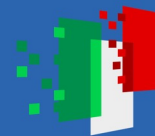




Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



terabit

PRACE-Italy use case

Workflow implementation

Giorgio Bolzon Elisa Zuccolo

Fabio Pitari Lucia Rodriguez

Bologna, 26 Giugno 2024

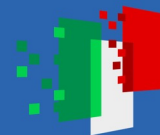
TeRABIT Workshop



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



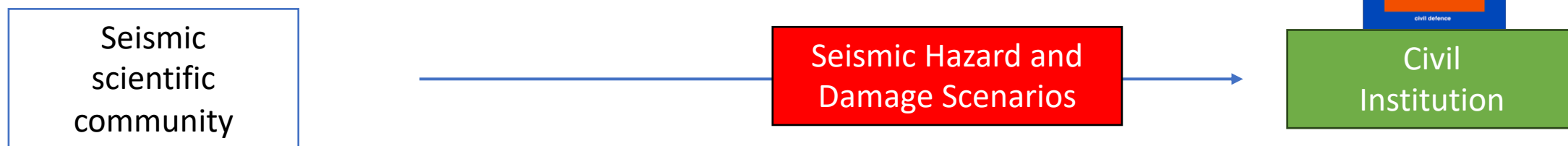
Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



OUTLINE

- Technical presentation of PRACE-Italy seismic use case
- Definition of the problem and set up of the workflow
- Solutions adopted within PRACE-Italy infrastructure
- DEMO
- Evolutions & strategies

FROM SCIENTIFIC PURPOSE TO THE WORKFLOW

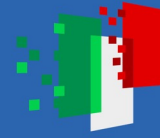


Challenge: scale all that work up to URGENT COMPUTING level

- Collection of codes
- Several input data
- Human work
 - Clusters
 - Local servers
 - Local PCs



AUTOMATIC WORKFLOW



WORKFLOW SETUP

Black Box



It's only a computational problem!

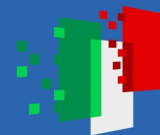
2h on 1000 cores



Problem 1 – AUTOMATIZATION !

Problem 2 – REDUCE THE TIME TO SOLUTION

Good solutions
Trade offs



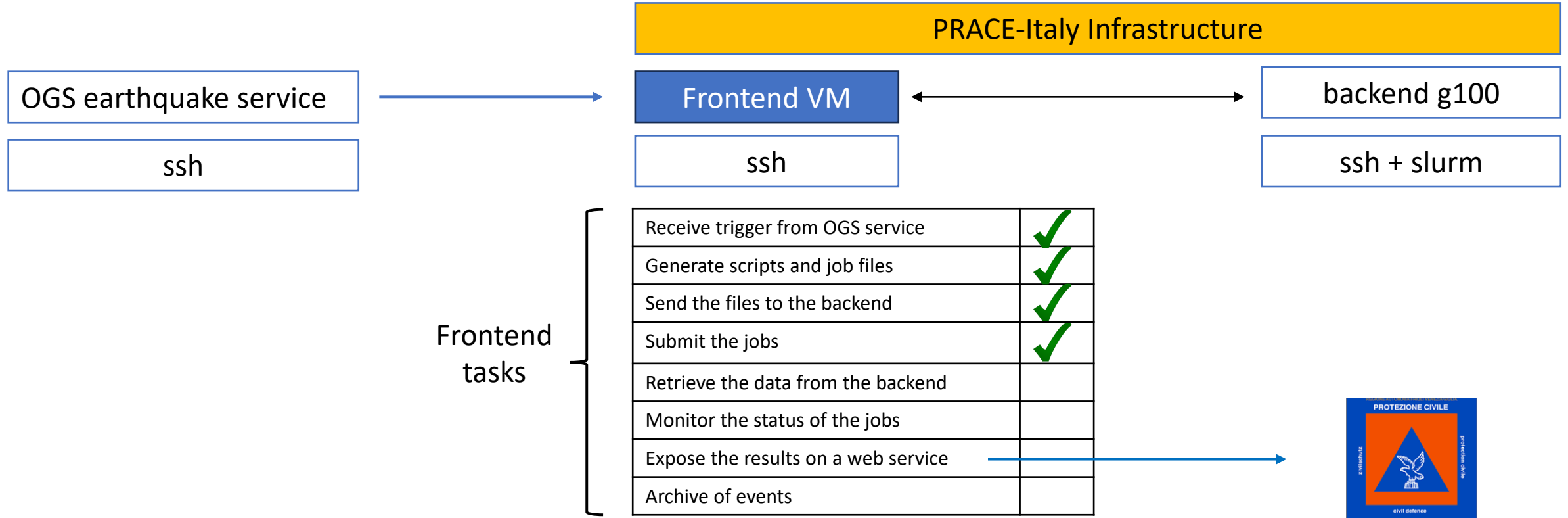
Problem 1 - AUTOMATIZATION

Replicate the workflow on a cluster, we started from g100@cineca

- Authentication ✓
- Authorization ✓
- Installations ✓
- Licenses ✓

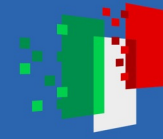


THE WORKFLOW



No interaction between user and backend

We need long-term ssh certificates !



DEMO

```
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] > ./earthquake.sh lon=12.8308,lat=46.3583,depth=7.0,mw=4.2,strike=98,dip=57,rake=132 no no g100
```

```
[CLOUD FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
```

```
[HPC CLUSTER - runs simulations]
```

JOBID	PARTITION	NAME	USER	ST	TIME	NODES	NODELIST(REASON)
-------	-----------	------	------	----	------	-------	------------------

0:21



DEMO

```
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] > ./earthquake.sh lon=12.8308,lat=46.3583,depth=7.0,mw=4.2,strike=98,dip=57,rake=132 no no g100
Triggering frontend for computation, using parameters lon=12.8308,lat=46.3583,depth=7.0,mw=4.2,strike=98,dip=57,rake=132
Creating settings.bubble using settings.g100
Running dispatcher.sh
PrEx Sun Jun 23 14:05:19 CEST 2024 > python3 create_input.py -e lon=12.8308,lat=46.3583,depth=7.0,mw=4.2,strike=98,dip=57,rake=132 -o rapids.ini.template -t no -mgf no
PrEx Sun Jun 23 14:05:19 CEST 2024 0< python3 create_input.py -e lon=12.8308,lat=46.3583,depth=7.0,mw=4.2,strike=98,dip=57,rake=132 -o rapids.ini.template -t no -mgf no [ok]
Now dispatch job*slurm and rapids.ini on backend
Running set_backend_environment
Creating /g100_scratch/userinternal/fpitari1/TeRABIT/2024-06-23_14:05:19 on remote backend fpitari1@login.g100.cineca.it

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
```

0:26



DEMO

```
756 100% 738.28kB/s 0:00:00 (xfr#4, to-chk=1/5)
job_C1_post_processing.slurm
773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818552 g100_usr_ A1UCSbsr fpitari1 PD 0:00 1 (None)
```

0:30



DEMO

```

job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
  JOBID PARTITION  NAME      USER ST    TIME  NODES NODELIST(REASON)
  13818555 g100_usr_ RunSPEED fpitari1 PD    0:00    20 (Dependency)
  13818553 g100_usr_ B1_UCSB  fpitari1 PD    0:00    10 (Dependency)
  13818554 g100_usr_ A2SPEEDi fpitari1 PD    0:00     1 (Dependency)
  13818556 g100_usr_ PostProc fpitari1 PD    0:00     1 (Dependency)
  13818552 g100_usr_ A1UCSBr  fpitari1 R    0:01     1 r520c10n03

```

0:31



DEMO

```

job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818555 g100_usr_ RunSPEED fpitari1 PD 0:00 20 (Dependency)
13818556 g100_usr_ PostProc fpitari1 PD 0:00 1 (Dependency)
13818553 g100_usr_ B1_UCSB fpitari1 R 0:01 10 r512c03n[02-03],r512c05n04,r512c06n02,r512c07n[01-04],r512n[001-002]
13818554 g100_usr_ A2SPEEDi fpitari1 R 0:01 1 r520c09n01

```

0:37



DEMO

```
job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818555 g100_usr_ RunSPEED fpitari1 PD 0:00 20 (Dependency)
13818556 g100_usr_ PostProc fpitari1 PD 0:00 1 (Dependency)
13818554 g100_usr_ A2SPEEDi fpitari1 R 0:07 1 r520c09n01
```

0:44



DEMO

```

job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818556 g100_usr_ PostProc fpitari1 PD 0:00 1 (Dependency)
13818555 g100_usr_ RunSPEED fpitari1 R 0:01 20 r512c03n[02-03],r512c05n04,r512c06n02,r512c07n[01-04],r512n[001-012]

```

0:45



DEMO

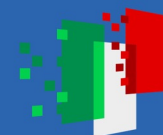
```

j..._C1_post_processing.slurm                                terabit_demo.mp4
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLUSTER FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
A1_ucsb_source.slurm
A2_speed_input.slurm
B1_run_ucsb.slurm
B2_run_speed.slurm
C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818556 g100_usr_ PostProc fpitari1 PD 0:00 1 (Dependency)
13818555 g100_usr_ RunSPEED fpitari1 R 2:45 20 r512c03n[02-03],r512c05n04,r512c06n02,r512c07n[01-04],r512n[001-012]
  
```

3:34



DEMO

```
job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLLOUD FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818556 g100_usr_PostProc fpitari1 R 0:01 1 r512c03n02
```

3:40



DEMO

```
job_C1_post_processing.slurm
773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm...
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm...
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm...
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
13818556 g100_usr_PostProc fpitari1 R 1:46 1 r512c03n02
```

5:25



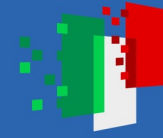
DEMO

```
job_C1_post_processing.slurm
 773 100% 754.88kB/s 0:00:00 (xfr#5, to-chk=0/5)
Submitting job_A1_ucsb_source.slurm..
Job ID: 13818552
Submitting job_B1_run_ucsb.slurm...
Job ID: 13818553
Submitting job_A2_speed_input.slurm..
Job ID: 13818554
Submitting job_B2_run_speed.slurm...
Job ID: 13818555
Submitting job_C1_post_processing.slurm..
Job ID: 13818556
[EARTHQUAKE DETECTION - triggers cloud frontend with the detected parameters] >

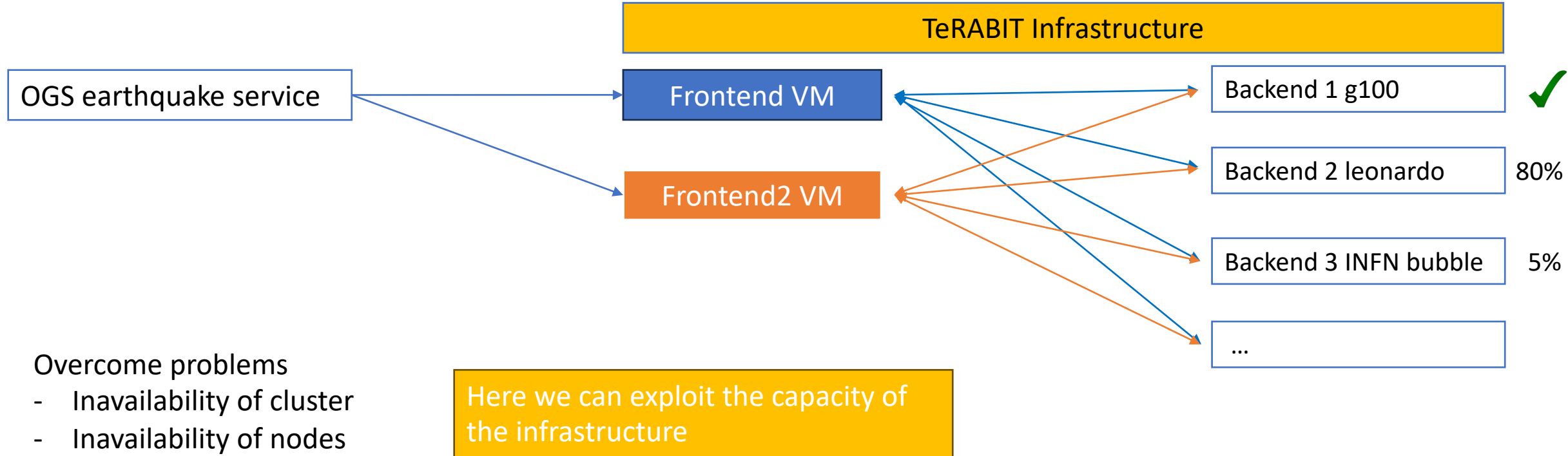
[CLoud FRONTEND - orchestrates inputs, outputs and submissions toward HPC backends]
job_A1_ucsb_source.slurm
job_A2_speed_input.slurm
job_B1_run_ucsb.slurm
job_B2_run_speed.slurm
job_C1_post_processing.slurm
rapids.ini
settings.ini

[HPC CLUSTER - runs simulations]
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
```

5:30



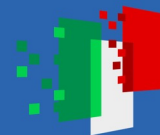
WORKFLOW EVOLUTION



Overcome problems

- Inavailability of cluster
- Inavailability of nodes

Role of frontend: simple orchestrator



Problem 2 – STRATEGIES to reduce the TIME TO SOLUTION

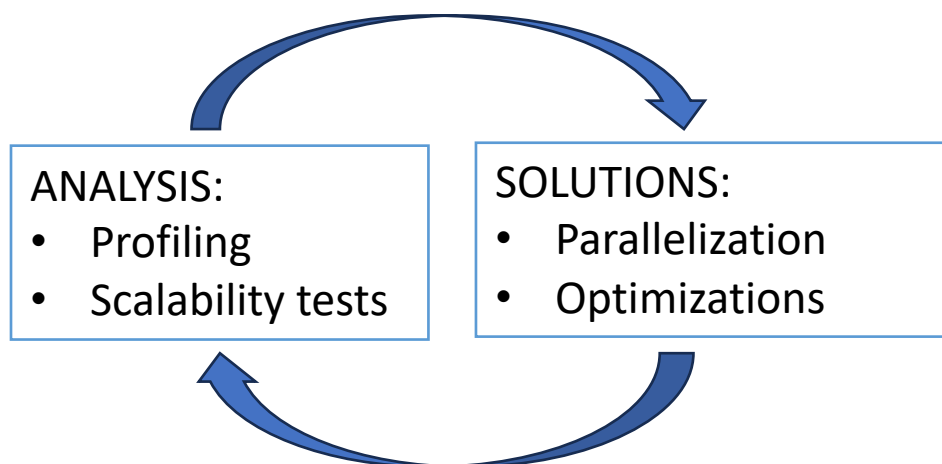
OUR URGENT COMPUTING PROBLEM HAS TWO MAIN CONSTRAINTS

No dedicated nodes

Time increases with:

- magnitude (external parameter)
- accuracy (model requirement)

Problem 2 – Strategy 1: profiling and parallelization

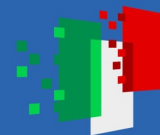


Parallelized code	Saved time
python postprocessing routine	1h
fortran processing routine	8-9 days

Scalability tests about SPEED:

Time hh:mm	cpus
02:30	480
01:20	960
01:05	1200

Very good scalability!



Problem 2 – Strategy 2: trade off about mesh

Problem

The generation of mesh for SPEED with CUBIT

- Serial version takes some hours
- Parallel version: to be investigated

Solution

Use a set of
precalculated meshes

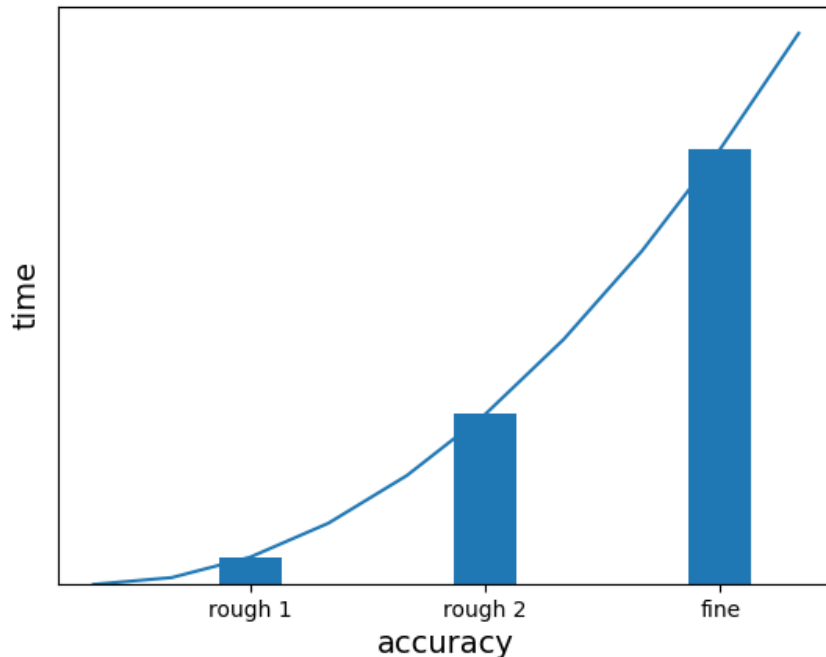
We have to put in place the system

- Scientific discussion
- Operational implementation



Problem 2 – Strategy 3: trade off about accuracy

Problem



Solution

Decrease the accuracy in order to get early results.

For example:

- without topography
- using a rough discretization for Green Functions

This is an interesting task for our frontend:

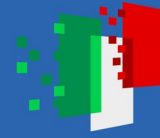
- to submit a sequence of jobs to the backends, from the roughest to the finest
- then publish the results as soon as they are ready, and choose the one that at the moment is considered the best



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Thank you

gbolzon@ogs.it, ezuccolo@ogs.it,
f.pitari@ Cineca.it,
l.rodriquezmunoz@ Cineca.it