Immersione profonda

ONTAP Select

NetApp
February 03, 2026

This PDF was generated from https://docs.netapp.com/it-it/ontap-select-9141/concept_stor_concepts_chars.html on February 03, 2026. Always check docs.netapp.com for the latest.

Sommario

Immersione profonda . . . 1
- Storage . . . 1
  - Storage: Concetti e caratteristiche generali . . . 1
  - Servizi RAID hardware per lo storage locale ONTAP Select . . . 6
  - Servizi RAID software per lo storage collegato in locale . . . 13
  - Configurazioni di array esterno e VSAN . . . 22
  - Aumentare la capacità di storage . . . 26
  - Supporto per l'efficienza dello storage . . . 29
- Networking . . . 32
  - Networking: Concetti e caratteristiche generali . . . 32
  - Configurazioni di rete a nodo singolo e multiplo . . . 34
  - Rete interna ed esterna ONTAP Select . . . 39
  - Configurazioni di rete supportate . . . 42
  - Configurazione di VMware vSphere vSwitch su ESXi . . . 43
  - Configurazione fisica dello switch . . . 52
  - Separazione dei dati e del traffico di gestione . . . 54
- Architettura ad alta disponibilità . . . 56
  - Configurazioni ad alta disponibilità . . . 56
  - HA RSM e aggregati mirrorati . . . 59
  - HA ulteriori dettagli . . . 62
- Performance . . . 65
  - Performance . . . 65
  - Performance: Storage SSD direct-attached ha Premium . . . 65

Immersione profonda

Storage

Storage: Concetti e caratteristiche generali

Scopri i concetti di storage generali applicabili all'ambiente ONTAP Select prima di esaminare i componenti di storage specifici.

Fasi della configurazione dello storage

Le principali fasi di configurazione dello storage host ONTAP Select includono:

- Prerequisiti di pre-implementazione
  - Assicurarsi che ciascun host dell'hypervisor sia configurato e pronto per un'implementazione ONTAP Select.
  - La configurazione comprende dischi fisici, controller e gruppi RAID, LUN e relativa preparazione della rete.
  - Questa configurazione viene eseguita al di fuori di ONTAP Select.
- Configurazione mediante l'utility di amministrazione dell'hypervisor
  - È possibile configurare alcuni aspetti dello storage utilizzando l'utility di amministrazione dell'hypervisor (ad esempio, vSphere in un ambiente VMware).
  - Questa configurazione viene eseguita al di fuori di ONTAP Select.
- Configurazione mediante l'utility di amministrazione di ONTAP Select Deploy
  - È possibile utilizzare l'utilità di amministrazione di Deploy per configurare i costrutti di storage logico principali.
  - Questa operazione viene eseguita in modo esplicito tramite comandi CLI o automaticamente dall'utility come parte di una distribuzione.
- Configurazione post-implementazione
  - Una volta completata l'implementazione di ONTAP Select, è possibile configurare il cluster utilizzando l'interfaccia utente di ONTAP o Gestore di sistema.
  - Questa configurazione viene eseguita al di fuori dell'implementazione di ONTAP Select.

Storage gestito e non gestito

Lo storage accessibile e controllato direttamente da ONTAP Select è lo storage gestito. Qualsiasi altro storage sullo stesso host hypervisor è uno storage non gestito.

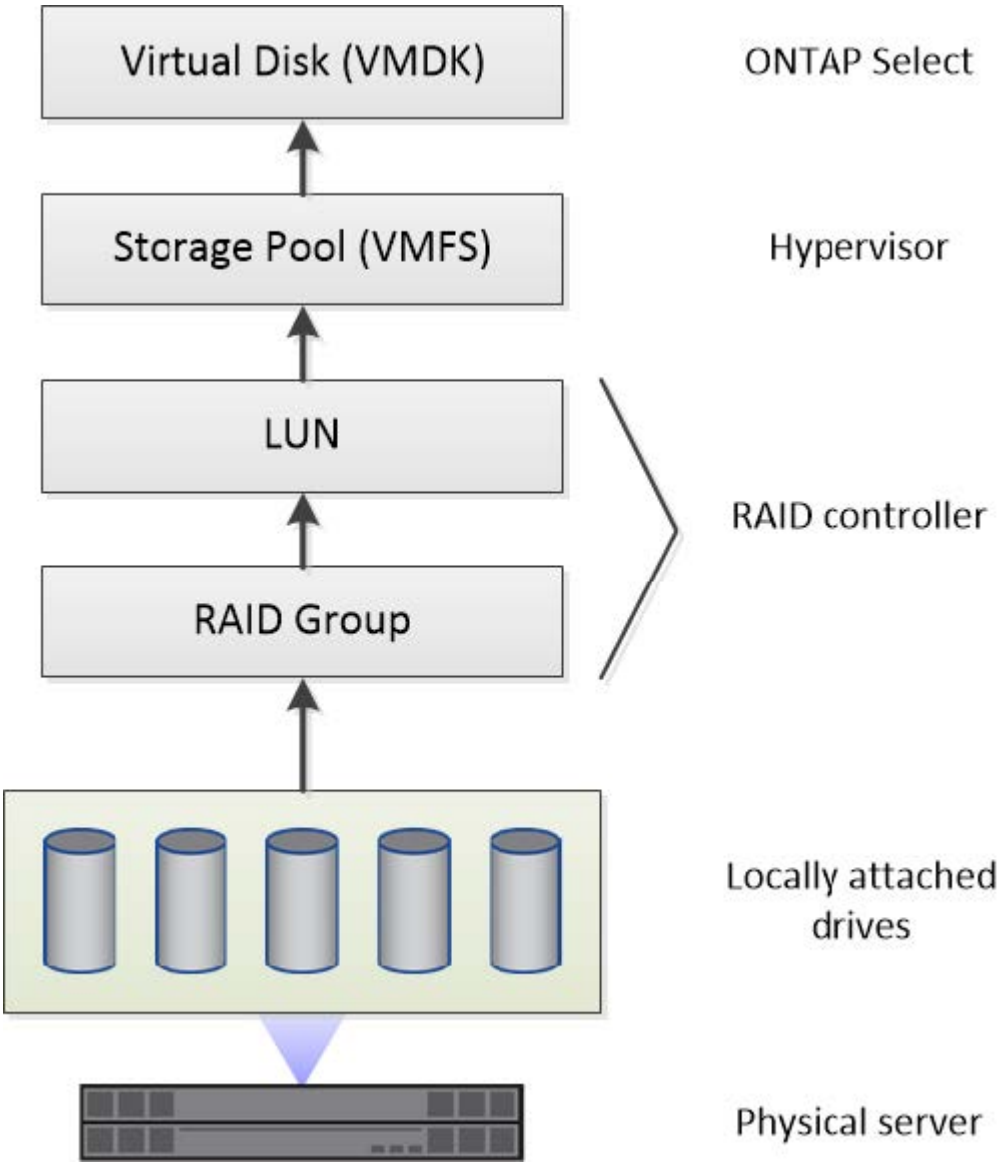
Storage fisico omogeneo

Tutti i dischi fisici che costituiscono lo storage gestito da ONTAP Select devono essere omogenei. Vale a dire che tutto l'hardware deve essere lo stesso per quanto riguarda le seguenti caratteristiche:

- TIPO (SAS, NL-SAS, SATA, SSD)
- Velocità (giri/min)

Immagine dell'ambiente di storage locale

Ogni host hypervisor contiene dischi locali e altri componenti di storage logico che possono essere utilizzati da ONTAP Select. Questi componenti di storage sono disposti in una struttura a più livelli, dal disco fisico.



Caratteristiche dei componenti dello storage locale

Esistono diversi concetti che si applicano ai componenti dello storage locale utilizzati in un ambiente ONTAP Select. Prima di prepararsi per un'implementazione di ONTAP Select, è necessario acquisire familiarità con questi concetti. Questi concetti sono organizzati in base alla categoria: Gruppi RAID e LUN, pool di storage e dischi virtuali.

Raggruppamento di dischi fisici in gruppi RAID e LUN

Uno o più dischi fisici possono essere collegati localmente al server host e disponibili per ONTAP Select. I dischi fisici vengono assegnati ai gruppi RAID, che vengono quindi presentati al sistema operativo host dell'hypervisor come una o più LUN. Ogni LUN viene presentato al sistema operativo host dell'hypervisor come disco rigido fisico.

Durante la configurazione di un host ONTAP Select, tenere presente quanto segue:

- Tutto lo storage gestito deve essere accessibile attraverso un singolo controller RAID
- A seconda del vendor, ciascun controller RAID supporta un numero massimo di dischi per gruppo RAID

Uno o più gruppi RAID

Ogni host ONTAP Select deve disporre di un singolo controller RAID. Creare un singolo gruppo RAID per ONTAP Select. Tuttavia, in alcune situazioni è possibile creare più di un gruppo RAID. Fare riferimento a. "Riepilogo delle Best practice".

Considerazioni sul pool di storage

Esistono diversi problemi relativi ai pool di storage di cui si dovrebbe essere a conoscenza durante la preparazione all'implementazione di ONTAP Select.

In un ambiente VMware, un pool di storage è sinonimo di un datastore VMware.

Pool di storage e LUN

Ogni LUN viene considerato come un disco locale sull'host dell'hypervisor e può essere parte di un pool di storage. Ogni pool di storage è formattato con un file system utilizzabile dal sistema operativo host dell'hypervisor.

È necessario assicurarsi che i pool di storage siano creati correttamente come parte di un'implementazione di ONTAP Select. È possibile creare un pool di storage utilizzando lo strumento di amministrazione dell'hypervisor. Ad esempio, con VMware è possibile utilizzare il client vSphere per creare un pool di storage. Il pool di storage viene quindi trasferito all'utilità di amministrazione di ONTAP Select Deploy.

Gestire i dischi virtuali su ESXi

Esistono diversi problemi relativi ai dischi virtuali di cui si dovrebbe essere a conoscenza durante la preparazione dell'implementazione di ONTAP Select.

Dischi virtuali e file system

Alla macchina virtuale ONTAP Select sono allocati più dischi virtuali. Ogni disco virtuale è in realtà un file contenuto in un pool di storage e viene gestito dall'hypervisor. ONTAP Select utilizza diversi tipi di dischi, principalmente dischi di sistema e dischi dati.

È inoltre necessario conoscere quanto segue in merito ai dischi virtuali:

- Il pool di storage deve essere disponibile prima di poter creare i dischi virtuali.
- I dischi virtuali non possono essere creati prima della creazione della macchina virtuale.
- Per creare tutti i dischi virtuali, è necessario affidarsi all'utilità di amministrazione di ONTAP Select Deploy (ovvero, un amministratore non deve mai creare un disco virtuale al di fuori dell'utilità di implementazione).

Configurazione dei dischi virtuali

I dischi virtuali sono gestiti da ONTAP Select. Vengono creati automaticamente quando si crea un cluster utilizzando l'utilità di amministrazione di Deploy.

Immagine dell'ambiente di storage esterno su ESXi

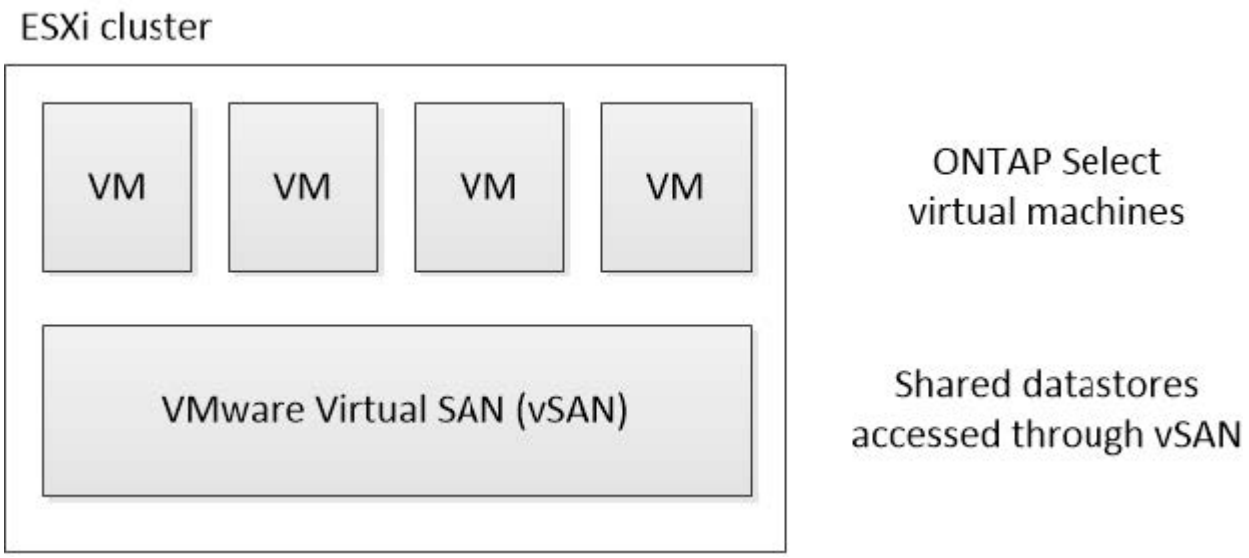
La soluzione vNAS di ONTAP Select consente a ONTAP Select di utilizzare datastore che risiedono su storage esterno all'host dell'hypervisor. È possibile accedere agli archivi dati tramite la rete utilizzando VMware vSAN o direttamente da uno storage array esterno.

ONTAP Select può essere configurato per utilizzare i seguenti tipi di datastore di rete VMware ESXi esterni all'host dell'hypervisor:

- VSAN (SAN virtuale)
- VMFS
- NFS

Datastore VSAN

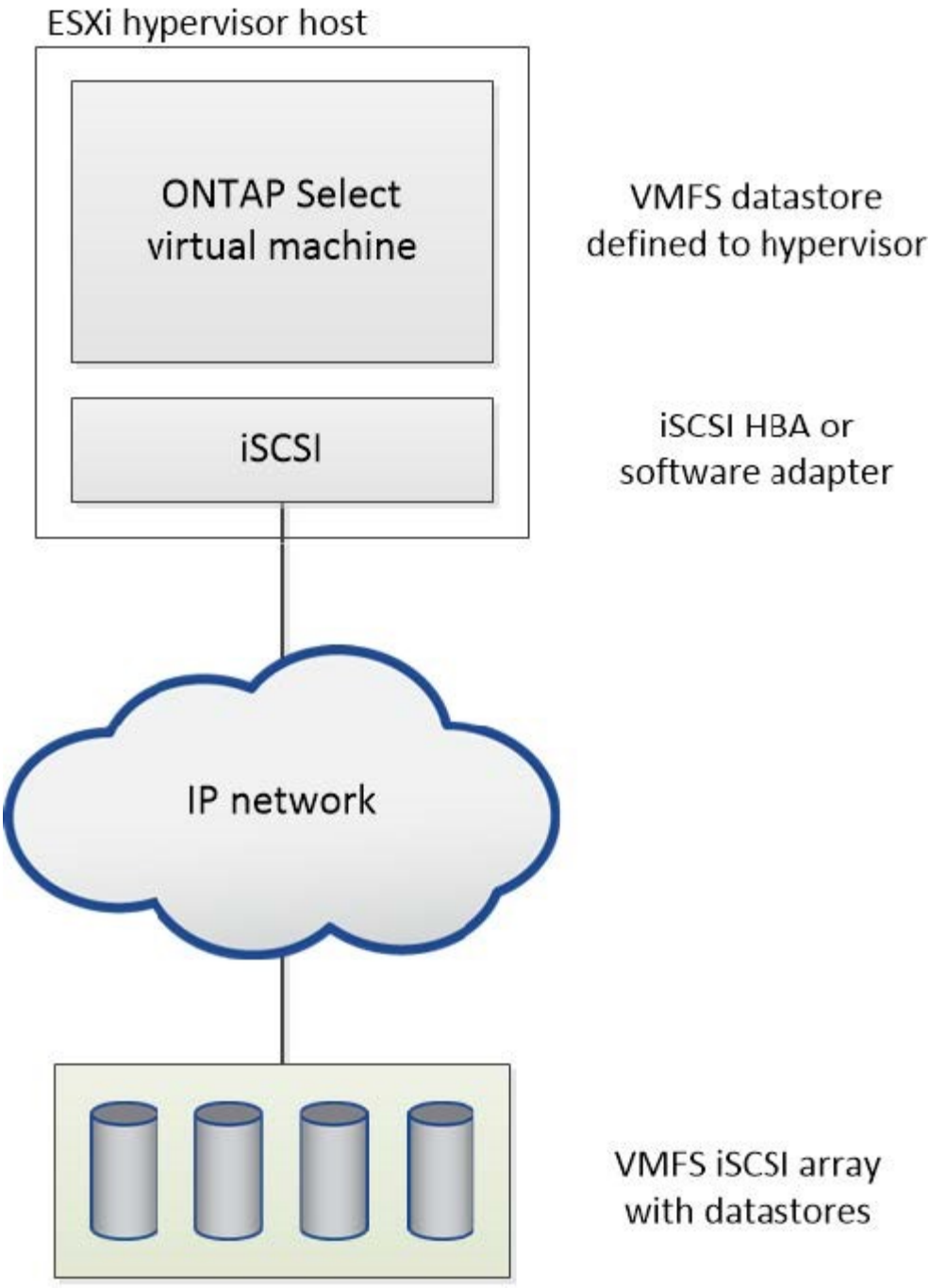
Ogni host ESXi può disporre di uno o più datastore VMFS locali. Normalmente questi datastore sono accessibili solo all'host locale. Tuttavia, VMware vSAN consente a ciascuno degli host di un cluster ESXi di condividere tutti i datastore del cluster come se fossero locali. La figura seguente illustra come vSAN crea un pool di datastore condivisi tra gli host nel cluster ESXi.



Datastore VMFS su storage array esterno

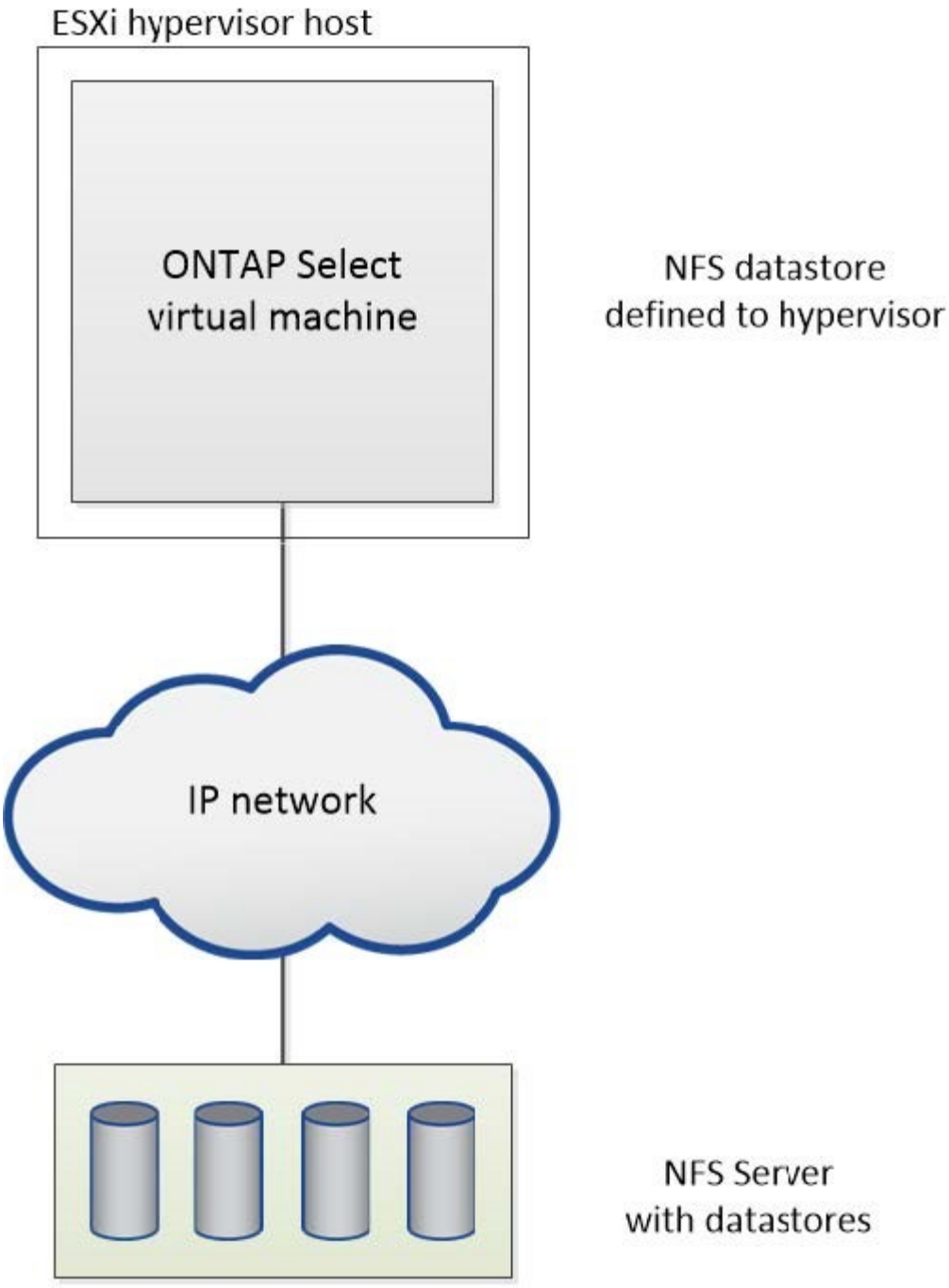
È possibile creare un datastore VMFS residente su un array di storage esterno. È possibile accedere allo storage utilizzando uno dei diversi protocolli di rete. La figura seguente illustra un datastore VMFS su un array di storage esterno a cui si accede utilizzando il protocollo iSCSI.

ONTAP Select supporta tutti gli array di storage esterni descritti nella documentazione relativa alla compatibilità di storage/SAN VMware, inclusi iSCSI, Fiber Channel e Fiber Channel over Ethernet.



Datastore NFS sullo storage array esterno

È possibile creare un datastore NFS che risiede su uno storage array esterno. L'accesso allo storage avviene tramite il protocollo di rete NFS. La figura seguente illustra un datastore NFS su storage esterno a cui si accede tramite l'appliance server NFS.



Servizi RAID hardware per lo storage locale ONTAP Select

Quando è disponibile un controller RAID hardware, ONTAP Select può spostare i servizi RAID nel controller hardware per aumentare le performance di scrittura e proteggere dai guasti fisici dei dischi. Di conseguenza, la protezione RAID per tutti i nodi all'interno del cluster ONTAP Select viene fornita dal controller RAID collegato localmente e non dal software RAID ONTAP.

Gli aggregati di dati ONTAP Select sono configurati per l'utilizzo di RAID 0 perché il controller RAID fisico sta fornendo striping RAID ai dischi sottostanti. Non sono supportati altri livelli RAID.

Configurazione del controller RAID per lo storage locale collegato

Tutti i dischi collegati localmente che forniscono a ONTAP Select lo storage di backup devono essere posizionati dietro un controller RAID. La maggior parte dei server commodity è dotata di diverse opzioni di controller RAID in diversi prezzi, ciascuno con diversi livelli di funzionalità. L'intento è quello di supportare il maggior numero possibile di queste opzioni, purché soddisfino determinati requisiti minimi sul controller.

Non è possibile scollegare dischi virtuali dalle macchine virtuali ONTAP Select che utilizzano la configurazione RAID hardware. La rimozione dei dischi è supportata solo per le macchine virtuali ONTAP Select che utilizzano la configurazione RAID software. Per ulteriori informazioni, vedere "Sostituire un'unità guasta in una configurazione RAID software ONTAP Select" .

Il controller RAID che gestisce i dischi ONTAP Select deve soddisfare i seguenti requisiti:

- Il controller RAID hardware deve disporre di un'unità di backup della batteria (BBU) o di una cache di scrittura con backup flash (FBWC) e supportare un throughput di 12 Gbps.
- Il controller RAID deve supportare una modalità in grado di resistere ad almeno uno o due guasti del disco (RAID 5 e RAID 6).
- La cache del disco deve essere impostata su Disabled (disattivata).
- Il criterio di scrittura deve essere configurato per la modalità di writeback con un fallback da scrivere in caso di guasto alla BBU o alla flash.
- Il criterio i/o per le letture deve essere impostato su cache.

Tutti i dischi collegati localmente che forniscono ONTAP Select con storage di backup devono essere collocati in gruppi RAID che eseguono RAID 5 o RAID 6. Per i dischi SAS e gli SSD, l'utilizzo di gruppi RAID fino a 24 dischi consente a ONTAP di sfruttare i vantaggi dello spread delle richieste di lettura in entrata su un numero più elevato di dischi. In questo modo si ottiene un significativo guadagno in termini di performance. Con le configurazioni SAS/SSD, i test delle performance sono stati eseguiti su configurazioni con LUN singola o su più LUN. Non sono state rilevate differenze significative, pertanto, per semplicità, NetApp consiglia di creare il minor numero di LUN necessari per supportare le esigenze di configurazione.

I dischi NL-SAS e SATA richiedono un set diverso di Best practice. Per motivi di performance, il numero minimo di dischi è ancora otto, ma la dimensione del gruppo RAID non deve essere superiore a 12 dischi. NetApp consiglia inoltre di utilizzare uno spare per gruppo RAID; tuttavia, è possibile utilizzare spare globali per tutti i gruppi RAID. Ad esempio, è possibile utilizzare due unità di riserva per ogni tre gruppi RAID, con ciascun gruppo RAID composto da otto a 12 unità.

L'estensione massima e la dimensione del datastore per le release ESX precedenti sono 64 TB, il che può influire sul numero di LUN necessari per supportare la capacità raw totale fornita da questi dischi a elevata capacità.

Modalità RAID

Molti controller RAID supportano fino a tre modalità operative, ciascuna delle quali rappresenta una differenza significativa nel percorso dei dati preso dalle richieste di scrittura. Queste tre modalità sono le seguenti:

- WriteThrough. Tutte le richieste di i/o in entrata vengono scritte nella cache del controller RAID e quindi immediatamente inviate al disco prima di confermare la richiesta all'host.
- Writearound. Tutte le richieste di i/o in entrata vengono scritte direttamente su disco, eludendo la cache del controller RAID.
- Sriveback. Tutte le richieste di i/o in entrata vengono scritte direttamente nella cache del controller e

riconnesse immediatamente all'host. I blocchi di dati vengono trasferiti su disco in modo asincrono utilizzando il controller.

La modalità di writeback offre il percorso dati più breve, con riconoscimento i/o che si verifica immediatamente dopo l'ingresso dei blocchi nella cache. Questa modalità offre la latenza più bassa e il throughput più elevato per carichi di lavoro misti in lettura/scrittura. Tuttavia, senza la presenza di una BBU o di una tecnologia flash non volatile, gli utenti corrono il rischio di perdere i dati se il sistema subisce un'interruzione dell'alimentazione durante il funzionamento in questa modalità.

ONTAP Select richiede la presenza di una batteria di backup o di un'unità flash; pertanto, possiamo essere sicuri che i blocchi memorizzati nella cache vengano scaricati sul disco in caso di questo tipo di guasto. Per questo motivo, è necessario che il controller RAID sia configurato in modalità writeback.

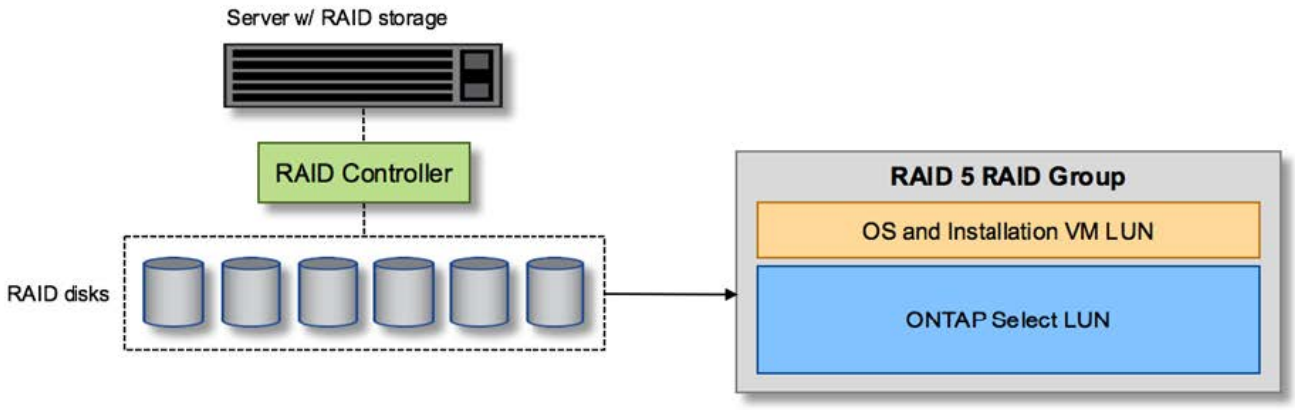
Dischi locali condivisi tra ONTAP Select e il sistema operativo

La configurazione del server più comune è quella in cui tutti gli spindle collegati localmente si trovano dietro un singolo controller RAID. È necessario eseguire il provisioning di almeno due LUN: Uno per l'hypervisor e uno per la macchina virtuale ONTAP Select.

Ad esempio, si consideri un HP DL380 g8 con sei dischi interni e un singolo controller RAID Smart Array P420i. Tutti i dischi interni sono gestiti da questo controller RAID e nessun altro storage è presente nel sistema.

La figura seguente mostra questo stile di configurazione. In questo esempio, nessun altro storage è presente nel sistema; pertanto, l'hypervisor deve condividere lo storage con il nodo ONTAP Select.

Configurazione del LUN del server con solo spindle gestiti da RAID



Il provisioning dei LUN del sistema operativo dallo stesso gruppo RAID di ONTAP Select consente al sistema operativo dell'hypervisor (e a qualsiasi macchina virtuale client fornita da tale storage) di beneficiare della protezione RAID. Questa configurazione impedisce che un guasto a un disco singolo causi il blocco dell'intero sistema.

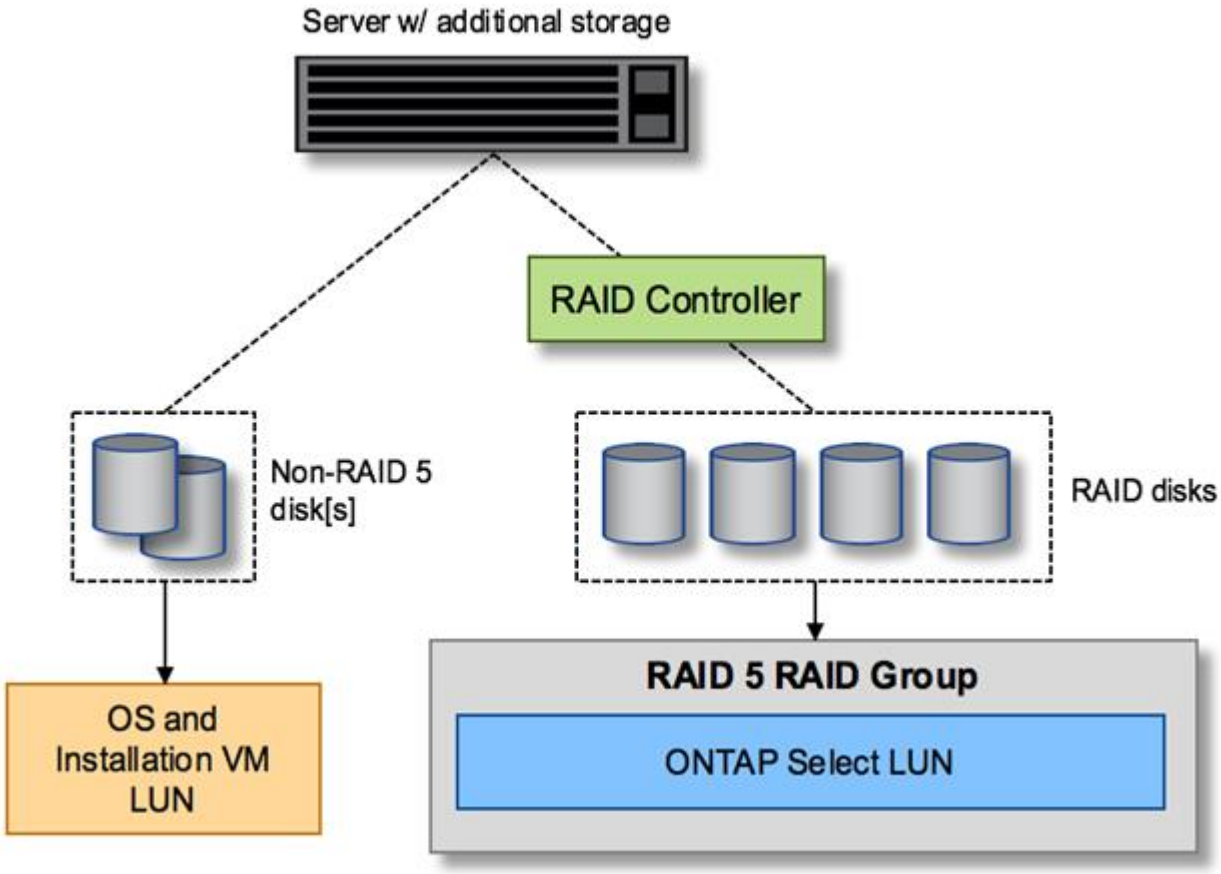
Dischi locali suddivisi tra ONTAP Select e sistema operativo

L'altra possibile configurazione fornita dai vendor di server prevede la configurazione del sistema con più RAID o controller di dischi. In questa configurazione, un set di dischi viene gestito da un controller di dischi, che potrebbe offrire o meno servizi RAID. Un secondo set di dischi è gestito da un controller RAID hardware in grado di offrire servizi RAID 5/6.

Con questo stile di configurazione, il set di spindle che si trovano dietro il controller RAID in grado di fornire i servizi RAID 5/6 deve essere utilizzato esclusivamente dalla macchina virtuale ONTAP Select. A seconda della capacità di storage totale in gestione, è necessario configurare gli spindle dei dischi in uno o più gruppi RAID e in una o più LUN. Questi LUN vengono quindi utilizzati per creare uno o più datastore, con tutti i datastore protetti dal controller RAID.

Il primo set di dischi è riservato al sistema operativo dell'hypervisor e a qualsiasi macchina virtuale client che non utilizza lo storage ONTAP, come illustrato nella figura seguente.

Configurazione del LUN del server su sistemi RAID/non RAID misti



LUN multipli

Esistono due casi in cui è necessario modificare le configurazioni di un singolo gruppo RAID/LUN singolo. Quando si utilizzano dischi NL-SAS o SATA, le dimensioni del gruppo RAID non devono superare i 12 dischi. Inoltre, un singolo LUN può diventare più grande dei limiti di storage dell'hypervisor sottostante, sia per quanto riguarda le dimensioni massime dei singoli file system che per quelle dei pool di storage totali. Quindi, lo storage fisico sottostante deve essere suddiviso in più LUN per consentire la corretta creazione del file system.

Limiti del file system per macchine virtuali VMware vSphere

La dimensione massima di un datastore su alcune versioni di ESX è di 64 TB.

Se un server dispone di oltre 64 TB di storage collegati, potrebbe essere necessario eseguire il provisioning di più LUN, ciascuna inferiore a 64 TB. La creazione di più gruppi RAID per migliorare il tempo di ricostruzione RAID per i dischi SATA/NL-SAS comporta anche il provisioning di più LUN.

Quando sono necessarie più LUN, un importante punto da considerare è assicurarsi che queste LUN abbiano performance simili e coerenti. Ciò è particolarmente importante se tutte le LUN devono essere utilizzate in un singolo aggregato ONTAP. In alternativa, se un sottoinsieme di una o più LUN ha un profilo di performance nettamente diverso, si consiglia di isolare tali LUN in un aggregato ONTAP separato.

È possibile utilizzare più estensioni del file system per creare un singolo datastore fino alle dimensioni massime del datastore. Per limitare la capacità che richiede una licenza ONTAP Select, assicurarsi di specificare un limite di capacità durante l'installazione del cluster. Questa funzionalità consente a ONTAP Select di utilizzare (e quindi richiedere una licenza per) solo un sottoinsieme dello spazio in un datastore.

In alternativa, è possibile iniziare creando un singolo datastore su una singola LUN. Quando è necessario uno spazio aggiuntivo che richiede una licenza di capacità ONTAP Select più grande, è possibile aggiungere tale spazio allo stesso datastore di un'estensione, fino alle dimensioni massime del datastore. Una volta raggiunta la dimensione massima, è possibile creare nuovi datastore e aggiungerli a ONTAP Select. Entrambi i tipi di operazioni di estensione della capacità sono supportati e possono essere ottenuti utilizzando la funzionalità storage-add di ONTAP Deploy. Ogni nodo ONTAP Select può essere configurato per supportare fino a 400 TB di storage. Il provisioning della capacità da più datastore richiede un processo in due fasi.

La creazione iniziale del cluster può essere utilizzata per creare un cluster ONTAP Select che occupa parte o tutto lo spazio nel datastore iniziale. Una seconda fase consiste nell'eseguire una o più operazioni di aggiunta di capacità utilizzando datastore aggiuntivi fino a raggiungere la capacità totale desiderata. Questa funzionalità è illustrata in dettaglio nella sezione "Aumentare la capacità di storage".

> L'overhead VMFS non è pari a zero (vedere "VMware KB 1001618"), e il tentativo di utilizzare l'intero spazio riportato come libero da un datastore ha causato errori spuri durante le operazioni di creazione del cluster.

Un buffer del 2% viene lasciato inutilizzato in ciascun datastore. Questo spazio non richiede una licenza di capacità perché non viene utilizzato da ONTAP Select. ONTAP Deploy calcola automaticamente il numero esatto di gigabyte per il buffer, a condizione che non venga specificato un limite di capacità. Se viene specificato un limite massimo di capacità, tale dimensione viene applicata per prima. Se la dimensione del limite di capacità rientra nella dimensione del buffer, la creazione del cluster non riesce e viene visualizzato un messaggio di errore che specifica il parametro di dimensione massima corretto che può essere utilizzato come limite di capacità:

```
"InvalidPoolCapacitySize: Invalid capacity specified for storage pool
"ontap-select-storage-pool", Specified value: 34334204 GB. Available
(after leaving 2% overhead space): 30948"
```

VMFS 6 è supportato sia per le nuove installazioni che come destinazione di un'operazione di storage vMotion di un'implementazione ONTAP esistente o di una macchina virtuale ONTAP Select.

VMware non supporta gli aggiornamenti in-place da VMFS 5 a VMFS 6. Pertanto, Storage vMotion è l'unico meccanismo che consente a qualsiasi macchina virtuale di passare da un datastore VMFS 5 a un datastore VMFS 6. Tuttavia, il supporto per Storage vMotion con ONTAP Select e ONTAP Deploy è stato esteso per coprire altri scenari oltre allo scopo specifico di passare da VMFS 5 a VMFS 6.

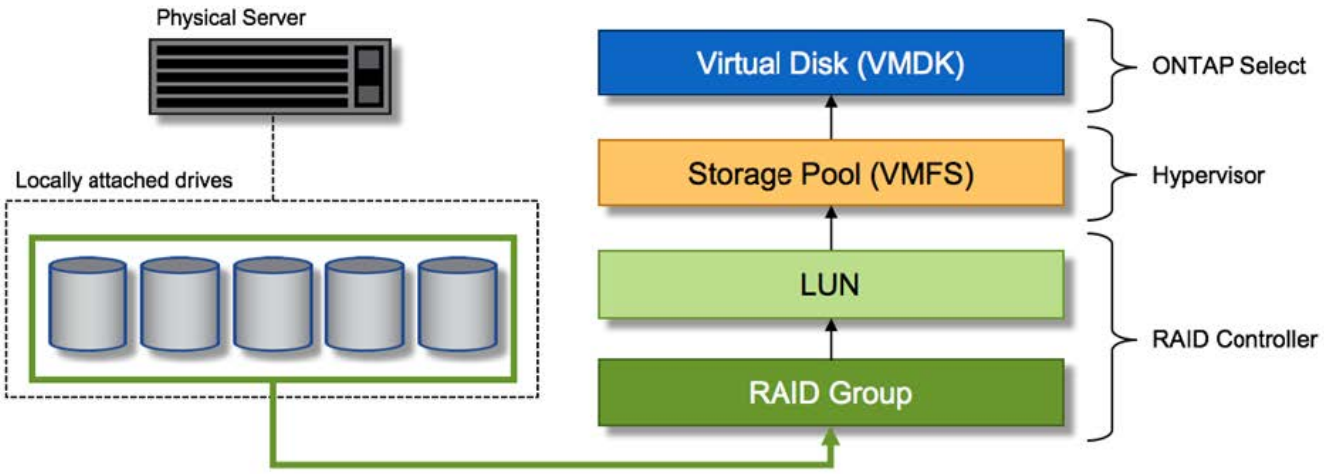
Dischi virtuali ONTAP Select

In primo luogo, ONTAP Select presenta ONTAP con un set di dischi virtuali forniti da uno o più pool di storage. ONTAP viene presentato con un set di dischi virtuali che considera fisici e la parte rimanente dello stack di storage viene astratta dall'hypervisor. La figura seguente mostra questa relazione in maggiore dettaglio,

evidenziando la relazione tra il controller RAID fisico, l'hypervisor e la macchina virtuale ONTAP Select.

- La configurazione del gruppo RAID e del LUN avviene all'interno del software del controller RAID del server. Questa configurazione non è necessaria quando si utilizzano array VSAN o esterni.
- La configurazione del pool di storage avviene dall'interno dell'hypervisor.
- I dischi virtuali sono creati e di proprietà di singole macchine virtuali; in questo esempio, ONTAP Select.

Mappatura da disco virtuale a disco fisico



Provisioning di dischi virtuali

Per offrire un'esperienza utente più ottimizzata, lo strumento di gestione ONTAP Select, ONTAP Deploy, esegue il provisioning automatico dei dischi virtuali dal pool di storage associato e li collega alla macchina virtuale ONTAP Select. Questa operazione viene eseguita automaticamente durante la configurazione iniziale e durante le operazioni di aggiunta dello storage. Se il nodo ONTAP Select fa parte di una coppia ha, i dischi virtuali vengono assegnati automaticamente a un pool di storage locale e mirror.

ONTAP Select suddivide lo storage collegato sottostante in dischi virtuali di uguali dimensioni, ciascuno non superiore a 16 TB. Se il nodo ONTAP Select fa parte di una coppia ha, vengono creati almeno due dischi virtuali su ciascun nodo del cluster e assegnati al plesso locale e mirror da utilizzare all'interno di un aggregato mirrorato.

Ad esempio, un ONTAP Select può assegnare un datastore o un LUN pari a 31 TB (lo spazio rimanente dopo l'implementazione della macchina virtuale e il provisioning dei dischi di sistema e root). Vengono quindi creati quattro dischi virtuali da ~7,75 TB e assegnati al plesso locale e mirror ONTAP appropriato.

> L'aggiunta di capacità a una macchina virtuale ONTAP Select probabilmente comporta VMDK di dimensioni diverse. Per ulteriori informazioni, consultare la sezione "Aumentare la capacità di storage". A differenza dei sistemi FAS, nello stesso aggregato possono esistere VMDK di dimensioni diverse. ONTAP Select utilizza uno stripe RAID 0 in tutti questi VMDK, il che consente di utilizzare completamente tutto lo spazio in ogni VMDK indipendentemente dalle sue dimensioni.

NVRAM virtualizzata

I sistemi NetApp FAS sono tradizionalmente dotati di una scheda PCI NVRAM fisica, una scheda dalle performance elevate contenente memoria flash non volatile. Questa scheda offre un significativo

miglioramento delle prestazioni di scrittura, consentendo a ONTAP di riconoscere immediatamente le scritture in entrata nel client. Può anche pianificare lo spostamento dei blocchi di dati modificati sui supporti di storage più lenti in un processo noto come destaging.

I sistemi commodity in genere non sono dotati di questo tipo di apparecchiatura. Pertanto, la funzionalità di questa scheda NVRAM è stata virtualizzata e inserita in una partizione sul disco di avvio del sistema ONTAP Select. È per questo motivo che il posizionamento del disco virtuale di sistema dell'istanza è estremamente importante. Questo è anche il motivo per cui il prodotto richiede la presenza di un controller RAID fisico con una cache resiliente per le configurazioni di storage locale collegato.

La NVRAM viene inserita nel proprio VMDK. La suddivisione della NVRAM nel proprio VMDK consente alla VM ONTAP Select di utilizzare il driver vNVMe per comunicare con il VMDK NVRAM. Richiede inoltre che la macchina virtuale ONTAP Select utilizzi la versione hardware 13, che è compatibile con ESX 6.5 e versioni successive.

Spiegazione del percorso dei dati: NVRAM e controller RAID

L'interazione tra la partizione di sistema NVRAM virtualizzata e il controller RAID può essere evidenziata al meglio attraverso il percorso dei dati intrapreso da una richiesta di scrittura quando entra nel sistema.

Le richieste di scrittura in entrata alla macchina virtuale ONTAP Select sono indirizzate alla partizione NVRAM della macchina virtuale. Al livello di virtualizzazione, questa partizione esiste all'interno di un disco di sistema ONTAP Select, un VMDK collegato alla macchina virtuale ONTAP Select. Al livello fisico, queste richieste vengono memorizzate nella cache del controller RAID locale, come tutte le modifiche dei blocchi mirate agli spindle sottostanti. Da qui, la scrittura viene ricondotto all'host.

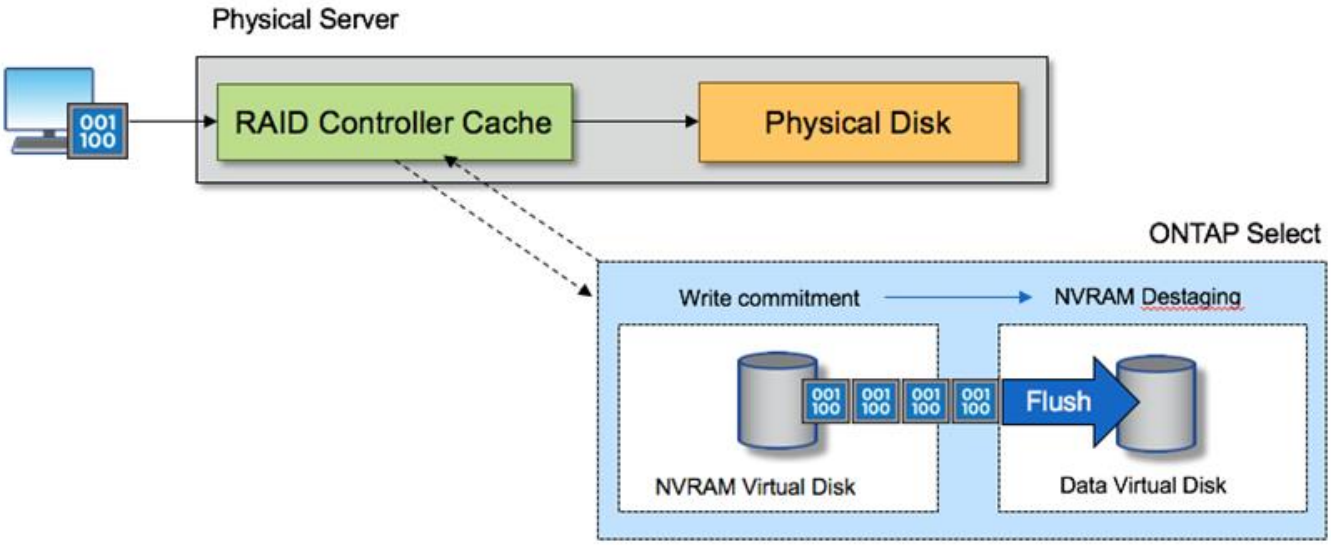
A questo punto, fisicamente, il blocco risiede nella cache del controller RAID, in attesa di essere trasferito su disco. Logicamente, il blocco risiede nella NVRAM in attesa di destaging sui dischi dati dell'utente appropriati.

Poiché i blocchi modificati vengono memorizzati automaticamente nella cache locale del controller RAID, le scritture in entrata nella partizione NVRAM vengono automaticamente memorizzate nella cache e periodicamente salvate nel supporto di storage fisico. Questo non deve essere confuso con il rossore periodico del contenuto NVRAM sui dischi dati ONTAP. Questi due eventi non sono correlati e si verificano in orari e frequenze differenti.

La figura seguente mostra il percorso i/o richiesto da una scrittura in entrata. Evidenzia la differenza tra il layer fisico (rappresentato dalla cache e dai dischi del controller RAID) e il layer virtuale (rappresentato dalla NVRAM della macchina virtuale e dai dischi virtuali dei dati).

> Sebbene i blocchi modificati nella NVRAM VMDK siano memorizzati nella cache del controller RAID locale, la cache non è a conoscenza del costrutto della macchina virtuale o dei suoi dischi virtuali. Memorizza tutti i blocchi modificati nel sistema, di cui la NVRAM è solo una parte. Sono incluse le richieste di scrittura associate all'hypervisor, se il provisioning viene eseguito dagli stessi spindle di backup.

Scritture in entrata su ONTAP Select VM



La partizione NVRAM è separata sul proprio VMDK. Che VMDK sia collegato utilizzando il driver vNVME disponibile nelle versioni ESX della versione 6.5 o successiva. Questa modifica è più significativa per le installazioni ONTAP Select con RAID software, che non beneficiano della cache del controller RAID.

Servizi RAID software per lo storage collegato in locale

Il RAID software è un livello di astrazione RAID implementato all'interno dello stack software ONTAP. Fornisce le stesse funzionalità del livello RAID all'interno di una piattaforma ONTAP tradizionale come FAS. Il livello RAID esegue i calcoli di parità dei dischi e fornisce protezione da guasti a singoli dischi all'interno di un nodo ONTAP Select.

Indipendentemente dalle configurazioni RAID hardware, ONTAP Select offre anche un'opzione RAID software. Un controller RAID hardware potrebbe non essere disponibile o essere indesiderabile in alcuni ambienti, ad esempio quando ONTAP Select viene implementato su un hardware commodity con fattore di forma ridotto. Software RAID espande le opzioni di implementazione disponibili per includere tali ambienti. Per abilitare il RAID software nel tuo ambiente, ecco alcuni punti da ricordare:

- È disponibile con una licenza Premium o Premium XL.
- Supporta solo dischi SSD o NVMe (richiede licenza Premium XL) per dischi root e dati ONTAP.
- Richiede un disco di sistema separato per la partizione di boot di ONTAP Select VM.
  - Scegliere un disco separato, SSD o NVMe, per creare un datastore per i dischi di sistema (NVRAM, scheda Boot/CF, coredump e Mediator in una configurazione multi-nodo).

Note

- I termini disco di servizio e disco di sistema vengono utilizzati in modo intercambiabile.
  - I dischi di servizio sono i VMDK che vengono utilizzati all'interno della macchina virtuale ONTAP Select per gestire diversi elementi come clustering, avvio e così via.
  - I dischi di servizio si trovano fisicamente su un singolo disco fisico (chiamato collettivamente disco fisico di servizio/sistema) visto dall'host. Il disco fisico deve contenere un datastore DAS. ONTAP

Deployment crea questi dischi di servizio per la macchina virtuale ONTAP Select durante l'implementazione del cluster.

- Non è possibile separare ulteriormente i dischi di sistema ONTAP Select tra più datastore o su più dischi fisici.
- Il RAID hardware non è obsoleto.

Configurazione RAID software per lo storage collegato in locale

Quando si utilizza il RAID software, l'assenza di un controller RAID hardware è ideale, ma se un sistema dispone di un controller RAID esistente, deve rispettare i seguenti requisiti:

- Il controller RAID hardware deve essere disattivato in modo che i dischi possano essere presentati direttamente al sistema (un JBOD). Questa modifica può in genere essere apportata nel BIOS del controller RAID
- In alternativa, il controller RAID hardware deve essere in modalità SAS HBA. Ad esempio, alcune configurazioni del BIOS consentono una modalità "AHCI" oltre a RAID, che può essere scelta per attivare la modalità JBOD. In questo modo viene attivato un pass-through, in modo che i dischi fisici possano essere visti come sono sull'host.

A seconda del numero massimo di dischi supportati dal controller, potrebbe essere necessario un controller aggiuntivo. Con la modalità HBA SAS, assicurarsi che il controller di i/o (HBA SAS) sia supportato con una velocità minima di 6 GB/s. Tuttavia, NetApp consiglia una velocità di 12 Gbps.

Non sono supportate altre configurazioni o modalità del controller RAID hardware. Ad esempio, alcuni controller consentono un supporto RAID 0 che può consentire artificialmente il passaggio dei dischi, ma le implicazioni possono essere indesiderate. Le dimensioni dei dischi fisici supportati (solo SSD) sono comprese tra 200 GB e 16 TB.

Gli amministratori devono tenere traccia dei dischi utilizzati dalla macchina virtuale ONTAP Select e impedire l'utilizzo involontario di tali dischi sull'host.

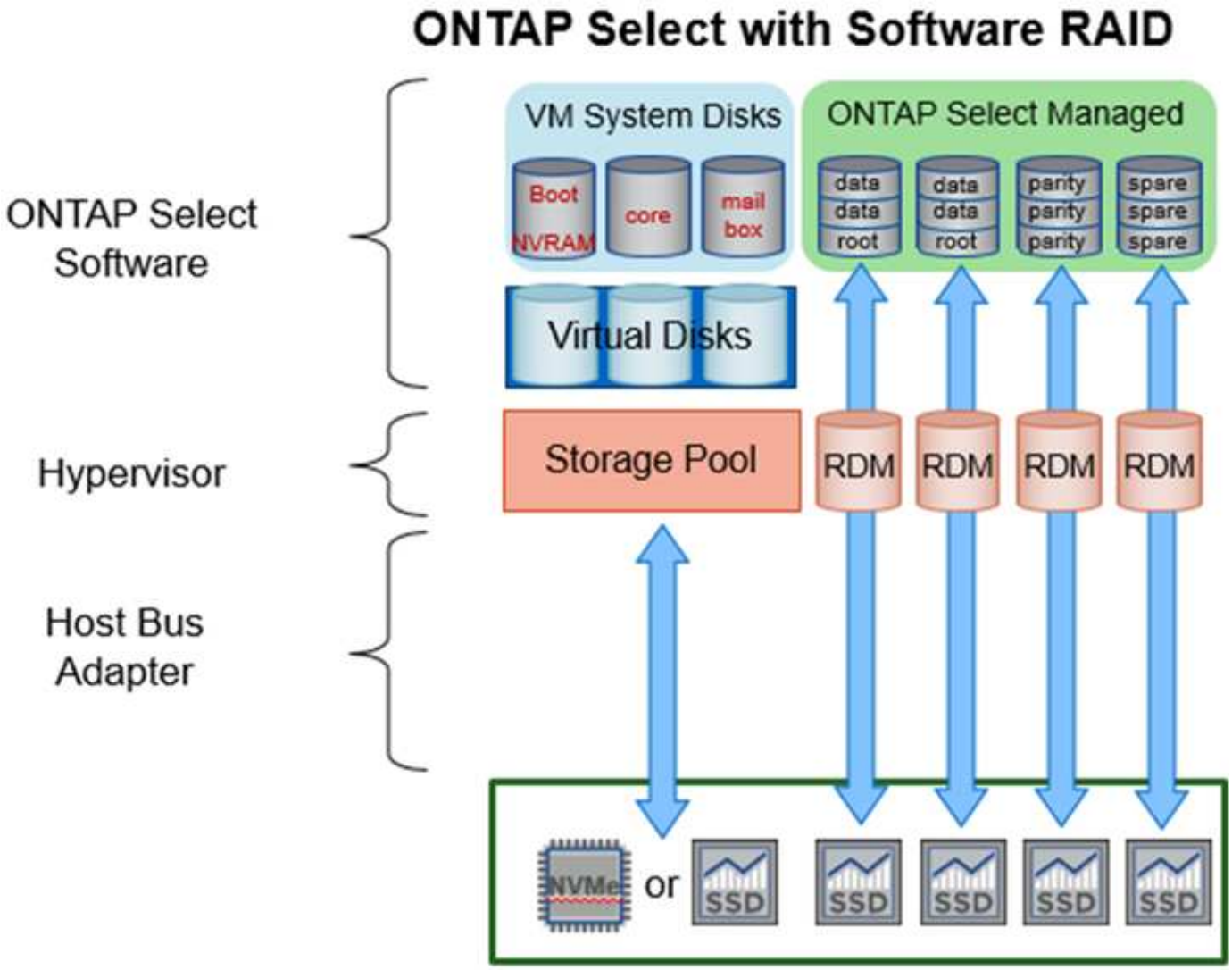
Dischi virtuali e fisici ONTAP Select

Per le configurazioni con controller RAID hardware, la ridondanza del disco fisico viene fornita dal controller RAID. ONTAP Select presenta uno o più VMDK da cui l'amministratore di ONTAP può configurare gli aggregati di dati. Questi VMDK vengono stripati in un formato RAID 0 perché l'utilizzo del RAID software ONTAP è ridondante, inefficiente e inefficace a causa della resilienza fornita a livello hardware. Inoltre, i VMDK utilizzati per i dischi di sistema si trovano nello stesso datastore dei VMDK utilizzati per memorizzare i dati dell'utente.

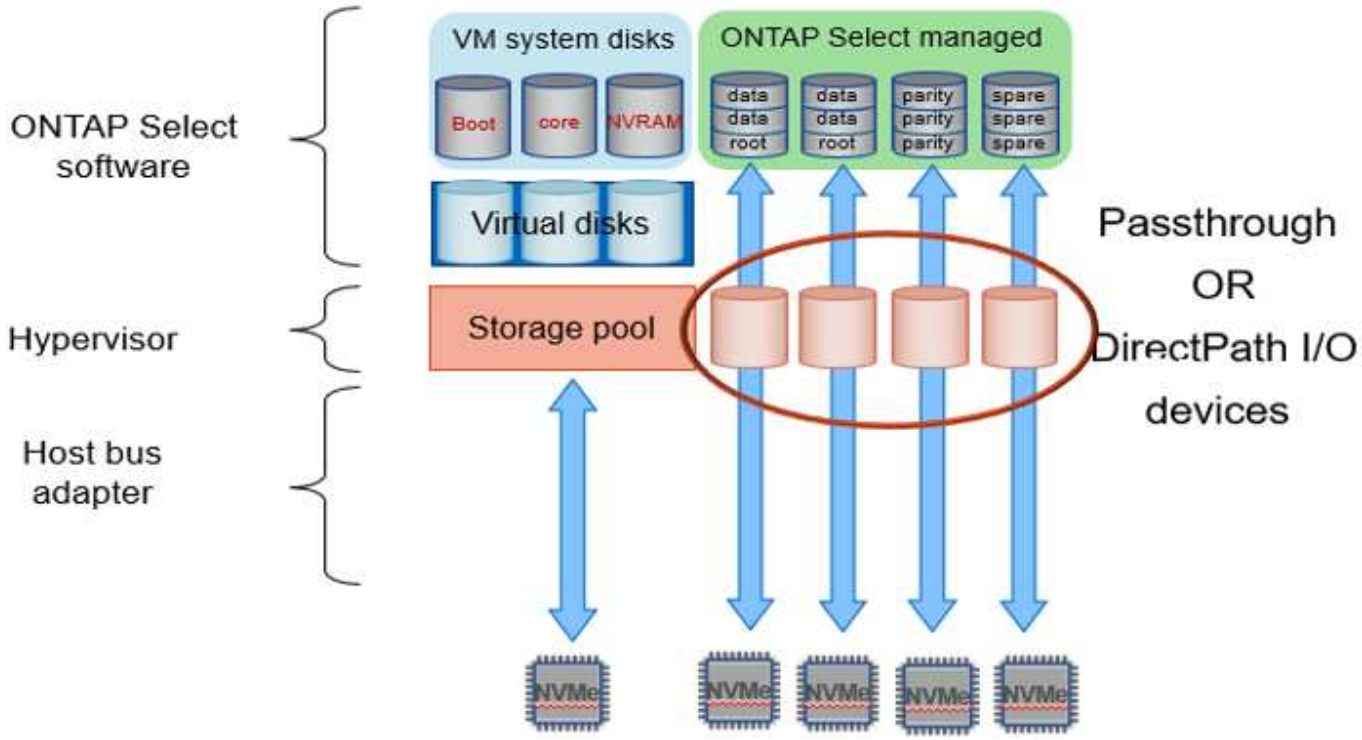
Quando si utilizza il RAID software, ONTAP Deploy presenta ONTAP Select con un set di dischi virtuali (VMDK) e dischi fisici mappature dei dispositivi raw (RDM) per SSD e dispositivi di i/o pass-through o DirectPath per NVMes.

Le seguenti figure mostrano questa relazione in maggiore dettaglio, evidenziando la differenza tra i dischi virtualizzati utilizzati per le macchine virtuali interne di ONTAP Select e i dischi fisici utilizzati per memorizzare i dati dell'utente.

RAID software ONTAP Select: Utilizzo di dischi e RDM virtualizzati



I dischi di sistema (VMDK) risiedono nello stesso datastore e sullo stesso disco fisico. Il disco virtuale NVRAM richiede un supporto rapido e duraturo. Pertanto, sono supportati solo gli archivi dati di tipo NVMe e SSD.



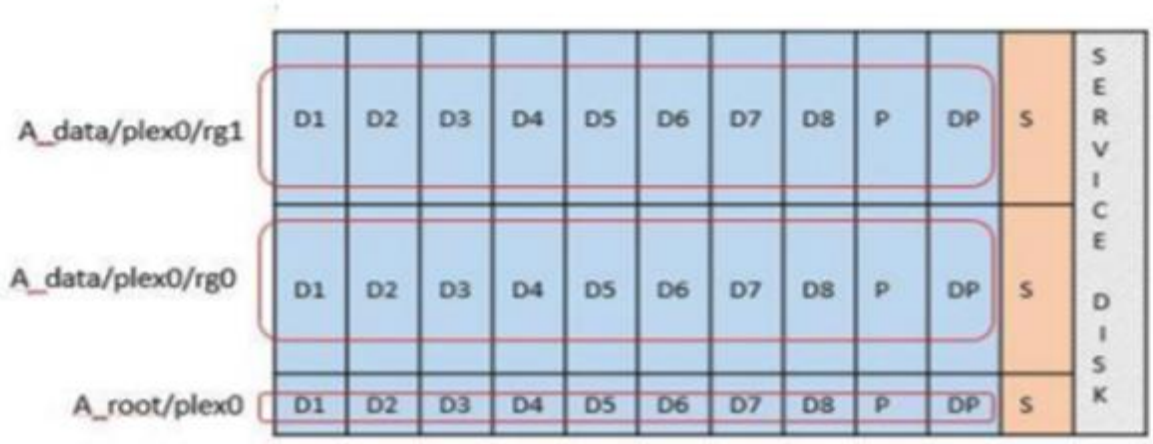
I dischi di sistema (VMDK) risiedono nello stesso datastore e sullo stesso disco fisico. Il disco virtuale NVRAM richiede un supporto rapido e duraturo. Pertanto, sono supportati solo gli archivi dati di tipo NVMe e SSD. Quando si utilizzano unità NVMe per i dati, il disco di sistema deve essere anche un dispositivo NVMe per motivi di performance. Un buon candidato per il disco di sistema in una configurazione All NVMe è una SCHEDA INTEL Optane.

Con la release corrente, non è possibile separare ulteriormente i dischi di sistema ONTAP Select tra più datastore o più dischi fisici.

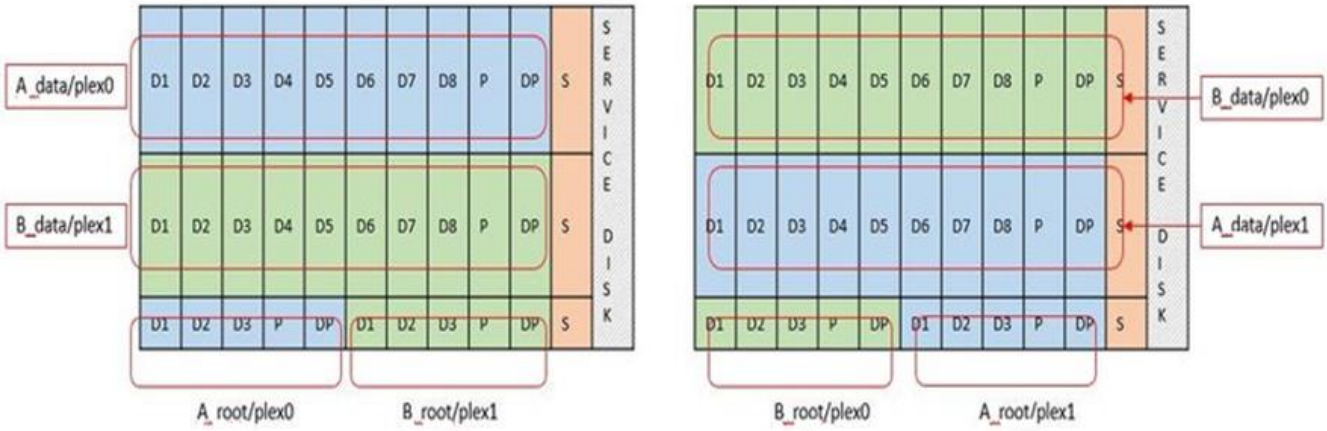
Ogni disco dati è diviso in tre parti: Una piccola partizione root (stripe) e due partizioni di uguali dimensioni per creare due dischi dati visibili all'interno della macchina virtuale ONTAP Select. Le partizioni utilizzano lo schema Root Data Data (RD2), come mostrato nelle figure seguenti, per un cluster a nodo singolo e per un nodo in una coppia ha.

`P` indica un disco di parità. `DP` indica un disco a parità doppia e. `S` indica un disco spare.

Partizione del disco RDD per cluster a nodo singolo



Partizione dei dischi RDD per cluster a più nodi (coppie ha)



Il software ONTAP RAID supporta i seguenti tipi di RAID: RAID 4, RAID-DP e RAID-TEC. Si tratta degli stessi costrutti RAID utilizzati dalle piattaforme FAS e AFF. Per il provisioning root, ONTAP Select supporta solo RAID 4 e RAID-DP. Quando si utilizza RAID-TEC per l'aggregato di dati, la protezione generale è RAID-DP. ONTAP Select ha utilizza un'architettura shared-nothing che replica la configurazione di ciascun nodo sull'altro nodo. Ciò significa che ciascun nodo deve memorizzare la propria partizione root e una copia della partizione root del peer. Poiché un disco dati ha una singola partizione root, il numero minimo di dischi dati varia a seconda che il nodo ONTAP Select faccia parte di una coppia ha o meno.

Per i cluster a nodo singolo, tutte le partizioni dei dati vengono utilizzate per memorizzare i dati locali (attivi). Per i nodi che fanno parte di una coppia ha, viene utilizzata una partizione di dati per memorizzare i dati locali (attivi) per quel nodo e la seconda partizione di dati per eseguire il mirroring dei dati attivi dal peer ha.

Dispositivi Passthrough (io DirectPath) vs RDM (Raw Device Maps)

VMware ESX attualmente non supporta i dischi NVMe come Raw Device Maps. Affinché ONTAP Select assuma il controllo diretto dei dischi NVMe, i dischi NVMe devono essere configurati in ESX come dispositivi pass-through. Si noti che la configurazione di un dispositivo NVMe come dispositivo pass-through richiede il supporto del BIOS del server e si tratta di un processo di interruzione, che richiede un riavvio dell'host ESX. Inoltre, il numero massimo di dispositivi pass-through per host ESX è 16. Tuttavia, ONTAP Deploy limita questa operazione a 14. Questo limite di 14 dispositivi NVMe per nodo ONTAP Select indica che una configurazione All NVMe fornirà una densità di IOPS molto elevata (IOPS/TB) a scapito della capacità totale. In

alternativa, se si desidera una configurazione dalle performance elevate con una capacità di storage superiore, la configurazione consigliata è una macchina virtuale ONTAP Select di grandi dimensioni, una scheda INTEL Optane per il disco di sistema e un numero nominale di unità SSD per lo storage dei dati.

> Per trarre il massimo vantaggio dalle performance di NVMe, prendere in considerazione le grandi dimensioni delle macchine virtuali ONTAP Select.

Esiste un'ulteriore differenza tra i dispositivi pass-through e gli RDM. Gli RDM possono essere mappati a una macchina virtuale in esecuzione. I dispositivi Passthrough richiedono un riavvio della macchina virtuale. Ciò significa che qualsiasi procedura di sostituzione o espansione della capacità del disco NVMe (aggiunta del disco) richiederà un riavvio della macchina virtuale ONTAP Select. La sostituzione dei dischi e l'espansione della capacità (aggiunta dei dischi) sono determinate da un workflow in ONTAP Deploy. ONTAP Deploy gestisce il reboot ONTAP Select per cluster a nodo singolo e failover/failback per coppie ha. Tuttavia, è importante notare la differenza tra l'utilizzo di unità dati SSD (non sono richiesti riavvio/failover ONTAP Select) e l'utilizzo di unità dati NVMe (è necessario riavviare/failover ONTAP Select).

Provisioning di dischi fisici e virtuali

Per offrire un'esperienza utente più ottimizzata, ONTAP Deploy effettua il provisioning automatico dei dischi (virtuali) del sistema dal datastore specificato (disco fisico del sistema) e li collega alla macchina virtuale ONTAP Select. Questa operazione viene eseguita automaticamente durante la configurazione iniziale in modo che la macchina virtuale ONTAP Select possa avviarsi. Gli RDM vengono partizionati e l'aggregato root viene creato automaticamente. Se il nodo ONTAP Select fa parte di una coppia ha, le partizioni dei dati vengono assegnate automaticamente a un pool di storage locale e a un pool di storage mirror. Questa assegnazione avviene automaticamente durante le operazioni di creazione del cluster e di aggiunta dello storage.

Poiché i dischi dati sulla macchina virtuale ONTAP Select sono associati ai dischi fisici sottostanti, vi sono implicazioni in termini di prestazioni per la creazione di configurazioni con un numero maggiore di dischi fisici.

> Il tipo di gruppo RAID dell'aggregato root dipende dal numero di dischi disponibili. ONTAP Deploy sceglie il tipo di gruppo RAID appropriato. Se il nodo dispone di dischi sufficienti, utilizza RAID-DP, altrimenti crea un aggregato root RAID-4.

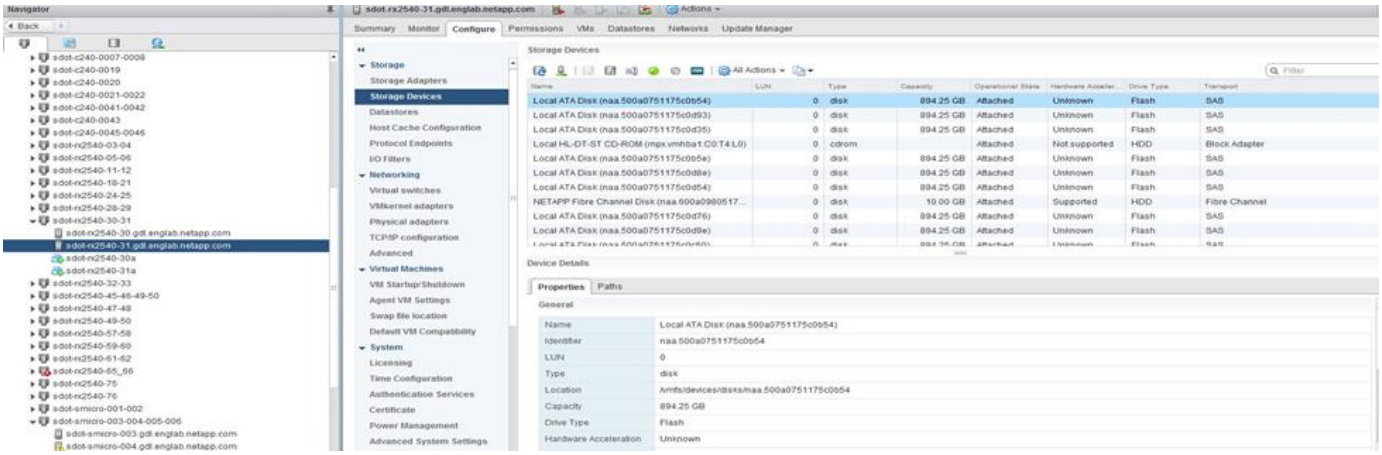
Quando si aggiunge capacità a una macchina virtuale ONTAP Select utilizzando RAID software, l'amministratore deve prendere in considerazione le dimensioni fisiche del disco e il numero di dischi necessari. Per ulteriori informazioni, consultare la sezione "Aumentare la capacità di storage".

Analogamente ai sistemi FAS e AFF, è possibile aggiungere solo dischi con capacità uguali o superiori a un gruppo RAID esistente. I dischi con capacità maggiore sono dimensionati correttamente. Se si stanno creando nuovi gruppi RAID, la dimensione del nuovo gruppo RAID deve corrispondere alla dimensione del gruppo RAID esistente per garantire che le prestazioni complessive dell'aggregato non peggiorino.

Associare un disco ONTAP Select al disco ESX corrispondente

I dischi ONTAP Select sono generalmente etichettati NET x.y. È possibile utilizzare il seguente comando ONTAP per ottenere l'UUID del disco:

```
<system name>::> disk show NET-1.1
Disk: NET-1.1
Model: Micron_5100_MTFD
Serial Number: 1723175C0B5E
UID:
*500A0751:175C0B5E*:00000000:00000000:00000000:00000000:00000000:00000000:
00000000:00000000
BPS: 512
Physical Size: 894.3GB
Position: shared
Checksum Compatibility: advanced_zoned
Aggregate: -
Plex: -This UID can be matched with the device UID displayed in the
'storage devices' tab for the ESX host
```



Nella shell ESXi, è possibile immettere il seguente comando per far lampeggiare il LED di un determinato disco fisico (identificato dal relativo naa.unique-id).

```
esxcli storage core device set -d <naa_id> -l=locator -L=<seconds>
```

Guasti multipli dei dischi quando si utilizza RAID software

È possibile che un sistema si trovi in una situazione in cui più dischi si trovano contemporaneamente in uno stato di guasto. Il comportamento del sistema dipende dalla protezione RAID aggregata e dal numero di dischi guasti.

Un aggregato RAID4 può sopravvivere a un guasto di un disco, un aggregato RAID-DP può sopravvivere a due guasti di disco e un aggregato RAID-TEC può sopravvivere a tre guasti di disco.

Se il numero di dischi guasti è inferiore al numero massimo di guasti supportato dal tipo RAID e se è disponibile un disco spare, il processo di ricostruzione viene avviato automaticamente. Se i dischi spare non sono disponibili, l'aggregato serve i dati in uno stato degradato fino all'aggiunta dei dischi spare.

Se il numero di dischi guasti è superiore al numero massimo di guasti supportato dal tipo RAID, il plex locale viene contrassegnato come failed e lo stato aggregato viene degradato. I dati vengono forniti dal secondo plex

residente sul partner ha. Ciò significa che tutte le richieste di i/o per il nodo 1 vengono inviate attraverso la porta di interconnessione del cluster e0e (iSCSI) ai dischi fisicamente ubicati sul nodo 2. Se anche il secondo plex non funziona, l'aggregato viene contrassegnato come non riuscito e i dati non sono disponibili.

Un plesso guasto deve essere cancellato e ricreato per poter riprendere il mirroring corretto dei dati. Si noti che un guasto a più dischi con conseguente degrado di un aggregato di dati comporta anche un degrado di un aggregato root. ONTAP Select utilizza lo schema di partizione RDD (root-data-data) per suddividere ogni disco fisico in una partizione root e due partizioni di dati. Pertanto, la perdita di uno o più dischi potrebbe avere un impatto su più aggregati, tra cui la radice locale o la copia dell'aggregato root remoto, nonché sull'aggregato di dati locale e la copia dell'aggregato di dati remoto.

```
C3111E67::> storage aggregate plex delete -aggregate aggr1 -plex plex1
Warning: Deleting plex "plex1" of mirrored aggregate "aggr1" in a non-
shared HA configuration will disable its synchronous mirror protection and
disable
         negotiated takeover of node "sti-rx2540-335a" when aggregate
"aggr1" is online.
Do you want to continue? {y|n}: y
[Job 78] Job succeeded: DONE

C3111E67::> storage aggregate mirror -aggregate aggr1
Info: Disks would be added to aggregate "aggr1" on node "sti-rx2540-335a"
in the following manner:
      Second Plex
        RAID Group rg0, 5 disks (advanced_zoned checksum, raid_dp)
                                                                Usable
Physical
          Position   Disk                      Type           Size
Size
          ---------- ------------------------- ---------- --------
--------
          shared     NET-3.2                   SSD               -
-
          shared     NET-3.3                   SSD               -
-
          shared     NET-3.4                   SSD         208.4GB
208.4GB
          shared     NET-3.5                   SSD         208.4GB
208.4GB
          shared     NET-3.12                  SSD         208.4GB
208.4GB

      Aggregate capacity available for volume use would be 526.1GB.
      625.2GB would be used from capacity license.
Do you want to continue? {y|n}: y

C3111E67::> storage aggregate show-status -aggregate aggr1
```

```
Owner Node: sti-rx2540-335a
 Aggregate: aggr1 (online, raid_dp, mirrored) (advanced_zoned checksums)
  Plex: /aggr1/plex0 (online, normal, active, pool0)
   RAID Group /aggr1/plex0/rg0 (normal, advanced_zoned checksums)
                                                                    Usable
Physical
     Position Disk                       Pool Type     RPM     Size
Size Status
     -------- -------------------------- ---- ----- ------ --------
-------- ----------
     shared   NET-1.1                     0   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-1.2                     0   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-1.3                     0   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-1.10                    0   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-1.11                    0   SSD        -  205.1GB
447.1GB (normal)
  Plex: /aggr1/plex3 (online, normal, active, pool1)
   RAID Group /aggr1/plex3/rg0 (normal, advanced_zoned checksums)
                                                                    Usable
Physical
     Position Disk                       Pool Type     RPM     Size
Size Status
     -------- -------------------------- ---- ----- ------ --------
-------- ----------
     shared   NET-3.2                     1   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-3.3                     1   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-3.4                     1   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-3.5                     1   SSD        -  205.1GB
447.1GB (normal)
     shared   NET-3.12                    1   SSD        -  205.1GB
447.1GB (normal)
10 entries were displayed..
```

Per testare o simulare guasti a uno o più dischi, utilizzare `storage disk fail -disk NET-x.y -immediate` comando. Se nel sistema è presente uno spare, l'aggregato inizierà a ricostruire. È possibile controllare lo stato della ricostruzione utilizzando il comando `storage aggregate show`. È possibile rimuovere il disco guasto simulato utilizzando ONTAP Deploy. Tenere presente che ONTAP ha contrassegnato il disco come `Broken`. Il disco non è effettivamente danneggiato e può essere aggiunto nuovamente utilizzando ONTAP Deploy. Per cancellare l'etichetta rotta, immettere i seguenti comandi nella CLI ONTAP Select:

```
set advanced
disk unfail -disk NET-x.y -spare true
disk show -broken
```

L'output dell'ultimo comando deve essere vuoto.

NVRAM virtualizzata

I sistemi NetApp FAS sono tradizionalmente dotati di una scheda PCI NVRAM fisica. Si tratta di una scheda dalle performance elevate contenente memoria flash non volatile che offre un significativo incremento delle prestazioni di scrittura. Ciò avviene concedendo a ONTAP la possibilità di riconoscere immediatamente le scritture in entrata nel client. Può anche pianificare lo spostamento dei blocchi di dati modificati su supporti di storage più lenti in un processo noto come destaging.

I sistemi commodity in genere non sono dotati di questo tipo di apparecchiatura. Pertanto, la funzionalità della scheda NVRAM è stata virtualizzata e inserita in una partizione sul disco di avvio del sistema ONTAP Select. È per questo motivo che il posizionamento del disco virtuale di sistema dell'istanza è estremamente importante.

Configurazioni di array esterno e VSAN

Le implementazioni di Virtual NAS (vNAS) supportano i cluster ONTAP Select su SAN virtuale (VSAN), alcuni prodotti HCI e tipi di datastore esterni. L'infrastruttura sottostante di queste configurazioni fornisce resilienza del datastore.

Il requisito minimo è che la configurazione sottostante sia supportata da VMware e debba essere elencata nei rispettivi HCL VMware.

Architettura vNAS

La nomenclatura vNAS viene utilizzata per tutte le configurazioni che non utilizzano DAS. Per i cluster ONTAP Select a più nodi, sono incluse le architetture per le quali i due nodi ONTAP Select nella stessa coppia ha condividono un singolo datastore (inclusi i datastore vSAN). I nodi possono anche essere installati su datastore separati dallo stesso array esterno condiviso. Ciò consente di ottenere efficienze dello storage lato array per ridurre l'impatto complessivo dell'intera coppia ONTAP Select ha. L'architettura delle soluzioni ONTAP Select vNAS è molto simile a quella di ONTAP Select su DAS con un controller RAID locale. Vale a dire che ogni nodo ONTAP Select continua a disporre di una copia dei dati del partner ha. Le policy di efficienza dello storage ONTAP hanno un ambito di nodo. Di conseguenza, è preferibile l'efficienza dello storage lato array, in quanto potenzialmente possono essere applicati tra set di dati da entrambi i nodi ONTAP Select.

È anche possibile che ogni nodo ONTAP Select in una coppia ha utilizzi un array esterno separato. Si tratta di una scelta comune quando si utilizza ONTAP Select MetroCluster SDS con storage esterno.

Quando si utilizzano array esterni separati per ciascun nodo ONTAP Select, è molto importante che i due array forniscano caratteristiche di performance simili a quelle della macchina virtuale ONTAP Select.

Architetture vNAS rispetto a DAS locale con controller RAID hardware

L'architettura vNAS è logicamente più simile all'architettura di un server con DAS e un controller RAID. In entrambi i casi, ONTAP Select consuma spazio datastore. Lo spazio del datastore viene intagliato nei VMDK e questi VMDK formano i tradizionali aggregati di dati ONTAP. L'implementazione di ONTAP garantisce che le VMDK siano dimensionate correttamente e assegnate al plesso corretto (nel caso di coppie ha) durante le operazioni di creazione del cluster e aggiunta dello storage.

Esistono due differenze principali tra vNAS e DAS con un controller RAID. La differenza più immediata è che vNAS non richiede un controller RAID. VNAS presuppone che l'array esterno sottostante fornisca la persistenza e la resilienza dei dati che un DAS con una configurazione del controller RAID potrebbe fornire. La seconda e più sottile differenza riguarda le prestazioni della NVRAM.

NVRAM vNAS

La NVRAM ONTAP Select è un VMDK. In altre parole, ONTAP Select emula uno spazio indirizzabile in byte (NVRAM tradizionale) su un dispositivo indirizzabile a blocchi (VMDK). Tuttavia, le prestazioni della NVRAM sono assolutamente critiche per le prestazioni complessive del nodo ONTAP Select.

Per le configurazioni DAS con un controller RAID hardware, la cache del controller RAID hardware funge da cache NVRAM de facto, perché tutte le scritture sulla NVRAM VMDK sono ospitate per la prima volta nella cache del controller RAID.

Per le architetture VNAS, ONTAP Deploy configura automaticamente i nodi ONTAP Select con un argomento di boot chiamato SIDL (Single Instance Data Logging). Quando questo argomento di boot è presente, ONTAP Select ignora la NVRAM e scrive il payload dei dati direttamente nell'aggregato di dati. La NVRAM viene utilizzata solo per registrare l'indirizzo dei blocchi modificati dall'operazione DI SCRITTURA. Il vantaggio di questa funzione è che evita una doppia scrittura: Una scrittura sulla NVRAM e una seconda scrittura quando la NVRAM viene destinata. Questa funzione è abilitata solo per vNAS perché le scritture locali nella cache del controller RAID hanno una latenza aggiuntiva trascurabile.

La funzione SIDL non è compatibile con tutte le funzioni di efficienza dello storage ONTAP Select. La funzione SIDL può essere disattivata a livello di aggregato utilizzando il seguente comando:

```
storage aggregate modify -aggregate aggr-name -single-instance-data
-logging off
```

Si noti che le prestazioni di scrittura sono influenzate se la funzione SIDL è disattivata. È possibile riattivare la funzione SIDL dopo aver disattivato tutte le policy di efficienza dello storage su tutti i volumi dell'aggregato:

```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

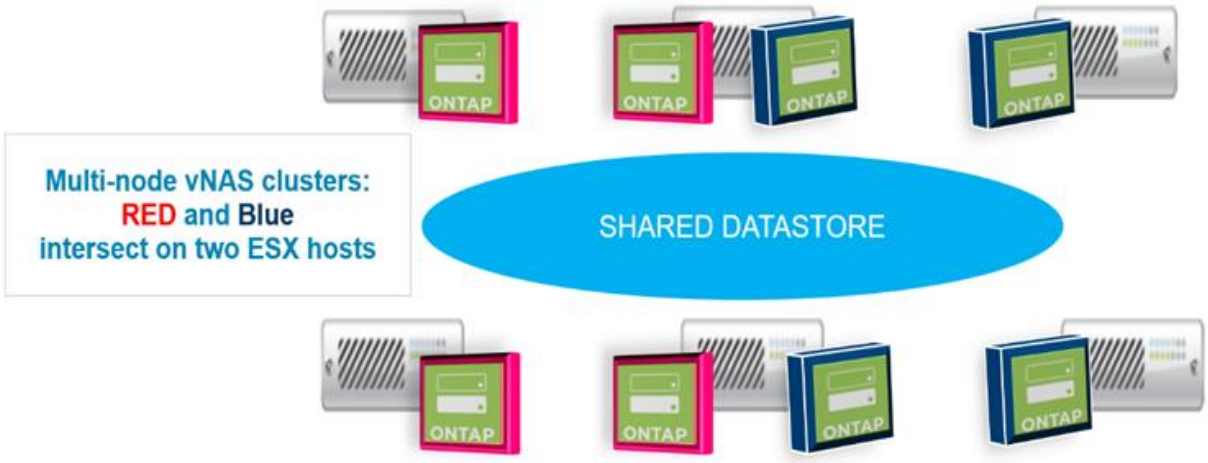
Collocare i nodi ONTAP Select quando si utilizza vNAS su ESXi

ONTAP Select include il supporto per cluster ONTAP Select a più nodi su storage condiviso. ONTAP Deploy consente la configurazione di più nodi ONTAP Select sullo stesso host ESX, purché tali nodi non facciano parte dello stesso cluster. Si noti che questa configurazione è valida solo per gli ambienti VNAS (datastore

condivisi). Le istanze multiple di ONTAP Select per host non sono supportate quando si utilizza lo storage DAS perché queste istanze competono per lo stesso controller RAID hardware.

L'implementazione di ONTAP garantisce che l'implementazione iniziale del cluster VNAS multinodo non posizioni più istanze di ONTAP Select dallo stesso cluster sullo stesso host. La figura seguente mostra un esempio di implementazione corretta di due cluster a quattro nodi che si intersecano su due host.

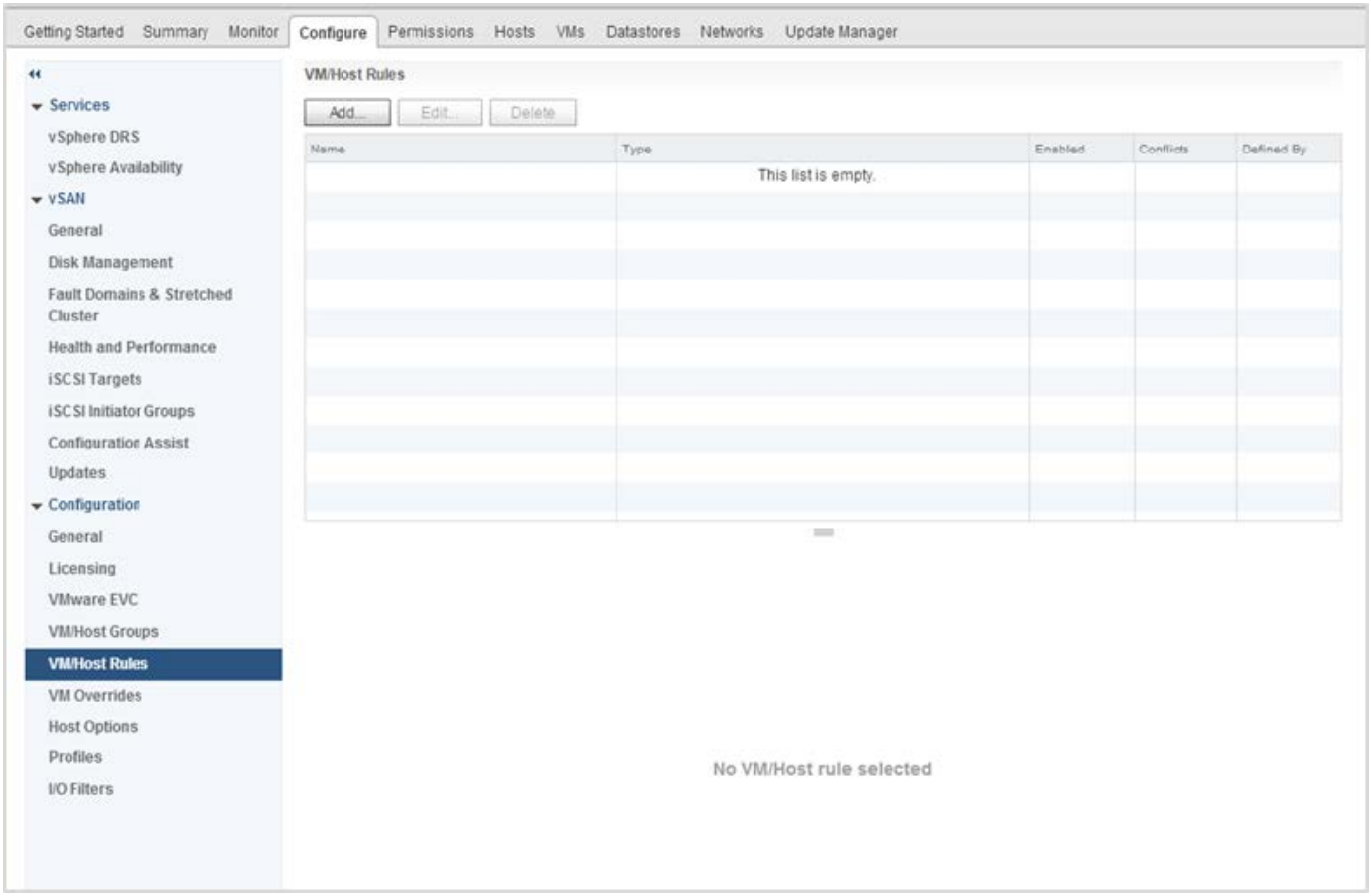
Implementazione iniziale di cluster VNAS multinazionali

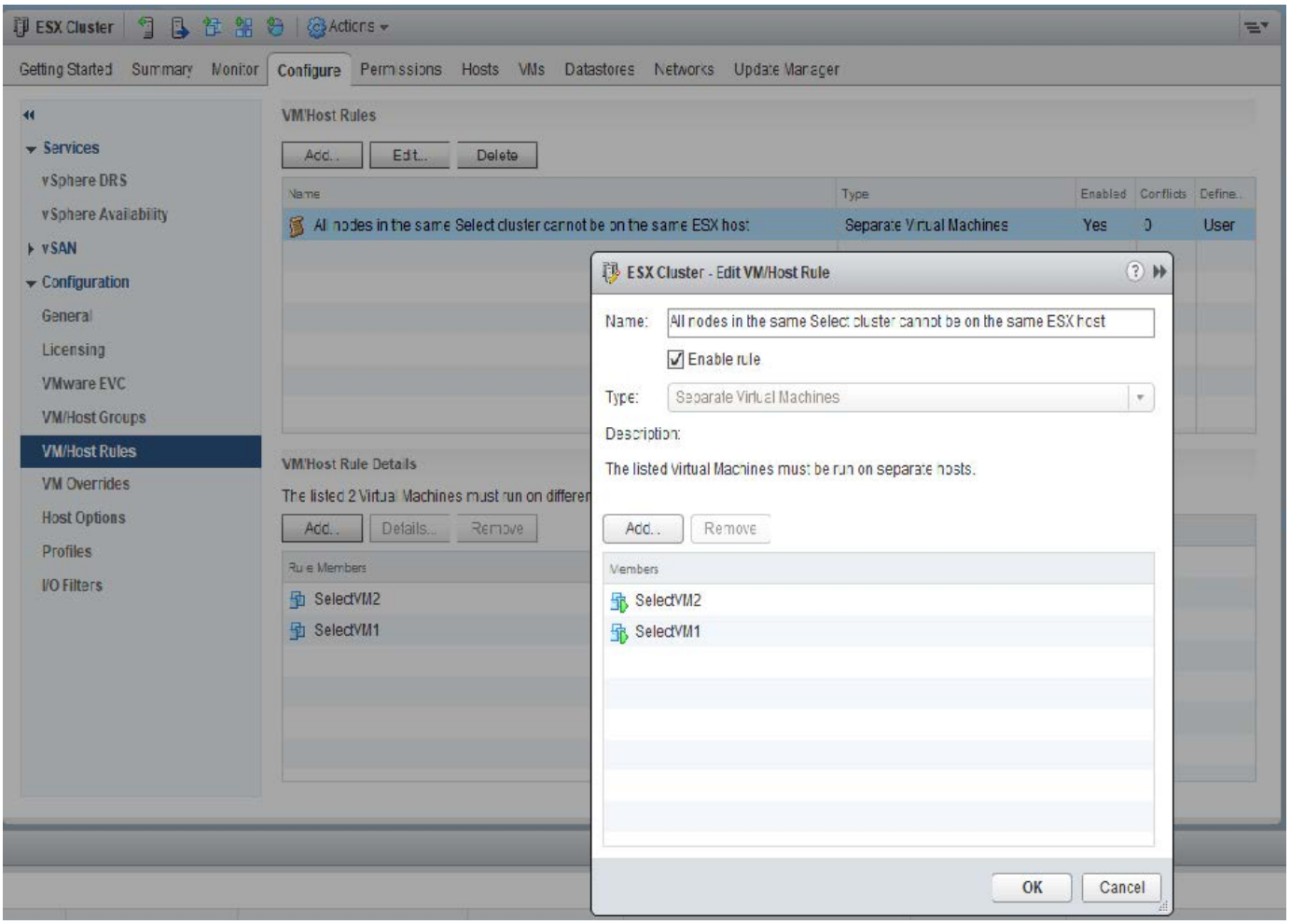


Dopo l'implementazione, è possibile migrare i nodi ONTAP Select tra gli host. Ciò potrebbe causare configurazioni non ottimali e non supportate per le quali due o più nodi ONTAP Select dello stesso cluster condividono lo stesso host sottostante. NetApp consiglia di creare manualmente le regole di affinità delle macchine virtuali in modo che VMware mantenga automaticamente la separazione fisica tra i nodi dello stesso cluster, non solo i nodi della stessa coppia ha.

Le regole di anti-affinità richiedono che DRS sia attivato sul cluster ESX.

Vedere l'esempio seguente su come creare una regola anti-affinità per le macchine virtuali ONTAP Select. Se il cluster ONTAP Select contiene più di una coppia ha, tutti i nodi del cluster devono essere inclusi in questa regola.


Getting Started
Summary
Monitor
Configure
Permissions
Hosts
VMs
Datastores
Networks
Update Manager
Services
vSphere DRS
vSphere Availability
vSAN
General
Disk Management
Fault Domains & Stretched Cluster
Health and Performance
iSCSI Targets
iSCSI Initiator Groups
Configuration Assist
Updates
Configuration
General
Licensing
VMware EVC
VM/Host Groups
VM/Host Rules
VM Overrides
Host Options
Profiles
I/O Filters
VM/Host Rules
Add...
Edit...
Delete
Name
Type
Enabled
Conflicts
Defined By
This list is empty.
No VM/Host rule selected



È possibile trovare due o più nodi ONTAP Select dello stesso cluster ONTAP Select sullo stesso host ESX per uno dei seguenti motivi:

- DRS non è presente a causa delle limitazioni di licenza di VMware vSphere o se DRS non è abilitato.
- La regola di affinità DRS viene ignorata perché ha la precedenza su un'operazione VMware ha o su una migrazione VM avviata dall'amministratore.

Tenere presente che ONTAP Deploy non esegue il monitoraggio proattivo delle posizioni delle macchine virtuali ONTAP Select. Tuttavia, un'operazione di refresh del cluster riflette questa configurazione non supportata nei log di implementazione di ONTAP:

Aumentare la capacità di storage

ONTAP Deploy può essere utilizzato per aggiungere e concedere in licenza storage aggiuntivo per ciascun nodo di un cluster ONTAP Select.

La funzionalità storage-add in ONTAP Deploy è l'unico modo per aumentare lo storage in gestione e non è supportata la modifica diretta della macchina virtuale ONTAP Select. La figura seguente mostra l'icona "+" che avvia la procedura guidata di aggiunta dello storage.

Le seguenti considerazioni sono importanti per il successo dell'operazione di espansione della capacità. L'aggiunta di capacità richiede la licenza esistente per coprire la quantità totale di spazio (esistente più nuovo). Un'operazione di aggiunta dello storage che comporta il superamento della capacità concessa in licenza da parte del nodo non riesce. È necessario installare prima una nuova licenza con capacità sufficiente.

Se la capacità aggiuntiva viene aggiunta a un aggregato ONTAP Select esistente, il nuovo pool di storage (datastore) dovrebbe avere un profilo di performance simile a quello del pool di storage esistente (datastore). Si noti che non è possibile aggiungere storage non SSD a un nodo ONTAP Select installato con una personalità simile a AFF (flash abilitata). Non è supportata anche la combinazione di DAS e storage esterno.

Se a un sistema viene aggiunto storage collegato localmente per fornire ulteriori pool di storage locali (DAS), è necessario creare un gruppo RAID e LUN (o LUN) aggiuntivi. Come per i sistemi FAS, è necessario assicurarsi che le prestazioni del nuovo gruppo RAID siano simili a quelle del gruppo RAID originale se si aggiunge nuovo spazio allo stesso aggregato. Se si crea un nuovo aggregato, il nuovo layout del gruppo RAID potrebbe essere diverso se le implicazioni delle performance per il nuovo aggregato sono ben comprese.

Il nuovo spazio può essere aggiunto allo stesso archivio dati come misura se le dimensioni totali dell'archivio dati non superano le dimensioni massime supportate. L'aggiunta di un'estensione del datastore all'archivio dati in cui ONTAP Select è già installato può essere eseguita in modo dinamico e non influisce sulle operazioni del nodo ONTAP Select.

Se il nodo ONTAP Select fa parte di una coppia ha, è necessario considerare alcuni problemi aggiuntivi.

In una coppia ha, ogni nodo contiene una copia speculare dei dati del partner. L'aggiunta di spazio al nodo 1 richiede l'aggiunta di una quantità di spazio identica al suo partner, il nodo 2, in modo che tutti i dati dal nodo 1 vengano replicati nel nodo 2. In altre parole, lo spazio aggiunto al nodo 2 come parte dell'operazione di aggiunta di capacità per il nodo 1 non è visibile o accessibile sul nodo 2. Lo spazio viene aggiunto al nodo 2 in modo che i dati del nodo 1 siano completamente protetti durante un evento ha.

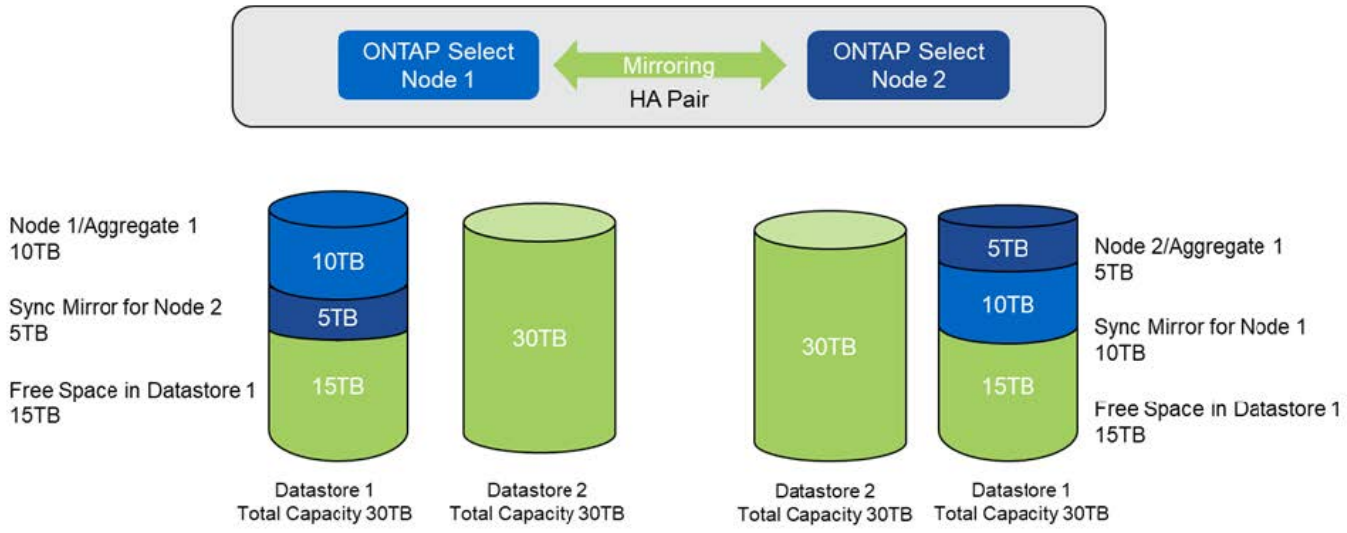
Vi è un'ulteriore considerazione per quanto riguarda le performance. I dati sul nodo 1 vengono replicati in modo sincrono sul nodo 2. Pertanto, le performance del nuovo spazio (datastore) sul nodo 1 devono corrispondere alle performance del nuovo spazio (datastore) sul nodo 2. In altre parole, l'aggiunta di spazio su entrambi i nodi, ma utilizzando diverse tecnologie di dischi o diverse dimensioni di gruppi RAID, può causare problemi di performance. Ciò è dovuto all'operazione RAID SyncMirror utilizzata per mantenere una copia dei dati sul nodo partner.

Per aumentare la capacità accessibile all'utente su entrambi i nodi di una coppia ha, è necessario eseguire due operazioni storage-add, una per ogni nodo. Ogni operazione storage-add richiede spazio aggiuntivo su entrambi i nodi. Lo spazio totale richiesto su ciascun nodo è uguale allo spazio richiesto sul nodo 1 più lo spazio richiesto sul nodo 2.

La configurazione iniziale prevede due nodi, ciascuno dei quali dispone di due datastore con 30 TB di spazio in ciascun datastore. ONTAP Deploy crea un cluster a due nodi, con ogni nodo che consuma 10 TB di spazio dal datastore 1. ONTAP Deploy configura ogni nodo con 5 TB di spazio attivo per nodo.

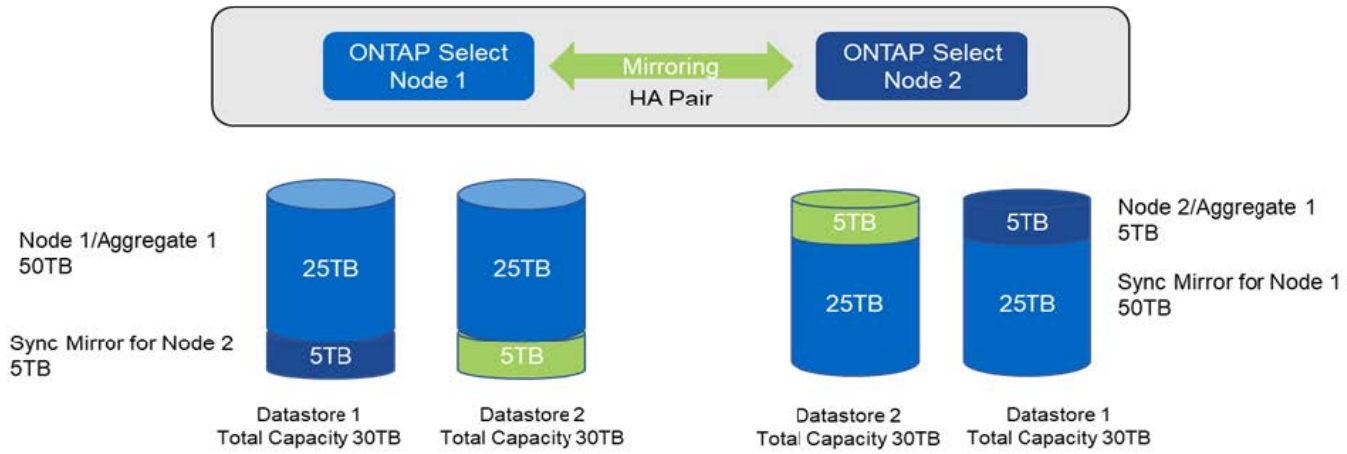
La figura seguente mostra i risultati di una singola operazione storage-add per il nodo 1. ONTAP Select utilizza ancora una quantità uguale di storage (15 TB) su ciascun nodo. Tuttavia, il nodo 1 ha più storage attivo (10 TB) rispetto al nodo 2 (5 TB). Entrambi i nodi sono completamente protetti, poiché ogni nodo ospita una copia dei dati dell'altro nodo. C'è spazio libero aggiuntivo nel datastore 1 e il datastore 2 è ancora completamente libero.

Distribuzione della capacità: Allocazione e spazio libero dopo una singola operazione di aggiunta dello storage



Due operazioni aggiuntive di storage-add sul nodo 1 consumano il resto del datastore 1 e una parte del datastore 2 (utilizzando il limite di capacità). La prima operazione storage-add consuma i 15 TB di spazio libero rimasti nel datastore 1. La figura seguente mostra il risultato della seconda operazione di aggiunta dello storage. A questo punto, il nodo 1 ha 50 TB di dati attivi in gestione, mentre il nodo 2 ha i 5 TB originali.

Distribuzione della capacità: Allocazione e spazio libero dopo due operazioni aggiuntive di storage-add per il nodo 1



La dimensione massima VMDK utilizzata durante le operazioni di aggiunta della capacità è di 16 TB. La dimensione massima VMDK utilizzata durante le operazioni di creazione del cluster continua ad essere di 8

TB. ONTAP Deploy crea VMDK dimensionati correttamente in base alla configurazione (un cluster a nodo singolo o a più nodi) e alla quantità di capacità aggiunta. Tuttavia, la dimensione massima di ciascun VMDK non deve superare 8 TB durante le operazioni di creazione del cluster e 16 TB durante le operazioni di aggiunta dello storage.

Aumenta la capacità di ONTAP Select con il software RAID

La procedura guidata storage-add può essere utilizzata in modo analogo per aumentare la capacità di gestione dei nodi ONTAP Select utilizzando il RAID software. La procedura guidata presenta solo i dischi DAS SDD disponibili e possono essere mappati come RDM alla macchina virtuale ONTAP Select.

Sebbene sia possibile aumentare la licenza di capacità di un singolo TB, quando si lavora con il software RAID, non è possibile aumentare fisicamente la capacità di un singolo TB. Analogamente all'aggiunta di dischi a un array FAS o AFF, alcuni fattori determinano la quantità minima di storage che può essere aggiunta in una singola operazione.

Si noti che in una coppia ha, l'aggiunta di storage al nodo 1 richiede che sia disponibile un numero identico di dischi anche sulla coppia ha del nodo (nodo 2). Sia i dischi locali che i dischi remoti vengono utilizzati da un'operazione storage-add sul nodo 1. In altre parole, i dischi remoti vengono utilizzati per garantire che il nuovo storage sul nodo 1 sia replicato e protetto sul nodo 2. Per aggiungere storage utilizzabile localmente sul nodo 2, è necessario che su entrambi i nodi siano disponibili un'operazione storage-add separata e un numero uguale di dischi.

ONTAP Select suddivide tutti i nuovi dischi nella stessa partizione root, dati e dati dei dischi esistenti. L'operazione di partizione avviene durante la creazione di un nuovo aggregato o durante l'espansione su un aggregato esistente. La dimensione della stripe di partizione root su ciascun disco è impostata in modo che corrisponda alla dimensione della partizione root esistente sui dischi esistenti. Pertanto, ciascuna delle due dimensioni uguali delle partizioni dati può essere calcolata come la capacità totale del disco meno la dimensione della partizione root divisa per due. La dimensione dello stripe della partizione root è variabile e viene calcolata durante la configurazione iniziale del cluster come segue. Lo spazio root totale richiesto (68 GB per un cluster a nodo singolo e 136 GB per le coppie ha) viene diviso tra il numero iniziale di dischi meno eventuali dischi di riserva e di parità. La dimensione dello stripe della partizione root viene mantenuta costante su tutti i dischi aggiunti al sistema.

Se si crea un nuovo aggregato, il numero minimo di dischi richiesto varia a seconda del tipo di RAID e se il nodo ONTAP Select fa parte di una coppia ha.

Se si aggiunge storage a un aggregato esistente, sono necessarie alcune considerazioni aggiuntive. È possibile aggiungere dischi a un gruppo RAID esistente, supponendo che il gruppo RAID non sia già al limite massimo. Anche le Best practice tradizionali di FAS e AFF per l'aggiunta di spindle ai gruppi RAID esistenti si applicano qui e la creazione di un hot spot sul nuovo spindle è un potenziale problema. Inoltre, è possibile aggiungere a un gruppo RAID esistente solo dischi con partizioni di dati uguali o superiori. Come spiegato in precedenza, la dimensione della partizione dei dati non corrisponde alla dimensione raw del disco. Se le partizioni dei dati aggiunte sono più grandi delle partizioni esistenti, le nuove unità sono di dimensioni corrette. In altre parole, una parte della capacità di ogni nuovo disco rimane inutilizzata.

È inoltre possibile utilizzare i nuovi dischi per creare un nuovo gruppo RAID come parte di un aggregato esistente. In questo caso, la dimensione del gruppo RAID deve corrispondere alla dimensione del gruppo RAID esistente.

Supporto per l'efficienza dello storage

ONTAP Select offre opzioni di efficienza dello storage simili alle opzioni di efficienza dello storage presenti negli array FAS e AFF.

Le implementazioni di ONTAP Select Virtual NAS (vNAS) che utilizzano VSAN all-flash o array flash generici devono seguire le Best practice per ONTAP Select con DAS (Direct-Attached Storage) non SSD.

Una personalità simile ad AFF viene attivata automaticamente nelle nuove installazioni, purché si disponga di storage DAS con unità SSD e una licenza Premium.

Con una personalità simile ad AFF, le seguenti funzioni se inline vengono attivate automaticamente durante l'installazione:

- Rilevamento dello zero pattern inline
- Deduplica inline del volume
- Deduplica del volume in background
- Compressione adattiva in linea
- Compaction dei dati inline
- Deduplica aggregata inline
- Deduplica in background aggregata

Per verificare che ONTAP Select abbia attivato tutti i criteri di efficienza dello storage predefiniti, eseguire il seguente comando su un volume appena creato:

```
<system name>::> set diag
Warning: These diagnostic commands are for use by NetApp personnel only.
Do you want to continue? {y|n}: y
twonode95IP15::*> sis config
Vserver:                                SVM1
Volume:                                 _export1_NFS_volume
Schedule                                -
Policy:                                 auto
Compression:                            true
Inline Compression:                     true
Compression Type:                       adaptive
Application IO Si                       8K
Compression Algorithm:                  lzopro
Inline Dedupe:                          true
Data Compaction:                        true
Cross Volume Inline Deduplication:      true
Cross Volume Background Deduplication:  true
```

Per gli aggiornamenti ONTAP Select dalla versione 9.6 e successive, è necessario installare ONTAP Select su storage SSD DAS con una licenza Premium. Inoltre, è necessario selezionare la casella di controllo **Enable Storage Efficiencies** (Abilita efficienza dello storage) durante l'installazione iniziale del cluster con l'implementazione di ONTAP. L'abilitazione di un aggiornamento post-ONTAP della personalità simile ad AFF quando le condizioni precedenti non sono state soddisfatte richiede la creazione manuale di un argomento di boot e il riavvio di un nodo. Per ulteriori informazioni, contatta il supporto tecnico.

Configurazioni per l'efficienza dello storage ONTAP Select

La seguente tabella riassume le varie opzioni di efficienza dello storage disponibili, attivate per impostazione predefinita o non attivate per impostazione predefinita, ma consigliate, a seconda del tipo di supporto e della licenza software.

Funzionalità di ONTAP Select	DAS SSD (Premium o Premium XL[1])	DAS HDD (tutte le licenze)	VNAS (tutte le licenze)
Rilevamento dello zero inline	Sì (impostazione predefinita)	Sì, abilitato dall'utente per volume	Sì, abilitato dall'utente per volume
Deduplica inline del volume	Sì (impostazione predefinita)	Non disponibile	Non supportato
Compressione inline 32K (compressione secondaria)	Sì, abilitato dall'utente per volume.	Sì, abilitato dall'utente per volume	Non supportato
Compressione inline 8K (compressione adattiva)	Sì (impostazione predefinita)	Sì, abilitato dall'utente per volume	Non supportato
Compressione in background	Non supportato	Sì, abilitato dall'utente per volume	Sì, abilitato dall'utente per volume
Scanner a compressione	Sì	Sì	Sì, abilitato dall'utente per volume
Compaction dei dati inline	Sì (impostazione predefinita)	Sì, abilitato dall'utente per volume	Non supportato
Scanner per la compattazione	Sì	Sì	Non supportato
Deduplica aggregata inline	Sì (impostazione predefinita)	N/A.	Non supportato
Deduplica del volume in background	Sì (impostazione predefinita)	Sì, abilitato dall'utente per volume	Sì, abilitato dall'utente per volume
Deduplica in background aggregata	Sì (impostazione predefinita)	N/A.	Non supportato

[Piccolo] 1^ONTAP Select 9.6 supporta una nuova licenza (premium XL) e una nuova dimensione della macchina virtuale (grande). Tuttavia, la macchina virtuale di grandi dimensioni è supportata solo per le configurazioni DAS che utilizzano il software RAID. Le configurazioni RAID hardware e vNAS non sono supportate con la grande macchina virtuale ONTAP Select nella release 9.6

Note sul comportamento dell'upgrade per le configurazioni SSD DAS

Dopo aver eseguito l'aggiornamento a ONTAP Select 9.6 o versioni successive, attendere `system node upgrade-revert show` per indicare che l'aggiornamento è stato completato prima di verificare i valori di efficienza dello storage per i volumi esistenti.

Su un sistema aggiornato a ONTAP Select 9.6 o versioni successive, un nuovo volume creato su un aggregato esistente o un aggregato appena creato ha lo stesso comportamento di un volume creato su una nuova implementazione. I volumi esistenti sottoposti all'aggiornamento del codice ONTAP Select presentano la maggior parte delle stesse policy di efficienza dello storage di un volume appena creato con alcune varianti:

Scenario 1

Se non sono state attivate policy di efficienza dello storage su un volume prima dell'aggiornamento, allora:

- Volumi con `space guarantee = volume` non sono abilitati la compaction dei dati inline, la deduplica inline aggregata e la deduplica in background aggregata. Queste opzioni possono essere attivate dopo l'aggiornamento.
- Volumi con `space guarantee = none` la compressione in background non è attivata. Questa opzione può essere attivata dopo l'aggiornamento.
- La policy di efficienza dello storage sui volumi esistenti viene impostata su auto dopo l'aggiornamento.

Scenario 2

Se alcune efficienze dello storage sono già abilitate su un volume prima dell'aggiornamento, allora:

- Volumi con `space guarantee = volume` dopo l'aggiornamento, non si vedono differenze.
- Volumi con `space guarantee = none` attivare la deduplica aggregata in background.
- Volumi con `storage policy inline-only` impostare la policy su auto.
- I volumi con policy di efficienza dello storage definite dall'utente non hanno alcun cambiamento nelle policy, ad eccezione dei volumi con `space guarantee = none`. Questi volumi hanno la deduplica in background aggregata attivata.

Networking

Networking: Concetti e caratteristiche generali

Acquisire familiarità con i concetti generali di rete applicabili all'ambiente ONTAP Select. Quindi, esplora le caratteristiche e le opzioni specifiche disponibili con i cluster a nodo singolo e multi-nodo.

Rete fisica

La rete fisica supporta un'implementazione del cluster ONTAP Select principalmente fornendo l'infrastruttura di switching di livello due sottostante. La configurazione relativa alla rete fisica include sia l'host dell'hypervisor che l'ambiente di rete con switch più ampio.

Opzioni NIC host

Ciascun host dell'hypervisor ONTAP Select deve essere configurato con due o quattro porte fisiche. La configurazione esatta scelta dipende da diversi fattori, tra cui:

- Se il cluster contiene uno o più host ONTAP Select
- Quale sistema operativo hypervisor viene utilizzato
- Come è configurato lo switch virtuale
- Se LACP viene utilizzato o meno con i collegamenti

Configurazione fisica dello switch

È necessario assicurarsi che la configurazione degli switch fisici supporti l'implementazione di ONTAP Select. Gli switch fisici sono integrati con gli switch virtuali basati su hypervisor. La configurazione esatta scelta

dipende da diversi fattori. Le considerazioni principali includono quanto segue:

- In che modo è possibile mantenere la separazione tra le reti interne ed esterne?
- Mantieni una separazione tra le reti di dati e di gestione?
- Come verranno configurate le VLAN di livello due?

Rete logica

ONTAP Select utilizza due diverse reti logiche, separando il traffico in base al tipo. In particolare, il traffico può fluire tra gli host all'interno del cluster, oltre che verso i client di storage e altre macchine all'esterno del cluster. Gli switch virtuali gestiti dagli hypervisor aiutano a supportare la rete logica.

Rete interna

Con un'implementazione di cluster multi-nodo, i singoli nodi ONTAP Select comunicano utilizzando una rete "interna" isolata. Questa rete non è esposta o disponibile al di fuori dei nodi nel cluster ONTAP Select.

La rete interna è presente solo con un cluster a più nodi.

La rete interna presenta le seguenti caratteristiche:

- Utilizzato per elaborare il traffico intra-cluster di ONTAP, tra cui:
  - Cluster
  - High Availability Interconnect (ha-IC)
  - RAID Synch Mirror (RSM)
- Rete single Layer-due basata su una VLAN
- Gli indirizzi IP statici vengono assegnati da ONTAP Select:
  - Solo IPv4
  - DHCP non utilizzato
  - Indirizzo link-local
- La dimensione MTU è di 9000 byte per impostazione predefinita e può essere regolata entro un intervallo di 7500-9000 (incluso)

Rete esterna

La rete esterna elabora il traffico tra i nodi di un cluster ONTAP Select, i client di storage esterni e le altre macchine. La rete esterna fa parte di ogni implementazione del cluster e presenta le seguenti caratteristiche:

- Utilizzato per elaborare il traffico ONTAP, tra cui:
  - Dati (NFS, CIFS, iSCSI)
  - Gestione (cluster e nodo; SVM opzionale)
  - Intercluster (opzionale)
- Supporto opzionale di VLAN:
  - Gruppo di porte dati
  - Gruppo di porte di gestione
- Indirizzi IP assegnati in base alle opzioni di configurazione dell'amministratore:

    - IPv4 o IPv6
- La dimensione MTU è di 1500 byte per impostazione predefinita (può essere regolata)

La rete esterna è presente con cluster di tutte le dimensioni.

Ambiente di rete di macchine virtuali

L'host dell'hypervisor offre diverse funzionalità di rete.

ONTAP Select si affida alle seguenti funzionalità esposte attraverso la macchina virtuale:

Porte delle macchine virtuali

Sono disponibili diverse porte per l'utilizzo da parte di ONTAP Select. Vengono assegnati e utilizzati in base a.
diversi fattori, inclusa la dimensione del cluster.

Switch virtuale

Il software dello switch virtuale all'interno dell'ambiente hypervisor, sia vSwitch (VMware) che Open vSwitch (KVM), unisce le porte esposte dalla macchina virtuale all'Ethernet fisica Porte NIC. È necessario configurare un vSwitch per ogni host ONTAP Select, a seconda della situazione ambiente.

Configurazioni di rete a nodo singolo e multiplo

ONTAP Select supporta configurazioni di rete a nodo singolo e a più nodi.

Configurazione di rete a nodo singolo

Le configurazioni ONTAP Select a nodo singolo non richiedono la rete interna ONTAP, in quanto non esiste traffico cluster, ha o mirror.

A differenza della versione multinodo del prodotto ONTAP Select, ogni macchina virtuale ONTAP Select contiene tre adattatori di rete virtuali, presentati alle porte di rete ONTAP e0a, e0b e e0c.

Queste porte vengono utilizzate per fornire i seguenti servizi: LIF di gestione, dati e intercluster.

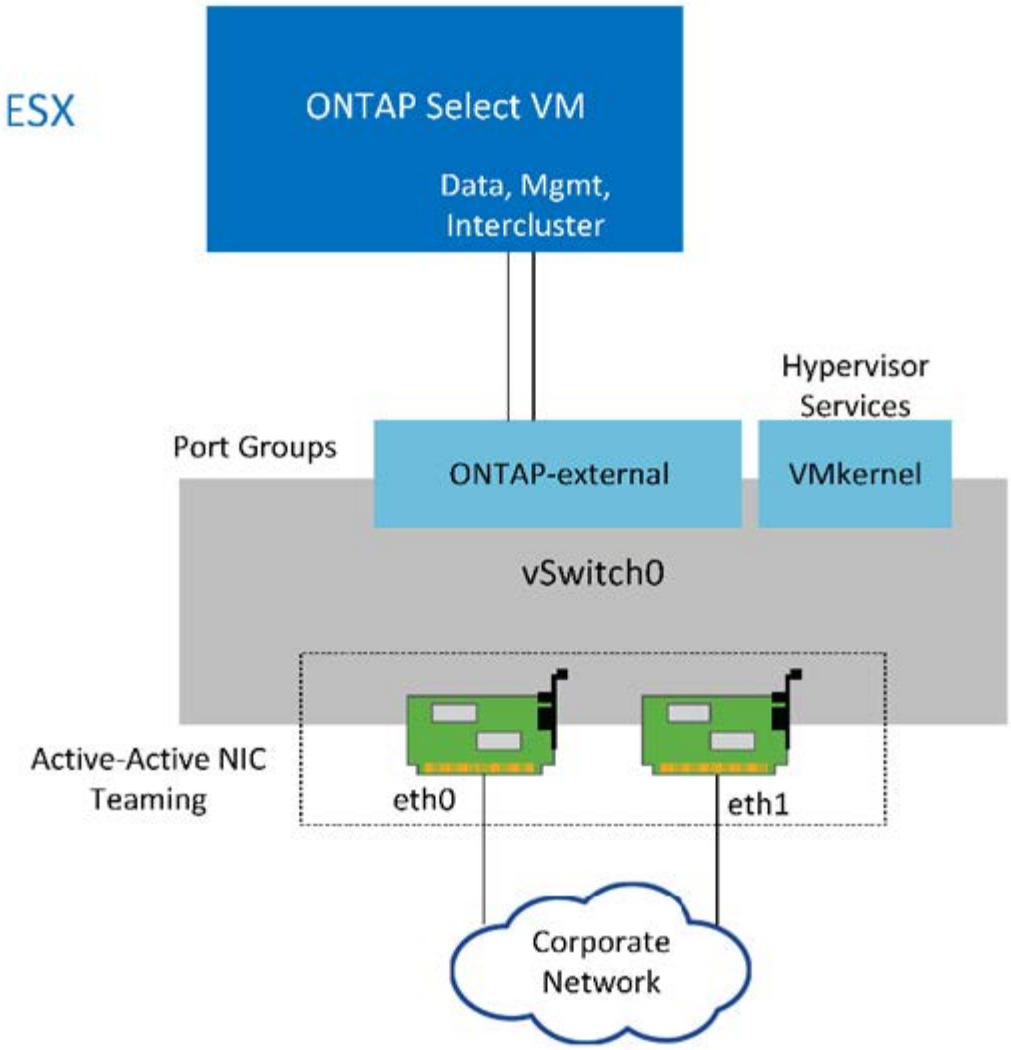
KVM

ONTAP Select può essere implementato come cluster a nodo singolo. L'host dell'hypervisor include uno switch virtuale che fornisce l'accesso alla rete esterna.

ESXi

La relazione tra queste porte e gli adattatori fisici sottostanti è illustrata nella seguente figura, che mostra un nodo del cluster ONTAP Select sull'hypervisor ESX.

Configurazione di rete del cluster ONTAP Select a nodo singolo



Anche se due adattatori sono sufficienti per un cluster a nodo singolo, è comunque necessario il raggruppamento delle NIC.

Assegnazione LIF

Come spiegato nella sezione relativa all'assegnazione LIF multinodo di questo documento, gli spazi IPvengono utilizzati da ONTAP Select per mantenere separato il traffico di rete del cluster dai dati e dal traffico di gestione. La variante a nodo singolo di questa piattaforma non contiene una rete cluster. Pertanto, non sono presenti porte nell'IPSpace del cluster.

Le LIF di gestione di cluster e nodi vengono create automaticamente durante l'installazione del cluster ONTAP Select. I restanti LIF possono essere creati dopo l'implementazione.

LIF di gestione e dati (e0a, e0b e e0c)

Le porte ONTAP e0a, e0b e e0c sono delegate come porte candidate per i LIF che trasportano i seguenti tipi di traffico:

- Traffico del protocollo SAN/NAS (CIFS, NFS e iSCSI)
- Traffico di gestione di cluster, nodi e SVM

- Traffico intercluster (SnapMirror e SnapVault)

Configurazione di rete multinode

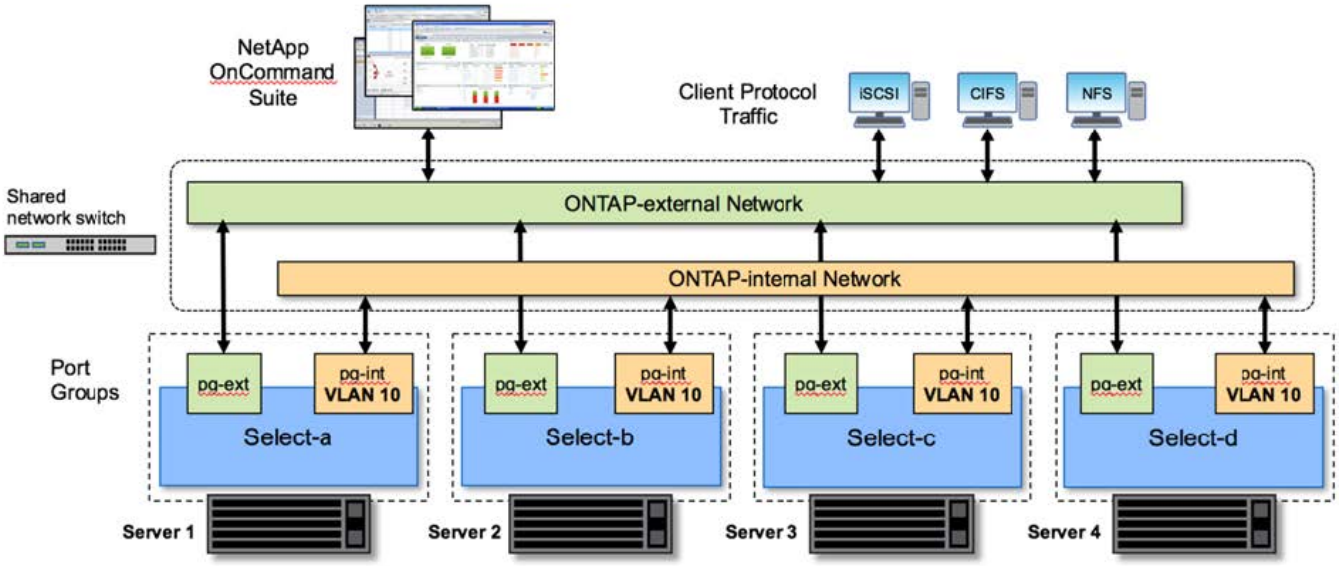
La configurazione di rete ONTAP Select multinodo è costituita da due reti.

Si tratta di una rete interna, responsabile della fornitura di servizi di replica interni e cluster, e di una rete esterna, responsabile della fornitura di servizi di accesso e gestione dei dati. L'isolamento end-to-end del traffico che scorre all'interno di queste due reti è estremamente importante per consentire la creazione di un ambiente adatto alla resilienza del cluster.

Queste reti sono rappresentate nella figura seguente, che mostra un cluster ONTAP Select a quattro nodi in esecuzione su una piattaforma VMware vSphere. I cluster a sei e otto nodi hanno un layout di rete simile.

Ogni istanza di ONTAP Select risiede su un server fisico separato. Il traffico interno ed esterno viene isolato utilizzando gruppi di porte di rete separati, assegnati a ciascuna interfaccia di rete virtuale e che consentono ai nodi del cluster di condividere la stessa infrastruttura dello switch fisico.

Panoramica di una configurazione di rete di cluster ONTAP Select multinode



Ogni macchina virtuale ONTAP Select contiene sette adattatori di rete virtuali presentati a ONTAP come un set di sette porte di rete, da e0a e0g. Sebbene ONTAP tratti questi adattatori come NIC fisiche, sono in realtà virtuali e mappati a un insieme di interfacce fisiche attraverso un layer di rete virtualizzato. Di conseguenza, ciascun server di hosting non richiede sei porte di rete fisiche.

L'aggiunta di adattatori di rete virtuali a ONTAP Select VM non è supportata.

Queste porte sono preconfigurate per fornire i seguenti servizi:

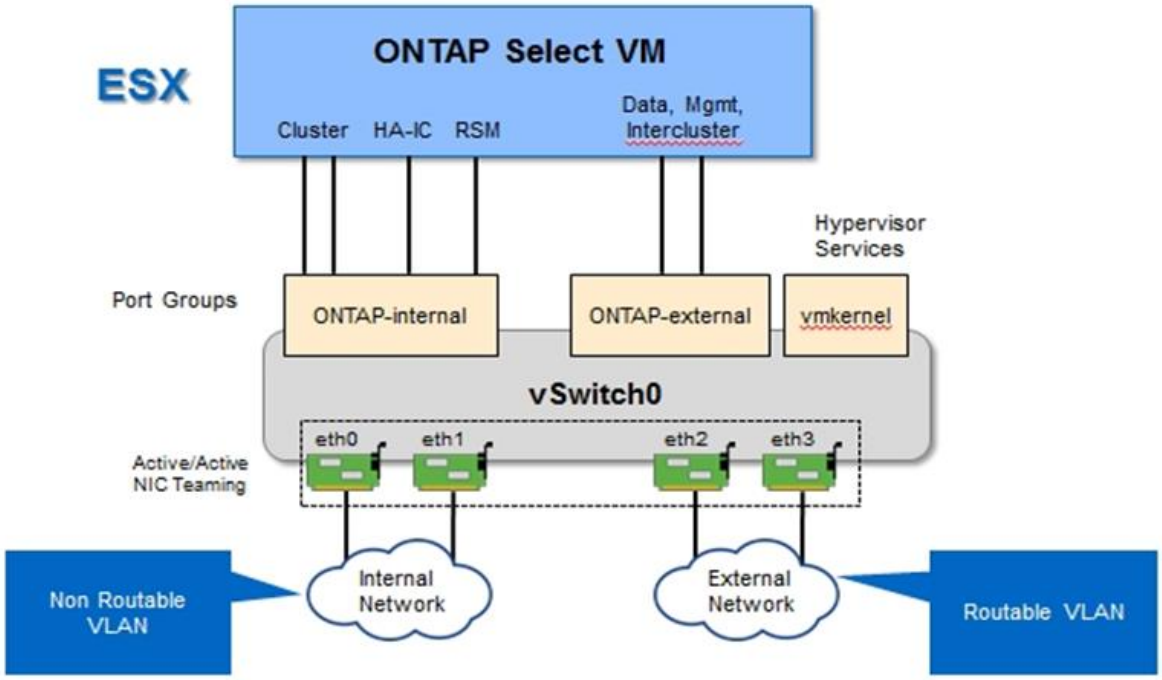
- e0a, e0b e e0g. LIF di gestione e dati
- e0c, e0d. LIF della rete del cluster
- e0e. RSM

- e0f. Interconnessione HA

Le porte e0a, e0b e e0g risiedono sulla rete esterna. Sebbene le porte da e0c a e0f eseguano diverse funzioni, nel loro insieme compongono la rete Select interna. Quando si prendono decisioni di progettazione della rete, queste porte devono essere collocate su una rete a singolo livello 2. Non è necessario separare questi adattatori virtuali su reti diverse.

La relazione tra queste porte e gli adattatori fisici sottostanti è illustrata nella figura seguente, che mostra un nodo del cluster ONTAP Select sull'hypervisor ESX.

Configurazione di rete di un singolo nodo che fa parte di un cluster ONTAP Select multinodo



La separazione del traffico interno ed esterno tra diverse schede di rete fisiche impedisce l'introduzione di latenze nel sistema a causa di un accesso insufficiente alle risorse di rete. Inoltre, l'aggregazione tramite il raggruppamento delle schede di rete garantisce che il guasto di una singola scheda di rete non impedisca al nodo del cluster ONTAP Select di accedere alla rispettiva rete.

Si noti che sia la rete esterna che i gruppi di porte di rete interna contengono tutti e quattro gli adattatori NIC in modo simmetrico. Le porte attive nel gruppo di porte di rete esterne sono le porte di standby nella rete interna. Al contrario, le porte attive nel gruppo di porte della rete interna sono le porte di standby nel gruppo di porte della rete esterna.

Assegnazione LIF

Con l'introduzione di IPspaces, i ruoli delle porte ONTAP sono stati deprecati. Come gli array FAS, i cluster ONTAP Select contengono sia un IPSpace predefinito che un IPSpace cluster. Posizionando le porte di rete e0a, e0b e e0g nell'IPSpace predefinito e le porte e0c e e0d nell'IPSpace del cluster, queste porte sono state sostanzialmente estruse dall'hosting di LIF che non appartengono. Le restanti porte all'interno del cluster ONTAP Select vengono utilizzate attraverso l'assegnazione automatica delle interfacce che forniscono servizi interni. Non sono esposti attraverso la shell ONTAP, come nel caso delle interfacce di interconnessione RSM e ha.

Non tutte le LIF sono visibili attraverso la shell dei comandi di ONTAP. L'interconnessione ha e le interfacce RSM sono nascoste da ONTAP e vengono utilizzate internamente per fornire i rispettivi servizi.

Le porte di rete e i LIF sono illustrati in dettaglio nelle sezioni seguenti.

LIF di gestione e dati (e0a, e0b e e0g)

Le porte ONTAP e0a, e0b e e0g sono delegate come porte candidate per le LIF che trasportano i seguenti tipi di traffico:

- Traffico del protocollo SAN/NAS (CIFS, NFS e iSCSI)
- Traffico di gestione di cluster, nodi e SVM
- Traffico intercluster (SnapMirror e SnapVault)

Le LIF di gestione di cluster e nodi vengono create automaticamente durante l'installazione del cluster ONTAP Select. I restanti LIF possono essere creati dopo l'implementazione.

LIF di rete del cluster (e0c, e0d)

Le porte ONTAP e0c e e0d sono delegate come porte home per le interfacce cluster. All'interno di ciascun nodo del cluster ONTAP Select, durante l'installazione di ONTAP vengono generate automaticamente due interfacce del cluster utilizzando gli indirizzi IP locali di collegamento (169.254.x.x).

A queste interfacce non possono essere assegnati indirizzi IP statici e non devono essere create interfacce cluster aggiuntive.

Il traffico di rete del cluster deve passare attraverso una rete Layer-2 a bassa latenza e non instradata. A causa del throughput del cluster e dei requisiti di latenza, si prevede che il cluster ONTAP Select si trovi fisicamente nelle vicinanze (ad esempio, multipack, singolo data center). Non è supportata la creazione di configurazioni di cluster stretch a quattro, sei o otto nodi mediante la separazione di nodi ha su una WAN o su distanze geografiche significative. È supportata una configurazione a due nodi allungata con un mediatore.

Per ulteriori informazioni, consultare la sezione "Best practice di ha (MetroCluster SDS) con due nodi estesi".

Per garantire il massimo throughput per il traffico di rete del cluster, questa porta di rete è configurata per l'utilizzo di frame jumbo (da 7500 a 9000 MTU). Per un corretto funzionamento del cluster, verificare che i frame jumbo siano attivati su tutti gli switch fisici e virtuali di upstream che forniscono servizi di rete interni ai nodi del cluster ONTAP Select.

Traffico RAID SyncMirror (e0e)

La replica sincrona dei blocchi nei nodi partner ha avviene utilizzando un'interfaccia di rete interna che risiede sulla porta di rete e0e. Questa funzionalità si verifica automaticamente, utilizzando le interfacce di rete configurate da ONTAP durante l'installazione del cluster e non richiede alcuna configurazione da parte dell'amministratore.

La porta e0e è riservata da ONTAP per il traffico di replica interno. Pertanto, né la porta né il LIF host sono visibili nell'interfaccia utente di ONTAP o in Gestione sistema. Questa interfaccia è configurata per utilizzare un indirizzo IP locale di collegamento generato automaticamente e la riassegnazione di un indirizzo IP alternativo non è supportata. Questa porta di rete richiede l'utilizzo di frame jumbo (da 7500 a 9000 MTU).

Interconnessione HA (e0f)

Gli array NetApp FAS utilizzano hardware specializzato per trasferire le informazioni tra coppie ha in un cluster ONTAP. Tuttavia, gli ambienti software-defined non tendono a disporre di questo tipo di apparecchiatura (come i dispositivi InfiniBand o iWARP), pertanto è necessaria una soluzione alternativa. Sebbene siano state prese in considerazione diverse possibilità, i requisiti ONTAP inseriti nel trasporto di interconnessione richiedevano che questa funzionalità fosse emulata nel software. Di conseguenza, all'interno di un cluster ONTAP Select, la funzionalità dell'interconnessione ha (tradizionalmente fornita dall'hardware) è stata progettata nel sistema operativo, utilizzando Ethernet come meccanismo di trasporto.

Ogni nodo ONTAP Select è configurato con una porta di interconnessione ha, e0f. Questa porta ospita l'interfaccia di rete di interconnessione ha, che è responsabile di due funzioni principali:

- Mirroring del contenuto della NVRAM tra coppie ha
- Invio/ricezione di informazioni sullo stato ha e messaggi heartbeat di rete tra coppie ha

Il traffico DI interconnessione HA passa attraverso questa porta di rete utilizzando una singola interfaccia di rete, sovrapporendo frame RDMA (Remote Direct Memory Access) all'interno di pacchetti Ethernet.

In modo simile alla porta RSM (e0e), né la porta fisica né l'interfaccia di rete ospitata sono visibili agli utenti dall'interfaccia CLI di ONTAP o da Gestore di sistema. Di conseguenza, l'indirizzo IP di questa interfaccia non può essere modificato e lo stato della porta non può essere modificato. Questa porta di rete richiede l'utilizzo di frame jumbo (da 7500 a 9000 MTU).

Rete interna ed esterna ONTAP Select

Caratteristiche delle reti interne ed esterne ONTAP Select.

Rete interna ONTAP Select

La rete ONTAP Select interna, presente solo nella variante multinodo del prodotto, è responsabile della fornitura al cluster ONTAP Select di servizi di comunicazione cluster, interconnessione ha e replica sincrona. Questa rete include le seguenti porte e interfacce:

- **E0c, e0d.** hosting delle LIF di rete del cluster
- **E0e.** hosting della LIF RSM
- **E0f.** hosting della LIF di interconnessione ha

Il throughput e la latenza di questa rete sono fondamentali per determinare le performance e la resilienza del cluster ONTAP Select. L'isolamento della rete è necessario per la sicurezza del cluster e per garantire che le interfacce di sistema siano separate da altri tipi di traffico di rete. Pertanto, questa rete deve essere utilizzata esclusivamente dal cluster ONTAP Select.

Non è supportato l'utilizzo di Select Internal Network per il traffico diverso da Select Cluster, ad esempio il traffico di gestione o dell'applicazione. Sulla VLAN interna di ONTAP non possono essere presenti altre macchine virtuali o host.

I pacchetti di rete che attraversano la rete interna devono trovarsi su una rete Layer-2 con tag VLAN dedicata. Per eseguire questa operazione, eseguire una delle seguenti operazioni:

- Assegnazione di un gruppo di porte con tag VLAN alle NIC virtuali interne (da e0c a e0f) (modalità VST)
- Utilizzo della VLAN nativa fornita dallo switch upstream in cui la VLAN nativa non viene utilizzata per altri tipi di traffico (assegnazione di un gruppo di porte senza ID VLAN, ovvero in modalità EST)

In tutti i casi, il tagging VLAN per il traffico di rete interno viene eseguito all'esterno della macchina virtuale ONTAP Select.

Sono supportati solo vSwitch ESX standard e distribuiti. Altri switch virtuali o connettività diretta tra host ESX non sono supportati. La rete interna deve essere completamente aperta; NAT o firewall non sono supportati.

All'interno di un cluster ONTAP Select, il traffico interno e il traffico esterno vengono separati utilizzando oggetti di rete Layer-2 virtuali noti come gruppi di porte. L'assegnazione corretta di vSwitch a questi gruppi di porte è estremamente importante, soprattutto per la rete interna, che è responsabile della fornitura di cluster, interconnessione ha e servizi di replica mirror. Una larghezza di banda di rete insufficiente per queste porte di rete può causare un peggioramento delle performance e persino compromettere la stabilità del nodo del cluster. Pertanto, i cluster a quattro, sei e otto nodi richiedono che la rete ONTAP Select interna utilizzi una connettività da 10 GB; le NIC da 1 GB non sono supportate. Tuttavia, è possibile effettuare compromessi sulla rete esterna, poiché la limitazione del flusso di dati in entrata a un cluster ONTAP Select non influisce sulla sua capacità di operare in modo affidabile.

Un cluster a due nodi può utilizzare quattro porte da 1 GB per il traffico interno o una singola porta da 10 GB invece delle due porte da 10 GB richieste dal cluster a quattro nodi. In un ambiente in cui le condizioni impediscono al server di adattarsi a quattro schede NIC da 10 GB, è possibile utilizzare due schede NIC da 10 GB per la rete interna e due schede NIC da 1 GB per la rete ONTAP esterna.

Validazione e troubleshooting della rete interna

La rete interna di un cluster multinode può essere validata utilizzando la funzionalità di verifica della connettività di rete. Questa funzione può essere richiamata dalla CLI di distribuzione che esegue `network connectivity-check start` comando.

Eseguire il seguente comando per visualizzare l'output del test:

```
network connectivity-check show --run-id X (X is a number)
```

Questo strumento è utile solo per la risoluzione dei problemi della rete interna in un cluster Multinode Select. Lo strumento non deve essere utilizzato per risolvere i problemi relativi a cluster a nodo singolo (incluse le configurazioni vNAS), implementazione ONTAP su connettività ONTAP Select o problemi di connettività lato client.

La creazione guidata del cluster (parte dell'interfaccia grafica di implementazione di ONTAP) include il Network checker interno come passaggio opzionale disponibile durante la creazione di cluster a più nodi. Dato l'importante ruolo svolto dalla rete interna nei cluster a più nodi, facendo parte del workflow di creazione del

cluster questo passaggio migliora il tasso di successo delle operazioni di creazione del cluster.

A partire da ONTAP Deploy 2.10, la dimensione MTU utilizzata dalla rete interna può essere impostata tra 7,500 e 9,000. Il controllo della connettività di rete può essere utilizzato anche per verificare le dimensioni MTU comprese tra 7,500 e 9,000. Il valore MTU predefinito viene impostato sul valore dello switch della rete virtuale. Se nell'ambiente è presente un overlay di rete come VXLAN, tale valore predefinito deve essere sostituito con un valore inferiore.

Rete esterna ONTAP Select

La rete esterna ONTAP Select è responsabile di tutte le comunicazioni in uscita dal cluster e, pertanto, è presente sia nella configurazione a nodo singolo che in quella a più nodi. Sebbene questa rete non abbia i requisiti di throughput strettamente definiti della rete interna, l'amministratore deve fare attenzione a non creare colli di bottiglia di rete tra il client e ONTAP VM, perché i problemi di performance potrebbero essere erroneamente definiti come problemi ONTAP Select.

In modo simile al traffico interno, il traffico esterno può essere etichettato al vSwitch Layer (VST) e al External Switch Layer (EST). Inoltre, il traffico esterno può essere etichettato dalla macchina virtuale ONTAP Select stessa in un processo noto come VGT. Vedere la sezione "Separazione dei dati e del traffico di gestione" per ulteriori dettagli.

La tabella seguente evidenzia le principali differenze tra le reti interne ed esterne di ONTAP Select.

Guida di riferimento rapido per reti interne ed esterne

Descrizione	Rete interna	Rete esterna
Servizi di rete	Cluster HA/IC RAID SyncMirror (RSM)	Gestione dei dati Intercluster (SnapMirror e SnapVault)
Isolamento della rete	Obbligatorio	Opzionale
Dimensioni del frame (MTU)	da 7,500 a 9,000	1.500 (impostazione predefinita) 9.000 (supportato)
Assegnazione dell'indirizzo IP	Generato automaticamente	Definito dall'utente
Supporto DHCP	No	No

Raggruppamento NIC

Per garantire che le reti interne ed esterne dispongano delle caratteristiche di larghezza di banda e resilienza necessarie per garantire performance elevate e tolleranza agli errori, si consiglia il raggruppamento fisico delle schede di rete. Sono supportate configurazioni cluster a due nodi con un singolo collegamento da 10 GB. Tuttavia, la Best practice consigliata da NetApp consiste nell'utilizzare il raggruppamento delle schede di rete sulle reti interne ed esterne del cluster ONTAP Select.

Generazione di indirizzi MAC

Gli indirizzi MAC assegnati a tutte le porte di rete ONTAP Select vengono generati automaticamente dall'utilità di implementazione inclusa. L'utility utilizza un OUI (Organizationally Unique Identifier) specifico della piattaforma per garantire che non vi siano conflitti con i sistemi FAS. Una copia di questo indirizzo viene quindi memorizzata in un database interno all'interno della macchina virtuale per l'installazione di ONTAP Select (implementazione di ONTAP), per evitare riassegnazioni accidentali durante le implementazioni future dei nodi. L'amministratore non deve mai modificare l'indirizzo MAC assegnato a una porta di rete.

Configurazioni di rete supportate

Seleziona l'hardware migliore e configura la tua rete per ottimizzare performance e resilienza.

I vendor di server sanno che i clienti hanno esigenze diverse e che la scelta è fondamentale. Di conseguenza, quando si acquista un server fisico, sono disponibili numerose opzioni per prendere decisioni sulla connettività di rete. La maggior parte dei sistemi commodity viene fornita con diverse opzioni NIC che offrono opzioni a porta singola e multiporta con variazioni di velocità e throughput. Incluso il supporto per schede NIC a 25GB GB/s e 40GB GB/s con VMware ESX.

Poiché le prestazioni della macchina virtuale ONTAP Select sono legate direttamente alle caratteristiche dell'hardware sottostante, l'aumento del throughput della macchina virtuale selezionando NIC a velocità più elevata consente di ottenere un cluster dalle performance più elevate e un'esperienza utente migliore. È possibile utilizzare quattro NIC da 10 GB o due NIC a velocità superiore (25/40 GB/s) per ottenere un layout di rete dalle performance elevate. Sono supportate anche altre configurazioni. Per i cluster a due nodi, sono supportate 4 porte da 1 GB o 1 porta da 10 GB. Per i cluster a nodo singolo, sono supportate 2 porte da 1 GB.

Configurazioni minime di rete e consigliate

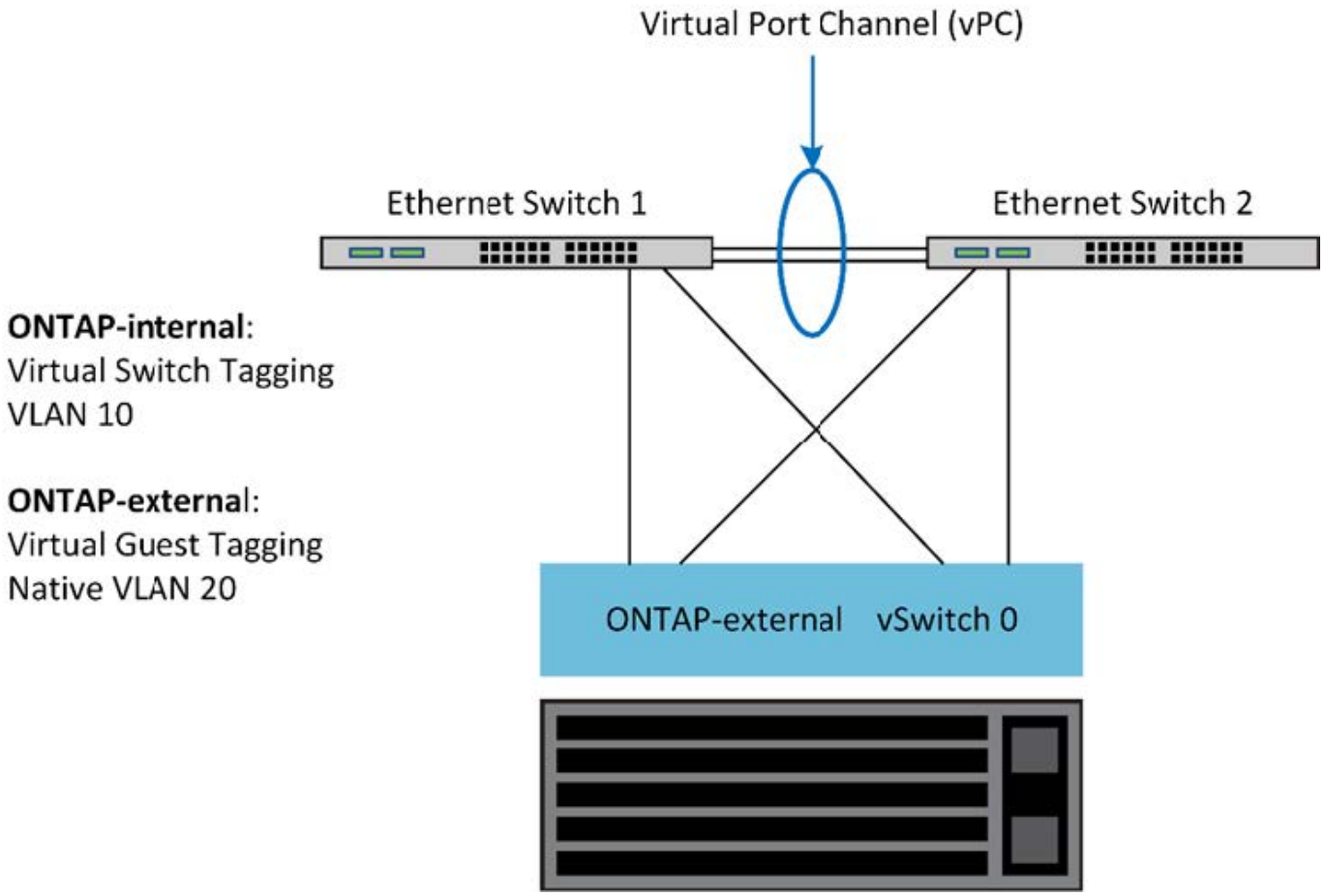
Sono supportate diverse configurazioni Ethernet in base alle dimensioni del cluster.

Dimensione del cluster	Requisiti minimi	Consiglio
Cluster a nodo singolo	2 x 1GbE	2 x 10GbE
Cluster a due nodi o SDS MetroCluster	4 x 1 GbE o 1 x 10 GbE	2 x 10GbE
cluster a 4/6/8 nodi	2 x 10GbE	4 x 10 GbE o 2 x 25/40 GbE

La conversione tra topologie di collegamento singolo e di collegamento multiplo in un cluster in esecuzione non è supportata a causa della possibile necessità di convertire tra diverse configurazioni di raggruppamento NIC richieste per ciascuna topologia.

Configurazione di rete con più switch fisici

Quando è disponibile hardware sufficiente, NetApp consiglia di utilizzare la configurazione multiswitch illustrata nella figura seguente, a causa della protezione aggiuntiva contro i guasti fisici degli switch.



Configurazione di VMware vSphere vSwitch su ESXi

Configurazione di ONTAP Select vSwitch e policy di bilanciamento del carico per configurazioni a due e quattro NIC.

ONTAP Select supporta l'utilizzo di configurazioni vSwitch standard e distribuite. I vSwitch distribuiti supportano i costrutti di aggregazione dei collegamenti (LACP). L'aggregazione dei collegamenti è una struttura di rete comune utilizzata per aggregare la larghezza di banda tra più adattatori fisici. LACP è uno standard indipendente dal vendor che fornisce un protocollo aperto per gli endpoint di rete che raggruppa gruppi di porte di rete fisiche in un singolo canale logico. ONTAP Select può funzionare con gruppi di porte configurati come gruppo di aggregazione di banda (LAG). Tuttavia, NetApp consiglia di utilizzare le singole porte fisiche come semplici porte di uplink (trunk) per evitare la configurazione LAG. In questi casi, le Best practice per vSwitch standard e distribuiti sono identiche.

Questa sezione descrive la configurazione di vSwitch e i criteri di bilanciamento del carico da utilizzare nelle configurazioni a due NIC e a quattro NIC.

Quando si configurano i gruppi di porte per l'utilizzo da parte di ONTAP Select, è necessario seguire le seguenti Best practice: Il criterio di bilanciamento del carico a livello di gruppo di porte è instradamento basato sull'ID della porta virtuale di origine. VMware consiglia di impostare STP su Portfast sulle porte dello switch collegate agli host ESXi.

Tutte le configurazioni vSwitch richiedono un minimo di due schede di rete fisiche raggruppate in un singolo raggruppamento di schede di rete. ONTAP Select supporta un singolo collegamento da 10 GB per cluster a due nodi. Tuttavia, è una Best practice di NetApp garantire la ridondanza hardware attraverso l'aggregazione NIC.

Su un server vSphere, i raggruppamenti NIC sono il costrutto di aggregazione utilizzato per raggruppare più adattatori di rete fisici in un singolo canale logico, consentendo la condivisione del carico di rete tra tutte le porte membri. È importante ricordare che i team NIC possono essere creati senza il supporto dello switch fisico. I criteri di bilanciamento del carico e failover possono essere applicati direttamente a un team NIC, che non conosce la configurazione dello switch upstream. In questo caso, i criteri vengono applicati solo al traffico in uscita.

I canali delle porte statiche non sono supportati con ONTAP Select. I canali abilitati LACP sono supportati con vSwitch distribuiti, ma l'utilizzo di LAG LACP può comportare una distribuzione del carico non uniforme tra i membri LAG.

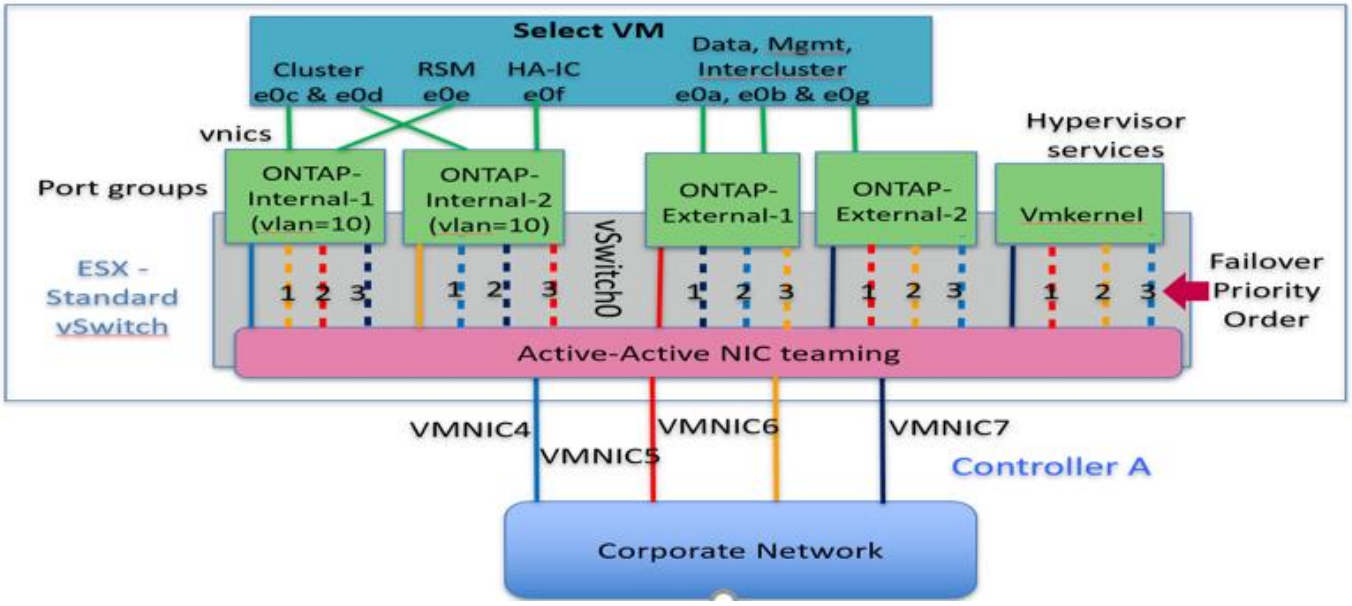
Per i cluster a nodo singolo, ONTAP Deploy configura la macchina virtuale ONTAP Select in modo che utilizzi un gruppo di porte per la rete esterna e lo stesso gruppo di porte o, facoltativamente, un gruppo di porte diverso per il traffico di gestione del cluster e dei nodi. Per i cluster a nodo singolo, è possibile aggiungere il numero desiderato di porte fisiche al gruppo di porte esterne come adattatori attivi.

Per i cluster a più nodi, ONTAP Deploy configura ogni macchina virtuale ONTAP Select in modo che utilizzi uno o due gruppi di porte per la rete interna e, separatamente, uno o due gruppi di porte per la rete esterna. Il traffico di gestione del cluster e dei nodi può utilizzare lo stesso gruppo di porte del traffico esterno o, facoltativamente, un gruppo di porte separato. Il traffico di gestione del cluster e dei nodi non può condividere lo stesso gruppo di porte con il traffico interno.

VSwitch standard o distribuito e quattro porte fisiche per nodo

È possibile assegnare quattro gruppi di porte a ciascun nodo di un cluster a più nodi. Ciascun gruppo di porte dispone di una singola porta fisica attiva e di tre porte fisiche di standby, come illustrato nella figura seguente.

VSwitch con quattro porte fisiche per nodo



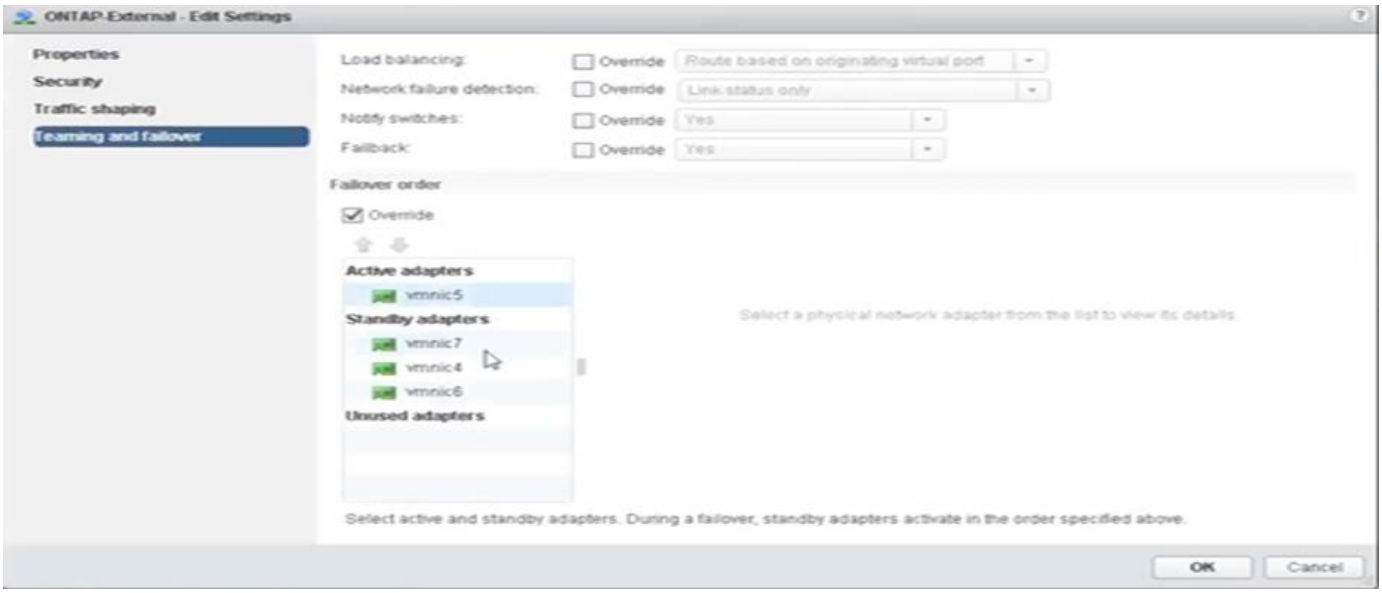
L'ordine delle porte nell'elenco di standby è importante. La tabella seguente fornisce un esempio della distribuzione fisica delle porte tra i quattro gruppi di porte.

Configurazioni minime di rete e consigliate

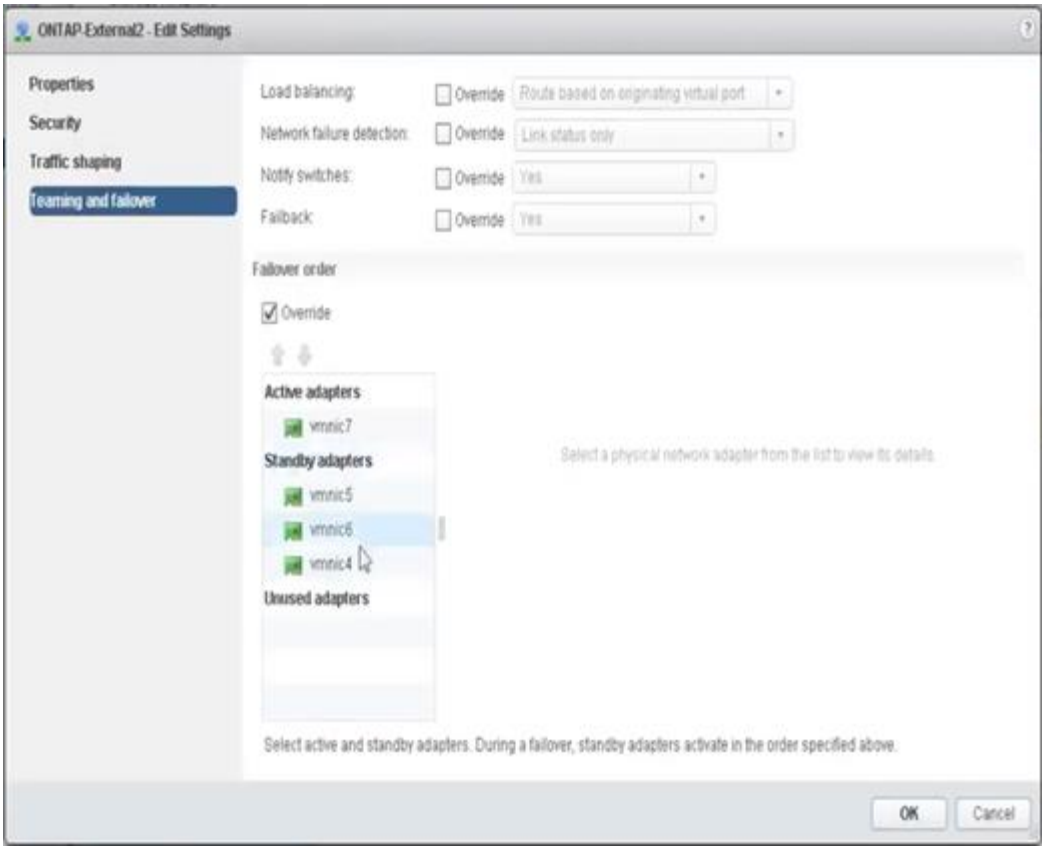
Gruppo di porte	Esterno 1	Esterni 2	Interno 1	Interno 2
Attivo	vmnic0	vmnic1	vmnic2	vmnic3
Standby 1	vmnic1	vmnic0	vmnic3	vmnic2
Standby 2	vmnic2	vmnic3	vmnic0	vmnic1
Standby 3	vmnic3	vmnic2	vmnic1	vmnic0

Le seguenti figure mostrano le configurazioni dei gruppi di porte di rete esterne dalla GUI di vCenter (ONTAP-External e ONTAP-External2). Si noti che gli adattatori attivi provengono da schede di rete diverse. In questa configurazione, vmnic 4 e vmnic 5 sono porte doppie sulla stessa scheda NIC fisica, mentre vmnic 6 e vminc 7 sono analogamente porte doppie su una scheda NIC separata (in questo esempio non vengono utilizzate le porte da 0 a 3). L'ordine degli adattatori di standby fornisce un failover gerarchico con le ultime porte della rete interna. L'ordine delle porte interne nell'elenco di standby viene similmente scambiato tra i due gruppi di porte esterne.

Parte 1: Configurazioni del gruppo di porte esterne ONTAP Select



Parte 2: Configurazioni dei gruppi di porte esterne ONTAP Select

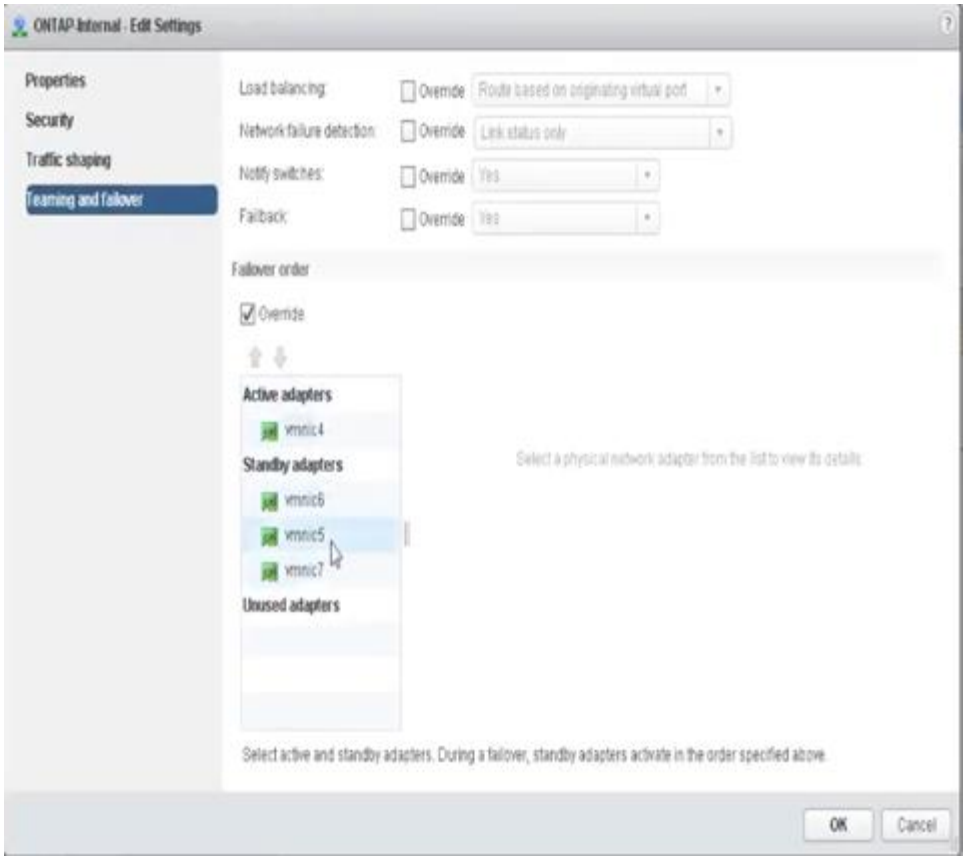


Per la leggibilità, le attività sono le seguenti:

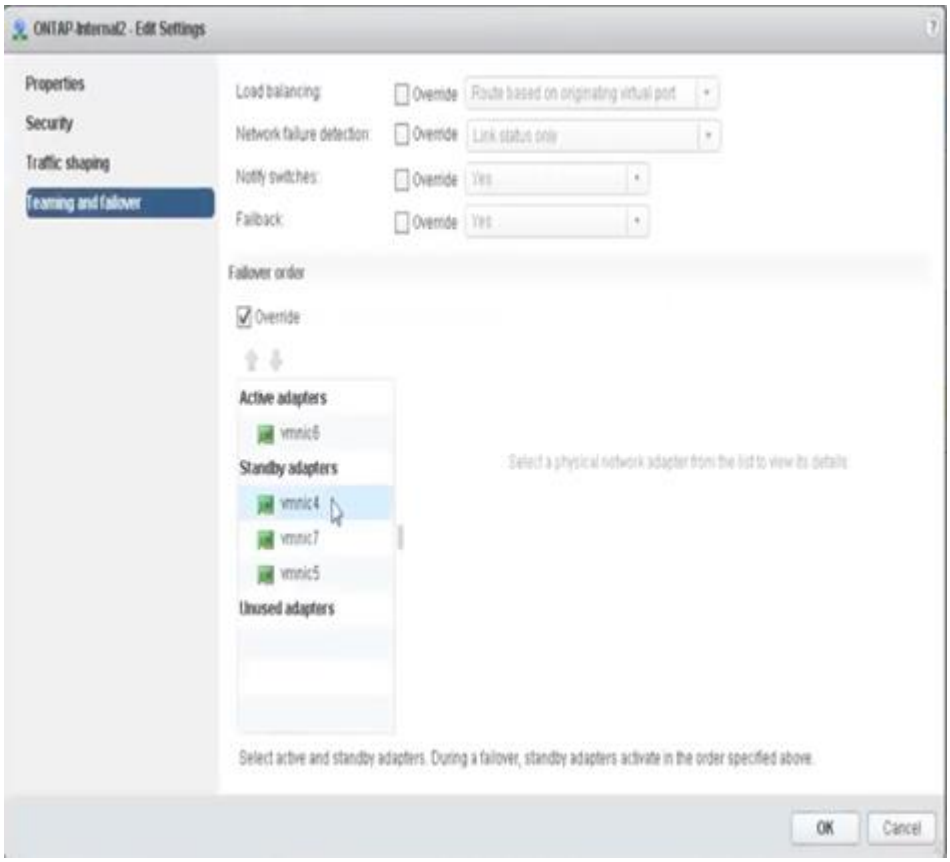
ONTAP-esterno	ONTAP-External2
Adattatori attivi: vmnic5 Adattatori di standby: vmnic7, vmnic4, vmnic6	Adattatori attivi: vmnic7 Adattatori di standby: vmnic5, vmnic6, vmnic4

Le seguenti figure mostrano le configurazioni dei gruppi di porte di rete interni (ONTAP-Internal e ONTAP-Internal2). Si noti che gli adattatori attivi provengono da schede di rete diverse. In questa configurazione, vmnic 4 e vmnic 5 sono porte doppie sullo stesso ASIC fisico, mentre vmnic 6 e vmnic 7 sono analogamente porte doppie su un ASIC separato. L'ordine degli adattatori di standby fornisce un failover gerarchico con le ultime porte della rete esterna. L'ordine delle porte esterne nell'elenco di standby viene similmente scambiato tra i due gruppi di porte interne.

Parte 1: Configurazioni del gruppo di porte interne ONTAP Select



Parte 2: Gruppi di porte interne ONTAP Select



Per la leggibilità, le attività sono le seguenti:

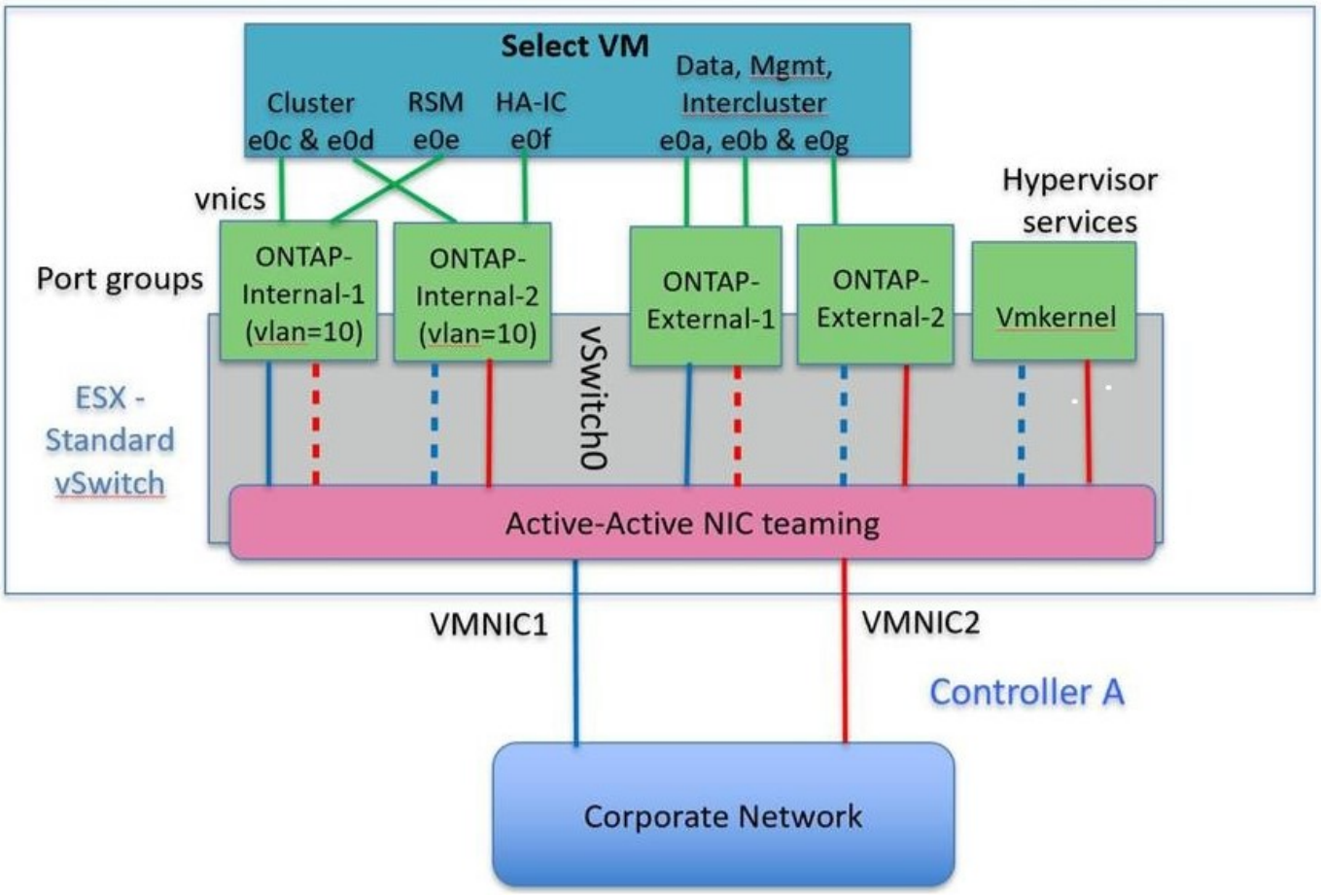
ONTAP-interno	ONTAP-Internal2
Adattatori attivi: vmnic4 Adattatori di standby: vmnic6, vmnic5, vmnic7	Adattatori attivi: vmnic6 Adattatori di standby: vmnic4, vmnic7, vmnic5

VSwitch standard o distribuito e due porte fisiche per nodo

Quando si utilizzano due schede di rete ad alta velocità (25 GB), la configurazione consigliata del gruppo di porte è concettualmente molto simile alla configurazione con quattro schede di rete da 10 GB. È necessario utilizzare quattro gruppi di porte anche quando si utilizzano solo due adattatori fisici. Le assegnazioni dei gruppi di porte sono le seguenti:

Gruppo di porte	Esterno 1 (e0a,e0b)	Interno 1 (e0c,e0e)	Interno 2 (e0d,e0f)	Esterno 2 (e0g)
Attivo	vmnic0	vmnic0	vmnic1	vmnic1
Standby	vmnic1	vmnic1	vmnic0	vmnic0

VSwitch con due porte fisiche ad alta velocità (25 GB) per nodo

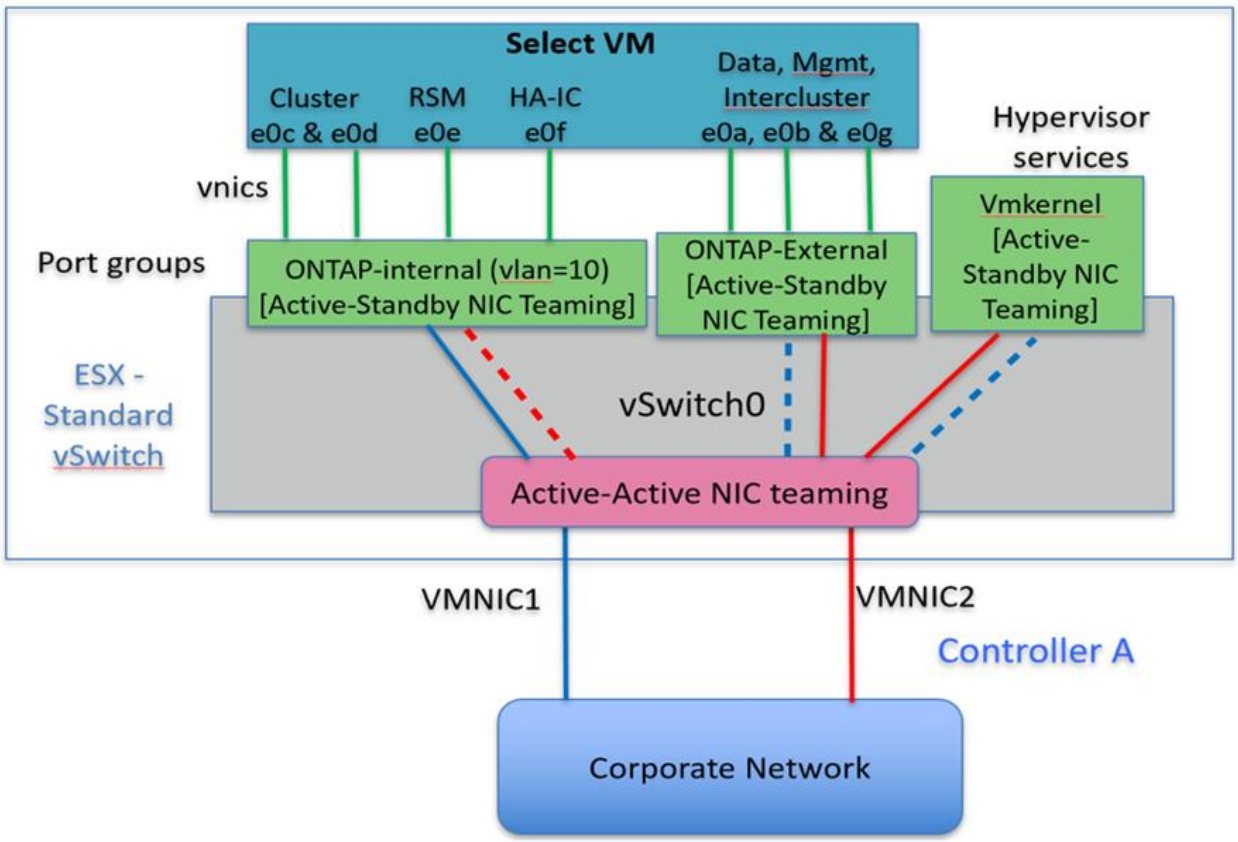


Quando si utilizzano due porte fisiche (10 GB o meno), ciascun gruppo di porte deve disporre di un adattatore attivo e di un adattatore di standby configurati uno di fronte all'altro. La rete interna è presente solo per i cluster ONTAP Select a più nodi. Per i cluster a nodo singolo, entrambi gli adattatori possono essere configurati come attivi nel gruppo di porte esterne.

Nell'esempio seguente viene illustrata la configurazione di vSwitch e dei due gruppi di porte responsabili della gestione dei servizi di comunicazione interni ed esterni per un cluster ONTAP Select multinodo. La rete esterna

può utilizzare LA RETE interna VMNIC in caso di interruzione della rete, in quanto le VMNIC della rete interna fanno parte di questo gruppo di porte e sono configurate in modalità standby. Il contrario è il caso della rete esterna. L'alternanza delle VMNIC attive e di standby tra i due gruppi di porte è fondamentale per il corretto failover delle VM ONTAP Select durante le interruzioni di rete.

VSwitch con due porte fisiche (10 GB o meno) per nodo

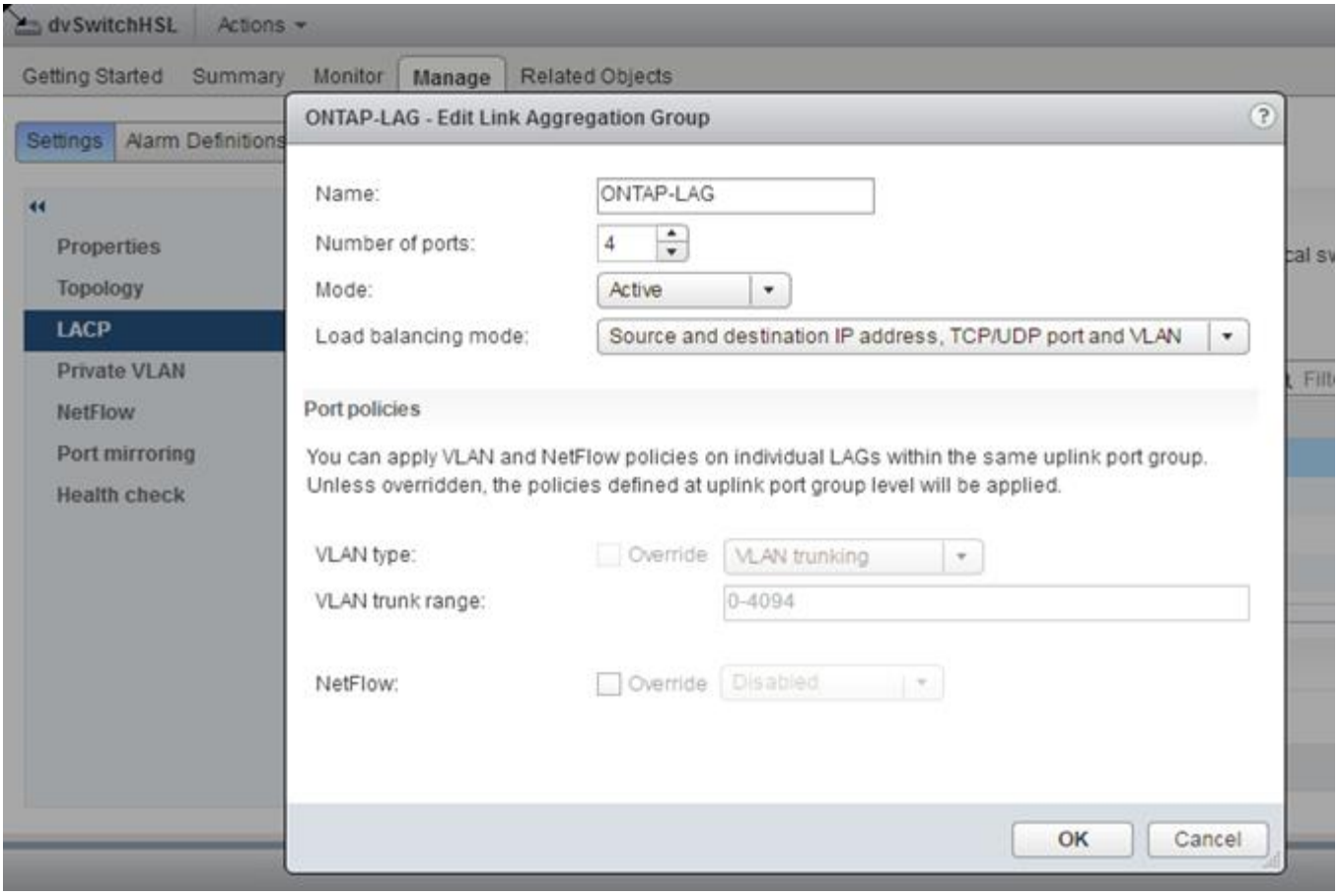


VSwitch distribuito con LACP

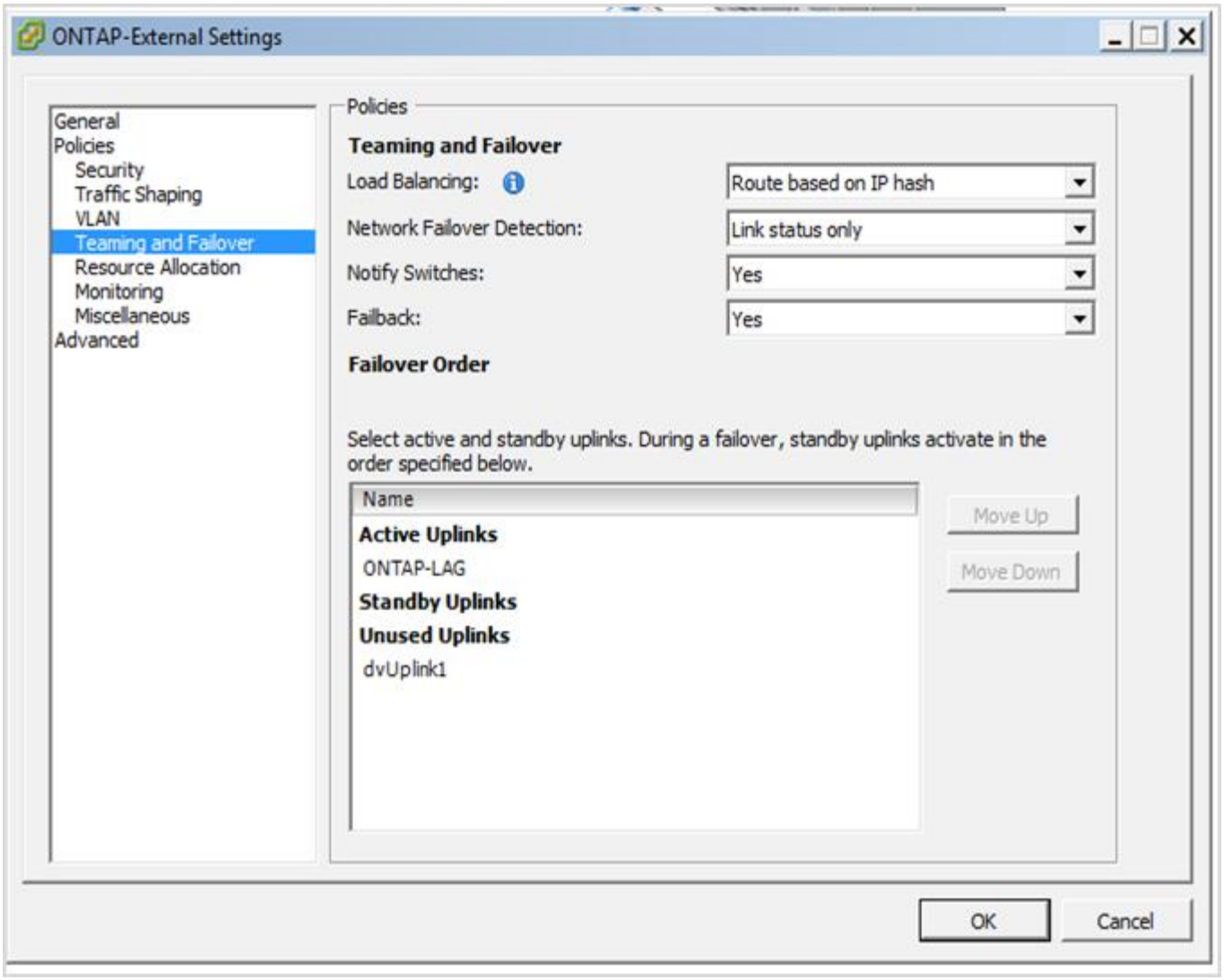
Quando si utilizzano vSwitch distribuiti nella configurazione, è possibile utilizzare LACP (anche se non si tratta di una procedura consigliata) per semplificare la configurazione di rete. L'unica configurazione LACP supportata richiede che tutti i VMNIC si trovino in un singolo LAG. Lo switch fisico uplink deve supportare una dimensione MTU compresa tra 7,500 e 9,000 su tutte le porte del canale. Le reti ONTAP Select interne ed esterne devono essere isolate a livello di gruppo di porte. La rete interna deve utilizzare una VLAN non instradabile (isolata). La rete esterna può utilizzare VST, EST o VGT.

Gli esempi seguenti mostrano la configurazione di vSwitch distribuito utilizzando LACP.

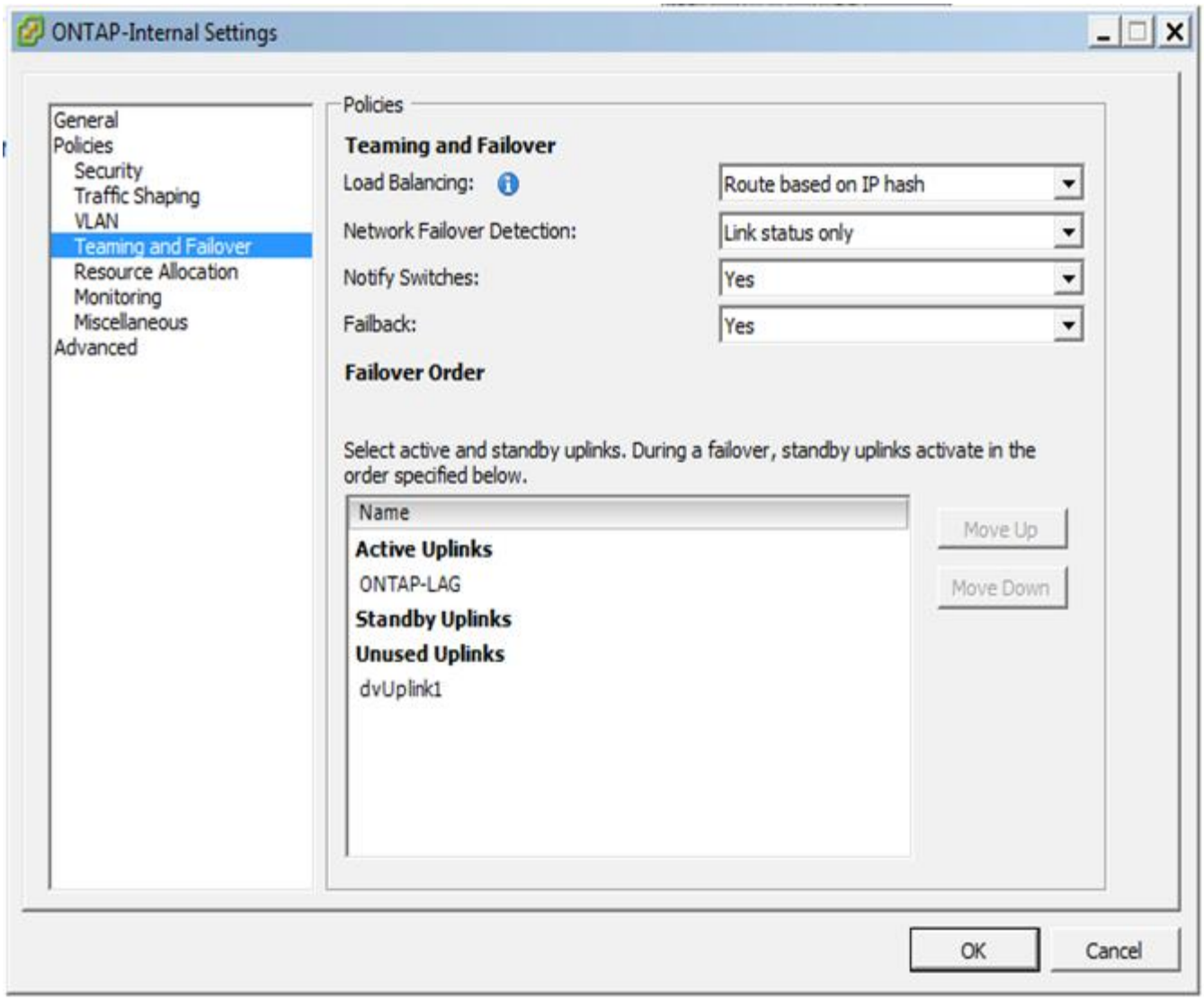
LAG properties when using LACP



Configurazioni di gruppi di porte esterne che utilizzano un vSwitch distribuito con LACP attivato



Configurazioni di gruppi di porte interne che utilizzano un vSwitch distribuito con LACP attivato



LACP richiede di configurare le porte dello switch upstream come canale di porta. Prima di attivare questa opzione su vSwitch distribuito, assicurarsi che un canale di porta abilitato LACP sia configurato correttamente.

Configurazione fisica dello switch

Dettagli sulla configurazione dello switch fisico in upstream basati su ambienti a switch singolo e multi-switch.

È necessario prendere attentamente in considerazione le decisioni relative alla connettività, dal livello di switch virtuale agli switch fisici. La separazione del traffico del cluster interno dai servizi dati esterni deve estendersi al livello di rete fisico di upstream attraverso l'isolamento fornito dalle VLAN di livello 2.

Le porte dello switch fisico devono essere configurate come porte trunkport. Il traffico esterno ONTAP Select può essere separato in più reti Layer-2 in uno dei due modi. Un metodo consiste nell'utilizzo di porte virtuali con tag VLAN ONTAP con un singolo gruppo di porte. L'altro metodo consiste nell'assegnare gruppi di porte separati in modalità VST alla porta di gestione e0a. È inoltre necessario assegnare le porte dati a e0b e e0c/e0g a seconda della release di ONTAP Select e della configurazione a nodo singolo o a più nodi. Se il traffico esterno viene separato su più reti Layer-2, le porte dello switch fisico uplink devono includere tali VLAN nell'elenco delle VLAN consentite.

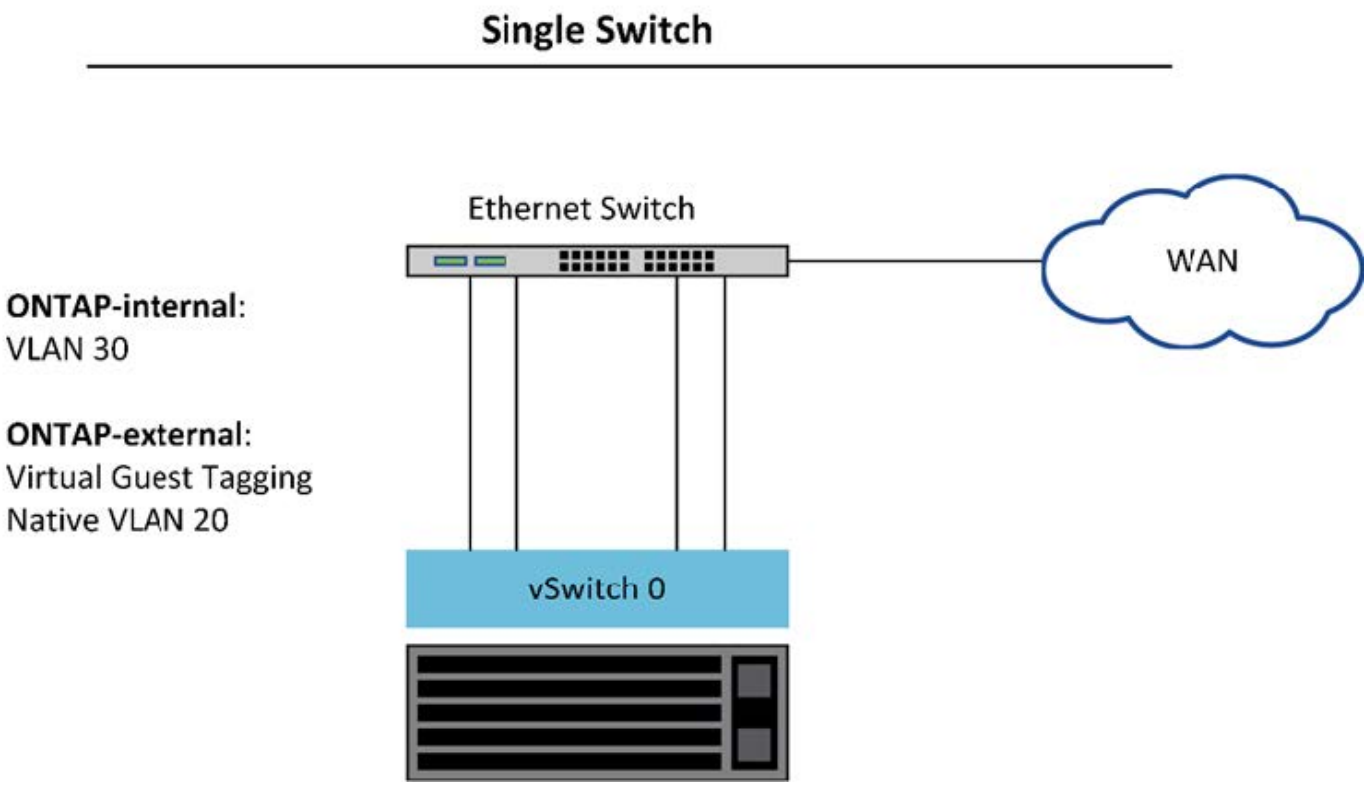
Il traffico di rete interno ONTAP Select si verifica utilizzando interfacce virtuali definite con indirizzi IP locali di collegamento. Poiché questi indirizzi IP non sono instradabili, il traffico interno tra i nodi del cluster deve fluire in una singola rete Layer-2. I route hop tra i nodi del cluster ONTAP Select non sono supportati.

Switch fisico condiviso

La figura seguente mostra una possibile configurazione dello switch utilizzata da un nodo in un cluster ONTAP Select a più nodi. In questo esempio, le NIC fisiche utilizzate dai vSwitch che ospitano i gruppi di porte di rete interna ed esterna vengono cablate allo stesso switch upstream. Il traffico dello switch viene mantenuto isolato utilizzando domini di broadcast contenuti all'interno di VLAN separate.

Per la rete interna ONTAP Select, il tagging viene eseguito a livello di gruppo di porte. Mentre nel seguente esempio viene utilizzato VGT per la rete esterna, VGT e VST sono supportati su tale gruppo di porte.

Configurazione di rete con switch fisico condiviso

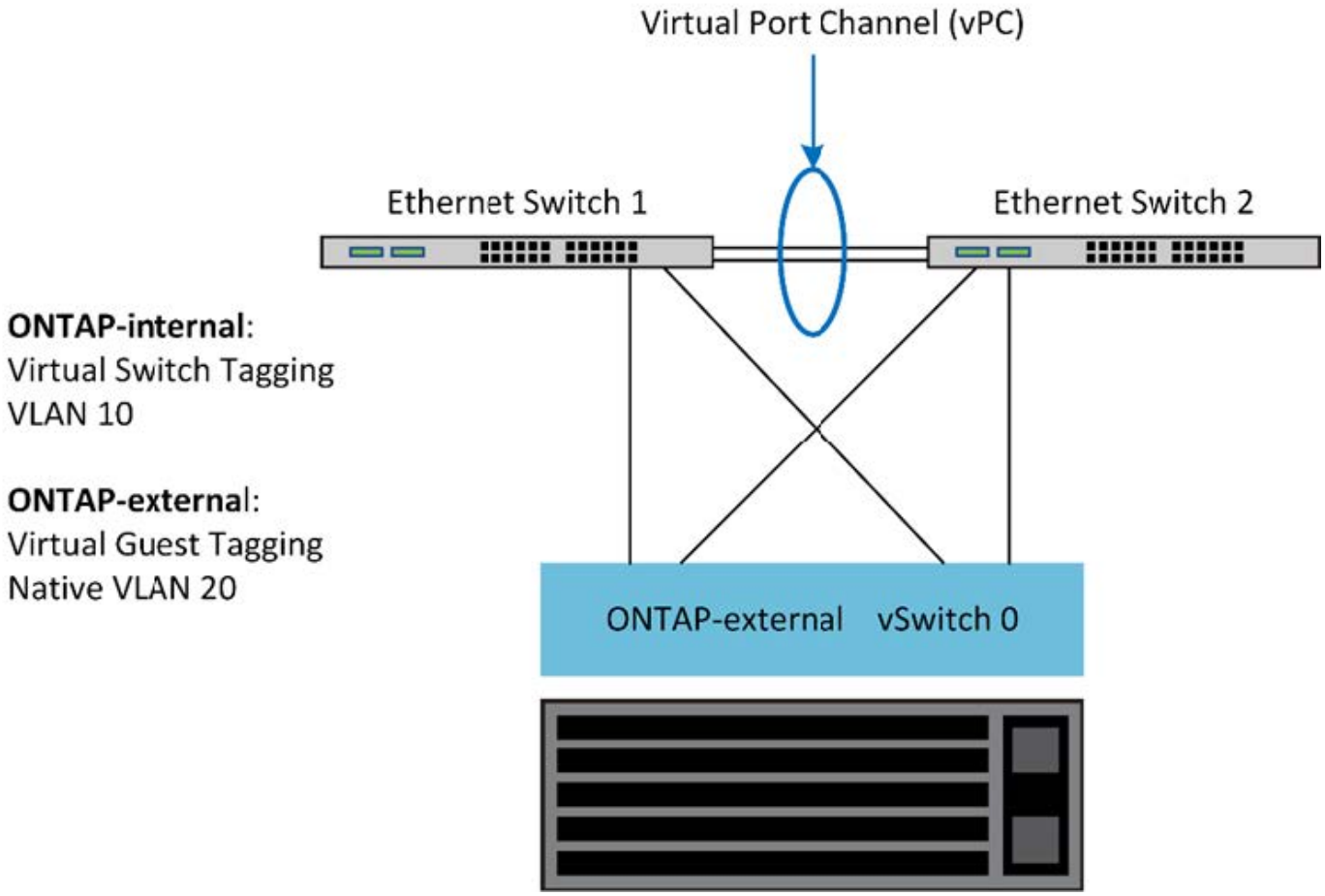


In questa configurazione, lo switch condiviso diventa un singolo punto di errore. Se possibile, è necessario utilizzare più switch per evitare che un guasto fisico dell'hardware causi un'interruzione della rete del cluster.

Switch fisici multipli

Quando è necessaria la ridondanza, è necessario utilizzare più switch di rete fisici. La figura seguente mostra una configurazione consigliata utilizzata da un nodo in un cluster ONTAP Select multinodo. Le schede di rete dei gruppi di porte interne ed esterne sono cablate in diversi switch fisici, proteggendo l'utente da un singolo guasto dello switch hardware. Tra gli switch viene configurato un canale di porte virtuali per evitare problemi di spanning tree.

Configurazione di rete con più switch fisici



Separazione dei dati e del traffico di gestione

Isolare il traffico dati e il traffico di gestione in reti Layer-2 separate.

Il traffico di rete esterno ONTAP Select è definito come traffico di dati (CIFS, NFS e iSCSI), gestione e replica (SnapMirror). All'interno di un cluster ONTAP, ogni stile di traffico utilizza un'interfaccia logica separata che deve essere ospitata su una porta di rete virtuale. Nella configurazione multinodo di ONTAP Select, queste porte sono indicate come e0a e e0b/e0g. Nella configurazione a nodo singolo, questi sono indicati come e0a e e0b/e0c, mentre le restanti porte sono riservate ai servizi cluster interni.

NetApp consiglia di isolare il traffico dati e il traffico di gestione in reti Layer-2 separate. Nell'ambiente ONTAP Select, questa operazione viene eseguita utilizzando i tag VLAN. Ciò può essere ottenuto assegnando un gruppo di porte con tag VLAN all'adattatore di rete 1 (porta e0a) per il traffico di gestione. È quindi possibile assegnare un gruppo di porte separato alle porte e0b e e0c (cluster a nodo singolo) e e0b e e0g (cluster a più nodi) per il traffico dati.

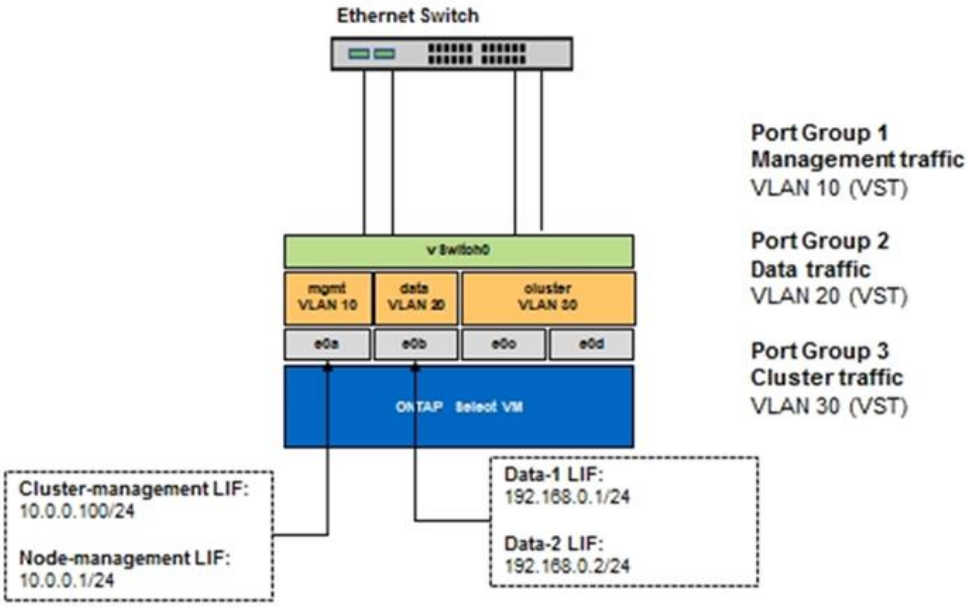
Se la soluzione VST descritta in precedenza in questo documento non è sufficiente, potrebbe essere necessario collocare i dati e le LIF di gestione sulla stessa porta virtuale. A tale scopo, utilizzare un processo noto come VGT, in cui il tagging VLAN viene eseguito dalla macchina virtuale.

> La separazione della rete di gestione e dei dati tramite VGT non è disponibile quando si utilizza l'utility di implementazione di ONTAP. Questa procedura deve essere eseguita al termine dell'installazione del cluster.

Esiste un'ulteriore avvertenza quando si utilizzano cluster VGT e a due nodi. Nelle configurazioni cluster a due nodi, l'indirizzo IP di gestione dei nodi viene utilizzato per stabilire la connettività con il mediatore prima che ONTAP sia completamente disponibile. Pertanto, solo il tagging EST e VST è supportato sul gruppo di porte mappato alla LIF di gestione dei nodi (porta e0a). Inoltre, se sia la gestione che il traffico dati utilizzano lo stesso gruppo di porte, solo EST/VST sono supportati per l'intero cluster a due nodi.

Sono supportate entrambe le opzioni di configurazione, VST e VGT. La figura seguente mostra il primo scenario, VST, in cui il traffico viene contrassegnato al livello vSwitch attraverso il gruppo di porte assegnato. In questa configurazione, le LIF di gestione di cluster e nodi vengono assegnate alla porta ONTAP e0a e contrassegnate con l'ID VLAN 10 attraverso il gruppo di porte assegnato. Le LIF dei dati vengono assegnate alla porta e0b e e0c o e0g e viene fornito l'ID VLAN 20 utilizzando un secondo gruppo di porte. Le porte del cluster utilizzano un terzo gruppo di porte e si trovano sull'ID VLAN 30.

Separazione di dati e gestione con VST

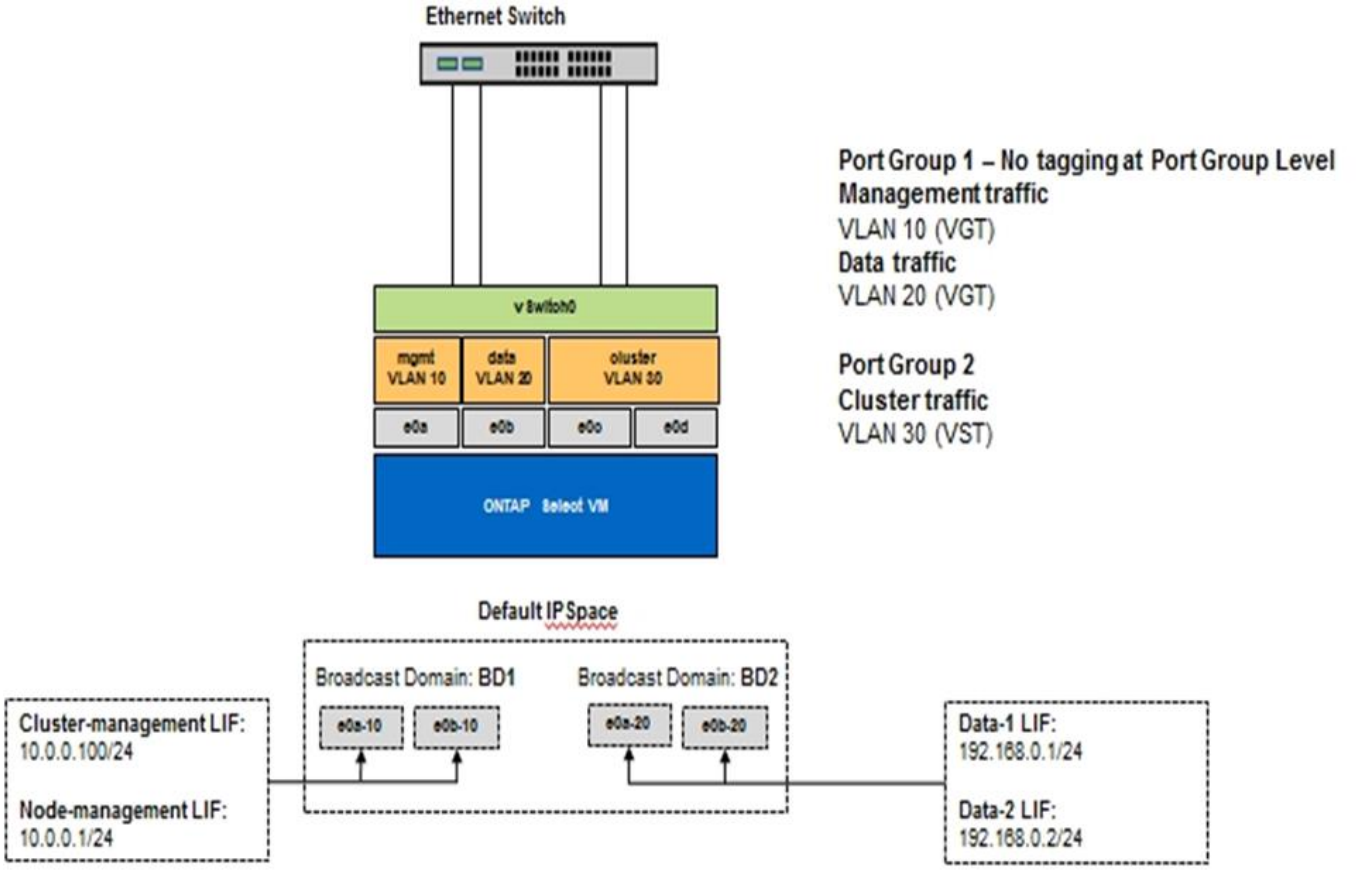


La figura seguente mostra il secondo scenario, VGT, in cui il traffico viene contrassegnato dalla macchina virtuale ONTAP utilizzando le porte VLAN che vengono collocate in domini di trasmissione separati. In questo esempio, le porte virtuali e0a-10/e0b-10/(e0c o e0g)-10 e e0a-20/e0b-20 sono posizionate sopra le porte delle macchine virtuali e0a e e0b. Questa configurazione consente di eseguire il tagging di rete direttamente all'interno di ONTAP, piuttosto che al livello vSwitch. Le LIF di gestione e dati sono collocate su queste porte virtuali, consentendo un'ulteriore suddivisione di livello 2 all'interno di una singola porta VM. La VLAN del cluster (ID VLAN 30) è ancora contrassegnata al gruppo di porte.

Note:

- Questo stile di configurazione è particolarmente utile quando si utilizzano più spazi IP. Raggruppare le porte VLAN in spazi IPpersonalizzati separati se si desidera un ulteriore isolamento logico e una maggiore multi-tenancy.
- Per supportare VGT, gli adattatori di rete host ESXi/ESX devono essere collegati alle porte trunk dello switch fisico. I gruppi di porte connessi allo switch virtuale devono avere l'ID VLAN impostato su 4095 per abilitare il trunking sul gruppo di porte.

Separazione dei dati e della gestione mediante VGT



Architettura ad alta disponibilità

Configurazioni ad alta disponibilità

Scopri le opzioni di alta disponibilità per selezionare la configurazione ha migliore per il tuo ambiente.

Anche se i clienti stanno iniziando a spostare i carichi di lavoro delle applicazioni dalle appliance di storage di livello Enterprise alle soluzioni basate su software eseguite su hardware commodity, le aspettative e le esigenze relative a resilienza e tolleranza agli errori non sono cambiate. Una soluzione ha che fornisce un RPO (Zero Recovery Point Objective) protegge il cliente dalla perdita di dati dovuta a un guasto da qualsiasi componente dello stack dell'infrastruttura.

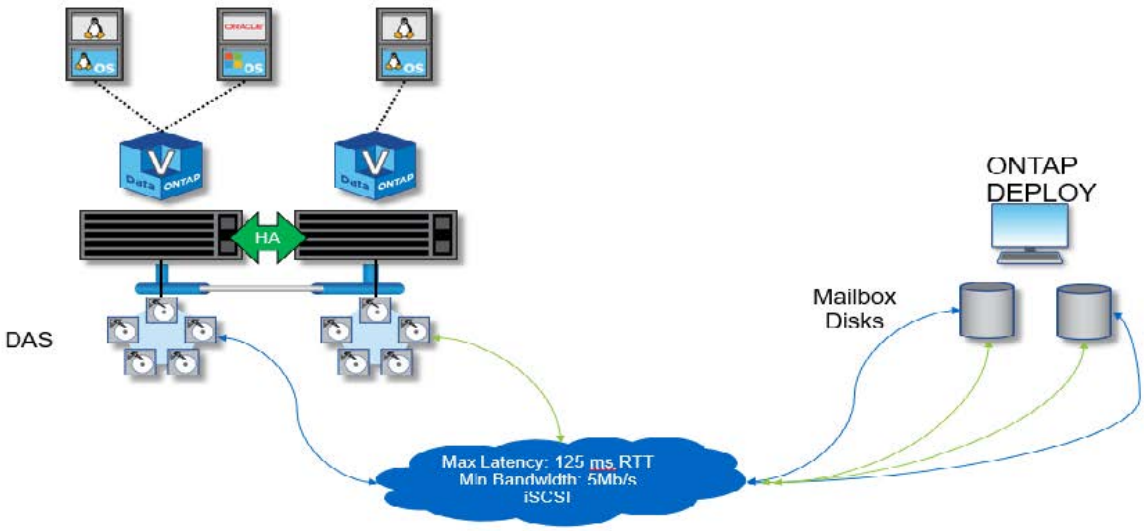
Gran parte del mercato SDS si basa sul concetto di storage shared-nothing, con la replica software che fornisce resilienza dei dati mediante l'archiviazione di più copie dei dati utente in diversi silos di storage. ONTAP Select si basa su questa premessa utilizzando le funzionalità di replica sincrona (RAID SyncMirror) fornite da ONTAP per memorizzare una copia aggiuntiva dei dati utente all'interno del cluster. Ciò si verifica nel contesto di una coppia ha. Ogni coppia ha memorizza due copie dei dati utente: Una sullo storage fornito dal nodo locale e una sullo storage fornito dal partner ha. All'interno di un cluster ONTAP Select, la replica sincrona e ha sono legate tra loro e le funzionalità dei due non possono essere disaccoppiate o utilizzate in modo indipendente. Di conseguenza, la funzionalità di replica sincrona è disponibile solo nell'offerta Multinode.

> In un cluster ONTAP Select, la funzionalità di replica sincrona è una funzione dell'implementazione ha, non una sostituzione dei motori di replica asincroni SnapMirror o SnapVault. La replica sincrona non può essere utilizzata indipendentemente da ha.

Esistono due modelli di implementazione di ONTAP Select ha: I cluster a più nodi (quattro, sei o otto nodi) e i cluster a due nodi. La caratteristica principale di un cluster ONTAP Select a due nodi è l'utilizzo di un servizio di mediatore esterno per risolvere gli scenari split-brain. La macchina virtuale ONTAP Deploy funge da mediatore predefinito per tutte le coppie ha a due nodi configurate.

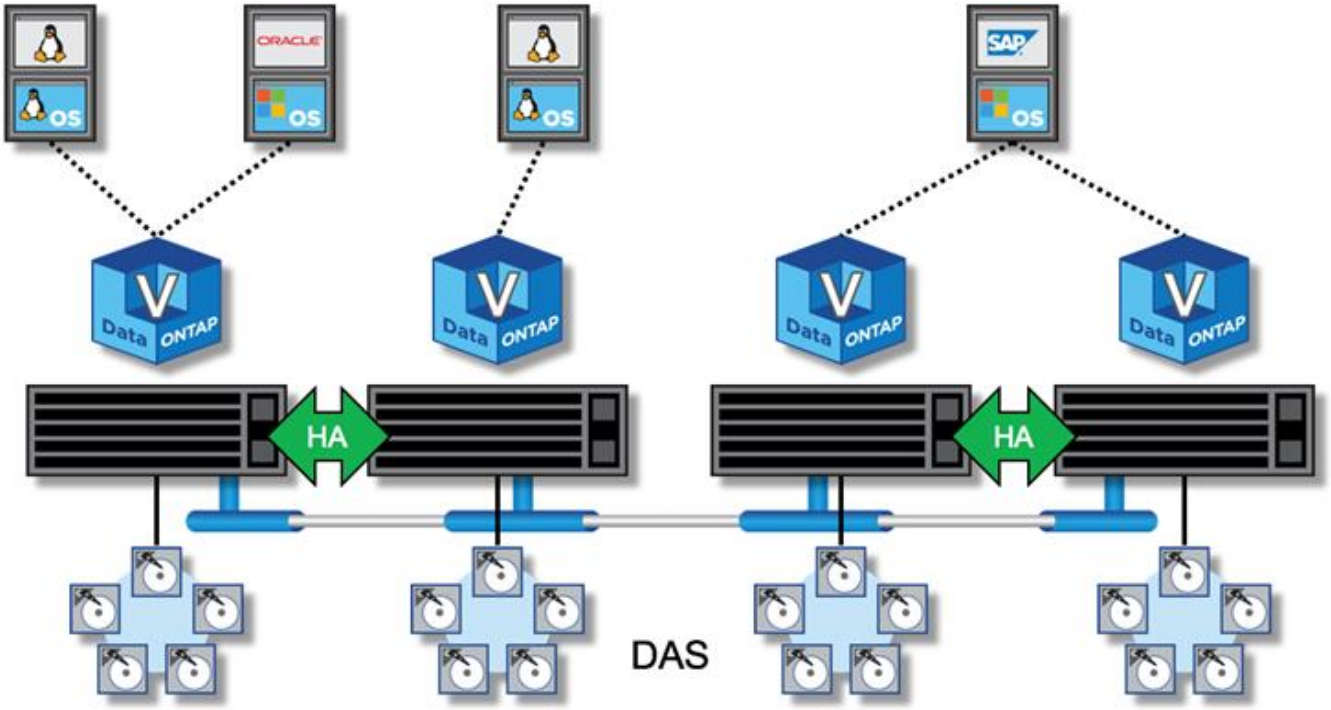
Le due architetture sono rappresentate nelle seguenti figure.

Cluster ONTAP Select a due nodi con mediatore remoto e storage locale



Il cluster ONTAP Select a due nodi è composto da una coppia ha e da un mediatore. All'interno della coppia ha, gli aggregati di dati su ciascun nodo del cluster vengono sottoposti a mirroring sincrono e, in caso di failover, non si verifica alcuna perdita di dati.

Cluster ONTAP Select a quattro nodi con storage locale



- Il cluster ONTAP Select a quattro nodi è composto da due coppie ha. I cluster a sei e otto nodi sono composti rispettivamente da tre e quattro coppie ha. All'interno di ciascuna coppia ha, gli aggregati di dati su ciascun nodo del cluster vengono sottoposti a mirroring sincrono e, in caso di failover, non si verifica alcuna perdita di dati.
- Quando si utilizza lo storage DAS, su un server fisico può essere presente una sola istanza di ONTAP Select. ONTAP Select richiede un accesso non condiviso al controller RAID locale del sistema ed è progettato per gestire i dischi collegati in locale, il che sarebbe impossibile senza la connettività fisica allo storage.

Ha a due nodi rispetto a ha a più nodi

A differenza degli array FAS, i nodi ONTAP Select di una coppia ha comunicano esclusivamente sulla rete IP. Ciò significa che la rete IP è un singolo punto di errore (SPOF) e la protezione da partizioni di rete e scenari di split-brain diventa un aspetto importante della progettazione. Il cluster a più nodi è in grado di sostenere guasti a nodo singolo perché il quorum del cluster può essere stabilito dai tre o più nodi sopravvissuti. Per ottenere lo stesso risultato, il cluster a due nodi si affida al servizio mediatore ospitato dalla VM di implementazione di ONTAP.

Il traffico di rete heartbeat tra i nodi ONTAP Select e il servizio ONTAP Deploy mediator è minimo e resiliente, in modo che la VM ONTAP Deploy possa essere ospitata in un data center diverso dal cluster a due nodi ONTAP Select.

> La macchina virtuale ONTAP Deploy diventa parte integrante di un cluster a due nodi quando funge da mediatore per quel cluster. Se il servizio mediatore non è disponibile, il cluster a due nodi continua a servire i dati, ma le funzionalità di failover dello storage del cluster ONTAP Select sono disattivate. Pertanto, il servizio ONTAP Deploy mediator deve mantenere una comunicazione costante con ciascun nodo ONTAP Select nella coppia ha. Una larghezza di banda minima di 5 Mbps e una latenza massima di andata e ritorno (RTT) di 125 ms sono necessarie per consentire il corretto funzionamento del quorum del cluster.

Se la macchina virtuale ONTAP Deploy che funge da mediatore non è temporaneamente o potenzialmente permanentemente disponibile, è possibile utilizzare una macchina virtuale ONTAP Deploy secondaria per ripristinare il quorum del cluster a due nodi. Ciò comporta una configurazione in cui la nuova macchina virtuale ONTAP Deploy non è in grado di gestire i nodi ONTAP Select, ma partecipa con successo all'algoritmo del quorum del cluster. La comunicazione tra i nodi ONTAP Select e la VM di implementazione ONTAP viene eseguita utilizzando il protocollo iSCSI su IPv4. L'indirizzo IP di gestione del nodo ONTAP Select è l'iniziatore e l'indirizzo IP della macchina virtuale di implementazione ONTAP è la destinazione. Pertanto, non è possibile supportare gli indirizzi IPv6 per gli indirizzi IP di gestione dei nodi quando si crea un cluster a due nodi. I dischi delle cassette postali ospitati ONTAP Deploy vengono creati automaticamente e mascherati agli indirizzi IP di gestione dei nodi ONTAP Select corretti al momento della creazione del cluster a due nodi. L'intera configurazione viene eseguita automaticamente durante l'installazione e non sono necessarie ulteriori azioni amministrative. L'istanza di implementazione di ONTAP che crea il cluster è il mediatore predefinito per quel cluster.

Se è necessario modificare la posizione del mediatore originale, è necessario eseguire un'azione amministrativa. È possibile ripristinare un quorum del cluster anche in caso di perdita della VM di implementazione ONTAP originale. Tuttavia, NetApp consiglia di eseguire il backup del database di implementazione ONTAP dopo ogni creazione di un'istanza del cluster a due nodi.

Ha a due nodi rispetto a ha a due nodi allungato (SDS MetroCluster)

È possibile estendere un cluster ha attivo/attivo a due nodi su distanze maggiori e potenzialmente posizionare ciascun nodo in un data center diverso. L'unica distinzione tra un cluster a due nodi e un cluster a due nodi (noto anche come SDS MetroCluster) è la distanza di connettività di rete tra i nodi.

Il cluster a due nodi è definito come un cluster per il quale entrambi i nodi si trovano nello stesso data center entro una distanza di 300 m. In generale, entrambi i nodi hanno uplink verso lo stesso switch di rete o insieme di switch di rete ISL (Interswitch link).

Per SDS MetroCluster a due nodi si intende un cluster per il quale i nodi sono fisicamente separati (stanze diverse, edifici diversi e data center diversi) di oltre 300 m. Inoltre, le connessioni uplink di ciascun nodo sono collegate a switch di rete separati. L'SDS MetroCluster non richiede hardware dedicato. Tuttavia, l'ambiente deve rispettare i requisiti di latenza (massimo 5 ms per RTT e 5 ms per jitter, per un totale di 10 ms) e distanza fisica (massimo 10 km).

MetroCluster SDS è una funzione premium e richiede una licenza Premium o Premium XL. La licenza Premium supporta la creazione di macchine virtuali di piccole e medie dimensioni, oltre a supporti HDD e SSD. La licenza Premium XL supporta anche la creazione di dischi NVMe.

> MetroCluster SDS è supportato sia con DAS (Local Attached Storage) che con vNAS (Shared Storage). Si noti che le configurazioni vNAS hanno solitamente una latenza innata più elevata a causa della rete tra la VM ONTAP Select e lo storage condiviso. Le configurazioni SDS di MetroCluster devono fornire un massimo di 10 ms di latenza tra i nodi, inclusa la latenza dello storage condiviso. In altre parole, solo la misurazione della latenza tra le Select VM non è adeguata, perché la latenza dello storage condiviso non è trascurabile per queste configurazioni.

HA RSM e aggregati mirrorati

Prevenzione della perdita di dati con RAID SyncMirror (RSM), aggregati mirrorati e percorso di scrittura.

Replica sincrona

Il modello ONTAP ha si basa sul concetto di partner ha. ONTAP Select estende questa architettura nel mondo dei server commodity non condivisi utilizzando la funzionalità RAID SyncMirror (RSM) presente in ONTAP per replicare i blocchi di dati tra i nodi del cluster, fornendo due copie dei dati utente distribuiti su una coppia ha.

Un cluster a due nodi con un mediatore può comprendere due data center. Per ulteriori informazioni, vedere la sezione "Best practice di ha (MetroCluster SDS) con due nodi estesi".

Aggregati mirrorati

Un cluster ONTAP Select è composto da due a otto nodi. Ogni coppia ha contiene due copie dei dati utente, con mirroring sincrono tra i nodi su una rete IP. Questo mirroring è trasparente per l'utente ed è una proprietà dell'aggregato di dati, configurato automaticamente durante il processo di creazione dell'aggregato di dati.

Tutti gli aggregati di un cluster ONTAP Select devono essere sottoposti a mirroring per la disponibilità dei dati in caso di failover di un nodo e per evitare un SPOF in caso di guasto hardware. Gli aggregati in un cluster ONTAP Select sono costruiti dai dischi virtuali forniti da ciascun nodo della coppia ha e utilizzano i seguenti dischi:

- Un set locale di dischi (fornito dal nodo ONTAP Select corrente)
- Un set di dischi mirrorati (contributo del partner ha del nodo corrente)

> I dischi locali e mirror utilizzati per creare un aggregato mirrorato devono avere le stesse dimensioni. Questi aggregati sono indicati come plesso 0 e plesso 1 (per indicare rispettivamente le coppie di mirror locale e remoto). I numeri effettivi del plex possono essere diversi nell'installazione.

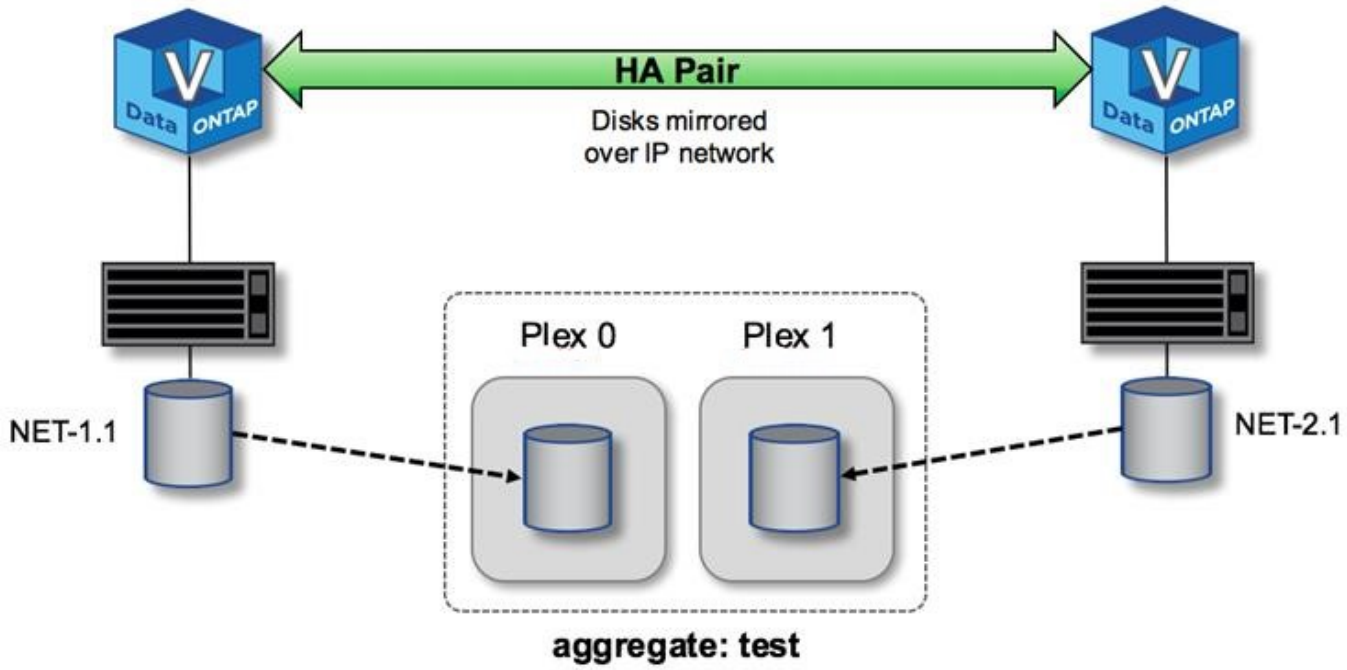
Questo approccio è fondamentalmente diverso dal modo in cui funzionano i cluster ONTAP standard. Questo vale per tutti i dischi root e dei dati all'interno del cluster ONTAP Select. L'aggregato contiene copie dei dati sia locali che mirror. Pertanto, un aggregato che contiene N dischi virtuali offre N/2 dischi di storage unico, perché la seconda copia dei dati risiede su dischi univoci.

La figura seguente mostra una coppia ha all'interno di un cluster ONTAP Select a quattro nodi. All'interno di questo cluster è presente un singolo aggregato (test) che utilizza lo storage di entrambi i partner ha. Questo aggregato di dati è composto da due set di dischi virtuali: Un set locale, fornito dal nodo del cluster proprietario di ONTAP Select (Plex 0), e un set remoto, fornito dal partner di failover (Plex 1).

Plex 0 è il bucket che contiene tutti i dischi locali. Plex 1 è il bucket che contiene dischi mirror o dischi responsabili della memorizzazione di una seconda copia replicata dei dati dell'utente. Il nodo proprietario dell'aggregato contribuisce ai dischi Plex 0 e il partner ha di tale nodo contribuisce ai dischi Plex 1.

Nella figura seguente è presente un aggregato mirrorato con due dischi. Il contenuto di questo aggregato viene mirrorato attraverso i nostri due nodi del cluster, con il disco locale NET-1.1 inserito nel bucket Plex 0 e il disco remoto NET-2.1 inserito nel bucket Plex 1. In questo esempio, il test aggregato è di proprietà del nodo del cluster a sinistra e utilizza il disco locale NET-1.1 e il disco mirror del partner ha NET-2.1.

Aggregato mirrorato ONTAP Select



Quando viene implementato un cluster ONTAP Select, tutti i dischi virtuali presenti nel sistema vengono assegnati automaticamente al plesso corretto, senza richiedere alcun passo aggiuntivo da parte dell'utente in merito all'assegnazione del disco. In questo modo si evita l'assegnazione accidentale dei dischi a un plesso non corretto e si ottiene una configurazione ottimale del disco mirror.

Percorso di scrittura

Il mirroring sincrono dei blocchi di dati tra i nodi del cluster e la necessità di non perdere dati in caso di guasto del sistema hanno un impatto significativo sul percorso richiesto da una scrittura in entrata durante la propagazione attraverso un cluster ONTAP Select. Questo processo si compone di due fasi:

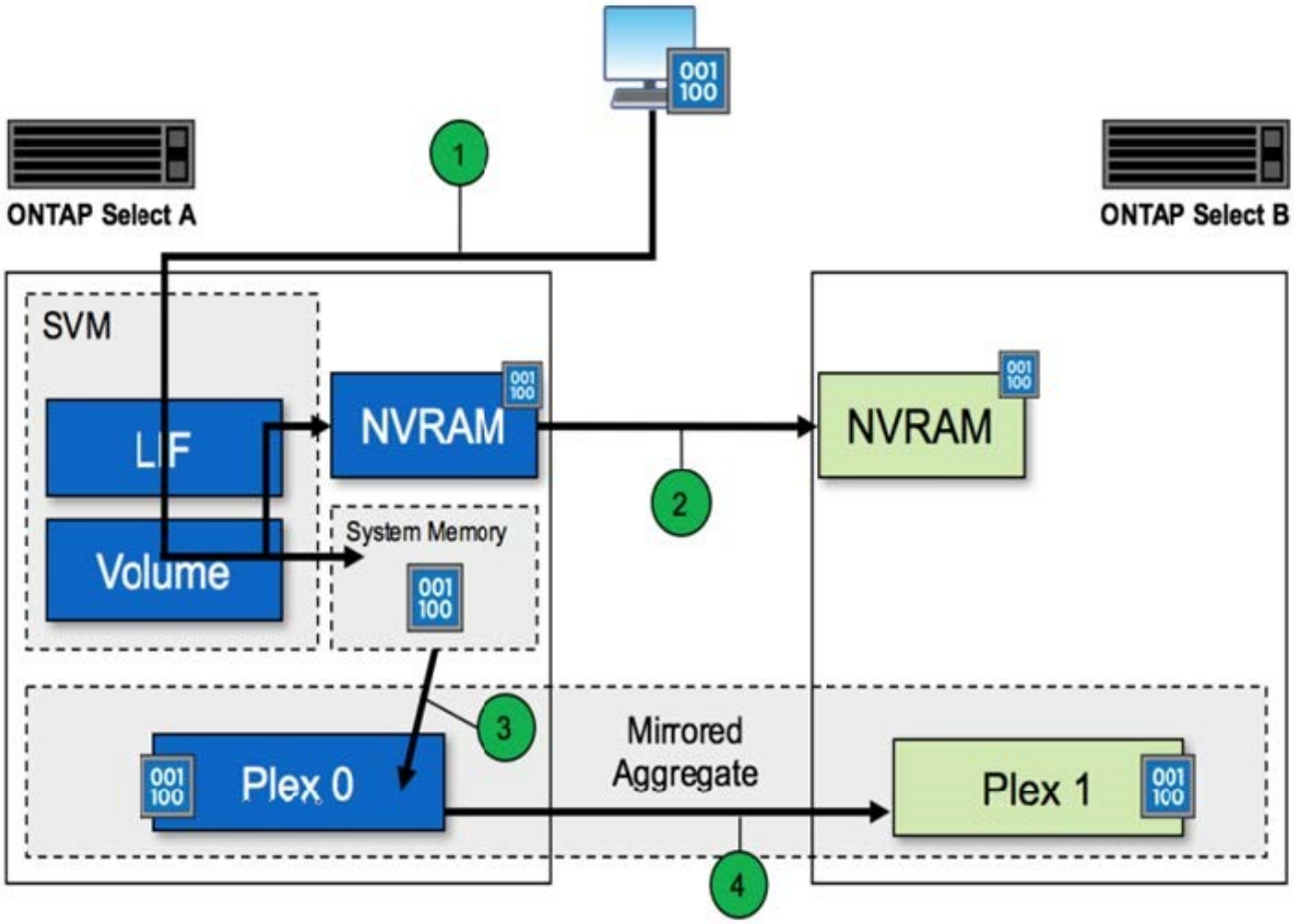
- Riconoscimento
- Destaging

Le scritture su un volume di destinazione avvengono su una LIF di dati e vengono assegnate alla partizione NVRAM virtualizzata, presente su un disco di sistema del nodo ONTAP Select, prima di essere riconnessi al client. In una configurazione ha, si verifica un'ulteriore fase, in quanto queste scritture NVRAM vengono immediatamente replicate al partner ha del proprietario del volume di destinazione prima di essere riconosciute. Questo processo garantisce la coerenza del file system nel nodo partner ha, in caso di guasto hardware nel nodo originale.

Una volta che la scrittura è stata impegnata nella NVRAM, ONTAP sposta periodicamente il contenuto di questa partizione sul disco virtuale appropriato, un processo noto come destaging. Questo processo si verifica una sola volta, sul nodo del cluster che possiede il volume di destinazione, e non si verifica sul partner ha.

La figura seguente mostra il percorso di scrittura di una richiesta di scrittura in entrata a un nodo ONTAP Select.

Workflow del percorso di scrittura ONTAP Select



La conferma di scrittura in entrata include i seguenti passaggi:

- Le scritture entrano nel sistema attraverso un'interfaccia logica di proprietà del nodo ONTAP Select A.
- Le scritture vengono assegnate alla NVRAM del nodo A e mirrorate al partner ha, il nodo B.
- Una volta che la richiesta di i/o è presente su entrambi i nodi ha, la richiesta viene quindi riconnessa al client.

Il destaging ONTAP Select dalla NVRAM all'aggregato di dati (ONTAP CP) include i seguenti passaggi:

- Le scritture vengono trasferite dalla NVRAM virtuale all'aggregato di dati virtuale.
- Il motore mirror replica in modo sincrono i blocchi su entrambi i plessi.

HA ulteriori dettagli

Heartbeat del disco HA, mailbox ha, heartbeat ha, failover ha e Giveback lavorano per migliorare la protezione dei dati.

Heartbeat dei dischi

Sebbene l'architettura ONTAP Select ha utilizzi molti dei percorsi di codice utilizzati dagli array FAS tradizionali, esistono alcune eccezioni. Una di queste eccezioni riguarda l'implementazione del heartbeat basato su disco, un metodo di comunicazione non basato sulla rete utilizzato dai nodi del cluster per impedire che l'isolamento della rete causi un comportamento split-brain. Uno scenario split-brain è il risultato del

partizionamento del cluster, in genere causato da errori di rete, per cui ciascun lato ritiene che l'altro sia inattivo e tenta di assumere il controllo delle risorse del cluster.

Le implementazioni di ha di livello Enterprise devono gestire questo tipo di scenario in modo corretto. ONTAP esegue questa operazione attraverso un metodo di heartbeat personalizzato e basato su disco. Questo è il lavoro della mailbox ha, una posizione sullo storage fisico che viene utilizzata dai nodi del cluster per passare i messaggi heartbeat. In questo modo, il cluster può determinare la connettività e definire il quorum in caso di failover.

Negli array FAS, che utilizzano un'architettura di storage ha condivisa, ONTAP risolve i problemi di split-brain nei seguenti modi:

- Riserve persistenti SCSI
- Metadati ha persistenti
- Stato HA inviato tramite interconnessione ha

Tuttavia, all'interno dell'architettura shared-nothing di un cluster ONTAP Select, un nodo è in grado di vedere solo il proprio storage locale e non quello del partner ha. Pertanto, quando la partizione di rete isola ciascun lato di una coppia ha, i metodi precedenti per determinare il quorum del cluster e il comportamento di failover non sono disponibili.

Sebbene non sia possibile utilizzare il metodo esistente di rilevamento ed evasione del cervello diviso, è comunque necessario un metodo di mediazione che rientri nei limiti di un ambiente senza condivisione. ONTAP Select estende ulteriormente l'infrastruttura di caselle postali esistente, consentendo all'IT di agire come metodo di mediazione in caso di partizione della rete. Poiché lo storage condiviso non è disponibile, la mediazione viene eseguita attraverso l'accesso ai dischi della mailbox su NAS. Questi dischi sono distribuiti in tutto il cluster, incluso il mediatore in un cluster a due nodi, utilizzando il protocollo iSCSI. Pertanto, un nodo del cluster può prendere decisioni di failover intelligenti in base all'accesso a questi dischi. Se un nodo può accedere ai dischi della mailbox di altri nodi al di fuori del proprio partner ha, è probabile che sia funzionante.

L'architettura della mailbox e il metodo di heartbeat basato su disco per risolvere i problemi di quorum del cluster e di split-brain sono i motivi per cui la variante multinodo di ONTAP Select richiede quattro nodi separati o un mediatore per un cluster a due nodi.

Pubblicazione della mailbox HA

L'architettura della cassetta postale ha utilizza un modello di message post. A intervalli ripetuti, i nodi del cluster pubblicano messaggi su tutti gli altri dischi della mailbox nel cluster, incluso il mediatore, indicando che il nodo è attivo e in esecuzione. All'interno di un cluster integro in qualsiasi momento, un singolo disco della mailbox su un nodo del cluster contiene messaggi inviati da tutti gli altri nodi del cluster.

A ciascun nodo del cluster Select è collegato un disco virtuale che viene utilizzato specificamente per l'accesso alla mailbox condivisa. Questo disco viene chiamato disco della mailbox del mediatore, perché la sua funzione principale è quella di agire come metodo di mediazione del cluster in caso di guasti al nodo o partizione di rete. Questo disco della mailbox contiene partizioni per ciascun nodo del cluster ed è montato su una rete iSCSI da altri nodi del cluster Select. Periodicamente, questi nodi pubblicano gli stati di integrità nella partizione appropriata del disco della mailbox. L'utilizzo di dischi di mailbox accessibili dalla rete distribuiti in tutto il cluster consente di dedurre lo stato dei nodi attraverso una matrice di raggiungibilità. Ad esempio, i nodi cluster A e B possono inviare alla mailbox del nodo cluster D, ma non alla mailbox del nodo C. Inoltre, il nodo del cluster D non può inviare alla mailbox del nodo C, quindi è probabile che il nodo C sia inattivo o isolato in rete e debba essere sostituito.

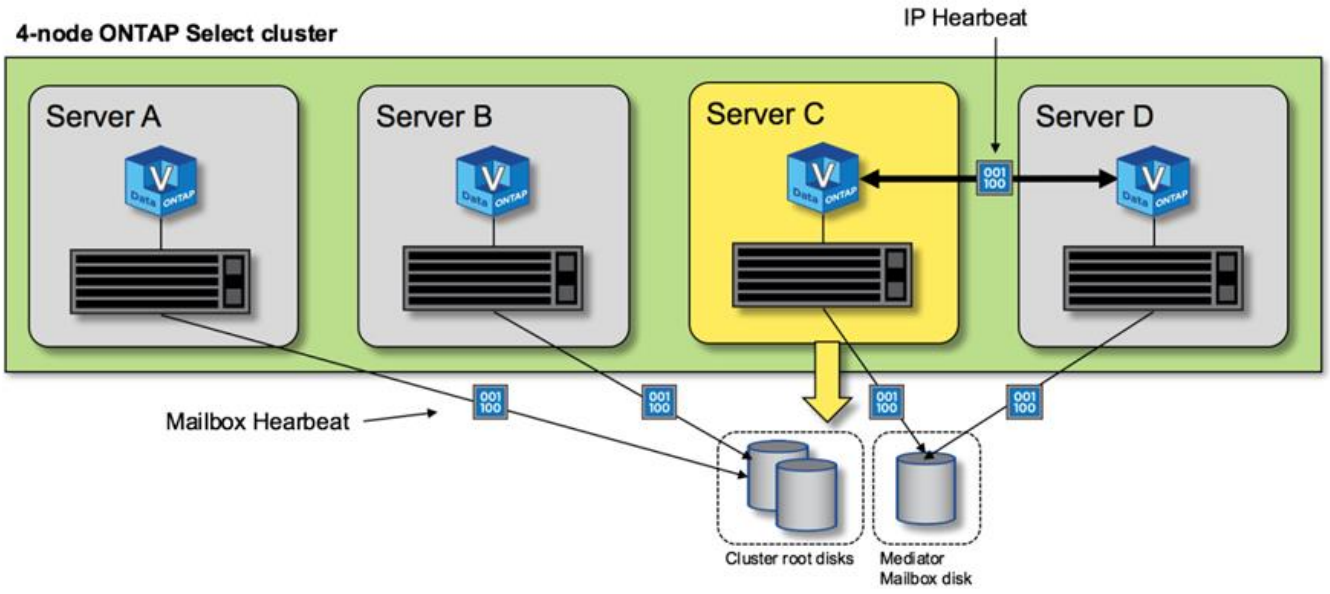
HA heartbeat

Come per le piattaforme NetApp FAS, ONTAP Select invia periodicamente messaggi heartbeat ha tramite l'interconnessione ha. All'interno del cluster ONTAP Select, questa operazione viene eseguita tramite una connessione di rete TCP/IP esistente tra i partner ha. Inoltre, i messaggi heartbeat basati su disco vengono passati a tutti i dischi di ha mailbox, inclusi i dischi di mediatore mailbox. Questi messaggi vengono trasmessi ogni pochi secondi e letti periodicamente. La frequenza con cui questi vengono inviati e ricevuti consente al cluster ONTAP Select di rilevare gli eventi di guasto ha entro circa 15 secondi, la stessa finestra disponibile sulle piattaforme FAS. Quando i messaggi heartbeat non vengono più letti, viene attivato un evento di failover.

La figura seguente mostra il processo di invio e ricezione di messaggi heartbeat sui dischi di interconnessione ha e mediatore dal punto di vista di un singolo nodo del cluster ONTAP Select, il nodo C.

> Gli heartbeat di rete vengono inviati tramite l'interconnessione ha al partner ha, nodo D, mentre gli heartbeat dei dischi utilizzano i dischi delle cassette postali in tutti i nodi del cluster, A, B, C e D.

Heartbeat ha in un cluster a quattro nodi: Stato stazionario



Failover E giveback HA

Durante un'operazione di failover, il nodo sopravvissuto assume le responsabilità di servizio dei dati per il nodo peer utilizzando la copia locale dei dati del partner ha. L'i/o del client può continuare senza interruzioni, ma le modifiche a questi dati devono essere replicate prima che possa verificarsi il giveback. Si noti che ONTAP Select non supporta un giveback forzato perché ciò causa la perdita delle modifiche memorizzate nel nodo sopravvissuto.

L'operazione di sincronizzazione viene attivata automaticamente quando il nodo riavviato si ricongiunge al cluster. Il tempo necessario per la sincronizzazione dipende da diversi fattori. Questi fattori includono il numero di modifiche da replicare, la latenza di rete tra i nodi e la velocità dei sottosistemi dei dischi su ciascun nodo. È possibile che il tempo necessario per la sincronizzazione superi la finestra di ritorno automatico di 10 minuti. In questo caso, è necessario un giveback manuale dopo la sincronizzazione. È possibile monitorare l'avanzamento della sincronizzazione utilizzando il seguente comando:

```
storage aggregate status -r -aggregate <aggregate name>
```

Performance

Performance

Le performance variano in base alla configurazione hardware.

Le prestazioni di un cluster ONTAP Select possono variare notevolmente a causa delle caratteristiche dell'hardware e della configurazione sottostante. La configurazione hardware specifica è il fattore più importante per le performance di una particolare istanza di ONTAP Select. Di seguito sono riportati alcuni fattori che influiscono sulle prestazioni di una specifica istanza di ONTAP Select:

- **Frequenza core**. In generale, è preferibile una frequenza maggiore.
- **Socket singolo rispetto a socket multiplo**. ONTAP Select non utilizza funzionalità multi-socket, ma l'overhead dell'hypervisor per il supporto delle configurazioni multi-socket rappresenta una certa deviazione nelle performance totali.
- **Configurazione della scheda RAID e driver dell'hypervisor associato**. Il driver predefinito fornito dall'hypervisor potrebbe dover essere sostituito dal driver del fornitore dell'hardware.
- **Tipo di disco e numero di dischi nei gruppi RAID**.
- **Versione dell'hypervisor e livello di patch**.

Performance: Storage SSD direct-attached ha Premium

Informazioni sulle performance per la piattaforma di riferimento.

Piattaforma di riferimento

Hardware ONTAP Select (Premium XL) (per nodo)

- FUJITSU PRIMERGY RX2540 M4:
  - CPU Intel® Xeon® Gold 6142b a 2.6 GHz
  - 32 core fisici (16 x 2 socket), 64 logici
  - 256 GB DI RAM
  - Unità per host: 24 SSD da 960 GB
  - ESX 6.5U1

Hardware del client

- 5 client NFSv3 IBM 3550m4

Informazioni di configurazione

- SW RAID 1 x 9 + 2 RAID-DP (11 dischi)
- 22+1 RAID-5 (RAID-0 in ONTAP) / NVRAM cache RAID
- Nessuna funzionalità di efficienza dello storage in uso (compressione, deduplica, copie Snapshot,

SnapMirror e così via)

La seguente tabella elenca il throughput misurato rispetto ai carichi di lavoro di lettura/scrittura su una coppia di nodi ONTAP Select ad alta disponibilità (ha) che utilizzano sia RAID software che RAID hardware. Le misurazioni delle performance sono state effettuate utilizzando il tool per la generazione del carico SIO.

Questi dati sulle performance si basano su ONTAP Select 9.6.

Risultati delle prestazioni per un cluster ONTAP Select a nodo singolo (parte di un'istanza di supporto a quattro nodi) su un SSD DAS (Direct-Attached Storage), con RAID software e RAID hardware

Descrizione	Lettura sequenziale 64 KiB	Scrittura sequenziale 64 KiB	Lettura casuale 8 KiB	Scrittura casuale 8 KiB	WR/ RD random (50/50) 8 KiB
ONTAP Select Large instance con RAID software DAS (SSD)	2171 MiBps	559 MIBps	954 MIBps	394 MIBps	564 MIBps
Media istanza di ONTAP Select con RAID software DAS (SSD)	2090 MIBps	592 MIBps	677 MIBps	335 MIBps	441 3 MiBps
Media istanza di ONTAP Select con RAID hardware DAS (SSD)	2038 MIBps	520 MIBps	578 MIBps	325 MIBps	399 MIBps

64 K di lettura sequenziale

Dettagli:

- I/o diretto SIO attivato
- 2 nodi
- 2 NIC dati per nodo
- 1 aggregato di dati per nodo (RAID hardware da 2 TB), (RAID software da 8 TB)
- 64 processi SIO, 1 thread per processo
- 32 volumi per nodo
- 1 x file per processo; i file sono da 12000MB ciascuno

64 K di scrittura sequenziale

Dettagli:

- I/o diretto SIO attivato
- 2 nodi

- 2 schede di interfaccia di rete dati (NIC) per nodo
- 1 aggregato di dati per nodo (RAID hardware da 2 TB), (RAID software da 4 TB)
- 128 SIO procs, 1 thread per proc
- Volumi per nodo: 32 (RAID hardware), 16 (RAID software)
- 1 x file per processo; i file sono da 30720 MB ciascuno

8K di lettura casuale

Dettagli:

- I/o diretto SIO attivato
- 2 nodi
- 2 schede di rete dati per nodo
- 1 aggregato di dati per nodo (RAID hardware da 2 TB), (RAID software da 4 TB)
- 64 processi SIO, 8 thread per processo
- Volumi per nodo: 32
- 1 x file per processo; i file sono da 12228 MB ciascuno

8K di scrittura casuale

Dettagli:

- I/o diretto SIO attivato
- 2 nodi
- 2 schede di rete dati per nodo
- 1 aggregato di dati per nodo (RAID hardware da 2 TB), (RAID software da 4 TB)
- 64 processi SIO, 8 thread per processo
- Volumi per nodo: 32
- 1 x file per processo; i file sono 8192 MB ciascuno

8K random 50% write 50% Read

Dettagli:

- I/o diretto SIO attivato
- 2 nodi
- 2 schede di rete dati per nodo
- 1 aggregato di dati per nodo (RAID hardware da 2 TB), (RAID software da 4 TB)
- 64 thread SIO proc208 per processo
- Volumi per nodo: 32
- 1 x file per processo; i file sono da 12228 MB ciascuno

Informazioni sul copyright

Copyright © 2026 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEQUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina http://www.netapp.com/TM sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.