

Roadmap per il computing

Giusy, Carmelo, Diego, Stefano, Andrea

Roadmap di novembre ancora valida

- Asset management
 - Marzo 2021, 1 FTE (farming)
- Ironic
 - Testbed al CNAF a settembre 2021, 1.5 FTE (farming, ds)
- Integrazione farm (HTC/HPC)
 - 2 mesi a trasferimento avvenuto, 1 FTE (farming, us)
- Espansione dinamica
 - Già possibile, 0.5 FTE per implementare tutti gli strumenti
- Orchestrazione
 - Testbed al CNAF ad aprile 2021, 1.5 FTE (farming, ds)

Batch Computing

- In attesa dell'effettiva migrazione intendiamo valutare e confrontare soluzioni possibili previste per calcolo batch
- Criteri di preferenza:
 - Migliori performances e/o capabilities
 - Semplicità di gestione ordinaria
- Due HTCondor testbed:
 - “stable” per cluster di produzione (test di configurazioni e operazioni “delicate”)
 - “latest” HTCondor release (test nuove features e soluzioni)

Test previsti (WN)

- WN come docker container su bare metal
 - Un container per host fisico (primi test iniziati)
 - Un container per CPU/RAM
 - Pro: job multithread usano core sulla stessa CPU e RAM «vicina»
 - Cons: Draining time medio per avere slot multicore raddoppia
- WN come POD K8S
 - Per ciascuna soluzione: verifica management, capabilities e performances rispetto a uso corrente (storage/network/shared fs)
 - Es: Aggiornamenti sw: nel running WN o nell'immagine, o entrambi
- Provisioning/Management: moduli puppet per provisioning e configurazione exec node
- Registry: gestione immagini docker, K8S

Altri Test

- HTCondor-CE
 - Token Id. Da iniziare subito e passare in produzione con la prima HTCondor stable release che lo supporti.
 - gateway verso risorse HTC/HPC (HTCondor/Slurm)
 - verso risorse GPU.
- Ironic (OpenStack) per bare metal provisioning
 - Pro: facilita condivisione in OpenStack di risorse HW (uso HV / WN)
 - Verificare: implicazioni per networking, storage, dipendenze da ulteriori servizi

Orchestratori

BATCH SYSTEM (LRMS)	HTCondor / Slurm	Provisioning/management di CPUs/GPUs/... a Job
COMPUTE NODE (CN)	Docker / K8S / libVirt / ...	Provisioning/management di WN a BATCH SYSTEM
BARE METAL (BM)	Ironic / Foreman / ...	Provisioning/management di Host as COMPUTE NODE

- Necessaria sinergia tra i differenti layer. Es. Come si coordina un security upgrade sw + kernel?
 1. LRMS: mettere in Drain I CN
 2. CN: aggiornare sw nell'immagine
 3. CN: Shutdown quando vuoti (ma lo sa il layer superiore!)
 4. BM: aggiornare immagine col nuovo kernel, bootloader;
 5. BM: reboot vuoti (ma lo sa il layer superiore)
- Una problematica simile esiste da tempo nel farming: coordinare drain, aggiornamenti e reboot di WN.
 - Risolta con suite di script prodotti in casa (es. *dynup*)
- Verosimilmente necessario ri-scrivere tool ad hoc per *orchestrare* interazione tra orchestratori