

WORK
SHOP
GARR
2023

NET
MAKERS

Resilienza, scalabilità e sostenibilità di soluzioni di private cloud per la ricerca



Politecnico
di Torino

Davide Stevano
Politecnico di Torino

Un confronto non convenzionale

In rete sono presenti molti articoli sul confronto tra infrastrutture open source basate su OpenStack e soluzioni commerciali VMware.

Alcuni non sono del tutto centrati perché comparano il prodotto **VMware vSphere** (una piattaforma di virtualizzazione) con **OpenStack** (insieme di strumenti che permette la realizzazione di cloud)

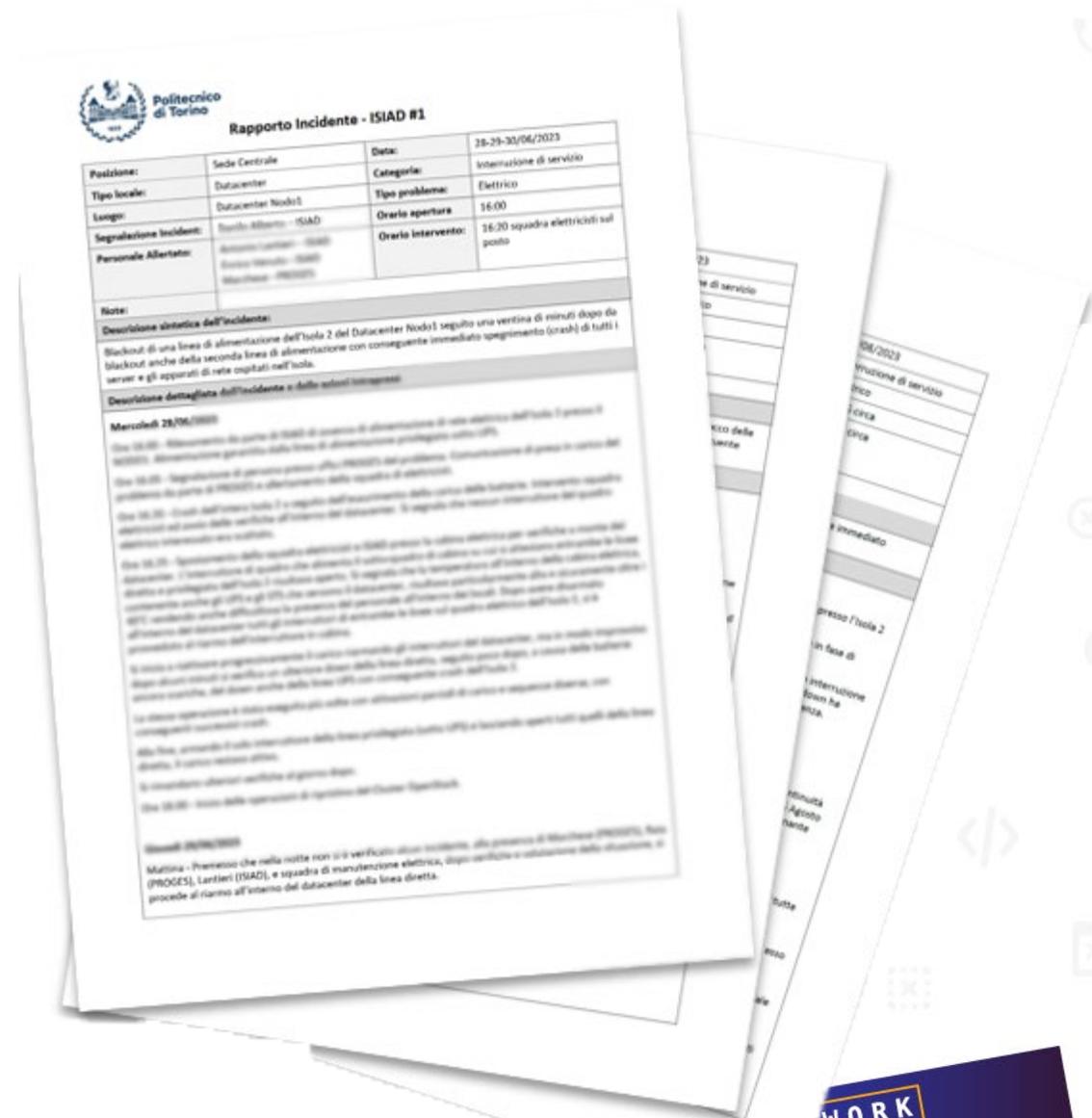
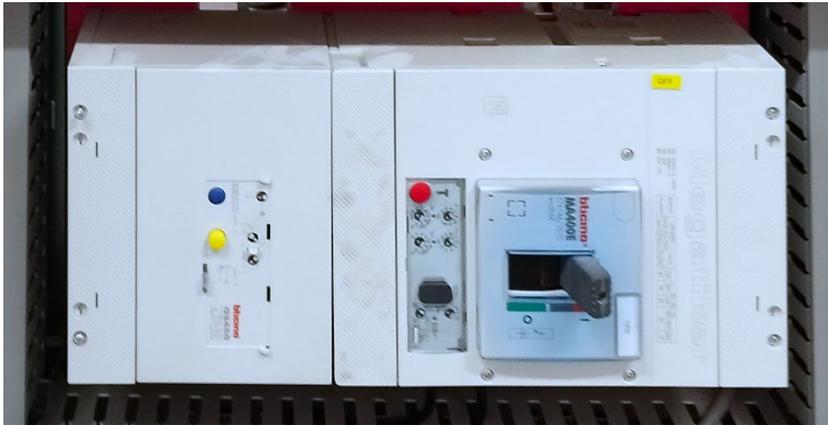
Altri sono presentazioni commerciali di una o dell'altra parte o redatti da «fanboy»

Punto di partenza del nostro confronto è un evento piuttosto inaspettato: **gravi problemi elettrici** hanno colpito un'isola del datacenter con impatto su due infrastrutture della ricerca



Gli eventi elettrici

Può un guasto ad interruttore in una cabina elettrica far rivalutare l'intero **processo di sviluppo del cloud privato** per la ricerca del Politecnico di Torino?



L'isola 2 nel datacenter primo nodo

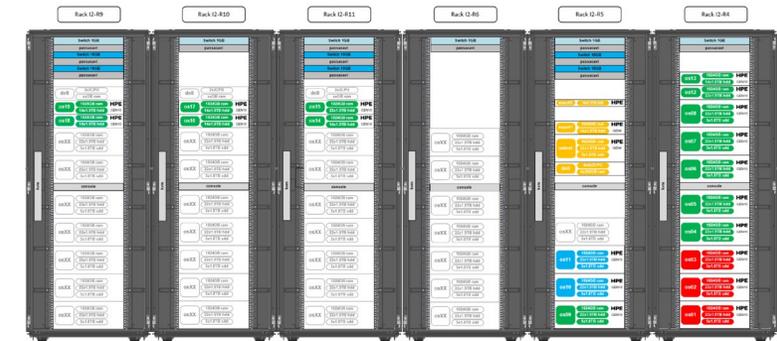
Cloud privato open source basato su **OpenStack** e regione federata GARR:

- 35 server iperconvergenti + 4 server fra mng e 2 istanze di test
- Circa 5000 vCPU, 30TB RAM, 650 HDD e 45 SSD

Un cluster basato su **VMware vSphere** e **vSAN**, con possibilità di backup usando il sw Veeam

- 5 server iperconvergenti
- 312 vCPU, 2.5TB RAM, 100 HDD e 20 SSD

Circa 15 server bare metal per progetti indipendenti



Gli incidenti elettrici

Dal 28 giugno al 1 ottobre, per problemi elettrici nella cabina che alimenta l'isola 2:

7 eventi di spegnimento totale o parziale nonostante alimentazioni ridondate, UPS e gruppo elettrogeno.

Nel primo evento ci sono state decine di up/down dell'alimentazione elettrica

- **42** ore in cui i server bare metal sono stati down
- **47** ore in cui il cluster vSphere vSAN è stato down
- **150** ore in cui il cluster OpenStack è stato down
- **330** ore impiegate per risolvere diversi problemi sulle singole VM OpenStack e renderle nuovamente disponibili agli utenti

In attesa della sostituzione dell'interruttore di cabina, i rack sono stati alimentati solo tramite le PDU

Danni: 2 DIMM memoria, 1 alimentatore, 5 server OS indisponibili per 90 giorni

Errori e bug

Tipologie di errori riscontrate:

- Alcuni dispositivi **senza la doppia alimentazione** non erano alimentati (kvm, console, luci corridoio caldo)
- Sovraccarichi locali sulle PDU causati dallo **sbilanciamento prolungato** del carico hanno provocato distacchi a rack o sue sezioni
- Corruzione **log cluster db mysql-innodb** di produzione (sulle due istanze di test problemi al cluster diversi)
- Le **azioni di recover definite nei charm** non hanno funzionato e hanno costretto ad un restore manuale del database
- **Volumi cinder in read-only** dopo la riaccensione non governata
- **Errori umani e di coordinamento** (per verificare che non ci fossero server in corto circuito, in vista dei lavori estivi alle cabine elettriche, è stato eseguito il test dell'UPS fino al suo completo scaricamento, dimenticando che fosse l'unica alimentazione attiva per i rack dell'isola)
- Perdita **configurazione di alcuni switch** nei rack
- **Firmware** schede di rete server: baco non documentato che rendeva isolate le VM

Punti da migliorare

- Documentazione snella, raggiungibile anche in caso di blackout, condivisa con i tutti i settori e uffici coinvolti
- Procedure di recover e riaccensione, schemi e script testate e sempre aggiornate dopo upgrade
- Riconsiderare l'iperconvergenza almeno sui nodi controller e network di OpenStack
- Verificare gli strumenti di monitoraggio per OS o implementarne di nuovi: Nagios e marginalmente Juju in alcune circostanze segnalano come «green» servizi che sono «up» ma non stanno funzionando e non rilevano malfunzionamenti delle VM. VMware vSphere già nelle versioni standard ha degli strumenti di monitoraggio che soddisfano queste esigenze. Sono presenti anche strumenti di diagnostica self-service (VMware Skyline Health)

The image displays several screenshots from IT management tools. The top left shows a file explorer with a 'Documentation' folder containing files like 'clone_instance_volume.md' and 'command_cli.md'. Below this is the vSphere Client interface, showing a 'VSAN' storage configuration with a list of hosts and their status. To the right is a 'Skyline Health' dashboard for VSAN, showing various health checks like 'vSAN cluster partition' and 'vSAN MTU check'. At the bottom left is a Nagios monitoring dashboard showing 'Current Network Status', 'Host Status Totals', and 'Service Status Totals'. At the bottom right is another vSphere Client screenshot showing 'vSphere HA' and 'Resource Allocation' settings.

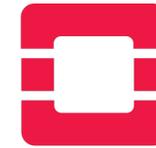


La provocazione

- In ognuno dei blackout, alla riaccensione il cluster VMware ha ripreso a funzionare spontaneamente. Hypervisor e VM non hanno richiesto alcun intervento umano.
- E se valutassimo di sostituire il cluster OpenStack attualmente in produzione con un sistema basato su VMware che possa fornire servizi simili?

vmware®

42h down



openstack®

150h down

some VMs down for 330h

Analisi dei bisogni degli utenti regione polito Openstack

- Attualmente la regione polito ospita circa 180 virtual machine. Dimensione media 16 vCPU e 64GB RAM.
- Il 15% delle VM attualmente utilizza SO Windows :
 - perché usano codice commerciale disponibile solo su tale sistema operativo
 - utenti hanno più dimestichezza con Windows che non con Linux
- Pochi progetti: attualmente meno di dieci
- Pochi utenti hanno una forte motivazione a sfruttare gli strumenti IaaS



Verificare la comunicazione con gli utenti: potremmo non valorizzare sufficientemente i vantaggi del private cloud rispetto ad una singola VM

VMware e il licensing

- La scelta del prodotto non è semplice, sul listino della convenzione «VMware Licenze Software Multibrand edizione 5» delle PA ci sono circa **1.300 voci**
- VMware ciclicamente applica politiche di rebranding sui prodotti integrando funzionalità. Tendenza a sostituire le licenze di proprietà con sottoscrizioni mensili o annuali
- VMware è contributor di OpenStack. Esiste anche una distribuzione OS di VMware anche se scarsamente utilizzata e mantenuta
- Nel dicembre 2019 VMware acquisisce un'azienda chiamata Pivotal
 - da lì a poco rilascia «Tanzu Community edition», una soluzione open source per creare cluster **k8s** basata su vSphere
 - a dicembre 2022 la ritira dando in cambio la possibilità di usare la versione commerciale fino a 100 core gratuitamente
- Curiosità: il software opensource **RabbitMQ** utilizzato in OpenStack è stato creato da Pivotal ed è attualmente mantenuto da VMware

La scelta

1. Per il soddisfare le esigenze del 90% degli utenti della nostra regione potrebbe essere sufficiente aggiungere alle licenze Crui vSphere unlimited la parte storage e rete virtuale (vSan + NSX). **NON** è un cloud privato
2. VMware Integrated OpenStack 7b (potrebbe essere un prodotto a breve in EOL visto che la nuova versione uscita a giugno 2023 è solo "Train")
3. VMware Cloud Foundation 4 Advanced: possibilità di avere VM e k8s multitenant

Vantaggi di utilizzare vSphere:

DRS bilancia il carico spostando le VM da un host all'altro

HA in caso guasto di un server le VM ripartono in automatico su un altro host

Lifecycle manager per gestione automatic di update host, driver server, tool VM

Site recovery manager componente opzionale

	vSphere "Crui"	Vsan ADV	NSX	Tanzu
1	+	+		
2	+	+	+	
3	compreso			

Costi

- Per calcolare i costi, impiegato il listino della convenzione Crui per università italiane. La cifra per le licenze vSphere enterprise + unlimited è calcolata come fosse la prima adesione e dipende da una tabella in base alle dimensioni dell'ateneo
- Sono state scelte le licenze di proprietà con 3 anni di supporto (obbligatorio per accedere alle nuove versioni)

	vSphere "Crui" Ent + unlim	vSAN ADV	NSX	Tanzu
1	dai 27 ai 110k	4,2k a cpu	5,2k a cpu	
2	dai 27 ai 110k	4,2k a cpu	5,2k a cpu	
3	31k a cpu			

Il costo per convertire il nostro attuale cluster OS (35 server da 4 CPU) a **VMware Cloud Foundation 4 Advanced** è di **4.340.000** euro con tre anni di supporto

Le altre due soluzioni, per una università come il Politecnico di Torino, costerebbero circa **1.350.000** euro, sempre con tre anni di supporto

Conclusioni

Valore infrastruttura **private Cloud OpenStack** 500.000€ + 50.000€ apparati network

Costo **annuale** considerando 5 anni di vita: 110.000€ + 50.000€ di energia → **160.000€**

Costi addizionali x passaggio a soluzione chiusa **vSphere**: 4.340.000€

Costo **annuale** addizionale: oltre 1.440.000€

Costo **annuale** totale: circa **1.600.000€**

Costo adozione **VM direttamente su Public Cloud** (Azure) → **1.600.000 €**

Cifra ottenuta impiegando vm per vCPU, RAM, spazio disco

- applicando uno sconto arbitrario del 64% sul costo «per use»
- non considerando i costi di «outband»
- non considerando i costi di storage I/O

vmware®

42h down

DRS - HA - Lifecycle
manager - Site recovery



openstack®

150h down



Grazie per l'ascolto

www.garr.it/domande
codice: 2318 9129

Ringraziamenti: Enrico Venuto e Paola Belmonte

WORK
SHOP
GARR
2023

**NET
MAKERS**

