Red Hat Summit

Connect

Il cluster compatto per container, VM e AI

Come in **EXTRAORDY** abbiamo pensato un cluster OpenShift Bare Metal, ultra compatto e con hardware specializzato (GPU). Un'unica piattaforma dove eseguire container e virtual machine, pronta per l'AI.

EXTRAORDY WeAreProject | Red Hat

Stefano Stagnaro

CTO
EXTRAORDY – WeAreProject

EXTRAORDY
WeAreProject

EXTRAORDY WeAreProject | Red Hat

Il dilemma dell'IT moderno

Le infrastrutture IT aziendali si trovano oggi a un bivio critico, caratterizzato da una frammentazione che genera inefficienza e rallenta l'innovazione.

Stack legacy

Hypervisor tradizionali come VMware vSphere per applicazioni monolitiche e database critici. Affidabile ma costoso e rigido.

Stack cloud-native

Container, microservizi e carichi AI/ML su bare metal. Agile e performante ma operativamente isolato.

Le conseguenze della frammentazione

TCO elevato

Duplicazione di costi hardware, licenze software, team specializzati e consumo energetico su entrambi gli stack.

Complessità operativa

Strumenti separati per monitoraggio, sicurezza, networking e backup aumentano rischi e carico di lavoro.

Freno all'innovazione

L'integrazione tra VM legacy e servizi AI containerizzati diventa complessa, rallentando la modernizzazione.

Il debito infrastrutturale impedisce alle aziende di realizzare il pieno potenziale dell'AI. I dati più preziosi per l'AI risiedono in sistemi legacy, ma portare accelerazione GPU e MLOps vicino a questi dati è estremamente difficile.

EXTRAORDY
WeAreProject

La visione di EXTRAORDY

Una Piattaforma unificata su OpenShift bare metal

Perché Red Hat OpenShift?

Piano di controllo unificato: gestione coerente di VM, container e serverless tramite OpenShift Virtualization

Operator Framework: automazione completa del ciclo di vita con operatori certificati

Leader Kubernetes enterprise: esperienza consistente dal data center all'edge

Perché bare metal?

Zero "Hypervisor Tax": accesso diretto a CPU, memoria e I/O per massime performance

Latenza minima: cruciale per inferenza AI real-time e database transazionali

Hardware avanzato: pieno utilizzo di VT-x, GPU e tecnologie come SR-IOV

Validazione delle performance

Zero Overhead: Elimina lo strato di virtualizzazione (hypervisor), dedicando il 100% delle risorse hardware ai carichi di lavoro.

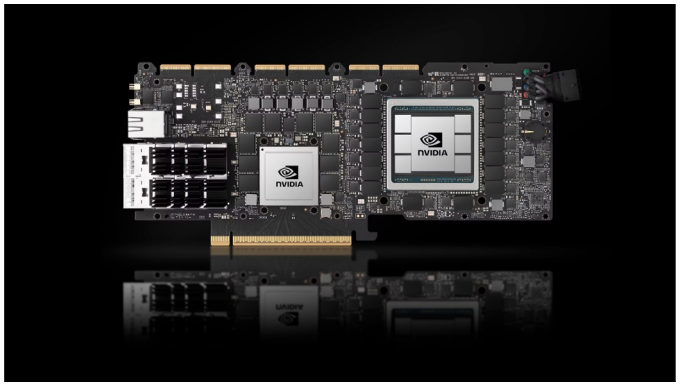
Latenza Minima: Fornisce accesso diretto a GPU, NVMe e rete, cruciale per carichi I/O intensivi e AI/ML.

Composable OpenShift: la nostra architettura rappresenta una composizione ottimizzata, pre-configurata per unificare VM, container e AI – massimizzando efficienza e posizionandoci all'avanguardia dell'innovazione.

EXTRAORDY

Pronti per l'AI

Accelerazione GPU per Ogni Workload



NVIDIA GPU Operator

Automazione end-to-end dello stack NVIDIA: driver kernel, container toolkit, device plugin e DCGM. Scoperta automatica GPU con NFD e scheduling intelligente tramite label specifiche.

AMD GPU Operator

Semplifica l'implementazione e la gestione degli acceleratori **AMD Instinct** GPU all'interno dei cluster Kubernetes. Consente una configurazione e un funzionamento fluidi dei carichi di lavoro accelerati da GPU, come il machine learning, l'Intelligenza Artificiale Generativa (Generative AI) e altre applicazioni GPU-intensive che sfruttano l'ecosistema **ROCm**.

Accelerazione flessibile multimodale

Container con OpenShift AI

È la piattaforma MLOps per costruire, addestrare e distribuire modelli AI/ML su scala in un ambiente ibrido.

VM con GPU Passthrough

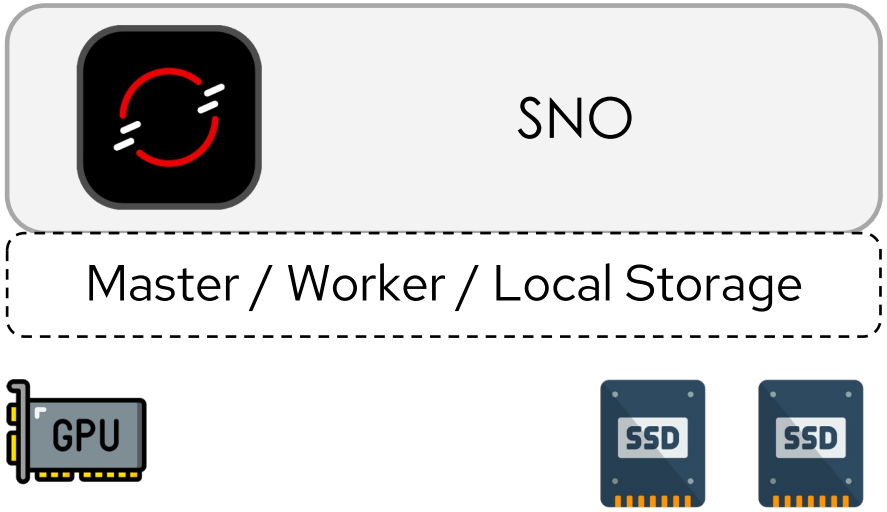
Assegnazione esclusiva di GPU intera via VFIO. Performance bare metal per training intensivo, HPC e workstation grafiche.

Condivisione avanzata:

- **NVIDIA MIG**: nativo/proprietario NVIDIA
- **SR-IOV**: standard ma non 100% implementato
- **Time-slicing**: sempre possibile

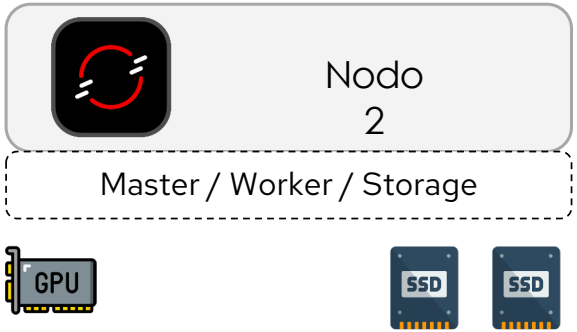
EXTRAORDY
WeAreProject

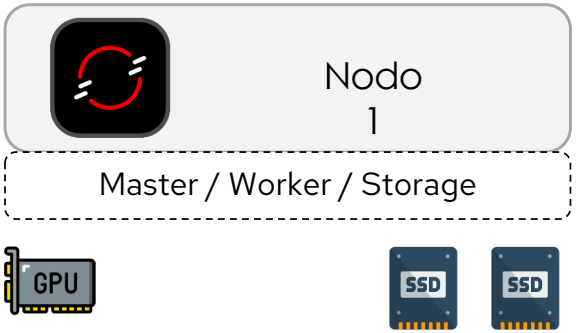
Architettura a singolo nodo

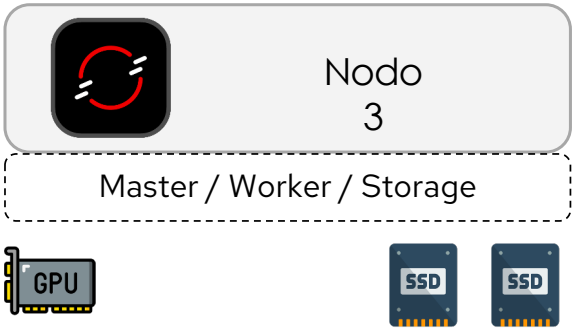


EXTRAORDY WeAreProject | Red Hat

Architettura a tre nodi

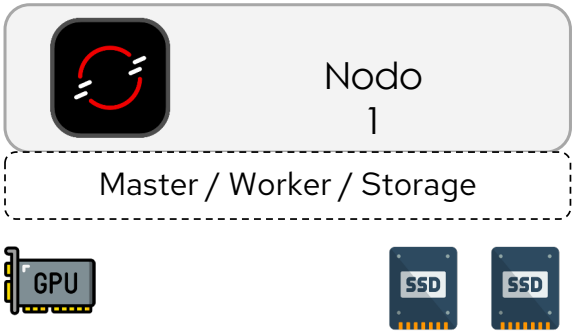


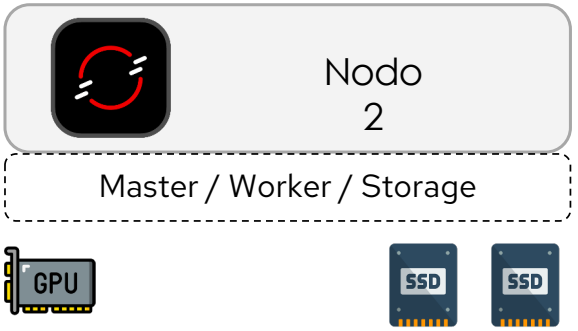




EXTRAORDY WeAreProject | Red Hat

Architettura 2+1





EXTRAORDY WeAreProject | Red Hat

Prova per un mese OpenShift

Per consentire alle Aziende di valutare efficacemente la validità della soluzione Red Hat OpenShift, **EXTRAORDY** offre un pacchetto “chiavi in mano” per effettuare una PoC della durata di un mese durante la quale:

Fase 0 - Setup Infrastrutturale (1 settimana, 5 gg lavorativi)

Fase 1 - Proof of Concept (2 settimane, 10 gg lavorativi)

- Condivisione dei criteri di successo per la valutazione della PoC
- Presentazione di OpenShift ed eventuali altre componenti (AI, Virtualization)
- Raccolta dei dati generati dall’esecuzione di test quali, ad esempio
  - Importazione VM da VMware
  - Operazioni VM (Live Migration, snapshot, clone)
  - Setup ambiente OpenShift AI e GPU
  - Backup & Restore
- Documentazione dei risultati con focus sulle eventuali criticità emerse

Fase 2 - “Free play” (2 settimane, 10 gg lavorativi)

L’infrastruttura viene lasciata al cliente come laboratorio “libero” al fine di provare eventuali scenari che non sono stati considerati all’interno della PoC. In questa fase, il supporto da parte di EXTRAORDY è limitato alle sole necessità tecniche.

EXTRAORDY WeAreProject | Red Hat

E dopo?

Al termine delle quattro settimane di Proof of Concept, si aprono due possibili scenari:

HERO

PoC conclusa con successo, si decide di procedere.

L'infrastruttura utilizzata può essere immediatamente convertita in un ambiente di produzione, previa sottoscrizione delle componenti software (per l'on-prem) o della controparte cloud.

ZERO

A prescindere dall'esito della PoC, si decide di non proseguire con il progetto.

Non c'è più alcun obbligo nei confronti di **EXTRAORDY** e si riceverà comunque copia della documentazione prodotta durante l'attività.

EXTRAORDY WeAreProject | Red Hat

Perché fidarsi di EXTRAORDY

Gli unici con oltre 20 anni di esperienza SOLO su **Red Hat**, siamo immediatamente operativi avendo implementato con successo l'installazione Bare Metal in vari contesti applicativi:

- ➔ Società di Gestione e Risparmio;
- ➔ Primario gruppo bancario del Nord Italia;
- ➔ Prestigioso Istituto di Ricerca Oncologica.

EXTRAORDY WeAreProject | Red Hat

Red Hat Summit

Connect

Compila il form →

Lavoriamo ad una **PoC** di virtualizzazione personalizzata per te insieme.

EXTRAORDY WeAreProject | Red Hat