

Un'infrastruttura di calcolo distribuita basata su HTCondor per esperimenti spaziali

S. Dal Pra (CNAF), M. Duranti (PG), V. Formato (RM2),
D. Lattanzio (CNAF), M. Mariotti (PG)
N. Mori (FI), A. Pascolini (CNAF), C. Pellegrino (CNAF).

Sommario

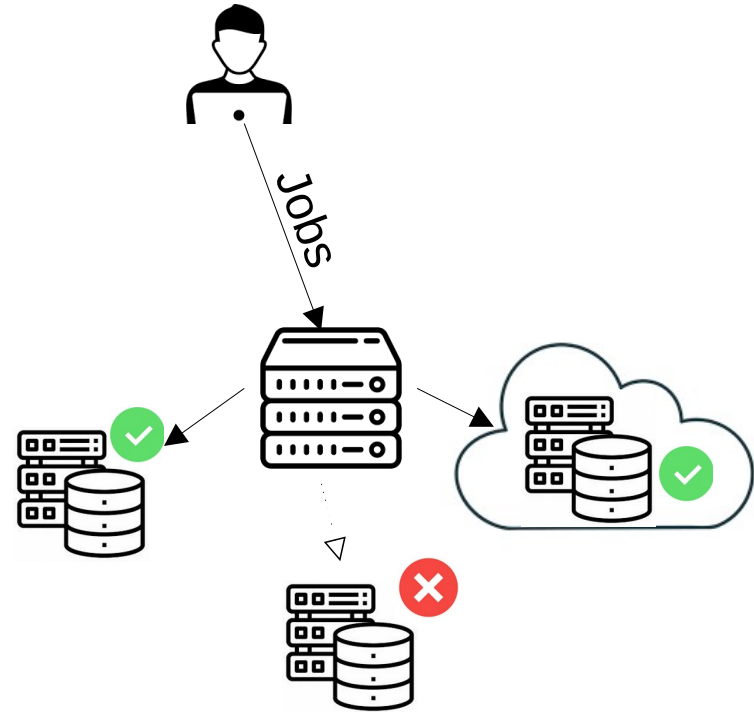
- Calcolo per esperimenti spaziali
- Caso di studio: HERD
- Infrastruttura e prototipaggio
- Sviluppi futuri

Introduzione

- Esperimenti “spazio” INFN con coinvolgimento ASI
 - AMS, EUCLID, FERMI, GAPS, HERD, IXPE, ...
- Risorse di calcolo fornite da INFN (T1)
- Farm di calcolo allo Space Science Data Center (SSDC) presso la sede ASI a Roma
 - Accordo attuativo N. 2021-43-HH.0 dell'Accordo Quadro ASI/INFN n. 2021-8-Q.0 Codice Unico di Progetto (CUP) F85F21006430005
- Possibilita' di risorse “non-WLCG”

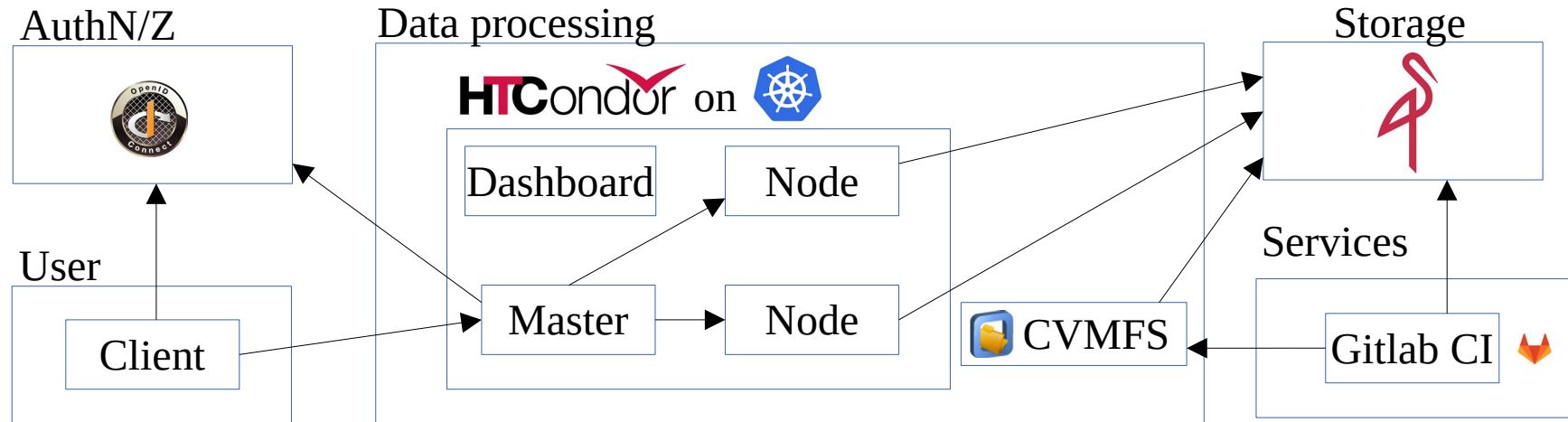
Introduzione

- Necessita' di sfruttare uniformemente ed efficientemente le risorse in maniera semplice per l'utente
 - Sottomettere job a un unico punto
 - Smistamento automatico sulle risorse attualmente disponibili
- Anche cloud / opportunistiche



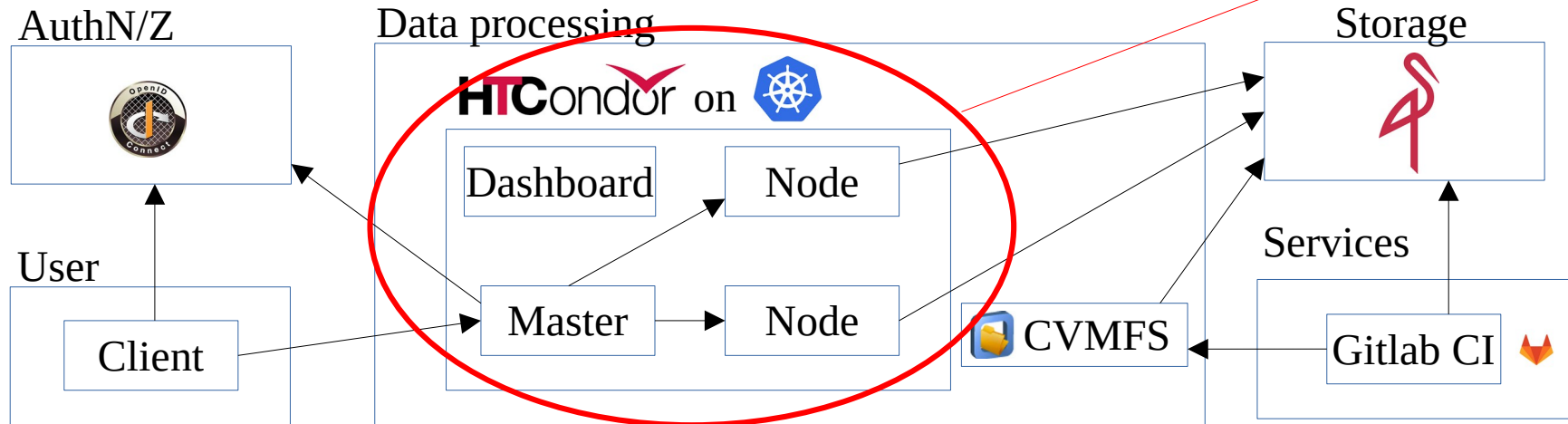
Use-case: HERD

- Misura diretta di raggi cosmici ad energie del PeV
- Computing model prototipale basato su Cloud
 - Cluster on-demand (DODAS), AuthN/Z via token, storage S3
 - Vedi precedenti Workshops + ACAT 2022 + CHEP 2023



Use-case: HERD

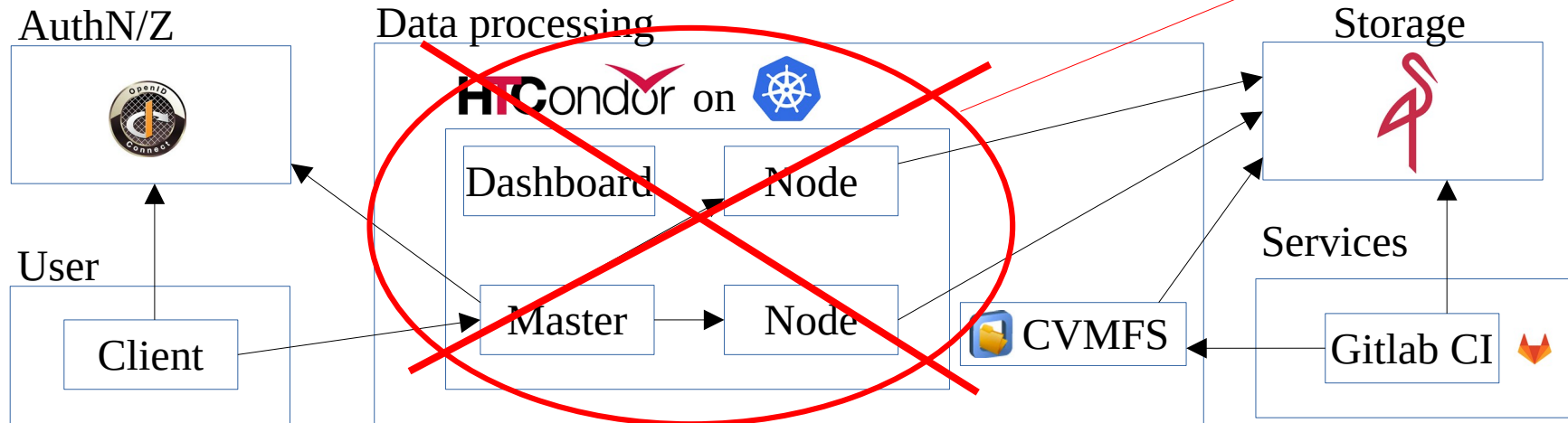
- Misura diretta di raggi cosmici ad energie del PeV
- Computing model prototipale basato su Cloud
 - Cluster on-demand (DODAS), AuthN/Z via token, storage
 - Vedi precedenti Workshops + ACAT 2022 + CHEP 2023



Use-case: HERD

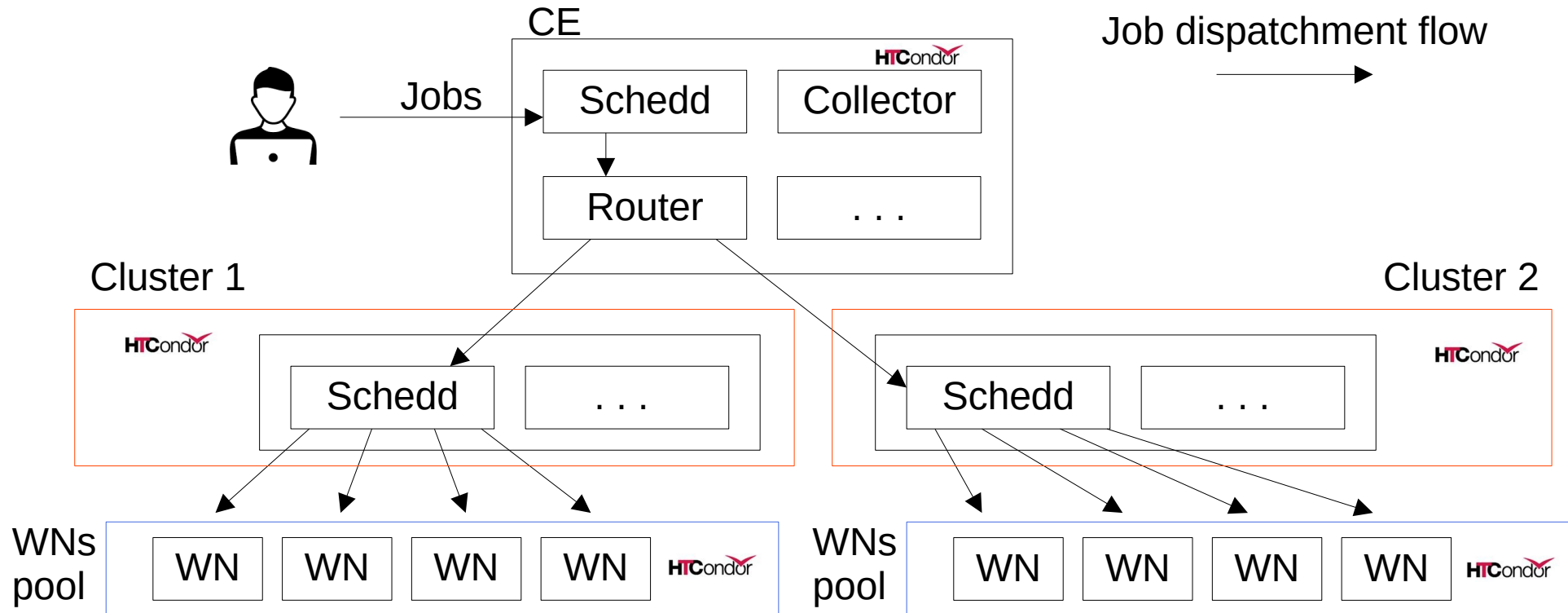
- Misura diretta di raggi cosmici ad energie del PeV
- Computing model prototipale basato su Cloud
 - Cluster on-demand (DODAS), AuthN/Z via token, storage
 - Vedi precedenti Workshops + ACAT 2022 + CHEP 2023

HTCondor-CE @CNAF
+
"Risorse"



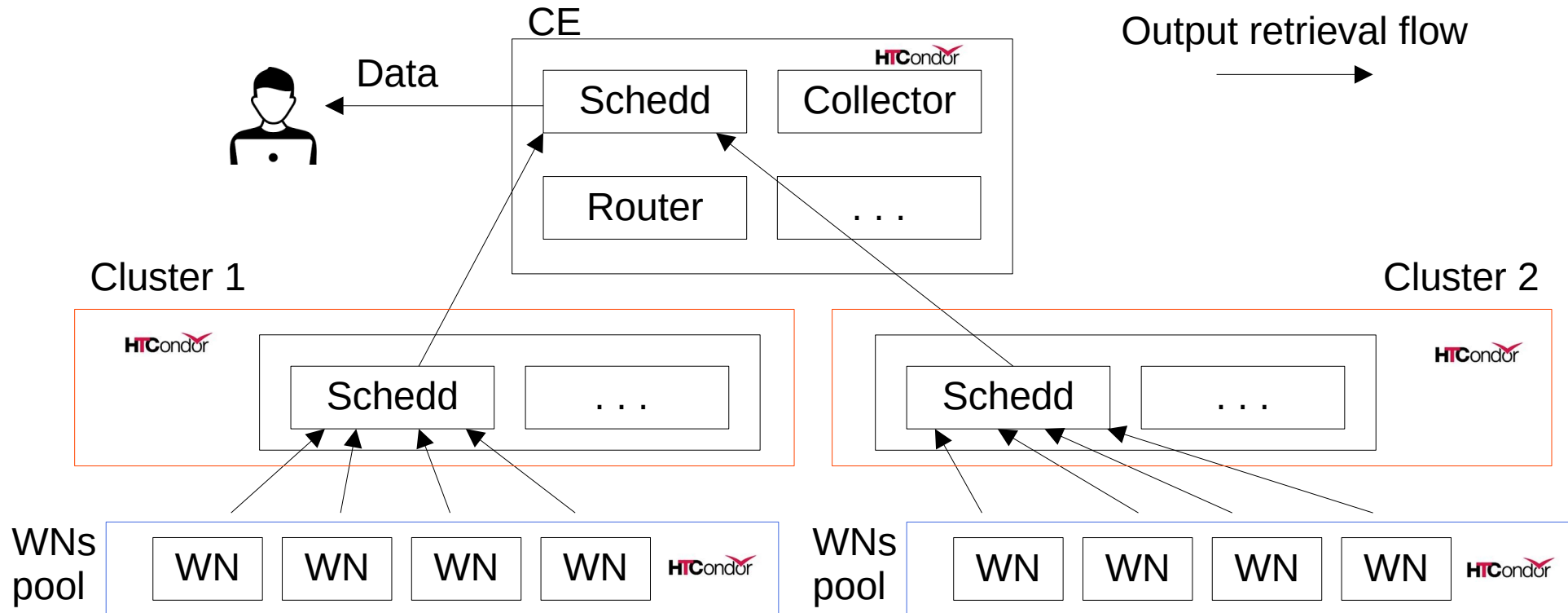
HTCondor-CE

- HTCondor Compute Entrypoint (CE)



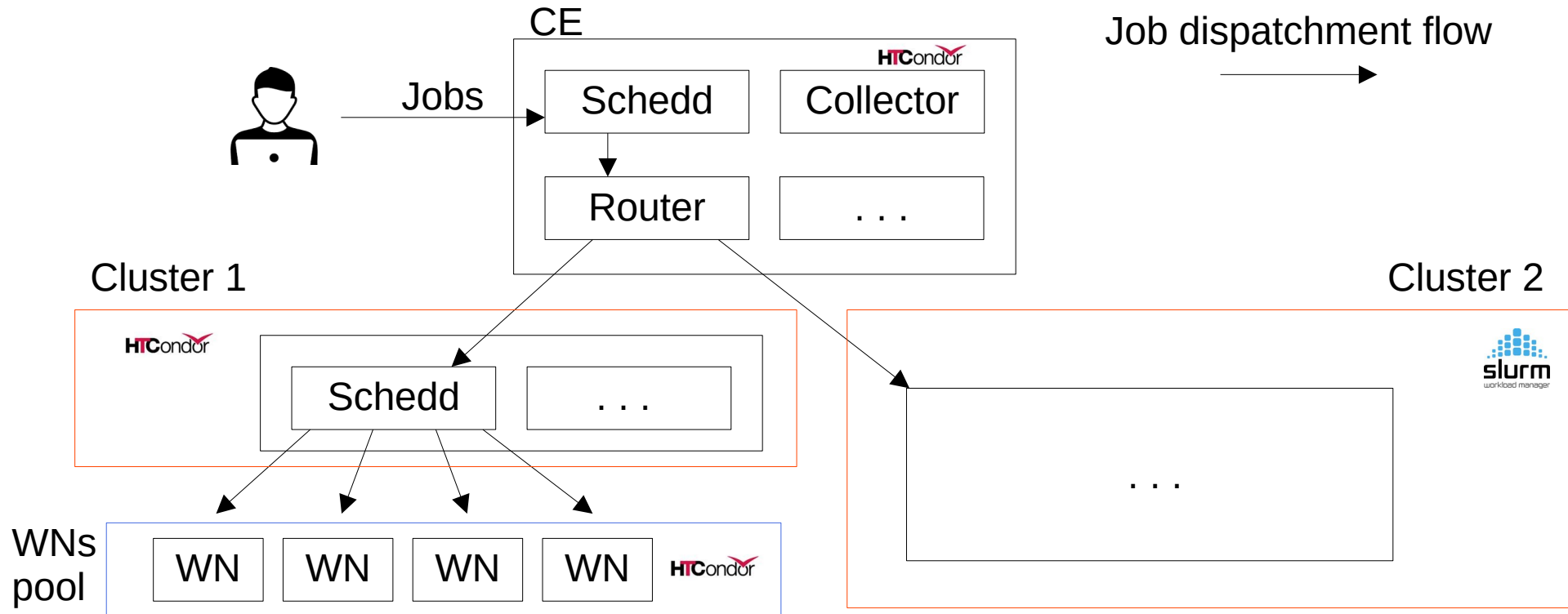
HTCondor-CE

- HTCondor Compute Entrypoint (CE)

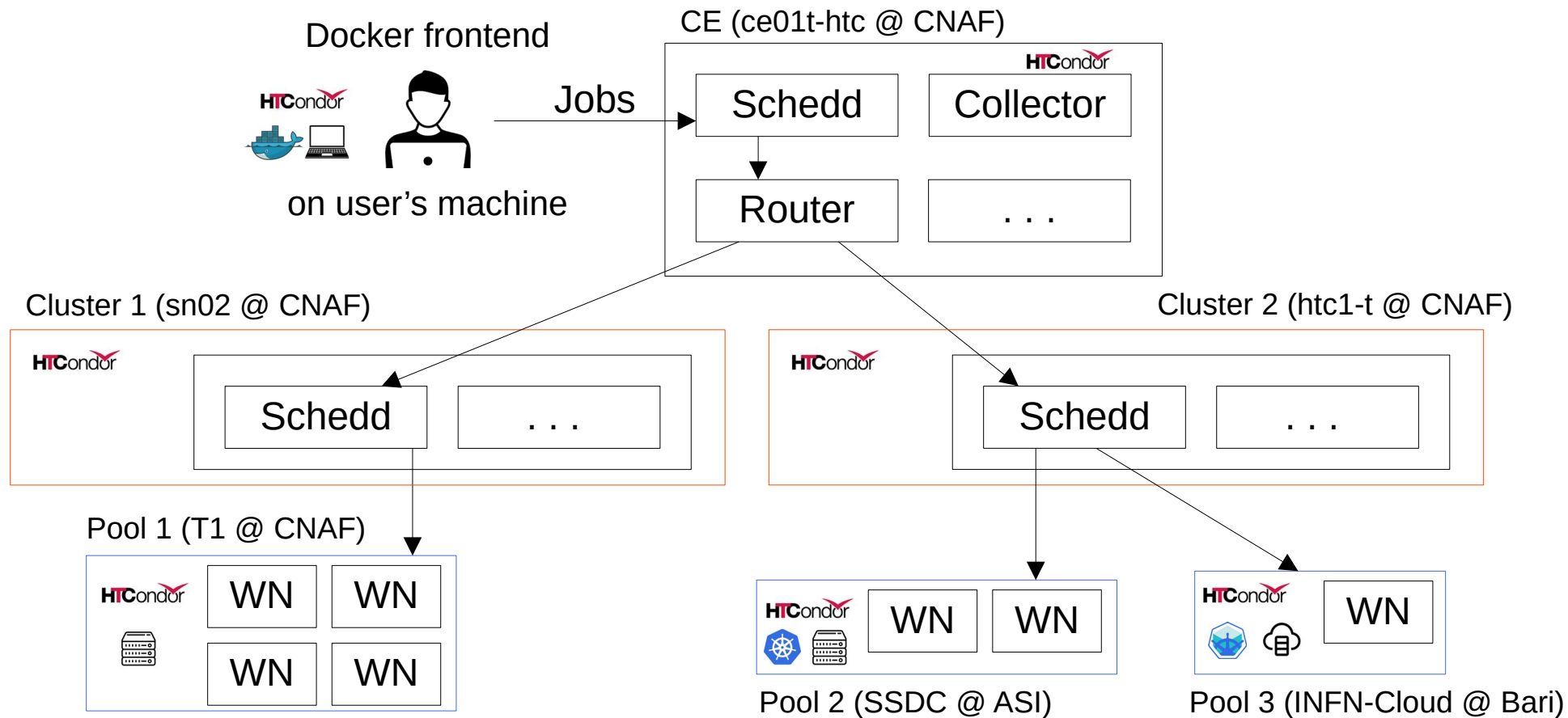


HTCondor-CE









- HTCondor Compute Entrypoint (CE)



Infrastruttura



Infrastruttura

- HTCondor worker nodes su bare metal (T1)  
- HTCondor worker nodes su k8s/minikube
 - Su k8s on bare metal (ASI)   
 - Su minikube on VM (Bari)   
 - Replicaset di containers con Condor startd + ambiente di runtime
 - Scaling manuale
 - Cluster token
 - CVMS daemonset
- SSDC farm:
 - 384 CPU
+
384 CPU con HT
 - 90 TB storage (cache)
 - Connessioni in fibra a 10 Gbps
 - k8s control plane a PG
 - Setup di test: 3 macchine

Infrastruttura

- AuthN/Z: IAM HERD
 - Access tokens via oidc-agent integrato nell'ambiente di frontend containerizzato
 - Refresh automatico via cron job per migliorare l'esperienza utente
 - Gestione autorizzazioni via gruppi IAM + client scopes
- Creazione e mapping dinamici degli utenti IAM/Condor
 - Subject (a853db99-539a...) → Condor user (herd####)
 - Plugin per HTCondor sviluppato ad hoc

Workflow utente

- Sottomissione job da frontend containerizzato
 - Script di lancio → submit file + recupero pre-signed URLs per I/O + librerie/configfiles utente
 - AuthN/Z via token tramite oidc-agent
- Esecuzione job:
 - Recupero input da S3 via pre-signed URL
 - Processing
 - Output su S3 via pre-signed URL
- Accesso POSIX a S3 da frontend via rclone mount
 - Per ispezione
 - AuthN/Z via token tramite oidc-agent

Sviluppi futuri

- Finalizzazione cluster SSDC + storage
- Tuning delle configurazioni di CE e schedd (regole di routing ecc.)
- Gestione long-lived tokens con MyToken
- Rework del data handling nei WNs
 - Pre-signed URLs → token
 - StoRM endpoints

Sviluppi futuri

- Definizione del modello di data storage
 - RUCIO
- Gestione librerie utente
 - CVMFS
 - Latenza ?
 - S3 (upload da frontend via POSIX, download su WN via S3 client + token)

Sviluppi (molto) futuri

- Realizzazione di un multi-tenant in SSDC
 - Risorse separate per ogni esperimento
- Scaling dinamico delle risorse per esperimento basato sulle richieste del momento (es. # job in coda)
 - Interazione k8s-condor
- Creazione di data caches nei siti (?)
- Integrazione di altri schedd
 - Risorse in cluster Condor già esistenti