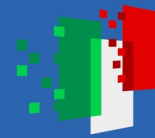




Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



terabit

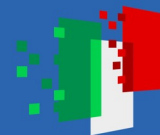
Earthquake location and tomography

Working Group 4

Luigi Zampa

Bologna, 26 giugno 2024

Workshop - TeRABIT



Working Groups: macro attività

- Identificazione del caso di studio



10 minuti

1.motivazione scientifica: realizzazione di un modello 3D del sottosuolo a partire da eventi sismici sia storici che acquisiti recentemente.

2.caratterizzazione computazionale (e.g., I/O, processing, constraints, profiling):

- È un sistema che va rilanciato ogni volta che nuovi eventi sismici vengono acquisiti o modificati;
- Richiede almeno 50 GB di RAM, per cui non è fattibile lato DAQ;
- Dati <3GB;
- Ho necessità di conoscere la provenienza dei dati (track dei dati).



10 minuti

Accesso alle risorse e servizi

- Autenticazione, Autorizzazione, UX&UI: richiesta affiliazione al Cineca.
- Ambiente di lavoro
 - Software: SLURM, workflow in python, codice in fortran;
 - Hardware: solo CPU, storage locale -> cluster g100



10 minuti

Problema da risolvere mantenendo la qualità del risultato

- Automatizzazione: check diff input iniziali, ogni volta che arrivano nuovi eventi;
- Riduzione del tempo di calcolo:
 - necessaria parallelizzazione del codice in fortran (attualmente mancante);
 - 300k fasi ~2h.



Istituto Nazionale di Fisica Nucleare

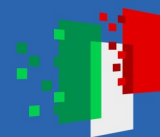


OGS

Istituto Nazionale
di Oceanografia
e di Geofisica
Sperimentale



CINECA



Working Groups: macro attività


20 minuti

- **Mappatura delle caratteristiche delle applicazioni**
 - Input data -> database relazione con metadata;
 - Localizzazione -> gpu;
 - Tomografia -> cpu;
 - Ray-tracing -> gpu;
 - Output data -> storage esposto al pubblico.


10 minuti

- **Porting dell'applicazione**
 - Applicazioni containerizzate;
 - Microservizi in cluster K8s (in cloud);
 - Storage federato in datalake;


20 minuti

- **Analisi scalabilità'**
 - Ogni evento è una soluzione seriale indipendente dalle altre, quindi si può pensare di parallelizzare (codice open source);
 - Supponiamo una scalabilità quasi-lineare con il numero di unità computazionali;
 - A seconda del modello di input, il numero di risorse può essere ridotto all'occorrenza.


10 minuti

- **Pre-produzione**
- **Produzione**



Istituto Nazionale di Fisica Nucleare



OGS

Istituto Nazionale
di Oceanografia
e di Geofisica
Sperimentale



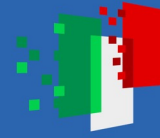
CINECA



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Thank you

izampa@ogs.it