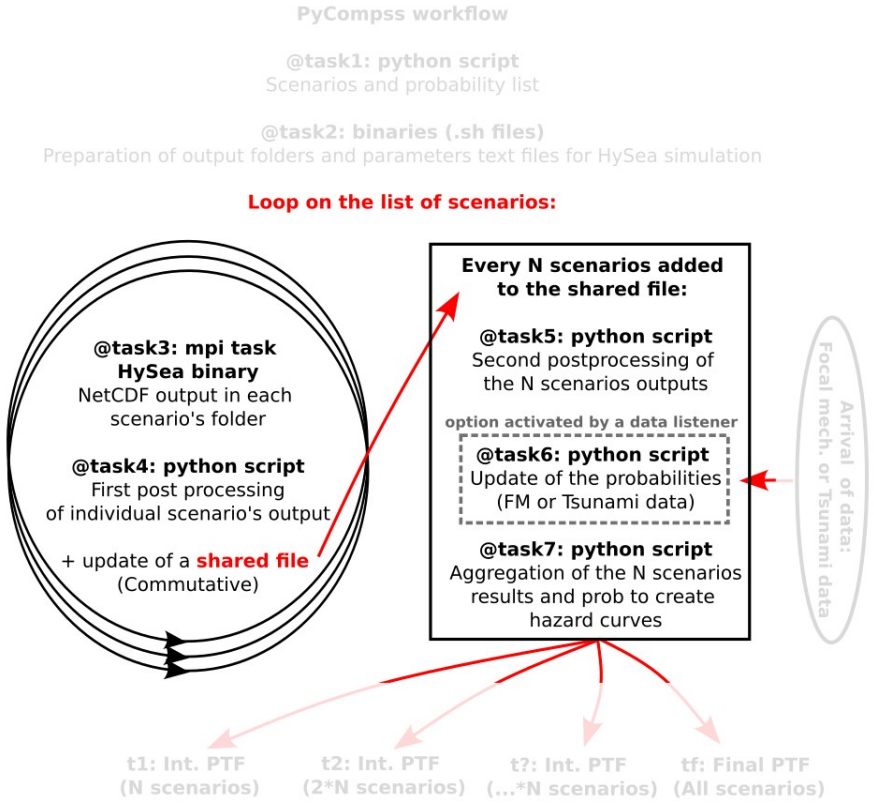
Pycompss workflow: Step2, HySea simulations



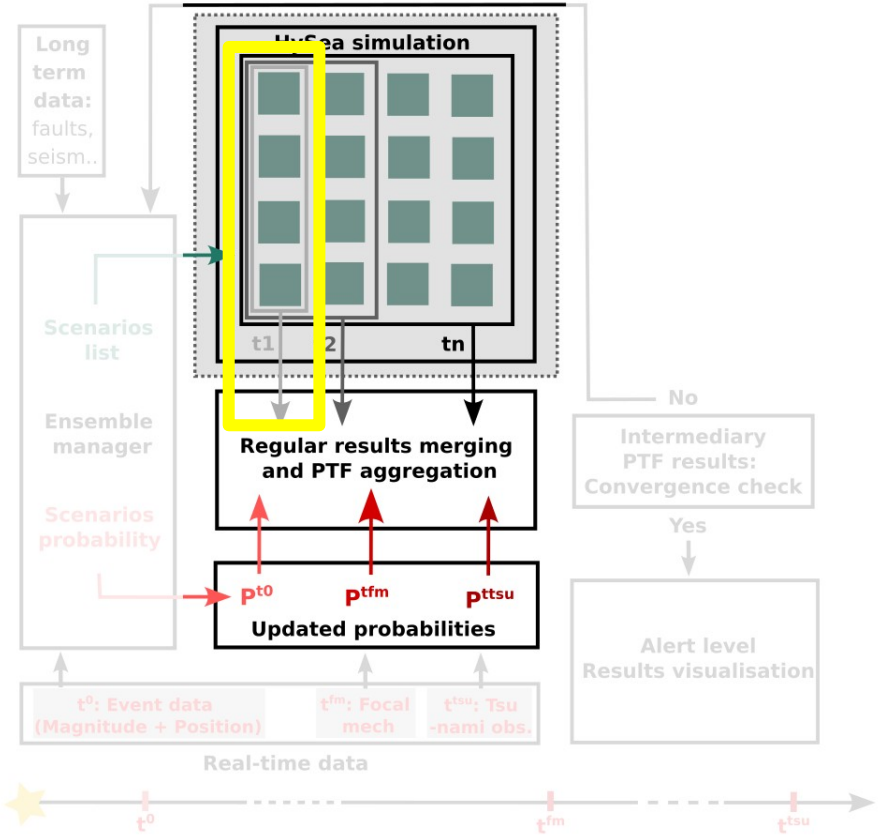


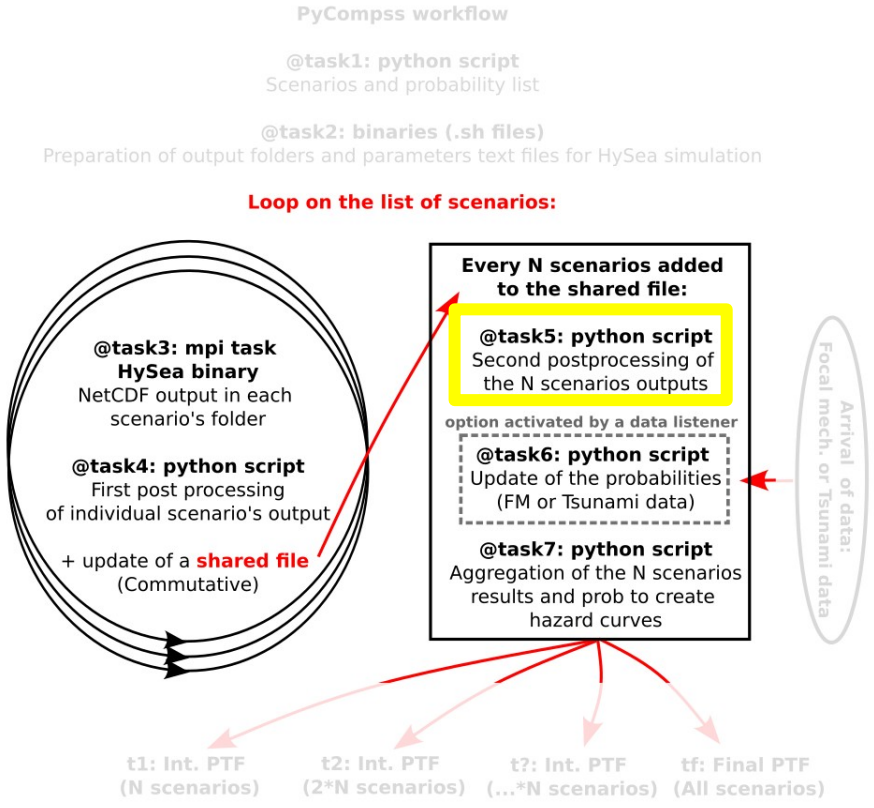
Pycompss workflow: Step3, post-processing of the output



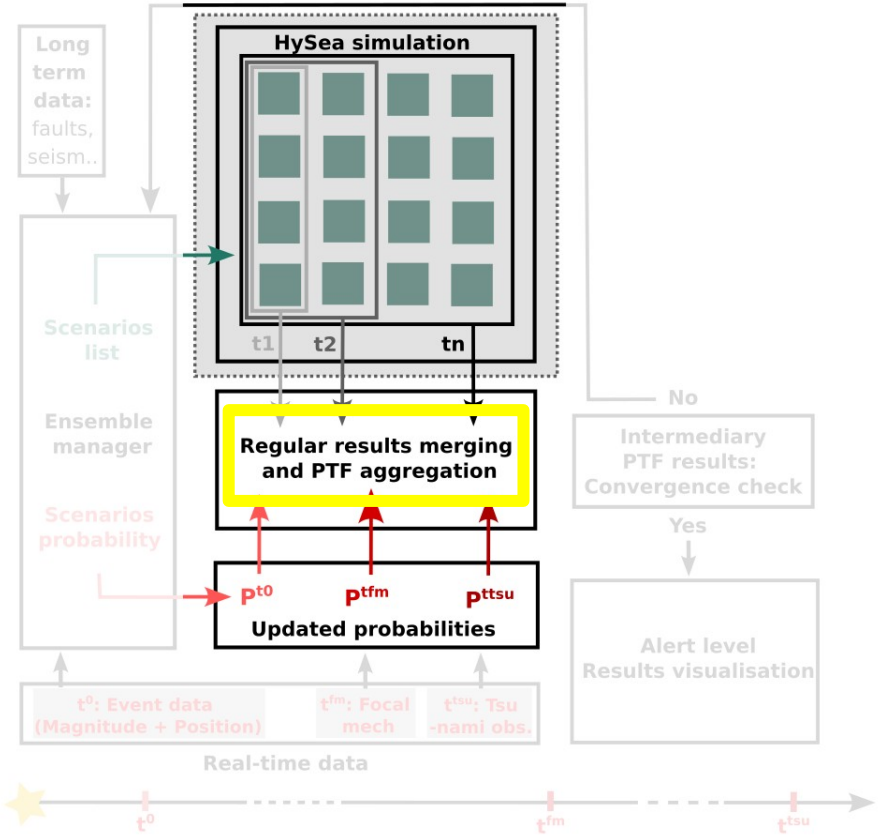


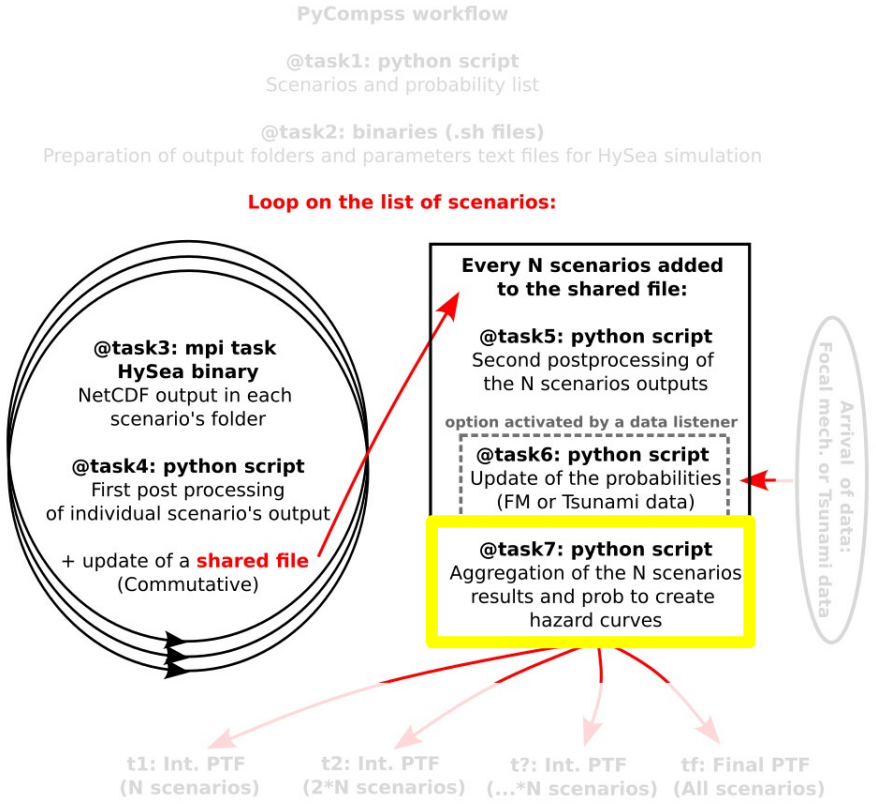
Pycompss workflow: Step3, post-processing of the output



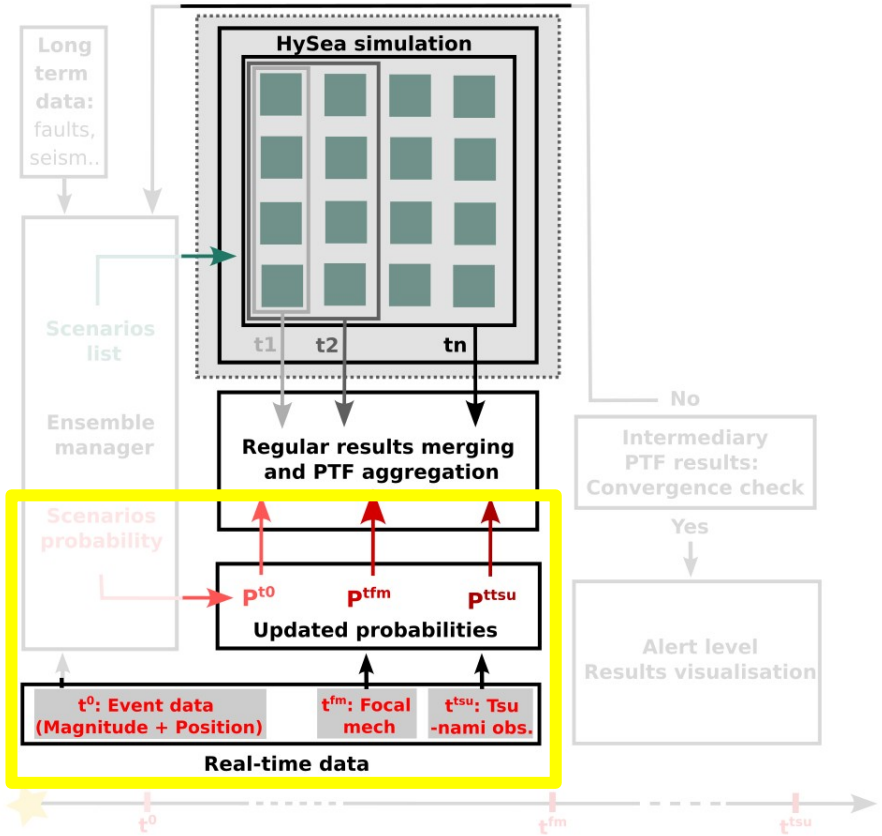


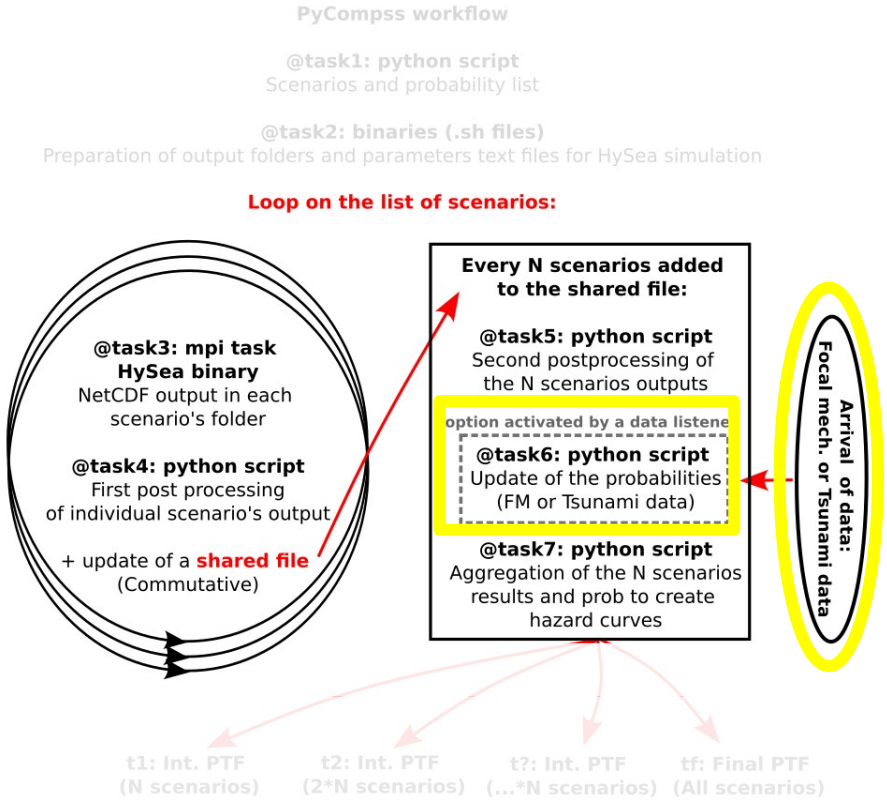
Pycompss workflow: Step3, post-processing of the output



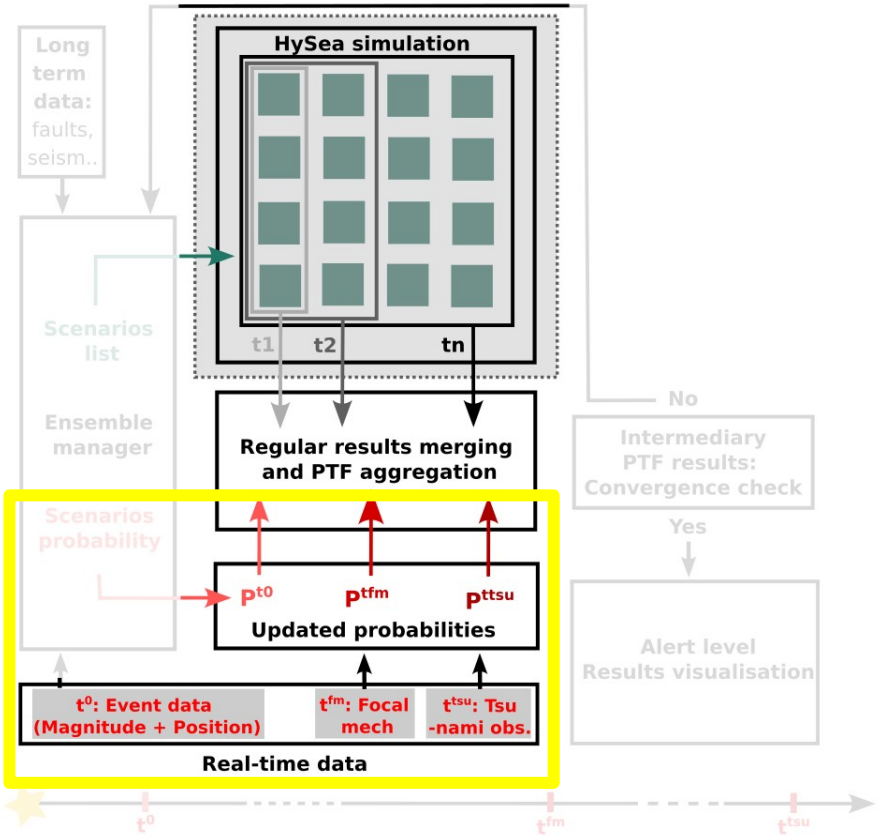


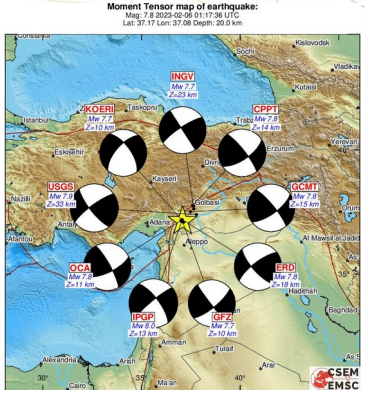
Pycompss workflow: Data-driven update of the PTF

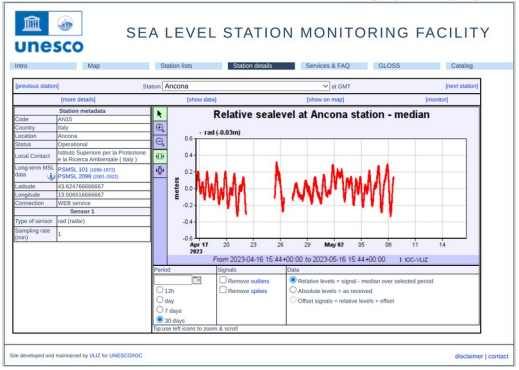


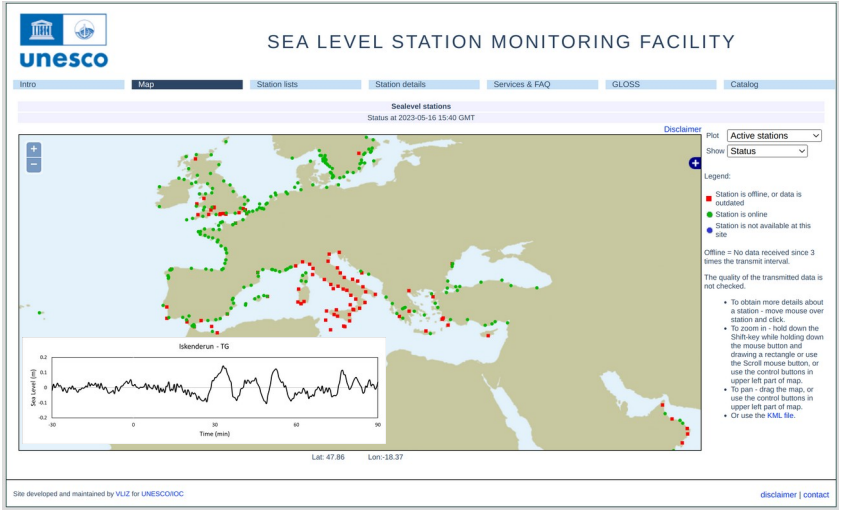


Pycompss workflow: Data-driven update of the PTF

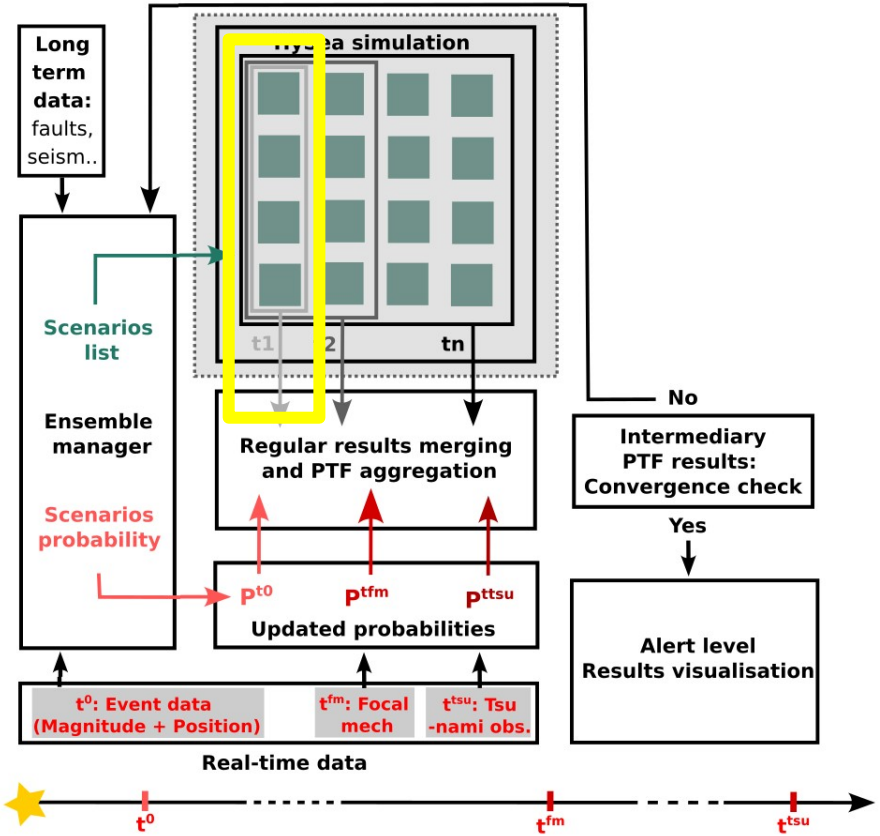


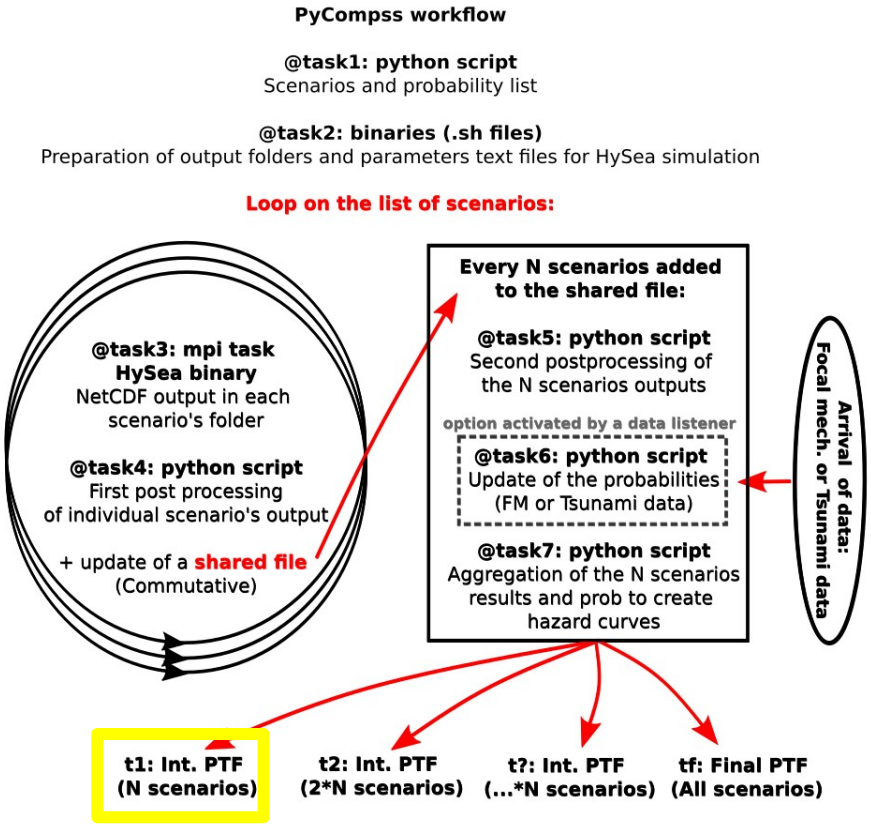




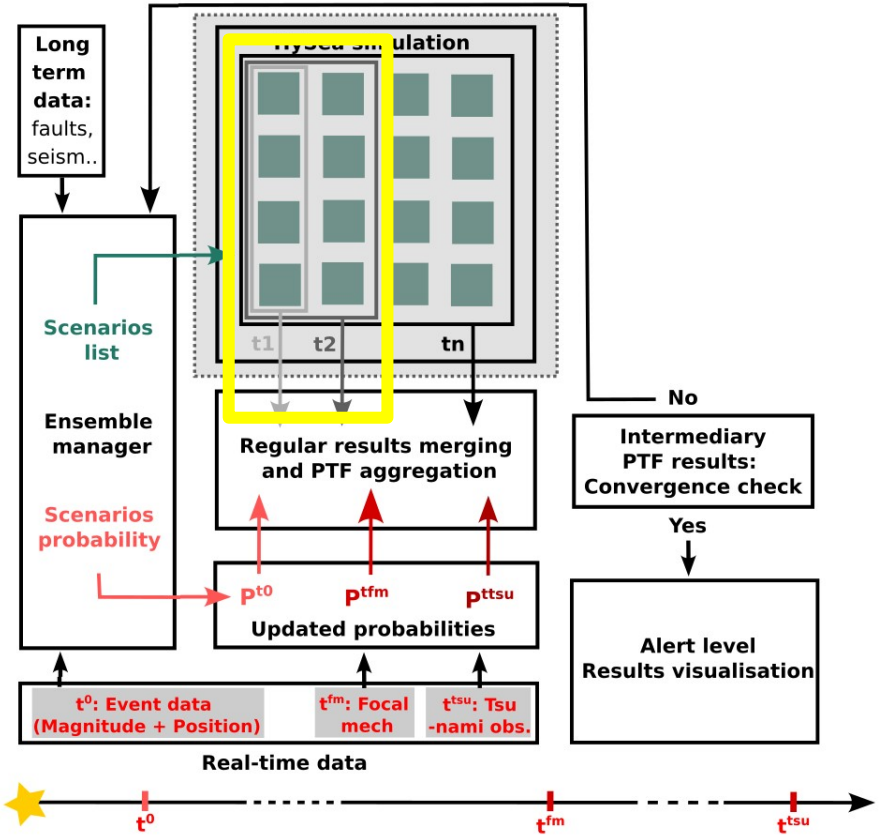


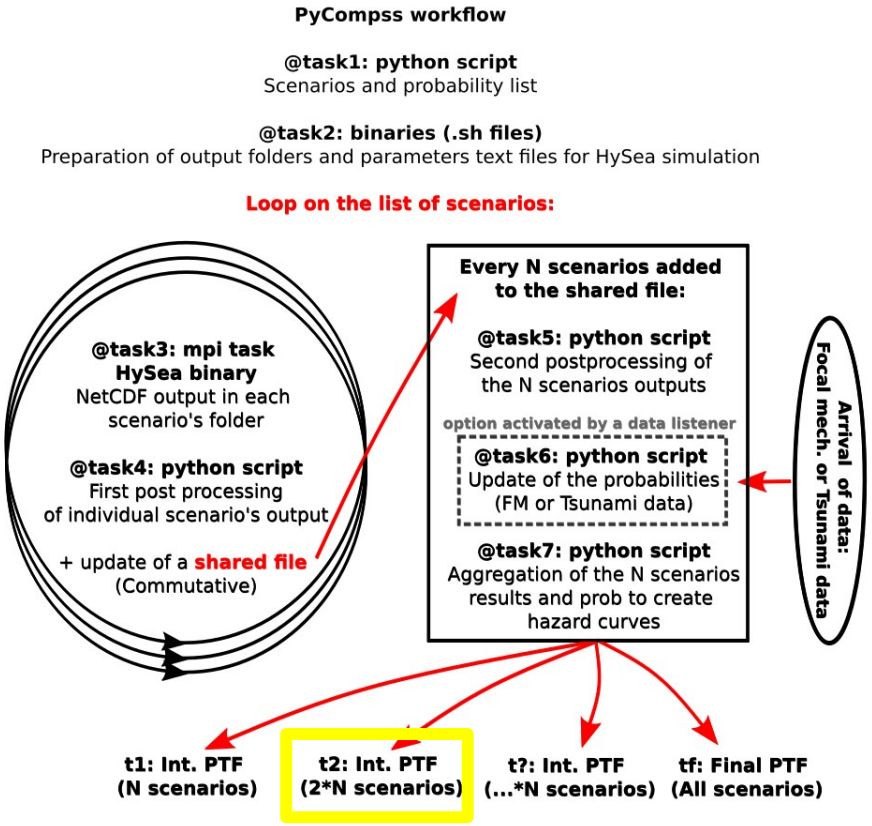
Pycompss workflow: Step3, intermediate evaluation of the PTF



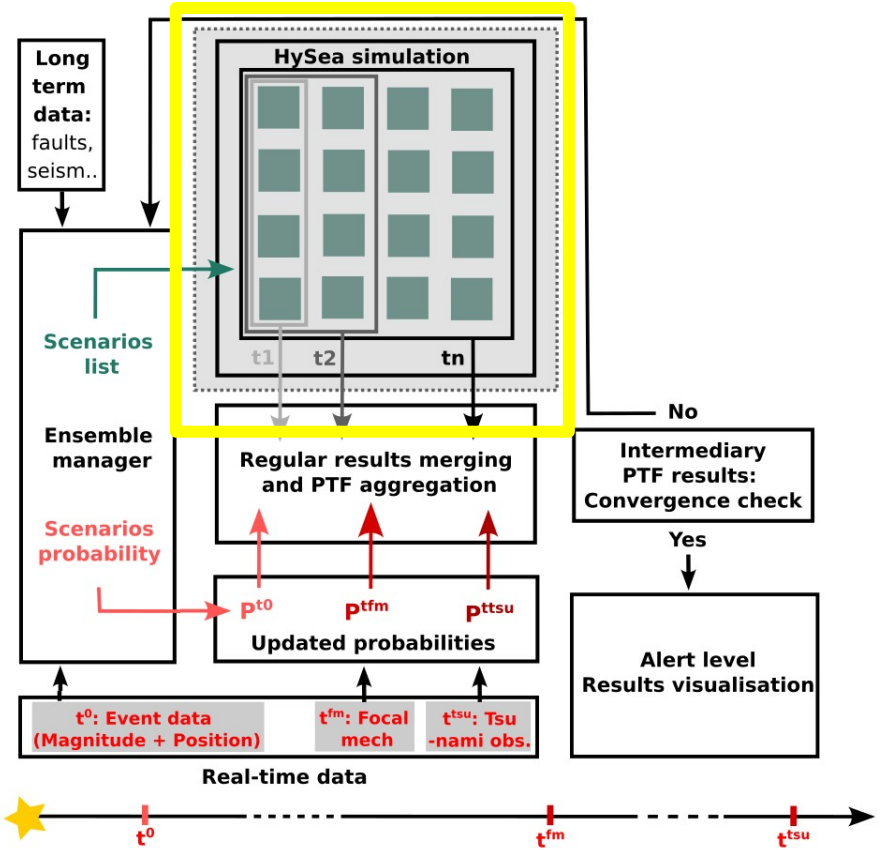


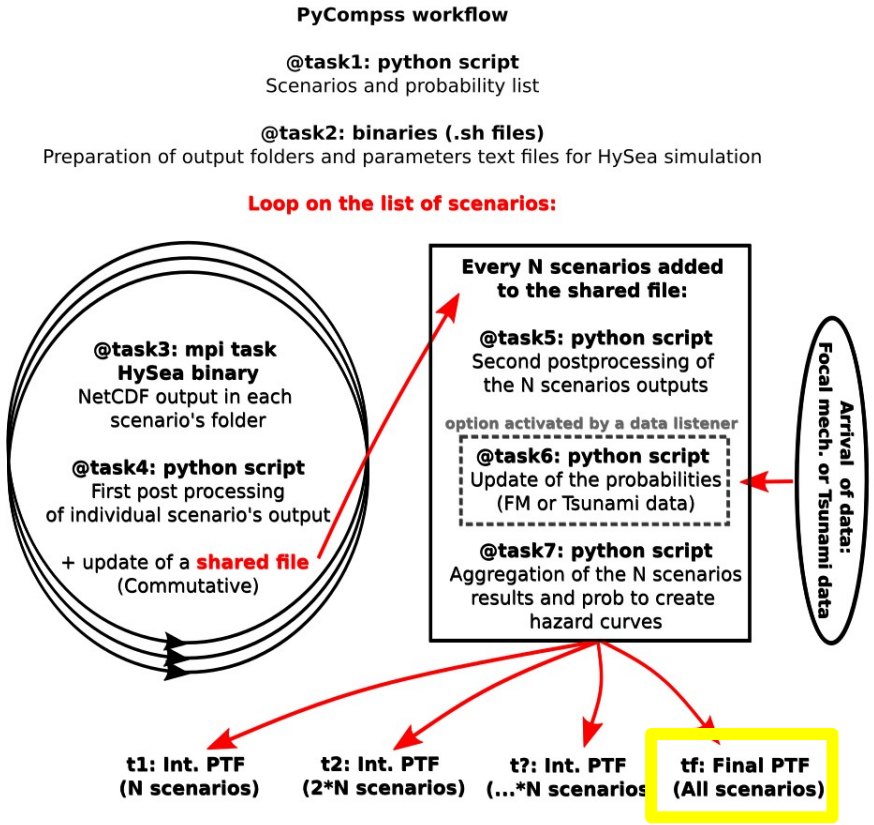
Pycompss workflow: Step3, intermediate evaluation of the PTF



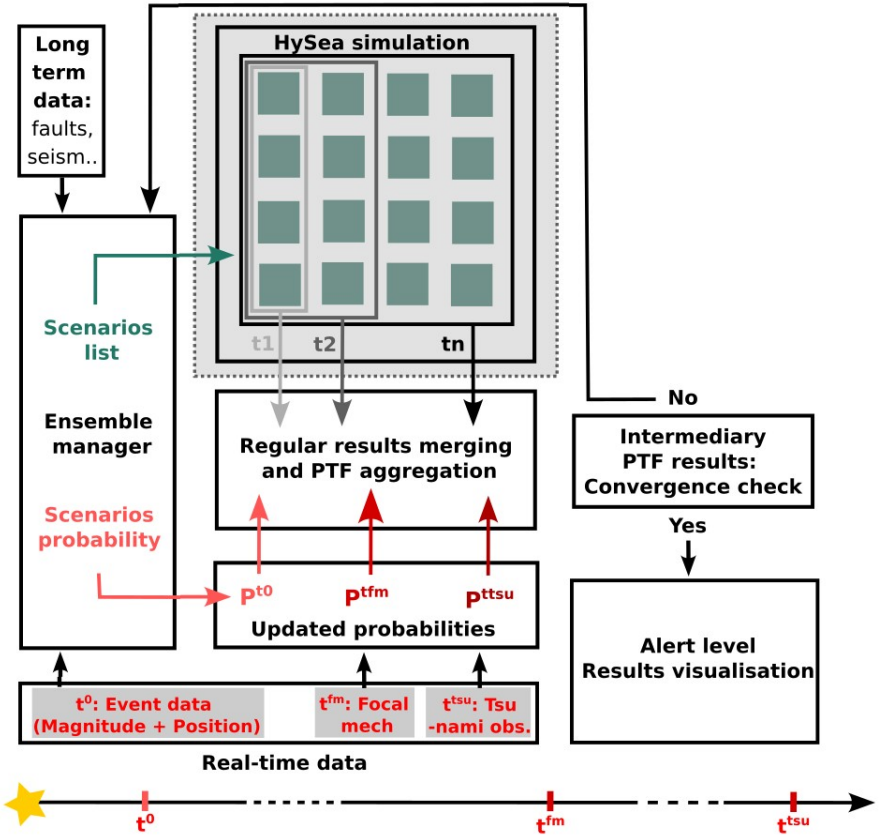


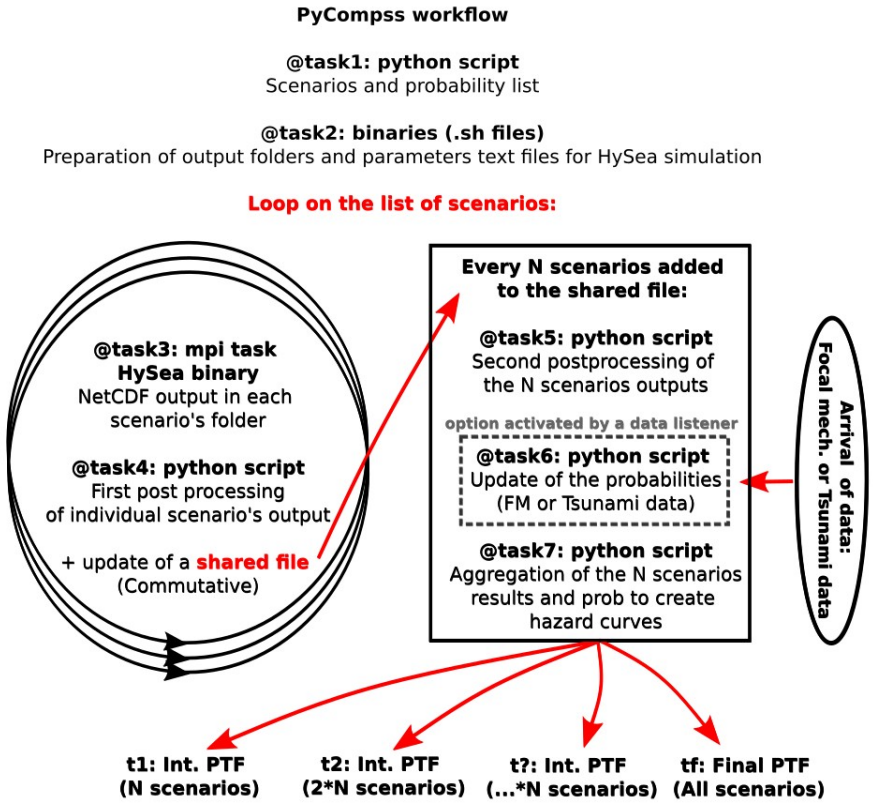
Pycompss workflow: Step3, intermediate evaluation of the PTF



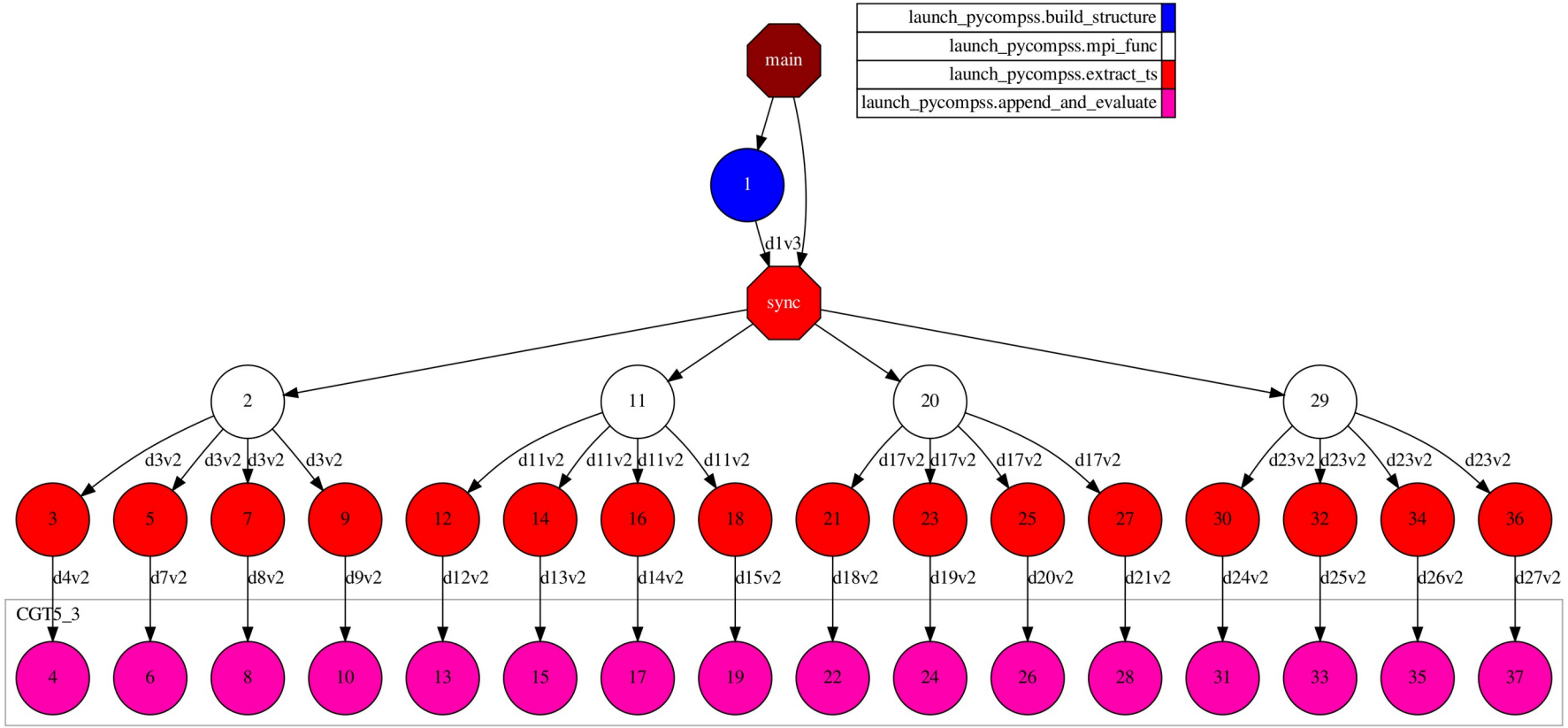


Pycompss workflow





Pycompss manager



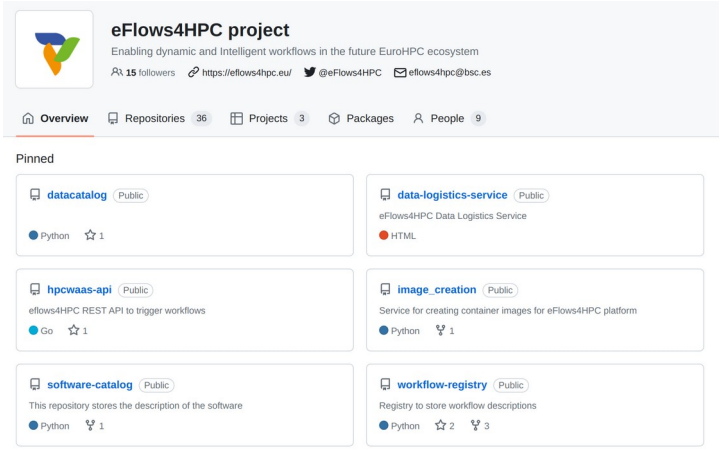
Portability of the workflow to any HPC system

eFlows4HPC software stack

Data Analytics and Storage:
→ Data analytics: Ophidia
→ Data INPUT Catalog
→ Data OUTPUT repository: B2drop

Portability and logistic of the workflow:
→ TOSCA description
→ Data pipelines, Data logistic service
→ Docker container

Workflow registery:
→ eFlows4HPC github



Use of HPC services for tsunami simulation

Tsunami-HySEA

- **Reduction of the ensemble size:**

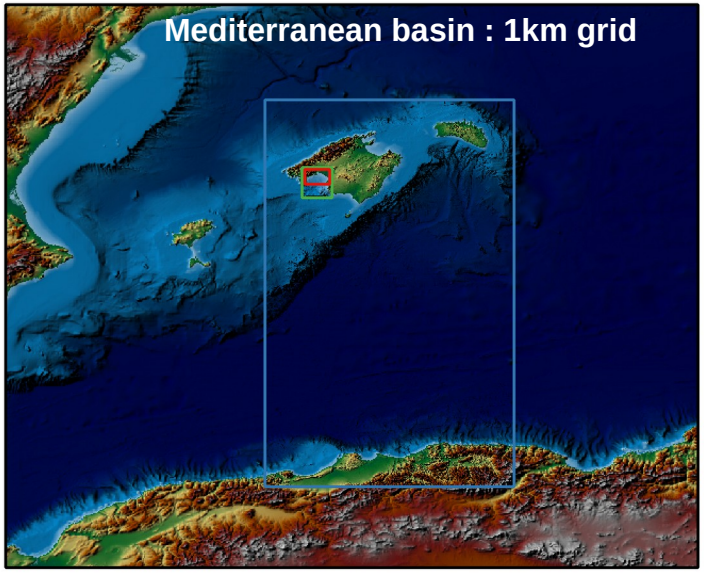
  Reduction from 50000 to 500 scenarios

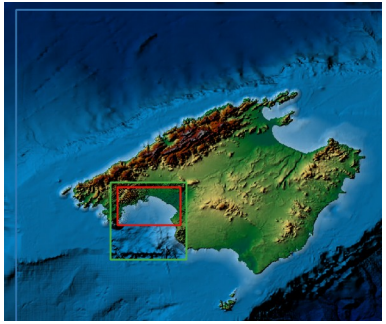
- **GPU FTRT code:**

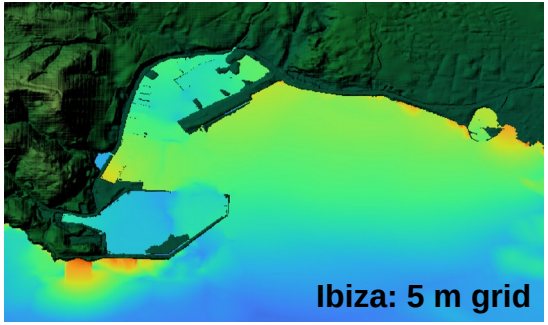
  Normal grid:
  1 simu ~ 2min GPU time
  WF for 50000 simu on 150 nodes → 2-3h
  WF for 500 simu on 150 nodes → 5-10min

  → **important in emergency context**

  High-resolution nested grids:
  1 simu ~ 1h GPU time
  WF for 500 simu on 150 nodes → 4-5h







Conclusion

Achievements

- PyCOMPSS manager
- Sampling and updating methods implemented
- Intermediate PTF delivery
- Workflow tested and validated for 4 different events (Boumerdes 2004, Kos-Bodrum 2017, Samos 2020, Turkey 2023)
- Ready to be ported outside BSC: containerization

Prospects

- Porting the workflow and testing it with the data logistic, streams and storage
- Integrating high-resolution grids and inundation calculation
- Testing to additional events and regions

Thank you

Cordrie L., Selva J., Ejarque J., Bernardi F., Tonini R., Sánchez Linares C., Macías J., S. Gibbons, F. Løvholt

eFlows4HPC partners: