Offloading e orchestrazione di workflow scientifici su risorse eterogenee

Giulio Bianchini, Diego Ciangottini, Massimo Sgaravatto, Daniele Spiga

Lucio Anderlini, Mauro Gattari, Rosa Petrini, Carlo Mancini Terracciano

Agenda

1. **Contesto**
   Testbed DataCloud WP6
   Le infrastrutture di calcolo integrate e stack tecnologico
2. **Interlink**
3. **Networking**
   Port forward e mesh network
4. **Casi d'uso**
   Kubeflow/GenAI, CI/CD Geant4 su HPC e Snakemake (AI_INFN)
5. **Conclusioni**
   Verso una federazione Cloud-HPC/HTC stabile, trasparente e riusabile.

Contesto - *Testbed per integrazione risorse*

INTEGRAZIONE DI RISORSE ETEROGENEE

Integrare risorse Cloud, HPC e HTC tramite meccanismo di **offloading**, garantendo all'utente finale accesso **trasparente** attraverso **k8s**

PoC CON CINECA

Attività di WP6 per consolidare Proof of Concept sviluppato con Cineca e TeRABIT per l'integrazione di LEONARDO

LAST YEAR

https://agenda.infn.it/event/45424/contributions/263545/

INFN DATACLOUD WP6

ONBOARDING COMUNITÀ

Selezionati casi d'uso più rilevanti tra quelli disponibili per **validazione** preliminare su ambiente di test

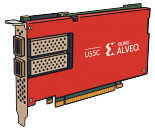
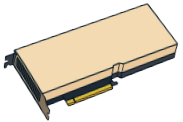
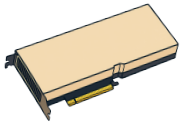
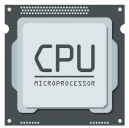
WP6 → WP1

Consolidamento SW, monitoring, documentazione per migrare verso la produzione e le operations in WP1

EVOLUTION

Contesto - *dove stiamo testando use case*

Obiettivo: mitigare la complessità delle risorse eterogenee, rendendo l'utilizzo dell'hardware semplice e immediato

	Risorse	Backend	Use case
1 **Bubble MiB**		docker	• **Kubeflow & GenAI (DSI)**
2 **Bubble HPC Padova**		slurm workload manager	• **Kubeflow & GenAI (DSI)** • **Snakemake (AI_INFN)** • **CI/CD - Geant4 (G4Med)**
3 **LEONARDO HPC**		slurm workload manager	• **Kubeflow & GenAI (DSI)** • **Snakemake (AI_INFN)**
4 **HTC CNAF (tier1)**		HTCondor	• **Snakemake (AI_INFN)**

eterogeneità

InterLink - *la tecnologia abilitante* *(quick recap)*

LAST YEAR

https://agenda.infn.it/event/45424/contributions/263545/

interLink

Esecuzione in offloading di workload containerizzati
Esecuzione di pod/container Kubernetes su backend remoti tramite virtual node e plugin.

Gestione del lifecycle del payload remoto
Gestione di sottomissione, esecuzione, stato e risultato del workload offloadato.

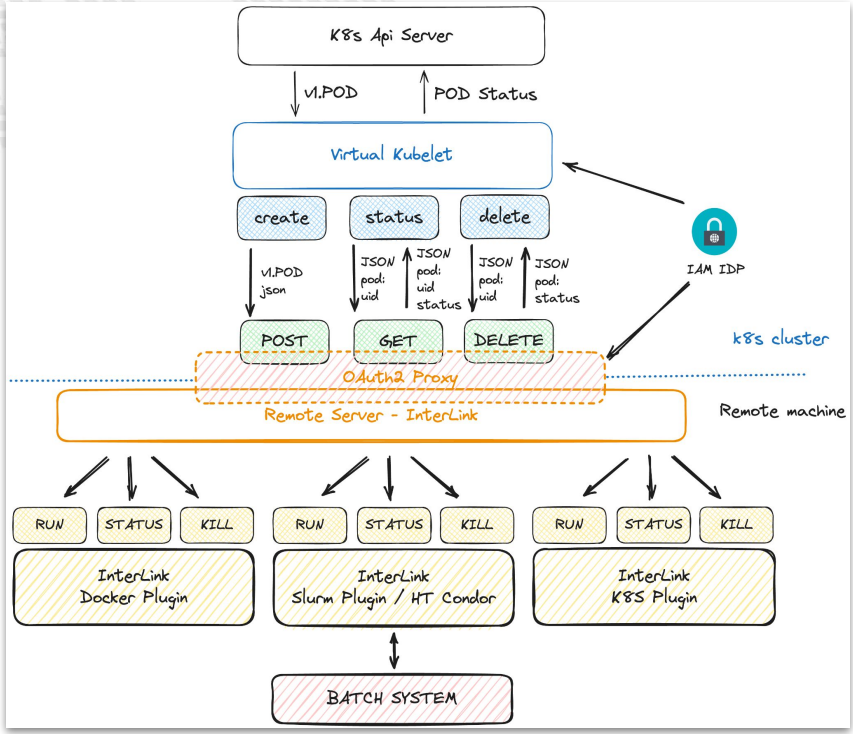
Annotations per definire risorse e requisiti site-specific
Parametri come CPU, RAM, GPU, walltime, nodi e mount configurabili via annotation; possibilità di passare opzioni specifiche del sito (account, partizione, ecc..)

Traduzione verso runtime container disponibili sul sito
Esecuzione di container Docker su HPC/Slurm tramite runtime compatibili, come Singularity/Apptainer.

Supporto a workload multi-container
Offloading di pod composti da più container applicativi o di supporto.

Plugin backend-specifici
Integrazione con backend diversi tramite plugin, ad esempio Slurm, HTCondor, Podman o k8s.

Separazione tra interfaccia Kubernetes e infrastruttura remota
L'utente "parla" k8s, mentre interLink gestisce le differenze del backend remoto.



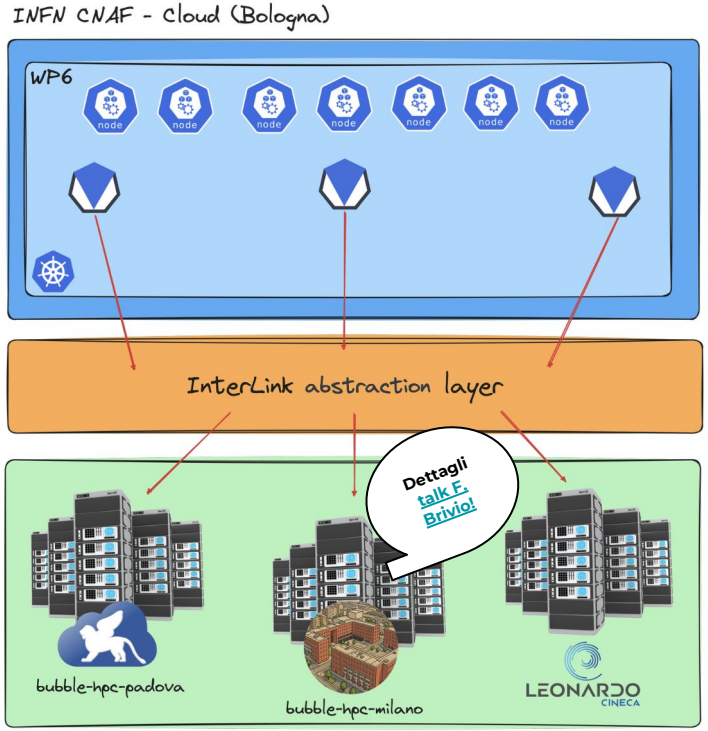
https://interlink-hq.github.io/interLink/

DONE

Punti aperti:
Networking pod-to-pod in-cluster: supporto alla comunicazione diretta tra pod all'interno dello stesso cluster

Contesto - *il testbed a wp6*

Il setup tecnologico e' basato su k8s e interLink per offloading trasparente di workload k8s-friendly su provider eterogenei.



Componente	Ruolo nel setup
Kubernetes / Virtual Kubelet	**Unico entrypoint**. Gestisce in modo trasparente esecuzione di pod su risorse reali e virtuali (remote).
InterLink abstraction layer	**Middleware**. Abilita l'offloading trasparente verso backend eterogenei come SLURM, HTCondor, Docker o altri cluster Kubernetes.
Risorse remote	**Risorse** HPC e HTC, acceleratori come GPU e FPGA, e siti distribuiti **accessibili come se fossero nodi del cluster k8s**

Networking - *perche' e' importante*

La connettività rappresenta un aspetto critico in specifici contesti, in particolare per la comunicazione di management in alcuni workflow complessi e per mitigare problematiche legate alla segregazione di rete (come limitazioni sull'outbound)

es. Pipeline multi-step

- connettività tra pod
- service discovery
- accesso ai servizi interni
- accesso a storage condiviso

BUT

- NO IP pubblico diretto ai nodi
- firewall restrittivi

- connessioni outbound only
- No porte in ingresso

Networking - *soluzioni implementate*

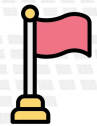
interLink

PORT FORWARD

values. YAML

```
network:
    enableTunnel: true
    tunnelImage: "ghcr.io/erebe/wstunnel:latest"
    wildcardDNS: "131.154.99.68.myip.cloud.infn.it"
```

Accesso sicuro ai servizi dei pod remoti tramite tunnel WebSocket, senza richiedere connettività VPN



Un semplice flag è sufficiente per abilitare trasparentemente il networking

MESH NETWORK

values. YAML

```
network:
    fullMesh: true
    tunnelImage: "ghcr.io/erebe/wstunnel:latest"
    wildcardDNS: "131.154.99.68.myip.cloud.infn.it"
```

Mesh networking abilita la comunicazione trasparente tra pod remoti e cluster K8S tramite WireGuard e wstunnel

Qual'e' la soluzione migliore? ***Dipende****, approcci guidati dalle necessità*

Criterio	Port Forward	Mesh Network
Connettività	Limitata alle porte esposte	Completa tra pod remoti e cluster
Complessità	Bassa / media	Alta
Caso d'uso ideale	Esporre web app, API o servizi puntuali	Integrare pod remoti come parte della rete k8s

Agenda

1. **Contesto**
   DataCloud WP6 e gli use cases selezionati.
   Le infrastrutture di calcolo integrate e stack tecnologico del testbed

2. **Interlink**

3. **Networking**
   Requisiti, limiti e soluzioni: port forward e mesh network.

4. **Casi d'uso**
   Esecuzione di servizi Kubeflow & GenAI (DSI), CI/CD - Geant4 (G4Med) e Snakemake (AI_INFN)

5. **Sviluppi futuri**
   Consolidamento del testbed, estensione a nuovi backend, networking e storage condiviso.

6. **Conclusioni**
   Verso una federazione Cloud-HPC/HTC stabile, trasparente e riusabile.

Contesto - *Use cases di alto livello*

Problema: workload scientifici sfruttano risorse computazionali eterogenee. Alcuni esempi

Kubeflow & GenAI (DSI) *

Fine-tuning e inference LLM, RAG su knowledge base, ...

GPU (es. H100)

CI/CD - Geant4 (G4Med) *

Pipeline automatizzate regression testing e CI per validazione continua di modelli su HPC

CSN5 **Molti core CPU**

Snakemake (AI_INFN) *

Orchestrazione DAG workflow ML

CSN5 **Risorse eterogenee**

Analisi interattiva

JupyterLab + scale out su HPC/HTC per analisi ad alto rate

Cloud + Burst HPC/HTC

MLOps & CI/CD

Experiment tracking, model registry e riproducibilità dei workflow ML dalla fase di sviluppo al deployment

Risorse eterogenee

DNN per fisica

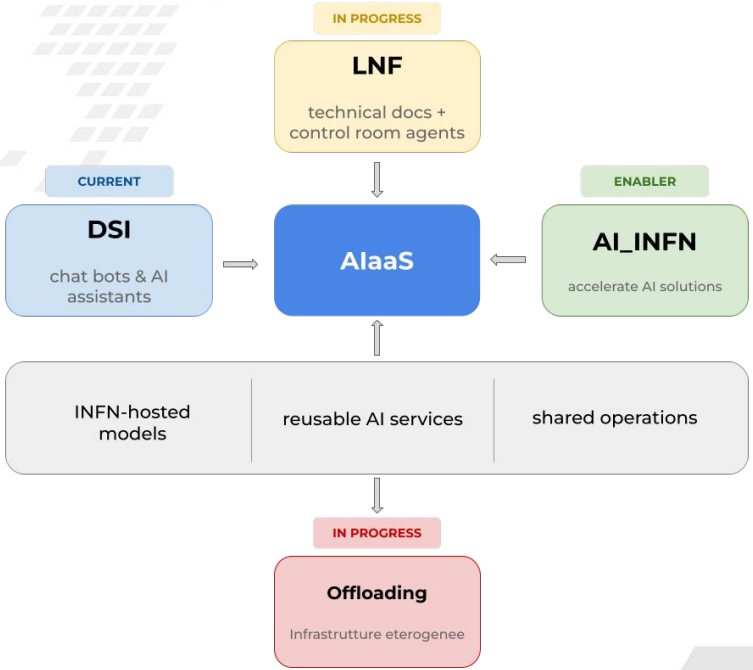
Training/Inference DNN per ricostruzioni tracce (LHCb, CMS), classificazioni eventi HEP, ...

GPU / FPGA

* *approfondiamo oggi!*

Kubeflow & GenAI (DSI) - 1

Kubeflow è una piattaforma integrata che semplifica l'intero ciclo di vita delle applicazioni AI, dalla creazione alla produzione, offrendo strumenti per sviluppo, automazione, inferenza e gestione dei modelli.

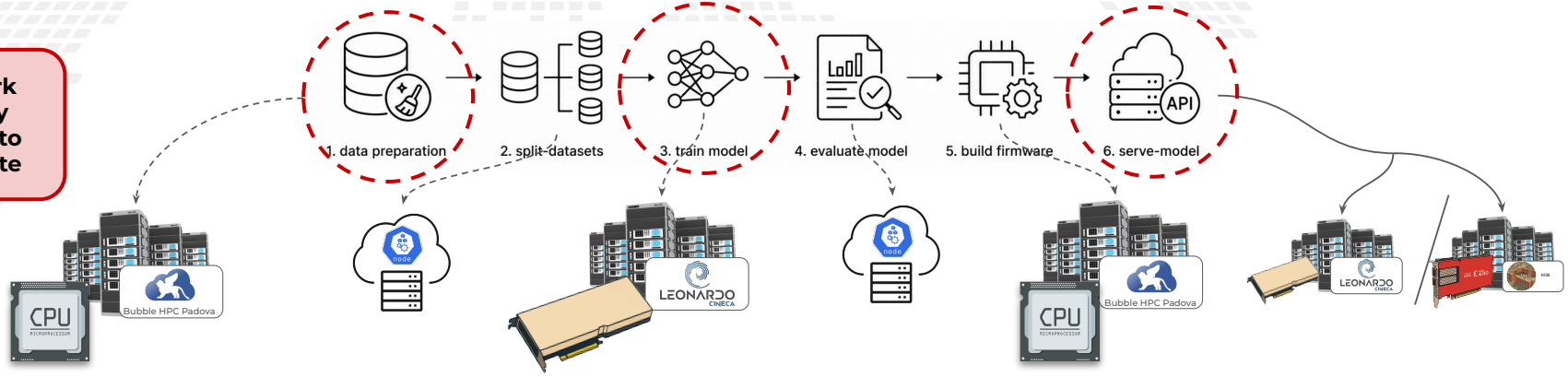


- **Piattaforma AI unica e condivisa**
  - accessibile a diverse comunità INFN e casi d'uso
- **Abilitare servizi AI riusabili**
  - chatbot, assistenti, semantic search, agenti intelligenti e workflow automation
- **Supportare l'intero ciclo di vita delle applicazioni AI**
  - dalla prototipazione allo sviluppo, fino all'inferenza e alla produzione.
- **Integrare modelli e conoscenza interna INFN**
  - tramite LLM on-prem, embeddings, RAG, basi documentali e knowledge base.
- **Garantire operazioni AI centralizzate e scalabili**
  - basate su INFN Cloud, k8s, GPU/CPU condivise, storage, autenticazione e multi-tenancy
- **Sfruttare infrastrutture computazionali eterogenee**
  - necessarie per eseguire workload AI complessi su HPC, SLURM, cluster remoti e risorse GPU esterne.

***Credits:** Mauro Gattari*
***CCR 2025:** https://agenda.infn.it/event/45424/contributions/263566/*

Pipeline ML e2e - Validazione di pipeline ML distribuite attraverso i vari VK e le diverse risorse eterogenee.

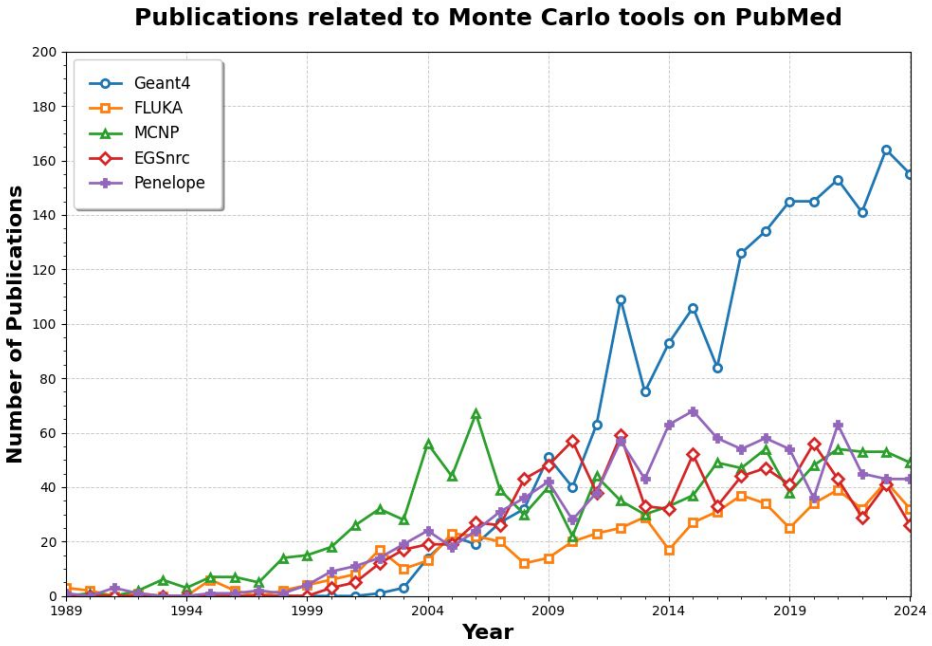
Network overlay elemento abilitante



KServe LLM inference - Validazione del serving di modelli LLM tramite KServe, con offloading del pod di inference verso risorse GPU.

- Serving di modelli LLM tramite KServe
- Offloading trasparente del pod di inference su GPU remote
- Accesso all'endpoint dal cluster Kubeflow
- Validazione di networking, service discovery
- **Base per servizi GenAI scalabili e riusabili**

G4Med - Regression Testing & CI/CD



Secondo PubMed, ***Geant4 è lo strumento MC più utilizzato per la ricerca nelle applicazioni mediche.***

G4MSBG - *Geant4 Medical Simulation Benchmarking Group*

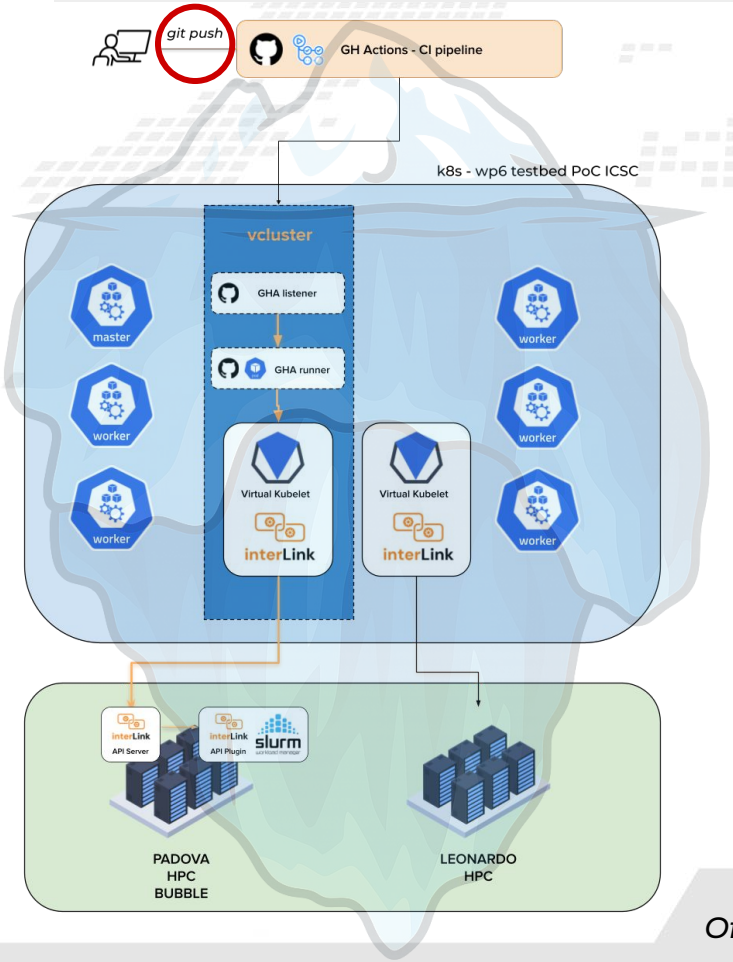
Un gruppo di membri della collaborazione Geant4 sviluppa, **mantiene ed esegue regolarmente una serie di test di benchmark per i modelli Geant4 di interesse per le applicazioni mediche**, che spaziano dai modelli di fisica elettromagnetica a quelli di fisica nucleare.

Requirements

- I modelli fisici di Geant4 devono essere validati continuamente
- Le simulazioni MC richiedono risorse computazionali elevate
- Necessità di garantire riproducibilità dei risultati su infrastrutture eterogenee

WP6-testbed - CI/CD Geant4 (G4Med)

Credits: *Carlo Mancini Terracciano*



Validazione continua di Geant4 su infrastrutture HPC integrate in un workflow GitHub Actions per eseguire automaticamente job di validazione a ogni commit o tag

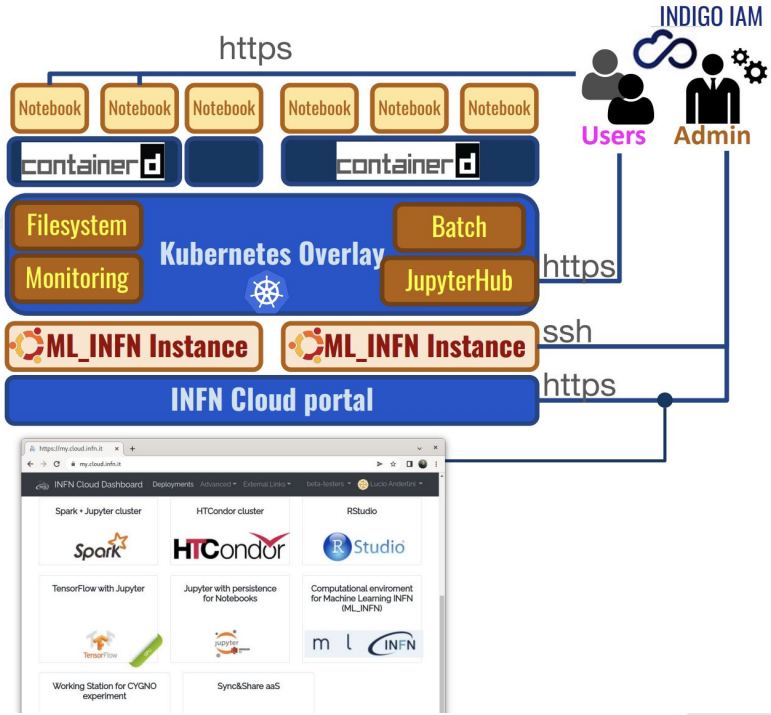
- **Pipeline CI/CD su GitHub Actions con container Apptainer/Singularity** per isolamento e riproducibilità
- **GitHub Runner self-hosted deployati come Pod Kubernetes** nel cluster k8s wp6
- **Offloading trasparente dei job verso la bolla HPC di Padova tramite InterLink e Virtual Kubelet**
- **Pod runner effimeri**: creati dinamicamente a ogni esecuzione e rilasciati al termine, con output salvati come artifact della pipeline

AI_INFN - leading use case

Credits: Lucio Anderlini

hub.ai.cloud.infn.it

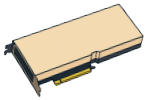
AI_INFN è una piattaforma k8s-based che integra risorse hardware eterogenee fornendo ai ricercatori INFN un ambiente unificato per sviluppo, training e deployment di modelli di ML



Alcuni punti chiave (dettagli in backup)

Ambienti interattivi per sviluppo ML/AI
- JupyterLab
- VSCode
- Desktop remoto con GPU

Accesso a risorse hardware eterogenee



- GPU: A100, RTX 5000, T4, A30
- FPGA: Xilinx U55c
- CPU/GPU/FPGA gestite in ambiente cloud/k8s

Autenticazione e accesso federato
- Login tramite AAI / INDIGO IAM

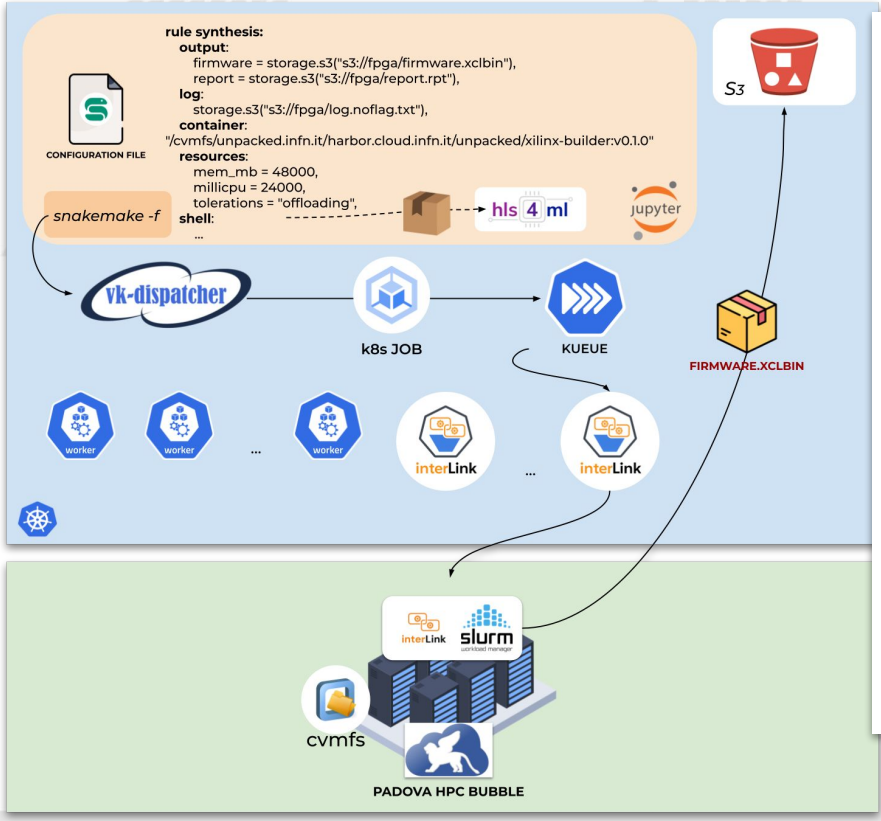
Ambienti software gestiti
- Immagini Docker personalizzabili
- Ambienti Conda isolati
- Container Apptainer per portabilità e riproducibilità

- **Batch jobs da sessioni interattive**
- **Offloading verso risorse esterne**
  - Esecuzione di container k8s su backend eterogenei tramite interLink
- **Workflow scientifici orchestrati**
  - Snakemake workflow

Source: https://agenda.infn.it/event/47736/contributions/278005

AI_INFN - Snakemake workflow - offload build firmware FPGA

Problema: La generazione del firmware (file binario per la programmazione di FPGA) è un processo computazionalmente intensivo, sia in termini di CPU e memoria, che richiede strumenti proprietari installati in locale e tempi di esecuzione elevati.



- L'utente lancia il workflow da JupyterLab usando **Snakemake**
- La regola Snakemake (**rule_synthesis**) descrive cosa eseguire: container/toolchain da usare, risorse richieste, output attesi e il payload (codice Python che usa hls4ml)
- Gli output sono su **S3**: firmware, report e log.
- **vk-dispatcher** prende la richiesta dalla sessione interattiva e la trasforma in un Kubernetes Job.
- **Kueue** gestisce la coda e lo scheduling del job.
- Se il job è marcato per offloading, Kueue lo manda verso un **vk interLink**.
- interLink inoltra il job al backend remoto, in questo caso la **bolla HPC di Padova**
- Firmware caricato su MinIO/S3 e accessibile dalla sessione interattiva.

Summary

- Estendendo il lavoro avviato con il PoC di Spoke0, WP6 ha implementato un testbed finalizzato a dimostrare la possibilità di fornire un modello di integrazione di risorse basato su:
  - interfaccia operativa uniforme (k8s)
  - meccanismo di offloading (VK & interLink).
- I casi d'uso validati (Kubeflow/GenAI, CI/CD Geant4 e AI_INFN/Snakemake) mostrano che lo stesso approccio può soddisfare scenari molto diversi: AI, simulazioni, workflow batch e servizi interattivi.
- La rete era una funzionalità mancante: le soluzioni implementate, dal port forwarding alla mesh network, abilitano casi d'uso progressivamente più complessi, inclusi applicazioni che richiedono connettività tra componenti (ad esempio piattaforme come Kubeflow).
- Sfida: validare finalizzazione. il lavoro svolto è finalizzato alla validazione di un modello sostenibile ed operabile per datacloud.

Grazie

BACKUP

Networking - soluzioni implementate

interLink

Port forward

Accesso sicuro ai servizi dei pod remoti tramite tunnel WebSocket, senza richiedere connettività VPN


```
network:
    enableTunnel: true
    tunnelImage: "ghcr.io/erebe/wstunnel:latest"
    wildcardDNS: "131.154.99.68.myip.cloud.infn.it"
```

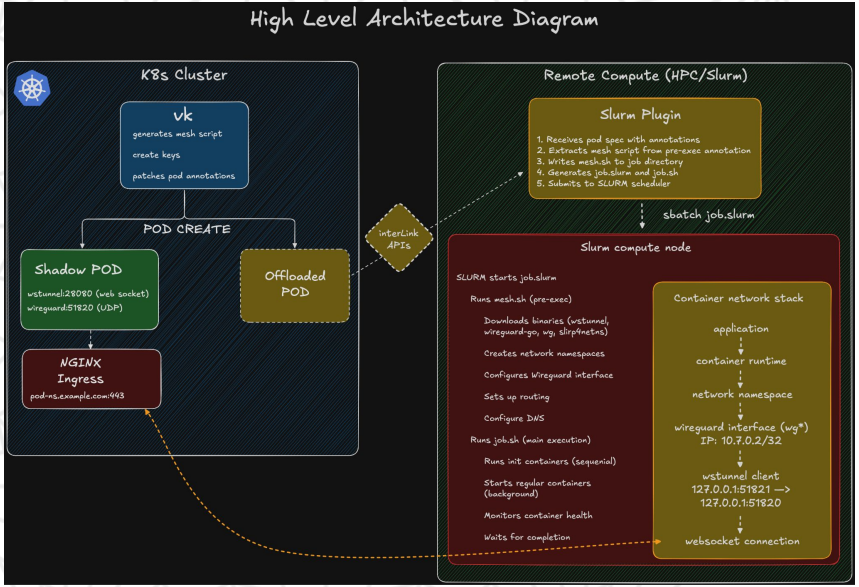
Qual'e' la soluzione migliore? Dipende...

Criterio	Port Forward	Mesh Network
Connettività	Limitata alle porte esposte	Completa tra pod remoti e cluster
Complessità	Bassa / media	Alta
Caso d'uso ideale	Esporre web app, API o servizi puntuali	Integrare pod remoti come parte della rete Kubernetes

Questo consente di affrontare piu casi d'uso

Mesh Network

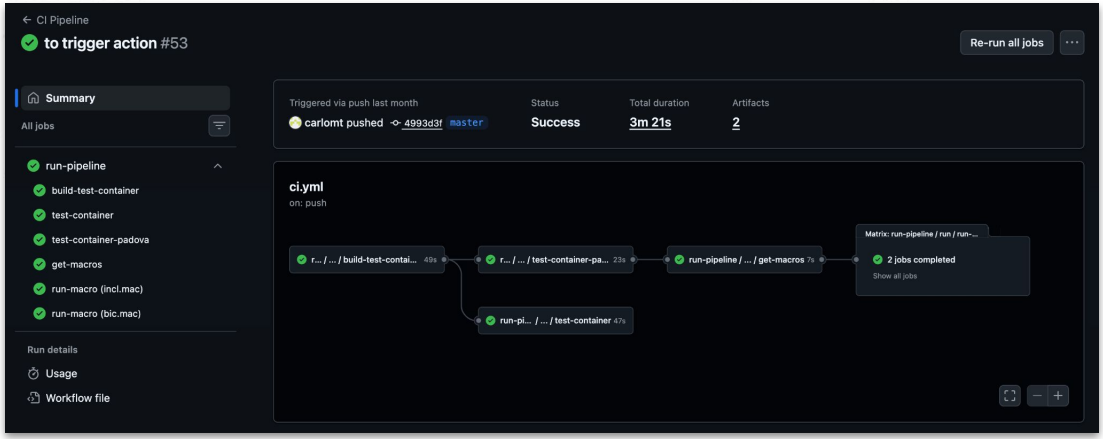
Mesh networking abilita la comunicazione trasparente tra pod remoti e cluster K8S tramite WireGuard e wstunnel

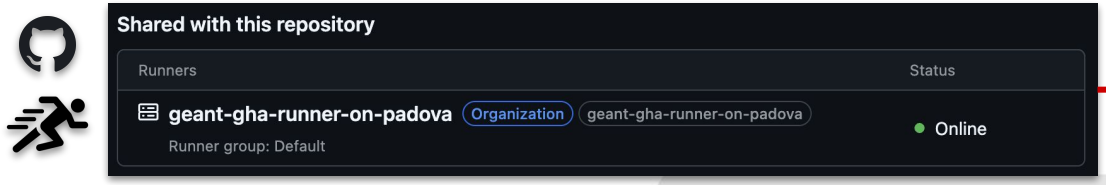


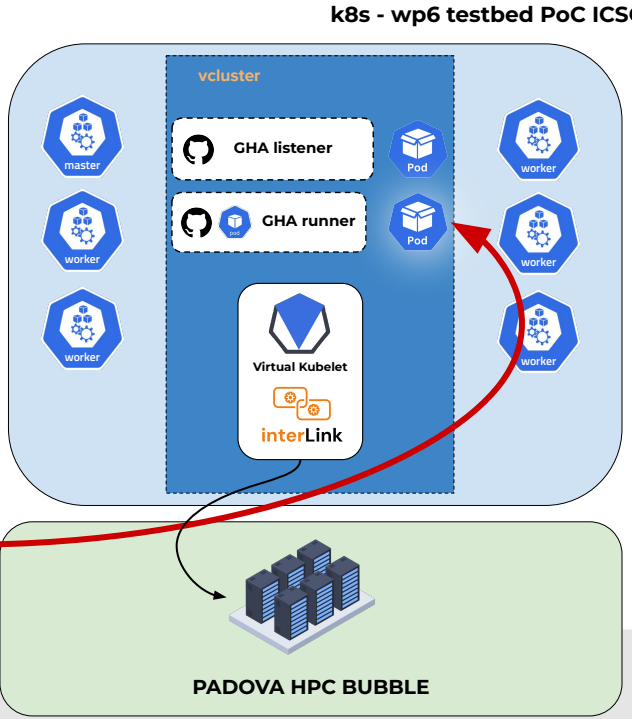
source: https://interlink-project.dev/docs/guides/mesh-network-configuration

WP6-testbed - G4Med CI/CD

LowEFrag (https://github.com/G4Med-test/LowEFrag) è un test di validazione riproducibile progettato per verificare quanto accuratamente Geant4, in particolare i modelli di frammentazione e produzione di frammenti nucleari, riproduca dati sperimentali rilevanti per applicazioni in ambito medico





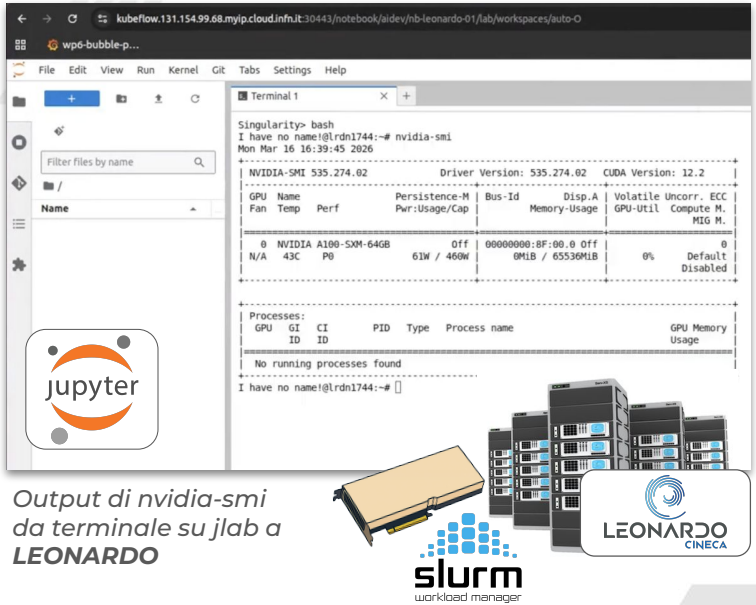


WP6-testbed - Kubeflow & GenAI (2)

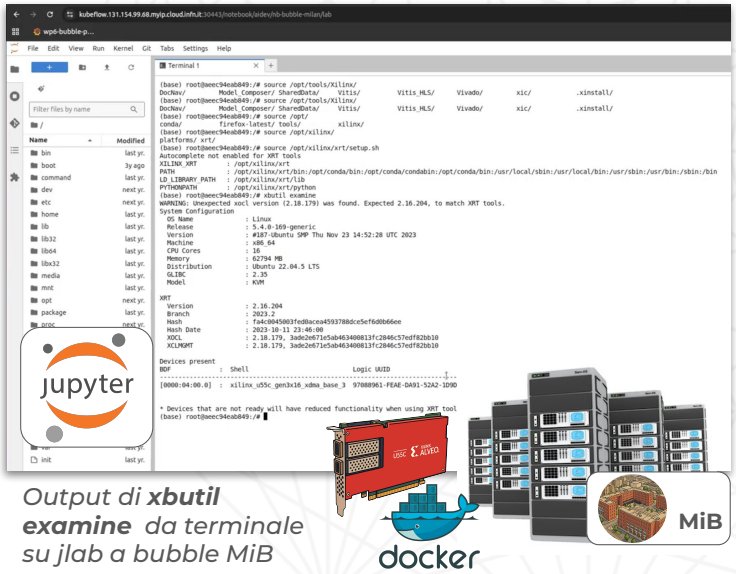
Kubeflow istanziato cluster k8s testbed WP6

I servizi istanziati dalla piattaforma possono essere eseguiti anche su risorse remote.

... quali servizi?



*Output di nvidia-smi da terminale su jlab a **LEONARDO***



*Output di **xbutil examine** da terminale su jlab a bubble MiB*

Dalla dashboard di Kubeflow e' possibile specificare le risorse ed, eventualmente, il VK dove eseguirlo

Full mesh abilitata

Supporto ad ISTIO

PVC in beta

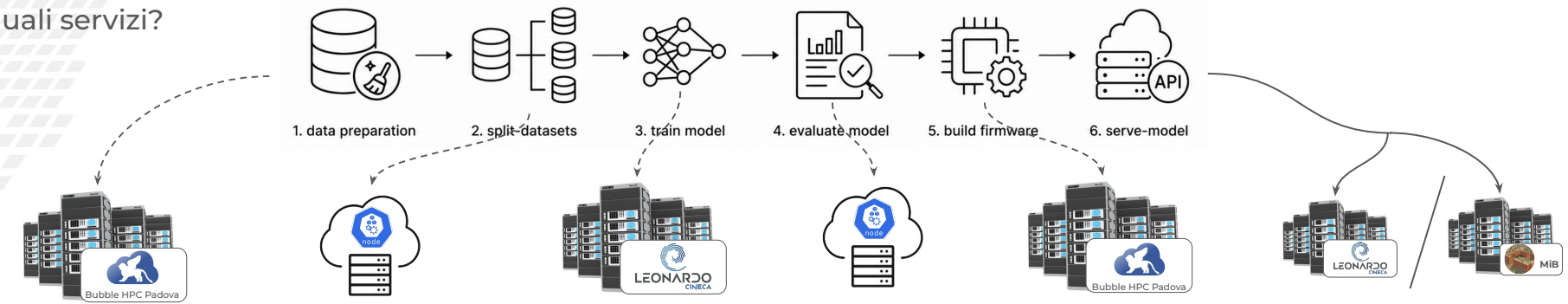
Credits: Mauro Gattari

WP6-testbed - Kubeflow & GenAI (3)

Kubeflow istanziato cluster k8s testbed WP6

I servizi istanziati dalla piattaforma possono essere eseguiti anche su risorse remote.

... quali servizi?




```
Name:                ml-pipeline-jvd41
Namespace:           aidev
ServiceAccount:      default-editor
Status:              Succeeded
Conditions:
 PodRunning          True
 Completed           True
Created:             Sun Apr 26 12:57:52 +0200 (5 hours ago)
Started:             Sun Apr 26 15:59:23 +0200 (2 hours ago)
Finished:            Sun Apr 26 18:43:44 +0200 (4 minutes ago)
Duration:            2 hours 44 minutes
Progress:            6/6
ResourcesDuration:   66h17m34s*(1 cpu),1796h4m11s*(100Mi memory)

STEP                   TEMPLATE              PODNAME                          DURATION  MESSAGE
 ✔ ml-pipeline-jvd41   ml-training-pipeline
 ├─✔ data-preparation  data-preparation      ml-pipeline-jvd41-2303692134     1m
 ├─✔ split-datasets(0) split-datasets        ml-pipeline-jvd41-3240582343     53s
 ├─✔ train-model       train-model           ml-pipeline-jvd41-2216260168     1m
 ├─✔ evaluate-model    evaluate-model        ml-pipeline-jvd41-1976162301     1m
 ├─✔ build-firmware    build-firmware        ml-pipeline-jvd41-3523754152     2h
 └─✔ serve-model       serve-model           ml-pipeline-jvd41-1508727617     27s
```

AI_INFN - leading use case

AI_INFN è una piattaforma k8s-based che integra risorse hardware eterogenee fornendo ai ricercatori INFN un ambiente unificato per sviluppo, training e deployment di modelli di ML

WP1

Cluster k8s su infn cloud
GPU A100, RTX 5000, T4, Xilinx U55C

Offloading via Interlink
Leonardo Cineca - HPC Bubble PD - CNAF Tier1 - ReCaS Bari

Batch opportunistico
Risorse idle del cluster usate per job CPU/GPU

Monitoring & Accounting
Prometheus, Grafana, OpenTelemetry

WP2

Hackathon avanzati
I and II AI-INFN Advanced Hackathon (Padova and Pavia)

E-learning
6h video-lezioni su ML per la ricerca INFN

MLOps con Kubeflow
DAG pipeline - experiment tracking- CI/CD su baltig

Snakemake workflow
Orchestrazione job AI su HPC eterogeneo

WP3

AI_INFN platform
Jupyterlab - VSCode - desktop remoto con GPU

Storage distribuito
JuiceFS - ORAS / Harbor - CVMFS - Ceph

Use case scientifici
Fisica HEP, radiologia, simulazioni mediche (G4Med)

Medical data platform
XNAT per imaging multimodale integrato con HPC@Pisa

WP4

FPGA in Cloud
Provisioning Xilinx U55C - Vivado M Edition - InterLink

Pipeline DNN -> FPGA
Pruning, quantizzazione, distillazione - auto encoder

Quantum computing
PennyLane - QUBO - Integrati nella piattaforma

IT4LIA
EuroHPC - Cineca next to INFN Tier 1 Bologna

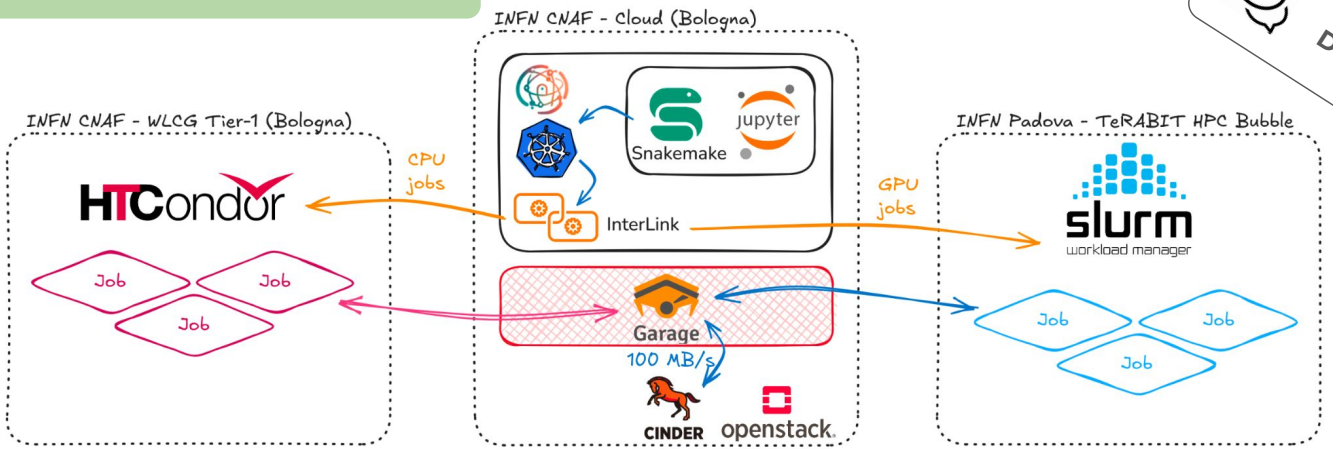
AI_INFN - Snakemake orchestrazione AI workloads

Sistema di workflow management utilizzato per definire, automatizzare e orchestrare pipeline computazionali.

interLink

Offloading di workload containerizzati verso siti di calcolo remoti in combinazione con Kueue per schedulare i job in base a risorse richieste e disponibilità dei siti.

GPU-Accelerated Simulation of 3D Diamond Sensors



Dettagli talk R. Petrini!!