深入探討
ONTAP Select

NetApp
April 01, 2025

This PDF was generated from https://docs.netapp.com/zh-tw/ontap-select-9121/concept_stor_concepts_chars.html on April 01, 2025. Always check docs.netapp.com for the latest.

目錄

深入探討 1
- 儲存設備 1
  - 儲存：一般概念與特性 1
  - ONTAP Select 本機附加儲存設備的硬體 RAID 服務 6
  - 本機附加儲存設備的軟體RAID服務 12
  - vSAN與外部陣列組態 20
  - 增加儲存容量 23
  - 儲存效率支援 26
- 網路 28
  - 網路：一般概念與特性 28
  - 單節點和多節點網路組態 30
  - 內部和外部網路ONTAP Select 35
  - 支援的網路組態 37
  - VMware vSphere vSwitch 組態 38
  - 實體交換器組態 47
  - 資料與管理流量分離 49
- 高可用度架構 51
  - 高可用度組態 51
  - HA RSM與鏡射Aggregate 53
  - HA其他詳細資料 56
- 效能 58
  - 效能 58
  - 效能：優質HA直接附加SSD儲存設備 59

深入探討

儲存設備

儲存：一般概念與特性

探索ONTAP Select 特定儲存元件之前、請先探索適用於整個環境的一般儲存概念。

儲存組態的各個階段

下列是支援整個系統的主要組態階段ONTAP Select ：

- 部署前的必要條件
  - 請確定每部Hypervisor主機均已設定完成、並已準備好ONTAP Select 進行支援。
  - 組態包括實體磁碟機、RAID控制器和群組、LUN、以及相關的網路準備。
  - 此組態是在ONTAP Select 不執行此功能的情況下執行。
- 使用Hypervisor管理員公用程式進行組態設定
  - 您可以使用Hypervisor管理公用程式（例如VMware環境中的vSphere）來設定儲存設備的特定層面。
  - 此組態是在ONTAP Select 不執行此功能的情況下執行。
- 組態設定：ONTAP Select 使用「整合部署」管理公用程式
  - 您可以使用部署管理公用程式來設定核心邏輯儲存架構。
  - 這可以透過CLI命令明確執行、也可以由公用程式在部署過程中自動執行。
- 部署後組態
  - 完成靜態部署之後、您可以使用CLI或System Manager來設定叢集。ONTAP Select ONTAP
  - 此組態是在ONTAP Select 不執行「整合」的情況下執行。

託管儲存設備與非託管儲存設備

由功能性資料管理的儲存設備ONTAP Select 可直接由功能性資料管理系統存取及控制。同一Hypervisor主機上的任何其他儲存設備均為未受管理的儲存設備。

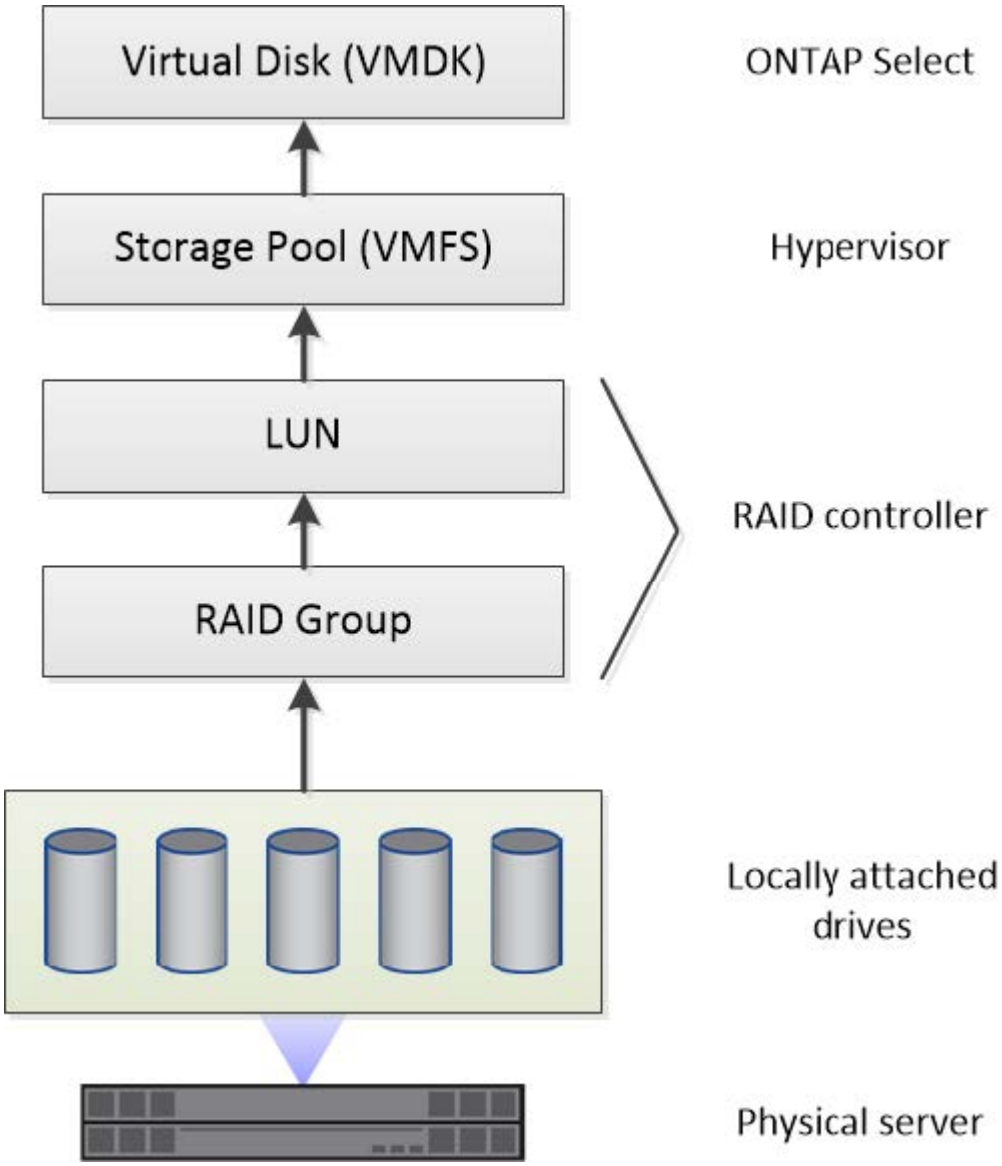
同質實體儲存設備

所有組成ONTAP Select 「不受影響的」託管儲存設備的實體磁碟機都必須是同質的。也就是所有硬體必須與下列特性相同：

- 類型（SAS、NL-SAS、SATA、SSD）
- 速度（RPM）

本機儲存環境的圖例

每個Hypervisor主機都包含本機磁碟和ONTAP Select 其他可由支援的邏輯儲存元件。這些儲存元件會從實體磁碟以分層結構排列。



本機儲存元件的特性

有幾個概念適用於ONTAP Select 在整個環境中使用的本機儲存元件。在準備ONTAP Select 進行一套不一致的部署之前、您應該先熟悉這些概念。這些概念是根據類別來排列：RAID群組和LUN、儲存資源池和虛擬磁碟。

將實體磁碟機分組為**RAID**群組和**LUN**

一或多個實體磁碟可在本機連接至主機伺服器、並可供ONTAP Select 使用。實體磁碟會指派給RAID群組、然後以一或多個LUN的形式呈現給Hypervisor主機作業系統。每個LUN會以實體硬碟的形式呈現給Hypervisor主機作業系統。

設定ONTAP Select 支援對象時、您應注意下列事項：

- 所有託管儲存設備都必須透過單一RAID控制器存取
- 視廠商而定、每個RAID控制器支援每個RAID群組的磁碟機數量上限

一個或多個**RAID**群組

每ONTAP Select 個支援的主機都必須有一個RAID控制器。您應該建立單一RAID群組ONTAP Select 以供使用。不過、在某些情況下、您可能會考慮建立多個RAID群組。請參閱 "最佳實務做法摘要"。

儲存資源池考量

在準備部署ONTAP Select 時、您應該注意幾個與儲存資源池有關的問題。

在VMware環境中、儲存資源池與VMware資料存放區同義。

儲存資源池和**LUN**

每個LUN在Hypervisor主機上都被視為本機磁碟、而且可以是一個儲存資源池的一部分。每個儲存資源池的格式都是以Hypervisor主機作業系統可以使用的檔案系統來設定。

您必須確定儲存資源池是在ONTAP Select 進行功能性部署時正確建立的。您可以使用Hypervisor管理工具建立儲存資源池。例如、使用VMware、您可以使用vSphere用戶端建立儲存資源池。然後、儲存資源池會傳遞到ONTAP Select 這個功能完善的系統管理公用程式。

管理虛擬磁碟

在準備部署ONTAP Select 時、您應該注意幾個與虛擬磁碟有關的問題。

虛擬磁碟與檔案系統

該支援將多個虛擬磁碟機分配給該虛擬ONTAP Select 機器。每個虛擬磁碟實際上都是儲存資源池中包含的檔案、並由Hypervisor維護。有多種類型的磁碟可供ONTAP Select 使用、主要是系統磁碟和資料磁碟。

您也應該瞭解下列關於虛擬磁碟的資訊：

- 在建立虛擬磁碟之前、儲存資源池必須可供使用。
- 無法在建立虛擬機器之前建立虛擬磁碟。
- 您必須仰賴ONTAP Select 這個資源管理公用程式來建立所有虛擬磁碟（也就是管理員絕不能在部署公用程式之外建立虛擬磁碟）。

設定虛擬磁碟

虛擬磁碟是ONTAP Select 由功能不完善的實體管理。當您使用部署管理公用程式建立叢集時、系統會自動建立這些程式。

外部儲存環境的圖例

利用此解決方案、支援使用位在Hypervisor主機外部儲存設備上的資料存放區。ONTAP Select ONTAP Select資料存放區可使用VMware vSAN透過網路存取、或直接從外部儲存陣列存取。

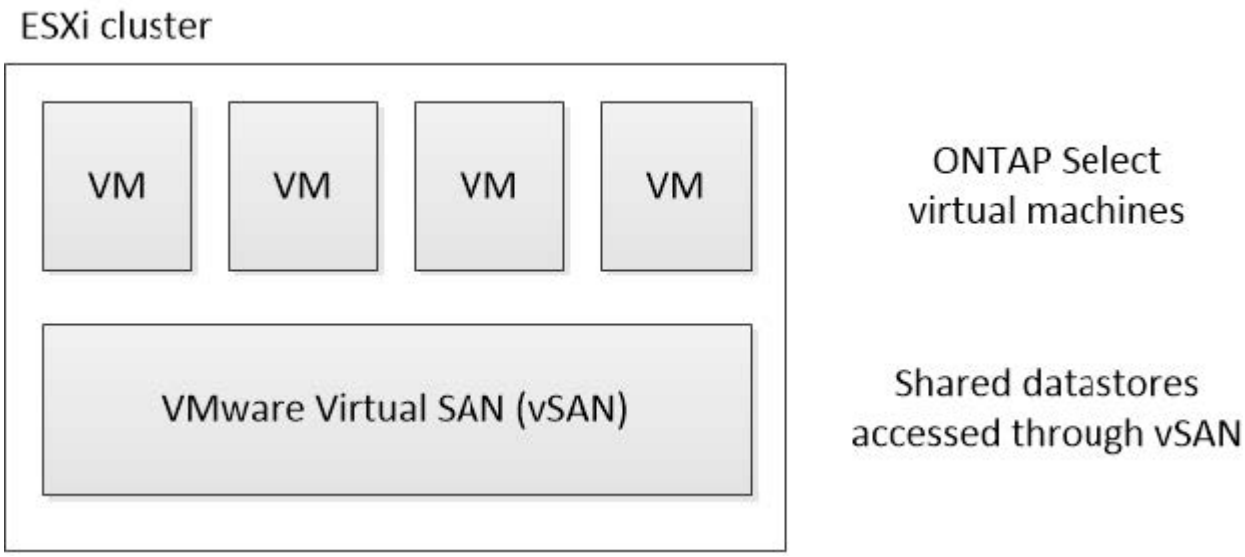
可設定使用Hypervisor主機外部的下列類型VMware ESXi網路資料存放區：ONTAP Select

- vSAN（虛擬SAN）
- VMFS

- NFS

vSAN資料存放區

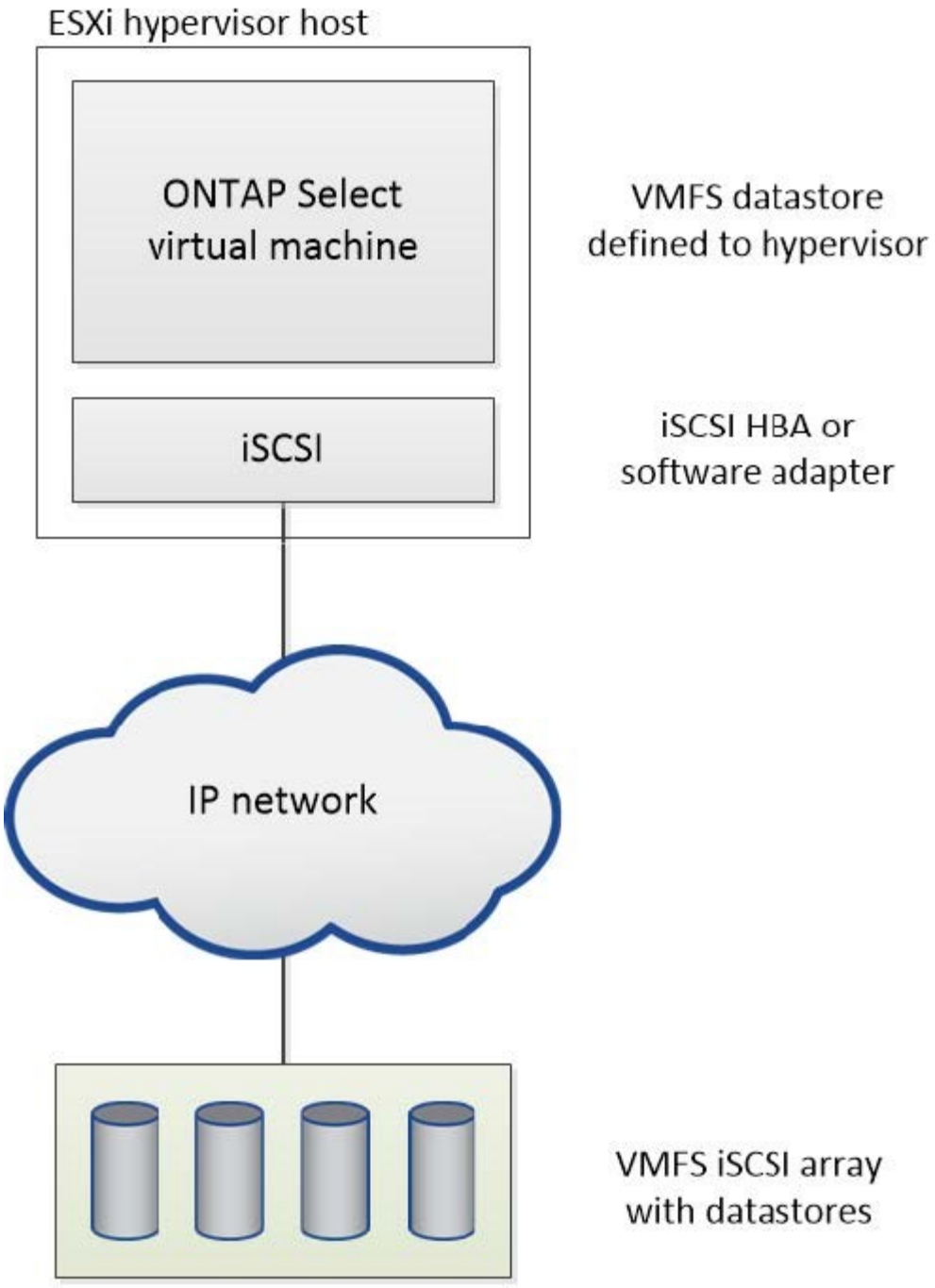
每個ESXi主機都可以擁有一個或多個本機VMFS資料存放區。通常這些資料存放區只能由本機主機存取。不過、VMware vSAN可讓ESXi叢集中的每個主機共用叢集中的所有資料存放區、就像它們是本機一樣。下圖說明vSAN如何建立ESXi叢集中主機之間共用的資料存放區集區。



外部儲存陣列上的**VMFS**資料存放區

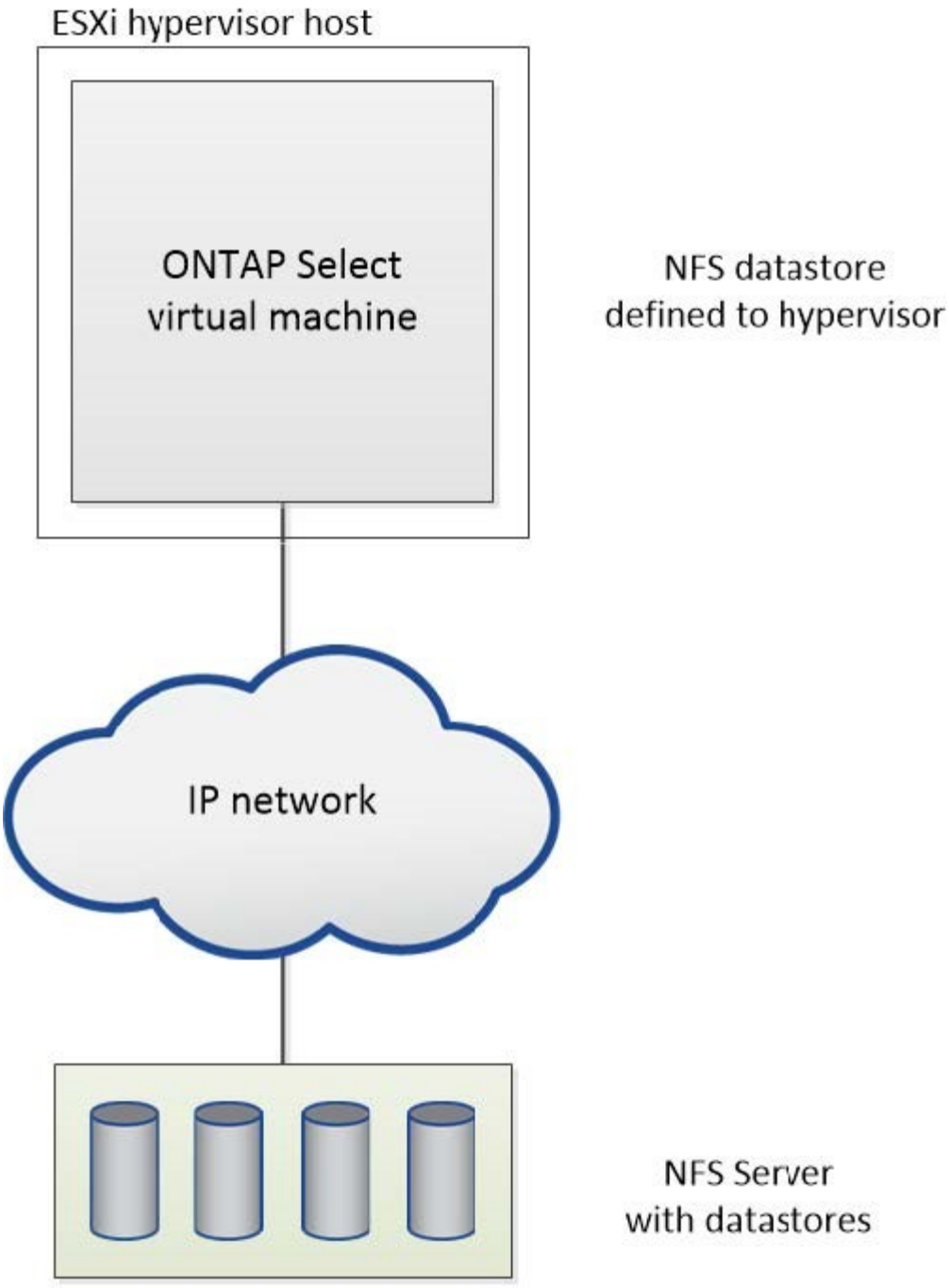
您可以建立位於外部儲存陣列上的VMFS資料存放區。儲存設備可使用多種不同的網路傳輸協定之一來存取。下圖說明使用iSCSI傳輸協定存取的外部儲存陣列上的VMFS資料存放區。

ONTAP Select 支援 VMware Storage/SAN 相容性文件中所述的所有外部儲存陣列、包括 iSCSI 、光纖通道和乙太網路光纖通道。



外部儲存陣列上的 **NFS** 資料存放區

您可以建立位於外部儲存陣列上的NFS資料存放區。使用NFS網路傳輸協定存取儲存設備。下圖說明透過NFS伺服器應用裝置存取的外部儲存設備上的NFS資料存放區。



ONTAP Select 本機附加儲存設備的硬體 RAID 服務

當有硬體RAID控制器可用ONTAP Select 時、即可將RAID服務移至硬體控制器、以提升寫入效能、並防止實體磁碟機故障。因此ONTAP Select 、針對整個叢集內所有節點的RAID保護是由本機附加的RAID控制器提供、而非ONTAP 透過支援此功能的軟體RAID提供。

由於實體RAID控制器為基礎磁碟機提供RAID分段功能、因此將支援使用RAID 0的實體資料集合體進行設定。ONTAP Select不支援其他RAID層級。

本機附加儲存設備的**RAID**控制器組態

所有提供ONTAP Select 支援儲存功能的本地附加磁碟、都必須位於RAID控制器後方。大部分的市售伺服器都有多種RAID控制器選項、可跨越多個價位、每個價位的功能層級各不相同。其目的是盡可能支援這些選項的數量、前提是這些選項符合控制器上的特定最低要求。

> 您無法從使用硬體 RAID 組態的 ONTAP Select VM 中分離虛擬磁碟。只有使用軟體 RAID 組態的 ONTAP Select VM 才支援分離磁碟。如需詳細資訊、請參閱 "在 ONTAP Select 軟體 RAID 組態中更換故障磁碟機" 。

管理ONTAP Select 此類磁碟的RAID控制器必須符合下列需求：

- 硬體RAID控制器必須具有電池備份單元（BBU/Flash備援寫入快取（FBWC）、並支援12Gbps的處理量。
- RAID控制器必須支援能夠承受至少一或兩個磁碟故障的模式（RAID 5和RAID 6）。
- 磁碟機快取必須設定為停用。
- 寫入原則必須設定為回寫模式、並在發生BBU/Flash故障時提供回寫功能。
- 讀取的I/O原則必須設定為快取。

所有提供ONTAP Select 支援儲存功能的本地附加磁碟、都必須放入執行RAID 5或RAID 6的RAID群組中。對於SAS磁碟機和SSD、使用最多24個磁碟機的RAID群組、ONTAP 讓NetApp能夠將傳入的讀取要求分散到更多的磁碟上。這樣做可大幅提升效能。使用SAS/SSD組態時、會針對單一LUN與多LUN組態進行效能測試。沒有發現顯著差異、因此為了簡化、NetApp建議您建立最少的LUN數量、以支援您的組態需求。

NL-SAS和SATA磁碟機需要一套不同的最佳實務做法。基於效能考量、磁碟的最小數量仍為8個、但RAID群組大小不應大於12個磁碟機。NetApp也建議每個RAID群組使用一個備援磁碟機；不過、所有RAID群組都可以使用全域備援磁碟機。例如、您可以每三個RAID群組使用兩個備援磁碟機、每個RAID群組包含八到十二個磁碟機。

> 舊版ESX的最大範圍和資料存放區大小為64TB、可能會影響支援這些大容量磁碟機所提供之總原始容量所需的LUN數量。

RAID 模式

許多RAID控制器最多可支援三種操作模式、每種模式代表寫入要求所採用的資料路徑有顯著差異。這三種模式如下：

- WriteThrough：所有傳入的I/O要求都會寫入RAID控制器快取、然後立即排清到磁碟、再將要求傳回主機。
- Writaround.所有傳入的I/O要求都會直接寫入磁碟、繞過RAID控制器快取。
- 回寫。所有傳入的I/O要求都會直接寫入控制器快取、並立即確認回主機。資料區塊會使用控制器以非同步方式排清到磁碟。

回寫模式提供最短的資料路徑、區塊進入快取後會立即出現I/O確認。此模式可為混合式讀取/寫入工作負載提供最低延遲和最高處理量。但是、如果沒有任何磁碟機或非揮發性Flash技術、使用者在這種模式下操作時、如果系統發生電源故障、就會有資料遺失的風險。

由於需要電池備份或Flash裝置、因此我們可以確信、快取的區塊會在發生此類故障時排清到磁碟。ONTAP Select因此、RAID控制器必須設定為回寫模式。

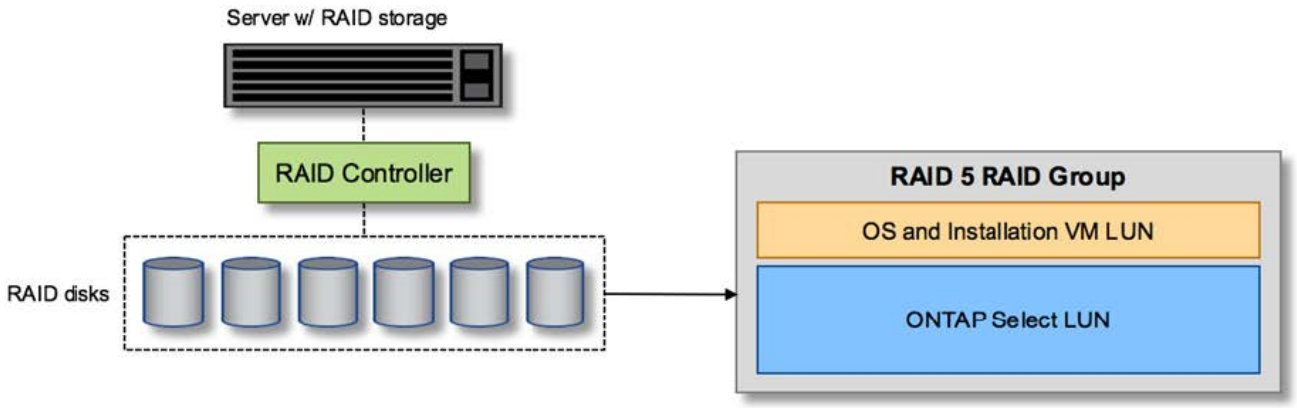
本地磁碟共享**ONTAP Select** 於支援支援的作業系統

最常見的伺服器組態是所有本機連接的磁碟位於單一RAID控制器後方的組態。您應該配置至少兩個LUN：一個用於Hypervisor、一個ONTAP Select 用於搭配使用。

例如、請考慮使用HP DL380 G8搭配六個內部磁碟機和一個Smart Array P420i RAID控制器。所有內部磁碟機均由此RAID控制器管理、且系統上沒有其他儲存設備。

下圖顯示此組態樣式。在此範例中、系統上沒有其他儲存設備、因此Hypervisor必須與ONTAP Select 該節點共用儲存設備。

伺服器**LUN**組態、僅使用**RAID**管理的磁碟



將OS LUN從ONTAP Select 同一個RAID群組配置為「支援」、可讓Hypervisor OS（以及任何也從該儲存設備配置的用戶端VM）受益於RAID保護。此組態可防止單一磁碟機故障導致整個系統關機。

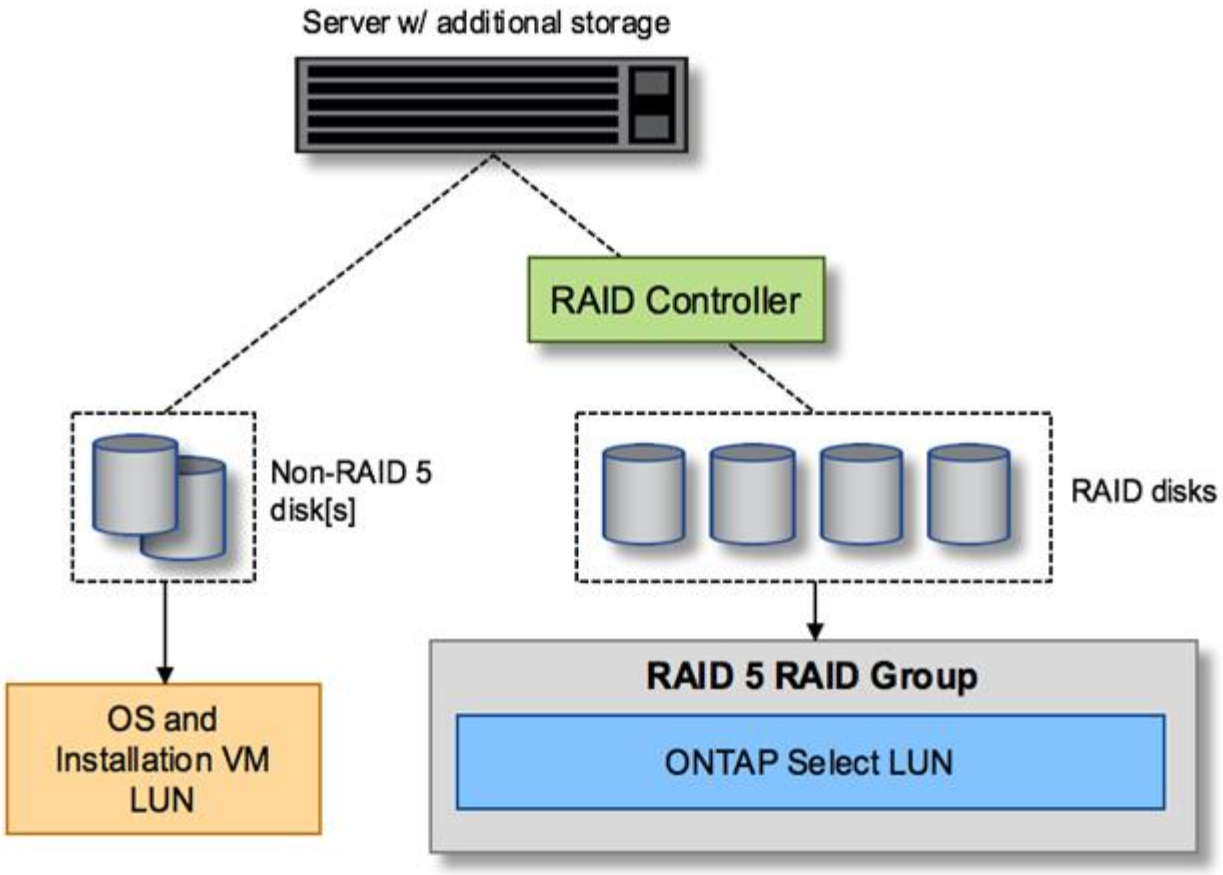
本機磁碟在**ONTAP Select** 不支援的地方和作業系統之間分割

伺服器廠商提供的其他可能組態包括使用多個RAID或磁碟控制器來設定系統。在此組態中、一組磁碟是由一個磁碟控制器管理、該磁碟控制器可能提供或可能不提供RAID服務。第二組磁碟由能夠提供RAID 5/6服務的硬體RAID控制器管理。

有了這種組態風格、位於RAID控制器後方可提供RAID 5/6服務的一組磁碟、應由ONTAP Select 該虛擬機器獨家使用。視所管理的總儲存容量而定、您應該將磁碟機設定為一個或多個RAID群組、以及一個或多個LUN。然後、這些LUN將用於建立一或多個資料存放區、所有資料存放區都受到RAID控制器的保護。

第一組磁碟保留給Hypervisor作業系統、以及任何不使用ONTAP 支援此功能的用戶端VM、如下圖所示。

混合式**RAID/**非**RAID**系統上的伺服器**LUN**組態



多個**LUN**

單一RAID群組/單一LUN組態必須變更的情況有兩種。使用NL-SAS或SATA磁碟機時、RAID群組大小不得超過12個磁碟機。此外、單一LUN可能會大於基礎Hypervisor儲存限制、無論是個別檔案系統範圍的最大大小或總儲存池的最大大小。然後必須將基礎實體儲存設備分割成多個LUN、才能成功建立檔案系統。

VMware vSphere虛擬機器檔案系統限制

某些ESX版本上的資料存放區大小上限為64TB。

如果伺服器所連接的儲存容量超過64TB、則可能需要配置多個LUN、每個LUN的容量都小於64TB。建立多個RAID群組來改善SATA/NL-SAS磁碟機的RAID重建時間、也會導致配置多個LUN。

當需要多個LUN時、最重要的考量是確保這些LUN的效能相似且一致。如果所有LUN都要用於單ONTAP 一的位向集合體、這點特別重要。或者、如果一個或多個LUN的子集具有明顯不同的效能設定檔、我們強烈建議您將這些LUN隔離在個別ONTAP 的「VMware Aggregate」中。

多個檔案系統範圍可用來建立單一資料存放區、最多可達資料存放區的最大大小。若要限制需要ONTAP Select 使用流通證的容量、請務必在叢集安裝期間指定容量上限。此功能僅允許ONTAP Select 使用（因此需要授權）資料存放區中空間的子集。

或者、您可以從在單一LUN上建立單一資料存放區開始著手。如果需要更多空間、需要更大ONTAP Select 的等量資料授權、則可將該空間新增至與某個範圍相同的資料存放區、最多可增加至資料存放區的最大大小。達到最大容量後、就能建立新的資料存放區並將其新增至ONTAP Select 功能區。這兩種類型的容量擴充作業均受到支援、並可透過ONTAP 使用「支援部署儲存新增功能」來達成。每ONTAP Select 個支援多達400TB儲存容量的

支援節點均可設定。從多個資料存放區配置容量需要兩個步驟的程序。

初始叢集建立可用於建立ONTAP Select 一個佔用初始資料存放區部分或全部空間的不實叢集。第二個步驟是使用其他資料存放區執行一或多個容量新增作業、直到達到所需的總容量為止。本節將詳細說明此功能 "增加儲存容量"。

VMFS負荷非零（請參閱 "VMware知識庫1001618"）、且嘗試使用資料存放區回報為可用的整個空間、導致叢集建立作業期間發生假錯誤。

每個資料存放區中有2%的緩衝區未使用。這個空間不需要容量授權、因為ONTAP Select 它不供人使用。只要未指定容量上限、即可自動計算緩衝區的確切GB數。ONTAP如果指定容量上限、則會先強制執行該大小。如果容量上限大小落在緩衝區大小內、叢集建立就會失敗、並顯示錯誤訊息、指出可用做容量上限的正確最大大小參數：

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

VMFS 6同時支援新安裝、也支援做為現有ONTAP 的VMware部署或ONTAP Select VMware VM Storage VMotion作業的目標。

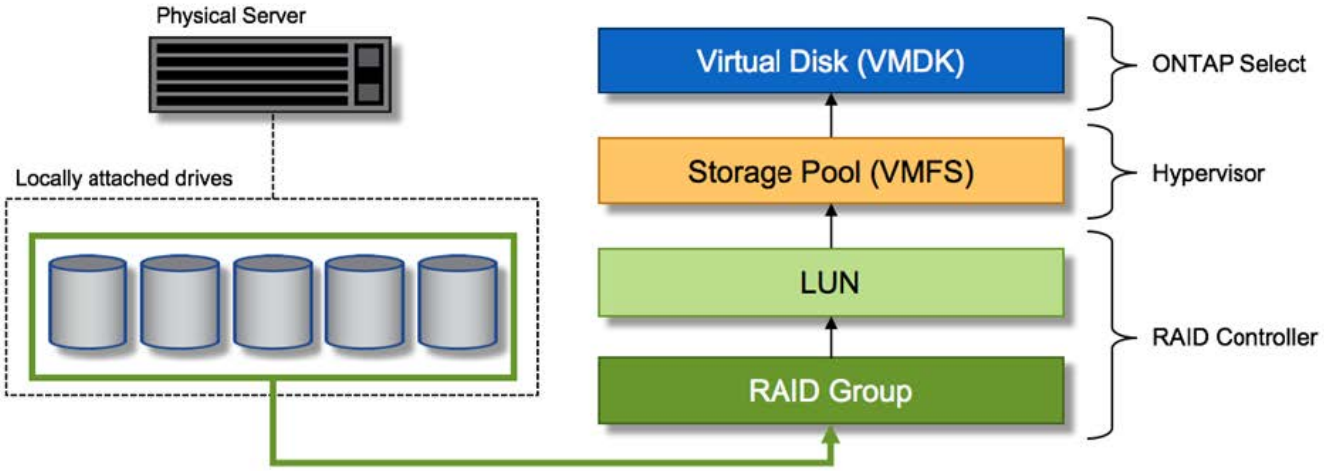
VMware不支援從VMFS 5就地升級至VMFS 6。因此、Storage VMotion是唯一允許任何VM從VMFS 5資料存放區移轉至VMFS 6資料存放區的機制。不過ONTAP Select 、除了ONTAP 從VMFS 5移轉至VMFS 6的特定目的之外、還擴大了對含VMware及VMware部署的Storage VMotion的支援、以涵蓋其他案例。

虛擬磁碟**ONTAP Select**

在其核心、ONTAP Select 透過ONTAP 一或多個儲存資源池配置的一組虛擬磁碟、呈現出一套功能完善的功能。提供一組虛擬磁碟、將其視為實體磁碟、而儲存堆疊的其餘部分則由Hypervisor抽象化。ONTAP下圖更詳細地顯示這種關係、強調實體RAID控制器、Hypervisor和ONTAP Select 不支援的VM之間的關係。

- RAID群組和LUN組態是從伺服器的RAID控制器軟體內部進行。使用VSAN或外部陣列時、不需要此組態。
- 儲存資源池組態是從Hypervisor內部進行。
- 虛擬磁碟是由個別VM所建立和擁有、ONTAP Select 在此範例中、由支援。

虛擬磁碟對實體磁碟對應



虛擬磁碟資源配置

為了提供更精簡的使用者體驗、ONTAP Select 我們的「更新」管理工具ONTAP 「還原部署」會自動從相關的儲存資源池配置虛擬磁碟、並將其附加至ONTAP Select 「更新」VM。這項作業會在初始設定和儲存新增作業期間自動執行。如果ONTAP Select 此節點是HA配對的一部分、則虛擬磁碟會自動指派給本機和鏡射儲存資源池。

將基礎附加儲存設備分割成大小相同的虛擬磁碟、每個磁碟不超過16TB。ONTAP Select如果ONTAP Select 此節點是HA配對的一部分、則每個叢集節點上至少會建立兩個虛擬磁碟、並指派給鏡射Aggregate中要使用的本機叢和鏡射叢。

例如ONTAP Select 、某個對象可以指派31 TB的資料存放區或LUN（部署VM後的剩餘空間、以及系統和根磁碟的資源配置）。然後建立四個~7.75TB虛擬磁碟、並指派給適當ONTAP 的鏡射本機叢和鏡射叢。

> 將容量新增至ONTAP Select 某個VMware可能會導致不同大小的VMDK。如需詳細資訊、請參閱一節 "增加儲存容量"。不同FAS 於VMware系統、不同大小的VMDK可存在於同一個集合體中。在這些VMDK上使用RAID 0等量磁碟區、無論其大小為何、都能充分利用每個VMDK中的所有空間。ONTAP Select

虛擬化**NVRAM**

NetApp FAS 支援系統通常裝有實體NVRAM PCI卡、這是一種高效能卡、內含非揮發性Flash記憶體。此卡可立即ONTAP 認可傳入寫入回用戶端的功能、大幅提升寫入效能。它也可以在稱為「減少需求」的程序中、將修改過的資料區塊排程回較慢的儲存媒體。

一般而言、市售系統並未安裝此類設備。因此、此NVRAM卡的功能已虛擬化、並放入ONTAP Select 了一個分區內的系統啟動磁碟。因此、放置執行個體的系統虛擬磁碟非常重要。這也是為什麼產品需要實體RAID控制器、並針對本機附加儲存組態提供彈性快取。

NVRAM位於自己的VMDK上。將NVRAM拆分成自己的VMDK、ONTAP Select 即可讓VMware使用vNVMe驅動程式與NVRAM VMDK通訊。此外、還需要ONTAP Select 使用與ESX 6.5及更新版本相容的硬體版本13。

資料路徑說明：**NVRAM**和**RAID**控制器

虛擬化NVRAM系統分割區與RAID控制器之間的互動、最好是在寫入要求進入系統時、透過資料路徑來強調顯示。

傳入ONTAP Select 的寫入要求會以VM的NVRAM分割區為目標。在虛擬化層、此分割區存在ONTAP Select 於一個連接ONTAP Select 到該VMware VM的VMware系統磁碟內。在實體層、這些要求會快取到本機RAID控制器、就像所有針對基礎磁碟的區塊變更一樣。從這裡、寫入作業會確認回傳給主機。

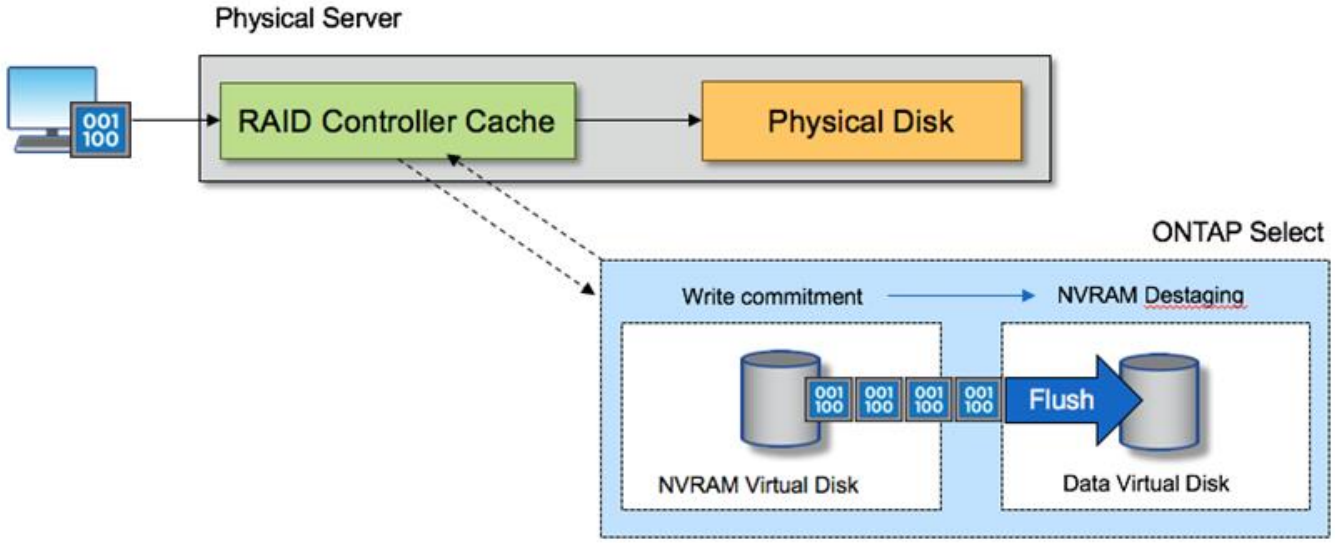
此時、實體區塊會駐留在RAID控制器快取中、等待排清到磁碟。邏輯上、區塊位於NVRAM中、等待將資料移轉至適當的使用者資料磁碟。

由於變更的區塊會自動儲存在RAID控制器的本機快取中、因此傳入的NVRAM分割區寫入作業會自動快取、並定期排清到實體儲存媒體。這不應與定期將NVRAM內容排清回ONTAP 還原至還原資料磁碟的做法相混淆。這兩個事件是不相關的、會在不同的時間和頻率發生。

下圖顯示傳入寫入所需的I/O路徑。它強調實體層（由RAID控制器快取和磁碟表示）與虛擬層（由VM的NVRAM和資料虛擬磁碟表示）之間的差異。

ⓘ 雖然NVRAM VMDK上變更的區塊會快取到本機RAID控制器快取中、但快取並不知道VM結構或其虛擬磁碟。它會將所有變更的區塊儲存在系統上、其中NVRAM只是其中的一部分。如果Hypervisor是從相同的備份磁碟配置、則這包括綁定至Hypervisor的寫入要求。

傳入寫入**ONTAP Select** 到**Sing VM**



ⓘ NVRAM磁碟分割區是在自己的VMDK上分隔。VMDK是使用ESX版本6.5或更新版本中提供的vNVME驅動程式來附加。這項變更對於ONTAP Select 使用軟體RAID進行的版本更新來說最重要、因為軟體RAID無法從RAID控制器快取中獲益。

本機附加儲存設備的軟體**RAID**服務

軟體RAID是一套RAID抽象層、實作於ONTAP 整套的功能。它提供的功能與傳統ONTAP的支援平台（例如FAS 、）內的RAID層相同。RAID層會執行磁碟機同位元檢查計算、並在ONTAP Select 一個節點內針對個別磁碟機故障提供保護。

不受硬體RAID組態的影響ONTAP Select 、而支援使用軟體RAID選項。硬體RAID控制器在某些環境中可能無法使用或不受歡迎、例如ONTAP Select 當將處理器部署在小型市售硬體上時。軟體RAID可擴充可用的部署選項、以納入這類環境。若要在您的環境中啟用軟體RAID、請記住以下幾點：

- 提供Premium或Premium XL授權。
- 它僅支援SSD或NVMe（需要Premium XL授權）磁碟機、用於ONTAP 支援以供支援的實體根磁碟和資料磁碟。
- 它需要獨立的系統磁碟、才能用於ONTAP Select 整個過程中的虛擬機器開機分割區。
  - 選擇獨立的磁碟（SSD或NVMe磁碟）、為系統磁碟建立資料存放區（在多節點設定中、可建立NVRAM、Boot/CF卡、coredump和資料管理器）。

附註

- 術語服務磁碟和系統磁碟可互換使用。
  - 服務磁碟是ONTAP Select 指在VMware支援中心內用於服務各種項目的VMDK、例如叢集、開機等。
  - 服務磁碟實體位於單一實體磁碟（統稱為服務/系統實體磁碟）上、從主機即可看到。該實體磁碟必須包含DAS資料存放區。在叢集部署期間、利用「部署」功能建立這些服務磁碟、以利實現不受故障的VM。ONTAP ONTAP Select
- 不可能在ONTAP Select 多個資料存放區或多個實體磁碟機之間進一步分隔出不支援的系統磁碟。
- 硬體RAID並未過時。

本機附加儲存設備的軟體**RAID**組態

使用軟體RAID時、不需要硬體RAID控制器是理想的選擇、但如果系統確實有現有的RAID控制器、則必須符合下列要求：

- 硬體RAID控制器必須停用、才能將磁碟直接呈現給系統（JBOD）。這項變更通常可在RAID控制器BIOS中進行
- 或硬體RAID控制器應處於SAS HBA模式。例如，某些BIOS組態除了RAID之外還允許使用「ahci」模式，而RAID則可選擇啟用JBOD模式。這會啟用傳遞、使實體磁碟機可視為主機上的實體磁碟機。

視控制器支援的磁碟機數量上限而定、可能需要額外的控制器。使用SAS HBA模式時、請確保IO控制器（SAS HBA）支援最低6Gb/s速度。不過、NetApp建議使用12Gbps速度。

不支援其他硬體RAID控制器模式或組態。例如、某些控制器允許RAID 0支援、這可能會人為地讓磁碟通過、但可能會造成不良影響。支援的實體磁碟大小（僅限SSD）介於200GB至16TB之間。

系統管理員必須追蹤ONTAP Select 哪些磁碟機正在由該名虛擬機器使用、並避免在主機上意外使用這些磁碟機。

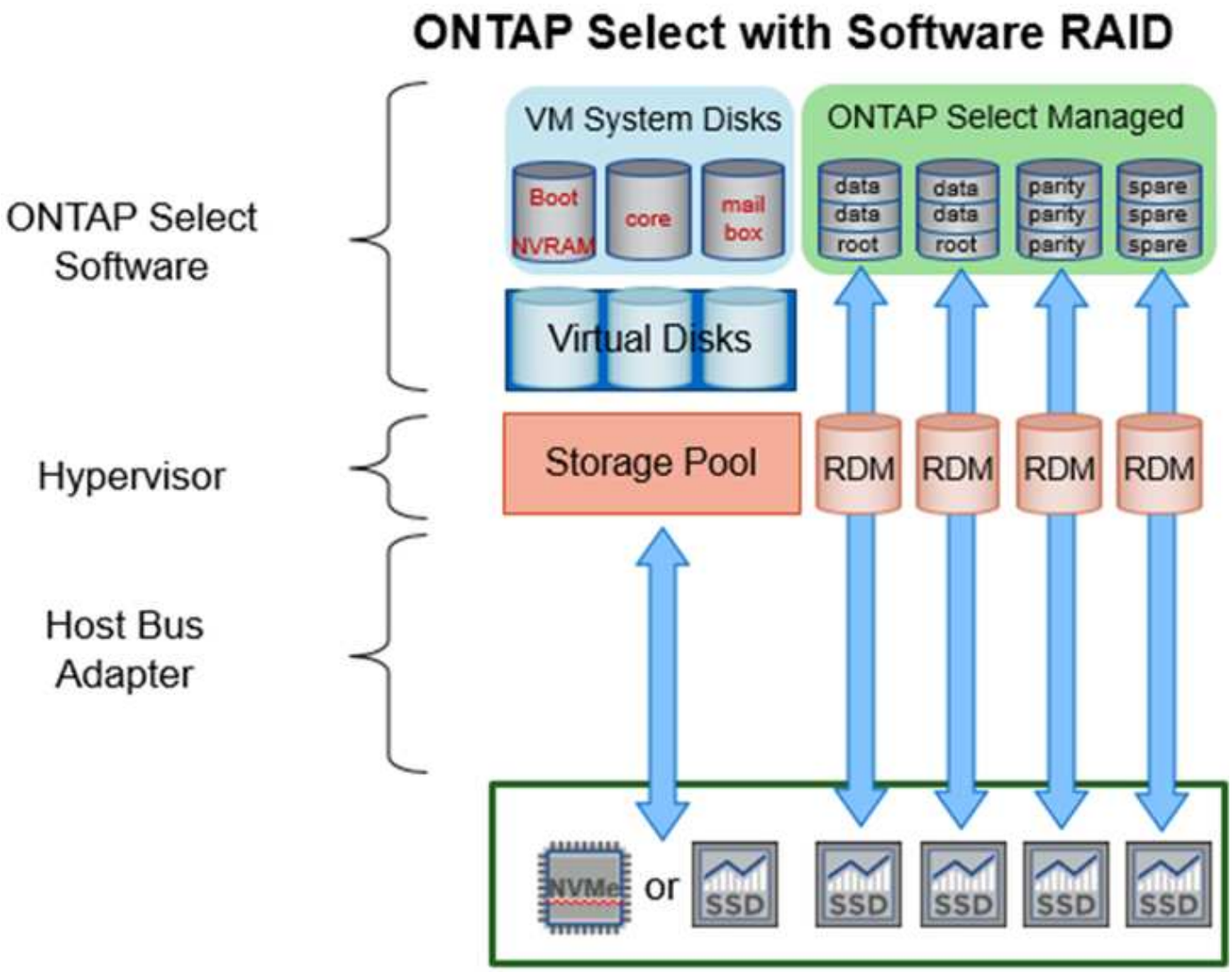
虛擬與實體磁碟**ONTAP Select**

對於使用硬體RAID控制器的組態、實體磁碟備援是由RAID控制器提供。提供一個或多個VMDK、供VMware管理員設定資料集合體。ONTAP Select ONTAP這些VMDK採用RAID 0格式等量分佈、因為使用ONTAP VMware軟體RAID是備援的、效率不彰、而且由於硬體層級提供恢復功能、因此效率不彰。此外、用於系統磁碟的VMDK與用於儲存使用者資料的VMDK位於同一個資料存放區。

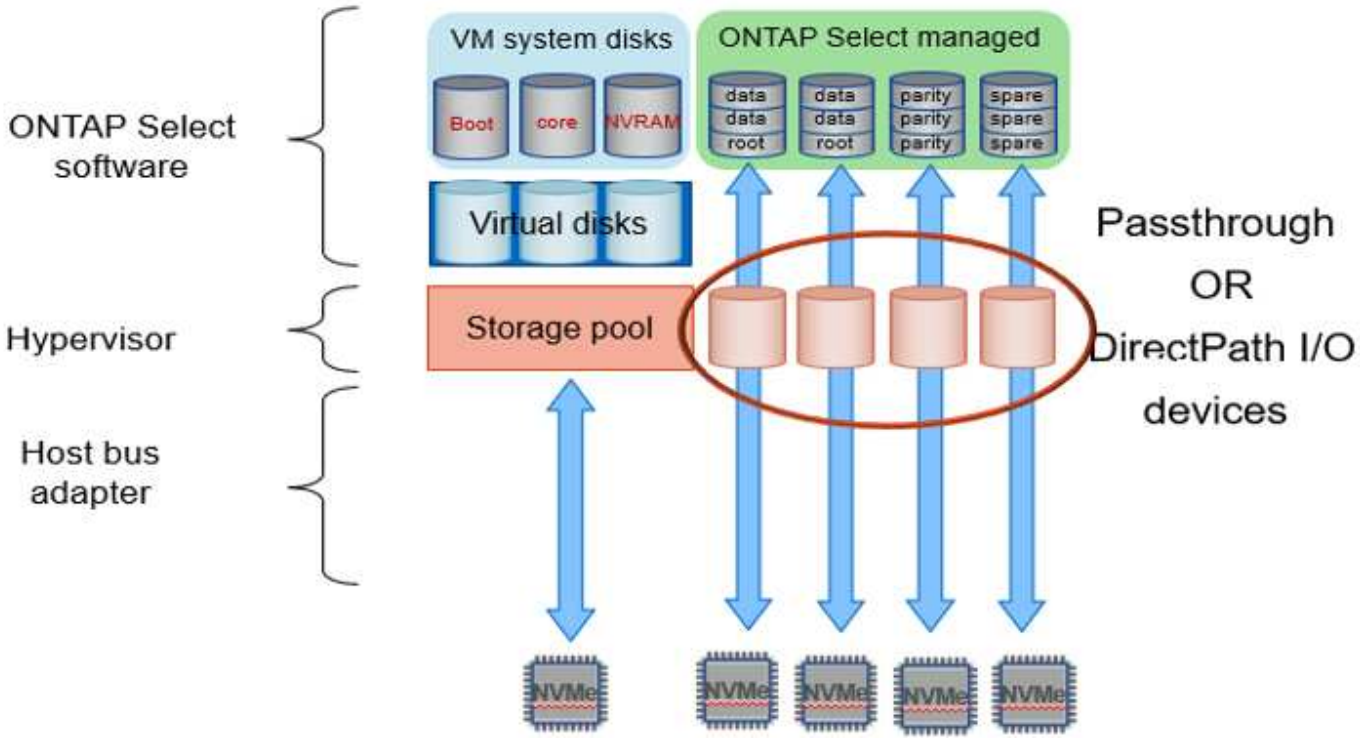
使用軟體RAID時ONTAP 、針對ONTAP Select SSD和用於NVMes的PassthThrough或DirectPath IO裝置、使用一套虛擬磁碟（VMDK）和實體磁碟原始裝置對應（RDM）時、即可呈現出一套功能強大的功能。

下圖更詳細地顯示這種關係、強調ONTAP Select 用來保存使用者資料的虛擬化磁碟與用來儲存使用者資料的實體磁碟之間的差異。

*支援軟體RAID：使用虛擬化磁碟和RDM * ONTAP Select



系統磁碟（VMDK）位於同一個資料存放區和同一個實體磁碟上。虛擬NVRAM磁碟需要快速且持久的媒體。因此、僅支援NVMe和SSD類型的資料存放區。



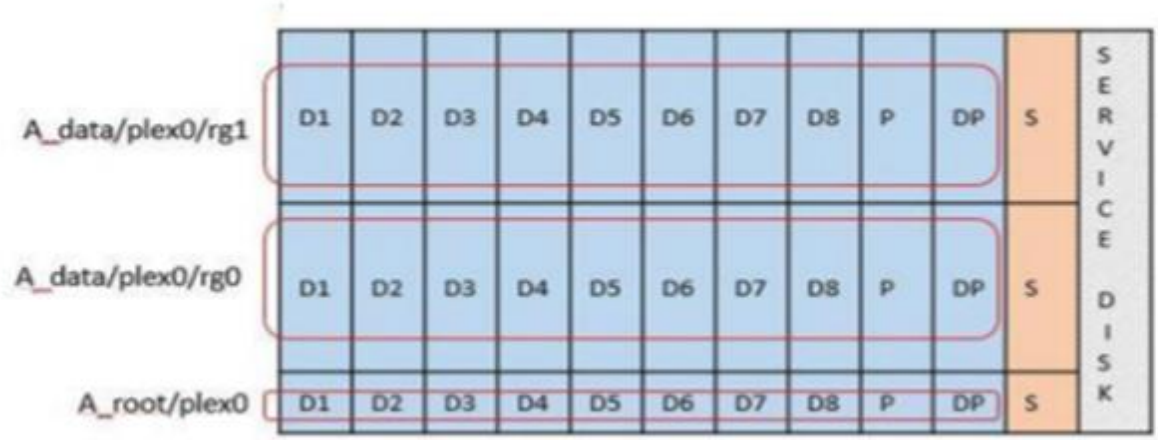
系統磁碟（VMDK）位於同一個資料存放區和同一個實體磁碟上。虛擬NVRAM磁碟需要快速且持久的媒體。因此、僅支援NVMe和SSD類型的資料存放區。在使用NVMe磁碟機進行資料時、基於效能考量、系統磁碟也應該是NVMe裝置。在All NVMe組態中、最適合使用系統磁碟的是Intel Optane卡。

在目前的版本中、ONTAP Select 不可能在多個資料存放區或多個實體磁碟機之間進一步分隔出系統磁碟。

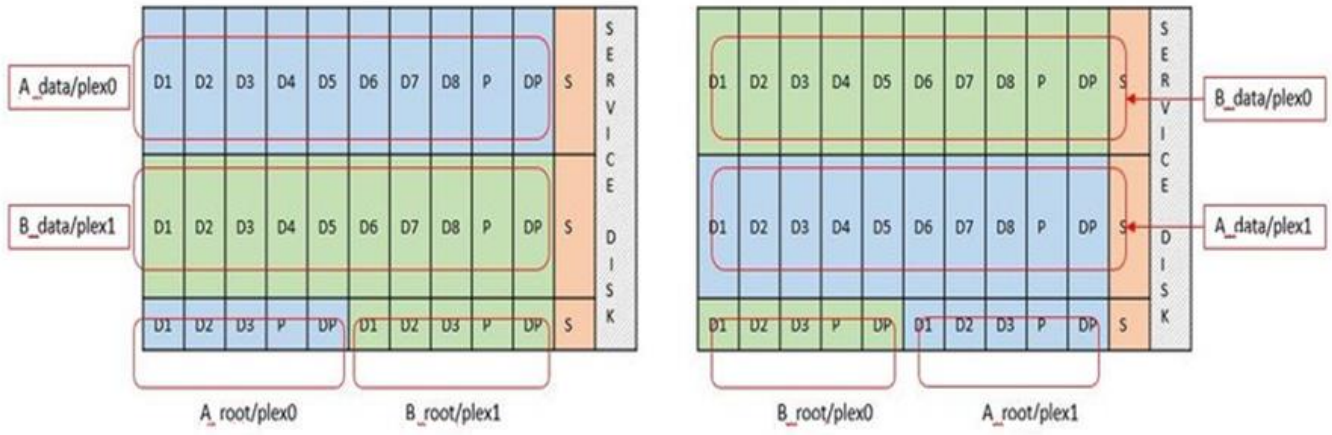
每個資料磁碟分為三個部分：一個小的根分割區（等量磁碟區）和兩個大小相同的分割區、以建立ONTAP Select 在該vme中可見的兩個資料磁碟。分割區使用根資料（RD2）架構、如下列圖表所示、適用於單一節點叢集和HA配對中的節點。

`P` 表示同位元檢查磁碟機。 `DP` 表示雙同位元檢查磁碟機和 `S` 表示備用磁碟機。

*單節點叢集*適用的RDD磁碟分割



適用於多節點叢集（**HA**配對）的**RDD**磁碟分割



支援下列RAID類型的軟件RAID：RAID 4、RAID-DP及ONTAP RAID-TEC這些都是FAS 由Same和AFF Same平台所使用的相同RAID架構。root資源配置ONTAP Select 僅支援RAID 4和RAID-DP。使用RAID-TEC 資料Aggregate時、整體保護是RAID-DP。NetApp HA使用無共享架構、將每個節點的組態複寫到另一個節點。ONTAP Select這表示每個節點都必須儲存其根分割區及其對等根分割區的複本。由於資料磁碟有單一根分割區、因此根據ONTAP Select 該節點是否屬於HA配對的一部分、資料磁碟的最小數量會有所不同。

對於單一節點叢集、所有資料分割區都會用來儲存本機（作用中）資料。對於屬於HA配對一部分的節點、會使用一個資料分割區來儲存該節點的本機（作用中）資料、而第二個資料分割區則是用來鏡射HA對等端的作用中資料。

PassthThrough（**DirectPath IO**）裝置與原始裝置對應（**RDM**）

VMware ESX目前不支援NVMe磁碟做為原始裝置對應。若要讓VMware直接控制NVMe磁碟、NVMe磁碟機必須在ESX中設定為傳遞裝置。ONTAP Select請注意、將NVMe裝置設定為傳遞裝置需要伺服器BIOS的支援、這是一種中斷運作的程序、需要重新啟動ESX主機。此外、每個ESX主機的最大通道裝置數為16。不過、只有將此限制在14個範圍內。ONTAP每ONTAP Select 個節點最多可有14個NVMe裝置、這表示所有NVMe組態都能提供極高的IOPs密度（IOPS/TB）、而犧牲總容量。或者、如果您需要較大儲存容量的高效能組態、建議使用的組態為大ONTAP Select 尺寸的VMware VM、系統磁碟的Intel Optane卡、以及資料儲存的名義SSD磁碟機數量。

若要充分發揮NVMe效能、請考慮使用大ONTAP Select 尺寸的VMware。

PassthThrough裝置與RDM之間還有其他差異。RDM可對應至執行中的VM。傳遞裝置需要重新開機VM。這表示任何NVMe磁碟機更換或容量擴充（新增磁碟機）程序都需要ONTAP Select 重新啟動以進行支援。磁碟機更換與容量擴充（磁碟機新增）作業是由ONTAP 在「支援」中的工作流程所驅動。藉由使用者可管理單一節點叢集的重開機功能、以及HA配對的容錯移轉/容錯回復功能。ONTAP ONTAP Select不過請務必注意使用SSD資料磁碟機（ONTAP Select 不需要進行重新開機/容錯移轉）和使用NVMe資料磁碟機（ONTAP Select 需要進行重新開機/容錯移轉）之間的差異。

實體與虛擬磁碟資源配置

為了提供更精簡的使用者體驗、ONTAP 利用此功能、即可自動從指定的資料存放區（實體系統磁碟）配置系統（虛擬）磁碟、並將其附加至ONTAP Select 還原VM。這項作業會在初始設定期間自動執行、ONTAP Select 以便讓支援的不中斷虛擬機器能夠開機。系統會分割RDM、並自動建立根Aggregate。如果ONTAP Select 此節點是HA配對的一部分、則會自動將資料分割指派給本機儲存資源池和鏡射儲存資源池。這項指派會在叢集建立作業和儲存新增作業期間自動進行。

由於ONTAP Select 在VMware上的資料磁碟與基礎實體磁碟相關聯、因此建立具有大量實體磁碟的組態會帶來效能影響。

> 根Aggregate的RAID群組類型取決於可用的磁碟數目。支援：部署選擇適當的RAID群組類型。ONTAP如果有足夠的磁碟配置給節點、則會使用RAID-DP、否則會建立RAID-4根Aggregate。

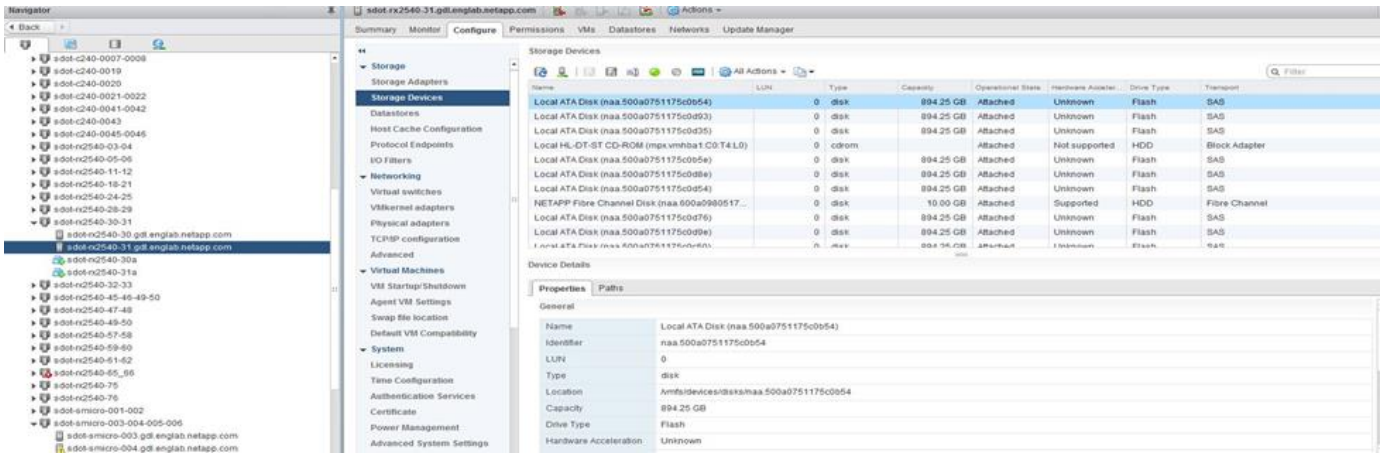
當使用軟體RAID將容量新增至ONTAP Select 某個物件時、系統管理員必須考量實體磁碟機大小和所需磁碟機數量。如需詳細資訊、請參閱一節 "增加儲存容量"。

與FAS 支援不相同AFF 的是、只有容量相等或更大的磁碟機可以新增至現有的RAID群組。容量較大的磁碟機大小適中。如果您要建立新的RAID群組、新的RAID群組大小應與現有的RAID群組大小相符、以確保整體的Aggregate效能不會降低。

將 **ONTAP Select** 磁碟與對應的 **ESX** 磁碟配對

通常將各個磁碟標示為NET x.y。ONTAP Select您可以使用下列ONTAP 的fuse命令來取得磁碟UUID：

```
<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
```



在ESXi Shell中、您可以輸入下列命令、以將特定實體磁碟的LED（以naa.unite-id識別）閃亮。

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

使用軟體**RAID**時發生多個磁碟機故障

系統可能會遇到多個磁碟機同時處於故障狀態的情況。系統的行為取決於Aggregate RAID保護和故障磁碟機的數量。

RAID4 Aggregate可在一次磁碟故障後繼續運作、RAID-DP Aggregate可在兩次磁碟故障時維持運作、RAID-TEC 而一個版本為可在三個磁碟故障後繼續運作的版本。

如果故障磁碟的數量少於RAID類型支援的最大故障數、而且有備用磁碟可用、重建程序就會自動啟動。如果備用磁碟無法使用、則Aggregate會以降級狀態提供資料、直到新增備用磁碟為止。

如果故障磁碟數量超過RAID類型支援的最大故障數、則本機叢會標示為故障、且Aggregate狀態會降級。資料是由位於HA合作夥伴的第二個叢提供。這表示任何節點1的I/O要求都會透過叢集互連連接埠e0e（iSCSI）傳送至實體位於節點2上的磁碟。如果第二個叢也失敗、則會將該集合體標示為故障、且資料無法使用。

必須刪除並重新建立故障的叢、才能恢復正確的資料鏡射。請注意、多磁碟故障導致資料集合降級、也會導致根集合體降級。使用root資料資料（RDD）分割架構、將每個實體磁碟分割成一個根分割區和兩個資料分割區。ONTAP Select因此、遺失一或多個磁碟可能會影響多個Aggregate、包括本機根或遠端根Aggregate的複本、以及本機資料Aggregate和遠端資料Aggregate的複本。

```
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..
```

若要測試或模擬一或多個磁碟機故障、請使用 `storage disk fail -disk NET-x.y -immediate` 命令。如果系統中有備援磁碟機、則會開始重建集合體。您可以使用命令檢查重建影像的狀態 `storage aggregate show`。您可以使用ONTAP 「還原部署」移除模擬的故障磁碟機。請注意、 ONTAP 已將磁碟機標示為 `Broken`。磁碟機實際上並未損壞、可以使用ONTAP還原部署來重新新增。若要清除損壞的標籤、請在ONTAP Select CLI中輸入下列命令：

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

最後一個命令的輸出應為空白。

虛擬化**NVRAM**

NetApp FAS 產品技術系統通常裝有實體NVRAM PCI卡。此卡為高效能卡、內含非揮發性快閃記憶體、可大幅提升寫入效能。它藉由授予ONTAP 功能來立即認可傳入寫入回用戶端。它也可以在稱為「減少需求」的程序中、將修改過的資料區塊排程回較慢的儲存媒體。

一般而言、市售系統並未安裝此類設備。因此、NVRAM卡的功能已虛擬化、並放入ONTAP Select 了一個分區內的系統啟動磁碟。因此、放置執行個體的系統虛擬磁碟非常重要。

vSAN與外部陣列組態

虛擬 NAS （ vNAS ）部署支援虛擬 SAN （ VSAN ）上的 ONTAP Select 叢集、部分 HCI 產品、以及外部陣列類型的資料存放區。這些組態的基礎架構可提供資料存放區恢復能力。

最低要求是基礎組態受VMware支援、並應列在各自的VMware HCL上。

VNAS架構

vNAS命名法適用於所有不使用DAS的設定。對於多節點ONTAP Select 的叢集、這包括ONTAP Select 相同HA配對中的兩個元節點共享單一資料存放區（包括vSAN資料存放區）的架構。節點也可以安裝在同一個共享外部陣列的不同資料存放區上。如此一來、陣列端的儲存效率就能降低ONTAP Select 整個堆高架構配對的整體佔用空間。利用本機RAID控制器建立ONTAP Select 的架構與ONTAP Select DAS上的支援架構非常相似。也就是說、每ONTAP Select 個節點都會繼續擁有HA合作夥伴資料的複本。以節點為範圍的實體儲存效率原則。ONTAP因此、陣列端的儲存效率較佳、因為它們可能會套用到來自ONTAP Select 兩個支援節點的資料集。

HA配對中的ONTAP Select 每個節點也可能使用獨立的外部陣列。這是搭配ONTAP Select MetroCluster 外部儲存設備使用NetApp SDS的常見選擇。

在每ONTAP Select 個節點使用不同的外部陣列時、這兩個陣列必須提供類似ONTAP Select 於該VMware VM的效能特性、這點非常重要。

vNAS架構在邏輯上最類似於採用DAS和RAID控制器的伺服器架構。在這兩種情況ONTAP Select 下、使用者都會佔用資料存放區空間。該資料存放區空間會分割成VMDK、而這些VMDK則是傳統ONTAP 的資訊區集合體。在叢集建立和儲存新增作業期間、利用VMware部署確保VMDK大小適當、並指派給正確的叢集（HA配對）ONTAP 。

使用RAID控制器的vNAS與DAS之間有兩大差異。最直接的差異在於vNAS不需要RAID控制器。VNAS假設基礎外部陣列提供的資料持續性和恢復能力、是採用RAID控制器設定的DAS所能提供的功能。第二個和更為細微的差異在於NVRAM效能。

vNAS NVRAM

此功能可讓您順利執行。ONTAP Select換句話說ONTAP Select 、VMware將位元組可定址空間（傳統NVRAM）模擬在區塊可定址裝置（VMDK）上方。不過、NVRAM的效能對於ONTAP Select 整個節點的效能絕對重要。

對於使用硬體RAID控制器的DAS設定、硬體RAID控制器快取會做為事實上的NVRAM快取、因為所有寫入NVRAM VMDK的作業都會先裝載在RAID控制器快取中。

針對VNAS架構、ONTAP Singfeploy會自動使用ONTAP Select 稱為「單一執行個體資料記錄」（SIDL）的開機引數來設定支援節點。當出現此開機引數時、ONTAP Select 即可略過NVRAM、並將資料有效負載直接寫入資料Aggregate。NVRAM僅用於記錄寫入作業所變更之區塊的位址。此功能的優點是避免重複寫入：一次寫入NVRAM、另一次寫入NVRAM。這項功能僅適用於vNAS、因為本機寫入RAID控制器快取的額外延遲可忽略不計。

SIDL功能與ONTAP Select 所有的功能不相容。您可以使用下列命令、在Aggregate層級停用SIDL功能：

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

請注意、如果關閉SIDL功能、寫入效能會受到影響。停用該集合中所有磁碟區的所有儲存效率原則之後、即可重新啟用SIDL功能：

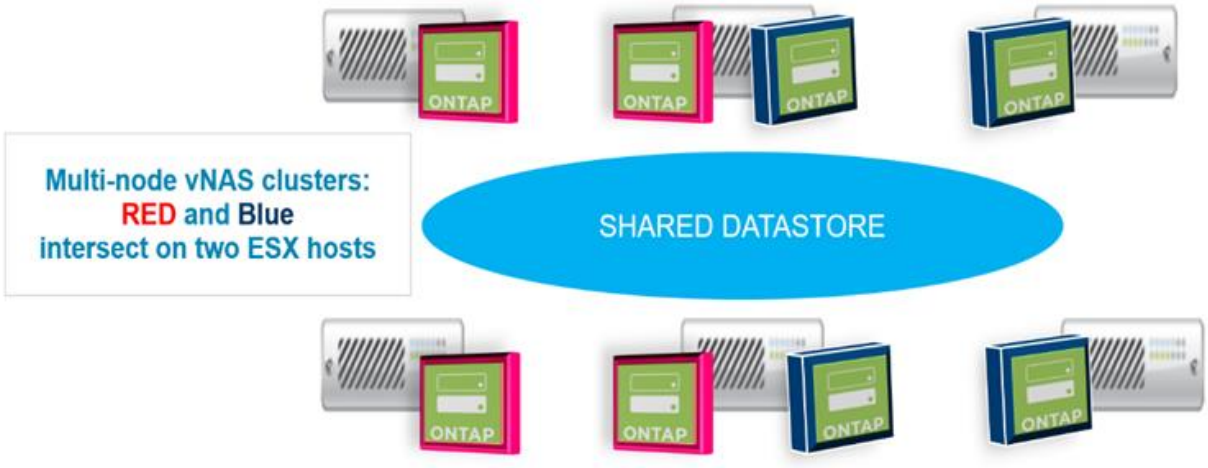
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

使用 vNAS 時組合 ONTAP Select 節點

支援共享儲存設備上的多節點元功能的叢集。ONTAP Select ONTAP Select只要這些節點不屬於同一個叢集、即可在同一部ESX主機上設定多個支援節點。ONTAP ONTAP Select請注意、此組態僅適用於VNAS環境（共享資料存放區）。使用DAS儲存設備時、每個主機不ONTAP Select 支援多個執行個體、因為這些執行個體競用相同的硬體RAID控制器。

藉由執行支援、可確保多節點VNAS叢集的初始部署不會將同一個叢集的多個實體執行個體置於同一部主機上。ONTAP ONTAP Select下圖顯示兩個四節點叢集在兩個主機上交叉的正確部署範例。

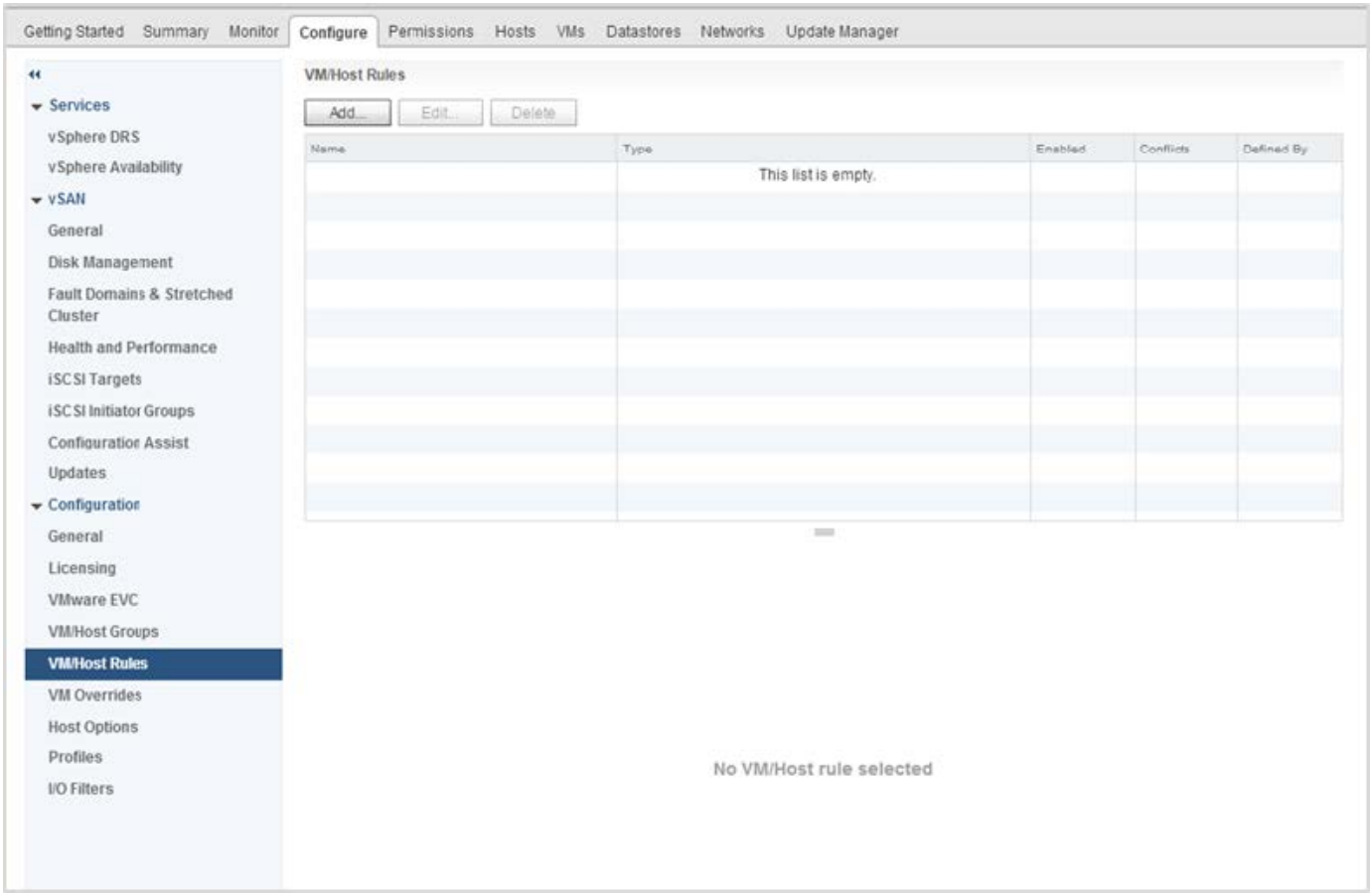
多節點VNAS叢集的初始部署

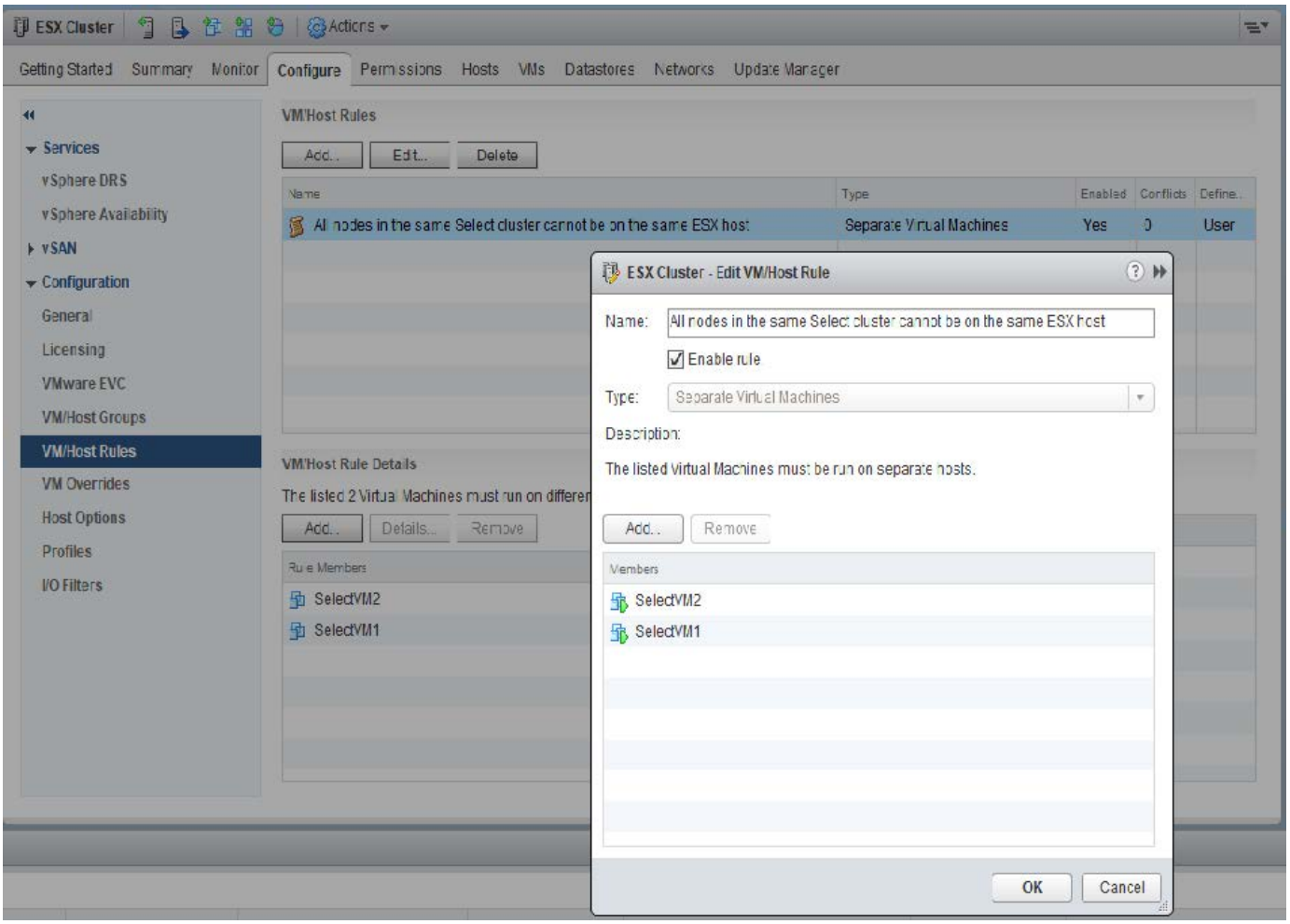


部署完成後ONTAP Select 、即可在主機之間移轉各個節點。這可能會導致非最佳和不受支援的組態、ONTAP Select 使來自同一個叢集的兩個或多個支援節點共用同一個基礎主機。NetApp建議手動建立VM反關聯規則、以便VMware自動在同一個叢集的節點之間維持實體分隔、而不只是同一個HA配對中的節點。

反關聯規則要求在ESX叢集上啟用DRS。

請參閱下列範例、瞭解如何為ONTAP Select 哪些虛擬機器建立反關聯規則。如果ONTAP Select 這個叢集包含多個HA配對、則叢集中的所有節點都必須包含在此規則中。





相同的VMware叢集有兩ONTAP Select 個以上的支援節點ONTAP Select 可能位於相同的ESX主機上、原因如下：

- 由於VMware vSphere授權限制或未啟用DRS、因此不存在DRS。
- 由於VMware HA作業或系統管理員啟動的VM移轉優先、因此會略過DRS反關聯規則。

請注意ONTAP 、不主動監控ONTAP Select 不支援的虛擬機器位置。不過、叢集重新整理作業會在ONTAP 「更新」記錄檔中反映出此不受支援的組態：

增加儲存容量

使用「部署」可為叢集中的每個節點新增及授權額外的儲存空間。ONTAP ONTAP Select

在支援的情況下、只有使用支援的儲存新增功能ONTAP 才能增加管理中的儲存容量、而且ONTAP Select 不支援直接修改此功能。下圖顯示啟動儲存新增精靈的「+」圖示。

下列考量對於容量擴充作業的成功非常重要。新增容量需要現有授權來涵蓋總空間量（現有加上新的）。導致節點超過授權容量的儲存新增作業會失敗。應先安裝具有足夠容量的新授權。

如果額外的容量已新增至現有ONTAP Select 的VMware Aggregate、則新的儲存資源池（資料存放區）應有類似於現有儲存資源池（資料存放區）的效能設定檔。請注意、無法將非SSD儲存設備新增至ONTAP Select 安裝有AFF特性（啟用Flash）的支援節點。也不支援混合DAS與外部儲存設備。

如果將本機附加儲存設備新增至系統以提供額外的本機（DAS）儲存池、則您必須建置額外的RAID群組和LUN（或LUN）。就像FAS 使用VMware系統一樣、如果您要在同一個集合體中新增空間、請務必確保新的RAID群組效能與原始RAID群組的效能相似。如果您要建立新的Aggregate、如果新的RAID群組配置對新Aggregate的效能影響已充分瞭解、則新的RAID群組配置可能會有所不同。

如果資料存放區的總大小未超過ESX支援的最大資料存放區大小、則可將新空間新增至該資料存放區、做為範圍。將資料存放區範圍新增至ONTAP Select 已安裝了資源庫的資料存放區、可以動態執行、不會影響ONTAP Select 到該節點的作業。

如果ONTAP Select 此節點是HA配對的一部分、則應考慮一些其他問題。

在HA配對中、每個節點都包含來自其合作夥伴的資料鏡射複本。若要將空間新增至節點1、必須將相同的空間新增至其合作夥伴節點2、才能將節點1的所有資料複寫至節點2。換句話說、在節點2上、新增至節點2的空間是節點1容量新增作業的一部分、因此無法在節點2上看到或存取。此空間會新增至節點2、以便在HA事件期間完全保護節點1資料。

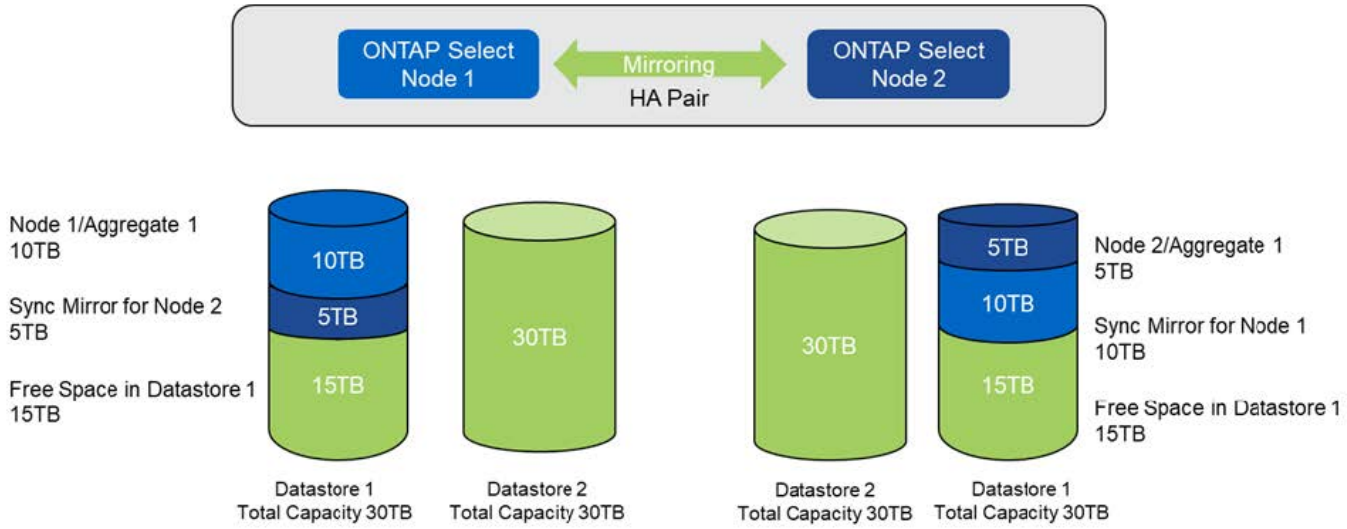
效能還有其他考量。節點1上的資料會同步複寫至節點2。因此、節點1上新空間（資料存放區）的效能必須符合節點2上新空間（資料存放區）的效能。換句話說、在兩個節點上新增空間、但使用不同的磁碟機技術或不同的RAID群組大小、可能會導致效能問題。這是因為使用RAID SyncMirror 功能來維護合作夥伴節點上的資料複本。

若要增加HA配對中兩個節點的使用者存取容量、必須執行兩項儲存新增作業、每個節點各執行一項。每項儲存新增作業都需要在兩個節點上增加空間。每個節點所需的總空間等於節點1所需的空間、加上節點2所需的空間。

初始設定包含兩個節點、每個節點都有兩個資料存放區、每個資料存放區有30TB的空間。藉由使用者可在資料存放區1中使用10TB的空間、藉由使用者可利用此功能來建立雙節點叢集。ONTAP利用每個節點5TB的使用中空間、即可設定各個節點的部署。ONTAP

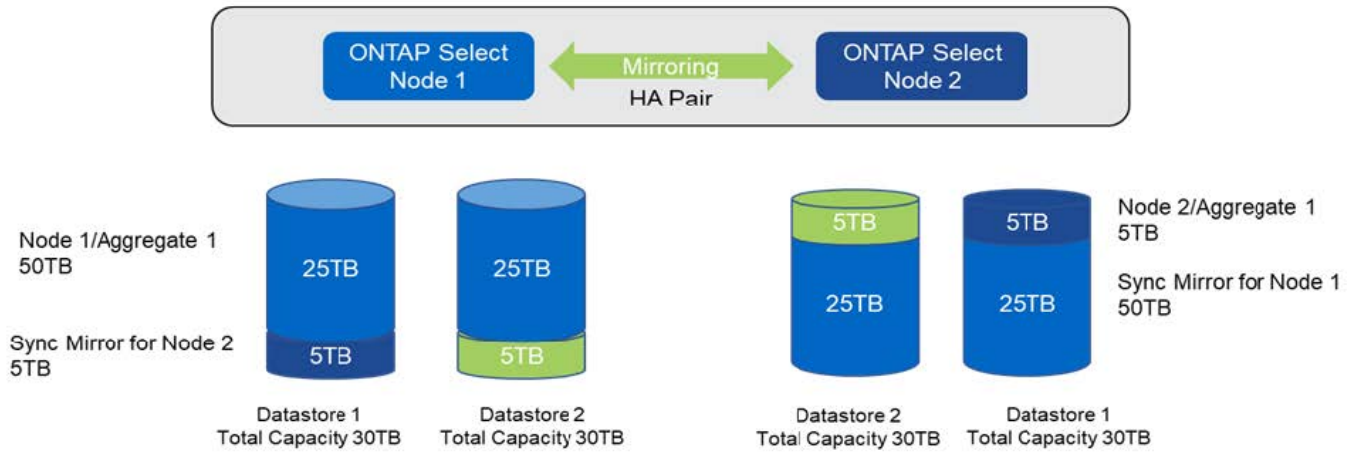
下圖顯示節點1單一儲存新增作業的結果。在每個節點上、仍然使用相同的儲存容量（15TB）ONTAP Select 。不過、節點1的主動式儲存設備（10TB）比節點2（5TB）多。當每個節點託管另一個節點的資料複本時、這兩個節點都受到完整保護。資料存放區1中還有額外的可用空間、而且資料存放區2仍完全可用。

容量分配：單一儲存新增作業後的配置與可用空間



節點1上的兩項額外儲存新增作業會消耗資料存放區1的其餘部分和資料存放區2的一部分（使用容量上限）。第一項儲存新增作業會佔用資料存放區1中的15TB可用空間。下圖顯示第二次儲存新增作業的結果。此時、節點1管理的作用中資料量為50TB、節點2則為原始5TB。

*容量發佈：在節點1*執行兩次額外的儲存新增作業之後、再進行配置與可用空間



容量新增作業期間使用的VMDK大小上限為16TB。叢集建立作業期間所使用的VMDK大小上限仍為8TB。根據您的組態（單節點或多節點叢集）和新增的容量量、使用VMware部署可建立大小正確的VMDK。ONTAP不過、在叢集建立作業期間、每個VMDK的最大大小不得超過8TB、在儲存新增作業期間、大小不得超過16TB。

利用軟體 **RAID** 增加 **ONTAP Select** 的容量

儲存新增精靈也可以使用ONTAP Select 軟體RAID來增加支援支援支援的支援能力。此精靈僅顯示可用的DAS SDD磁碟機、並可將其對應為RDM至ONTAP Select 該VM。

雖然容量授權可以增加單一TB、但在使用軟體RAID時、無法實際增加單一TB的容量。與將磁碟新增至FAS 某個實體或AFF 非實體陣列類似、某些因素決定了單一作業中可新增的最低儲存容量。

請注意、在HA配對中、若要將儲存設備新增至節點1、則節點的HA配對（節點2）上也需要相同數量的磁碟機。本機磁碟機和遠端磁碟都是由節點1上的單一儲存新增作業所使用。也就是說、遠端磁碟機可用來確保節點1上

的新儲存設備在節點2上進行複寫和保護。為了在節點2上新增本機可用的儲存設備、兩個節點上必須分別執行儲存新增作業和相同數量的磁碟機。

此功能可將任何新磁碟分割成與現有磁碟機相同的根磁碟、資料和資料分割區。ONTAP Select分割作業會在建立新的Aggregate期間或在現有Aggregate上擴充期間執行。每個磁碟上的根分割區等量磁碟區大小會設定為符合現有磁碟上現有的根分割區大小。因此、兩個等量資料分割區大小中的每一個都可以計算為磁碟總容量減去根分割區大小除以兩個。根分割區等量磁碟區大小是可變的、在初始叢集設定期間會依照下列方式進行計算。所需的根空間總計（單節點叢集為68GB、HA配對為136GB）會分成初始磁碟數、減去任何備用磁碟和同位元磁碟。根分割區等量磁碟區大小會維持不變、以固定於要新增至系統的所有磁碟機上。

如果您要建立新的Aggregate、所需的磁碟機數量下限會因RAID類型和ONTAP Select 該節點是否屬於HA配對的一部分而有所不同。

如果要將儲存設備新增至現有的Aggregate、則必須考量其他事項。假設RAID群組尚未達到最大限制、您可以將磁碟機新增至現有的RAID群組。在FAS 現有RAID群組中新增磁碟的傳統功能僅供使用、AFF 因此在新磁碟上建立熱點是可能的考量。此外、只能將資料分割區大小等於或大於相同的磁碟機新增至現有的RAID群組。如上所述、資料分割區大小與磁碟機原始大小不同。如果要新增的資料分割區大於現有的分割區、則新磁碟機的大小是適當的。換句話說、每個新磁碟機的一部分容量仍未使用。

也可以使用新磁碟機建立新的RAID群組、作為現有Aggregate的一部分。在此情況下、RAID群組大小應與現有的RAID群組大小相符。

儲存效率支援

提供的儲存效率選項類似於介紹在功能性和功能性上的儲存效率選項。ONTAP Select FAS AFF

使用All Flash VSAN或一般Flash陣列進行的虛擬NAS（vNAS）部署、應遵循非SSD直接附加儲存設備（DAS）的最佳實務做法ONTAP Select ONTAP Select 。

只要您有具備SSD磁碟機和優質授權的DAS儲存設備、就會在新安裝中自動啟用AFF型特性設定。

有了AFF般的特性設定、下列即時SE功能會在安裝期間自動啟用：

- 即時零模式偵測
- Volume內嵌重複資料刪除技術
- Volume背景重複資料刪除技術
- 可調適的即時壓縮
- 即時資料精簡
- 集合體即時重複資料刪除技術
- Aggregate背景資料刪除技術

若要驗ONTAP Select 證是否已啟用所有預設的儲存效率原則、請在新建立的磁碟區上執行下列命令：

```
<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
```

若要從9.6版或更新版本升級、您必須在具有優質授權的DAS SSD儲存設備上安裝支援功能ONTAP Select ONTAP Select 。此外、您必須在執行ONTAP 初始叢集安裝時勾選*啟用儲存效率*核取方塊、以使用「整合」功能進行叢集安裝。如果在未符合先前條件的情況下啟用AFF型特性設定ONTAP後升級、則需要手動建立開機引數和節點重新開機。如需詳細資訊、請聯絡技術支援部門。

儲存效率組態**ONTAP Select**

下表摘要列出各種可用的儲存效率選項、預設為啟用、或預設為未啟用、但建議使用、視媒體類型和軟體授權而定。

功能**ONTAP Select**	**DAS SSD**（優質或優質**xL1^**）	**DAS HDD**（所有授權）	**VNAS**（所有授權）
即時零偵測	是（預設）	是、由使用者以每個磁碟區為基礎來啟用	是、由使用者以每個磁碟區為基礎來啟用
Volume內嵌重複資料刪除技術	是（預設）	無法使用	不支援
32K即時壓縮（二線壓縮）	是、由使用者在每個磁碟區基礎上啟用。	是、由使用者以每個磁碟區為基礎來啟用	不支援
8K內嵌壓縮（調適性壓縮）	是（預設）	是、由使用者在每個磁碟區基礎上啟用	不支援
背景壓縮	不支援	是、由使用者在每個磁碟區基礎上啟用	是、由使用者以每個磁碟區為基礎來啟用
壓縮掃描儀	是的	是的	是、由使用者以每個磁碟區為基礎來啟用
即時資料精簡	是（預設）	是、由使用者在每個磁碟區基礎上啟用	不支援

功能**ONTAP Select**	**DAS SSD**（優質或優質**xL1^**）	**DAS HDD**（所有授權）	**VNAS**（所有授權）
壓實掃描器	是的	是的	不支援
集合體即時重複資料刪除技術	是（預設）	不適用	不支援
Volume背景重複資料刪除技術	是（預設）	是、由使用者在每個磁碟區基礎上啟用	是、由使用者以每個磁碟區為基礎來啟用
Aggregate背景資料刪除技術	是（預設）	不適用	不支援

[small）1^ONTAP Select 9.6支援新的授權（優質XL）和新的VM大小（大型）。不過、大型VM僅支體RAID的DAS組態。9.6 ONTAP Select版的大型版不支援硬體RAID和vNAS組態。

DAS SSD組態的升級行為注意事項

升級ONTAP Select 至32・6或更高版本後，請等待 `system node upgrade-revert show` 命令指出升級已完成、然後再驗證現有磁碟區的儲存效率值。

在升級至ONTAP Select S菜 系9.6或更新版本的系統上、在現有集合體或新建立的集合體上建立的新磁碟區、其行為與在全新部署中建立的磁碟區相同。進行ONTAP Select 程式碼更新的現有磁碟區、與新建立的磁碟區具有大部分相同的儲存效率原則、但有一些差異：

案例**1**

如果在升級之前未在磁碟區上啟用任何儲存效率原則、則：

- Volume 搭配 `space guarantee = volume` 不啟用即時資料壓縮、彙總即時重複資料刪除及彙總背景重複資料刪除。這些選項可在升級後啟用。
- Volume 搭配 `space guarantee = none` 未啟用背景壓縮。此選項可在升級後啟用。
- 現有磁碟區上的儲存效率原則會設定為在升級後自動執行。

案例 **2**

如果在升級之前、某個磁碟區已啟用部分儲存效率、則：

- Volume 搭配 `space guarantee = volume` 升級後沒有任何差異。
- Volume 搭配 `space guarantee = none` 開啟 Aggregate 背景重複資料刪除功能。
- Volume 搭配 `storage policy inline-only` 將其原則設為自動。
- 具有使用者定義儲存效率原則的磁碟區在原則上沒有變更、但磁碟區除外 `space guarantee = none`。這些磁碟區已啟用Aggregate背景重複資料刪除功能。

網路

網路：一般概念與特性

首先熟悉ONTAP Select 適用於整個過程的一般網路概念。然後探索單一節點和多節點叢集的特定特性和選項。

實體網路

實體網路主要透過ONTAP Select 提供底層的第二層交換基礎架構、來支援叢集式部署。與實體網路相關的組態包括Hypervisor主機和更廣泛的交換式網路環境。

主機NIC選項

每ONTAP Select 個物件Hypervisor主機都必須設定兩個或四個實體連接埠。您選擇的確切組態取決於多種因素、包括：

- 叢集是否包含一個或多ONTAP Select 個支援主機
- 所使用的Hypervisor作業系統
- 虛擬交換器的設定方式
- LACP是否與鏈路一起使用

實體交換器組態

您必須確定實體交換器的組態支援ONTAP Select 進行物件部署。實體交換器與Hypervisor型虛擬交換器整合。您選擇的確切組態取決於多種因素。主要考量包括：

- 您要如何在內部和外部網路之間保持區隔？
- 您是否會在資料與管理網路之間保持區隔？
- 如何設定第二層VLAN？

邏輯網路

使用兩個不同的邏輯網路、根據類型來分隔流量。ONTAP Select具體而言、流量可在叢集內的主機之間、以及儲存用戶端和叢集外的其他機器之間流通。Hypervisor管理的虛擬交換器有助於支援邏輯網路。

內部網路

透過多節點叢集部署、個別ONTAP Select 的元集節點會使用隔離的「內部」網路進行通訊。此網路不會在ONTAP Select 叢集節點外曝光或可用。

> ⓘ 內部網路僅存在於多節點叢集。

內部網路具有下列特性：

- 用於處理ONTAP 叢集內的各種流量、包括：
  - 叢集
  - 高可用度互連（HA-IC）
  - RAID同步鏡射（RSM）
- 單一第二層網路、以VLAN為基礎
- 靜態IP位址是由ONTAP Select 下列項目指派：
  - 僅限IPV4
  - 未使用DHCP

- 連結本機位址

- MTU大小預設為9000位元組、可在7500-9000範圍內調整（含）

外部網路

外部網路負責處理ONTAP Select VMware叢集節點與外部儲存用戶端及其他機器之間的流量。外部網路是每個叢集部署的一部分、具有下列特性：

- 用於處理ONTAP 不順暢的流量、包括：
  - 資料（NFS、CIFS、iSCSI）
  - 管理（叢集與節點；選用SVM）
  - 叢集間（選用）
- （可選）支援VLAN：
  - 資料連接埠群組
  - 管理連接埠群組
- 根據管理員的組態選項指派的IP位址：
  - IPV4或IPV6
- MTU大小預設為1500位元組（可調整）

外部網路存在各種規模的叢集。

虛擬機器網路環境

Hypervisor主機提供多項網路功能。

透過虛擬機器展示下列功能：ONTAP Select

虛擬機器連接埠

有多個連接埠ONTAP Select 可供使用。它們是根據指派和使用的包括叢集大小在內的幾個因素。

虛擬交換器

Hypervisor 環境中的虛擬交換器軟體、無論是 vSwitch （ VMware ）或開啟 vSwitch （ KVM ）、將虛擬機器所公開的連接埠與實體乙太網路連接 NIC 連接埠。您必須為每個 ONTAP Select 主機設定適當的 vSwitch 環境。

單節點和多節點網路組態

支援單一節點和多節點網路組態。ONTAP Select

單節點網路組態

單節點ONTAP Select 的不需要ONTAP 使用內部的不含功能性網路、因為沒有叢集、HA或鏡射流量。

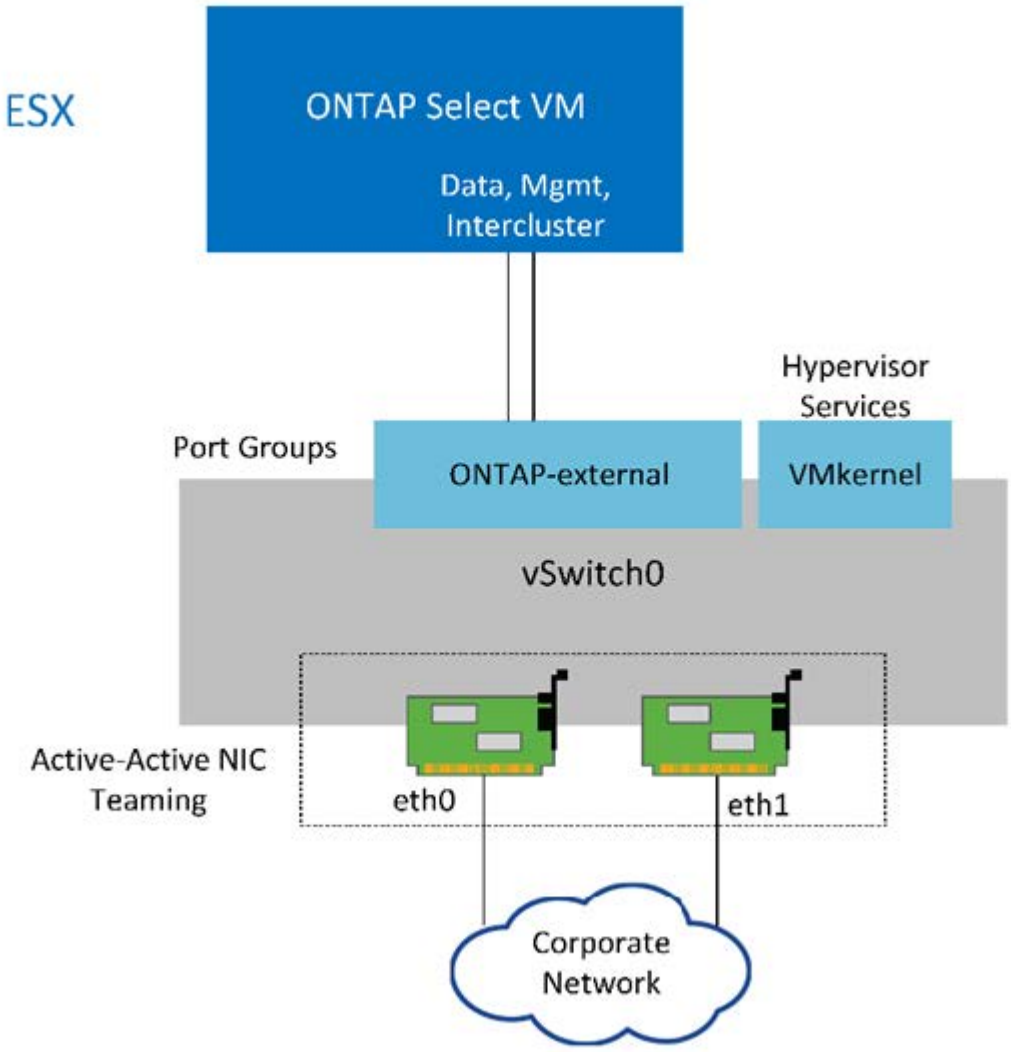
不同於ONTAP Select 多節點版本的產品、每ONTAP Select 個版本的不只包含三個虛擬網路介面卡、可呈現

給ONTAP 各個網路連接埠e0a、e0b和e0c。

這些連接埠可用來提供下列服務：管理、資料及叢集間的LIF。

下圖顯示這些連接埠與基礎實體介面卡之間的關係、說明ONTAP Select ESX Hypervisor上的一個VMware叢集節點。

單節點**ONTAP Select** 的網路組態不完整叢集



雖然兩個介面卡足以容納單節點叢集、但仍需要NIC群組。

LIF指派

如本文的「多節點LIF指派」一節所述、ONTAP Select 由NetApp使用IPspaces來將叢集網路流量與資料和管理流量分開。此平台的單節點機型不含叢集網路。因此、叢集IPspace中沒有連接埠。

叢集與節點管理生命期會在ONTAP Select 執行叢集功能時自動建立。其餘的生命階段可在部署後建立。

管理與資料生命量（**e0a**、**e0b**和**e0c**）

將連接埠e0a、e0b和e0c委派為裝載下列流量類型之LIF的候選連接埠：ONTAP

- SAN/NAS傳輸協定流量（CIFS、NFS和iSCSI）
- 叢集、節點和SVM管理流量
- 叢集間流量（SnapMirror和SnapVault

多節點網路組態

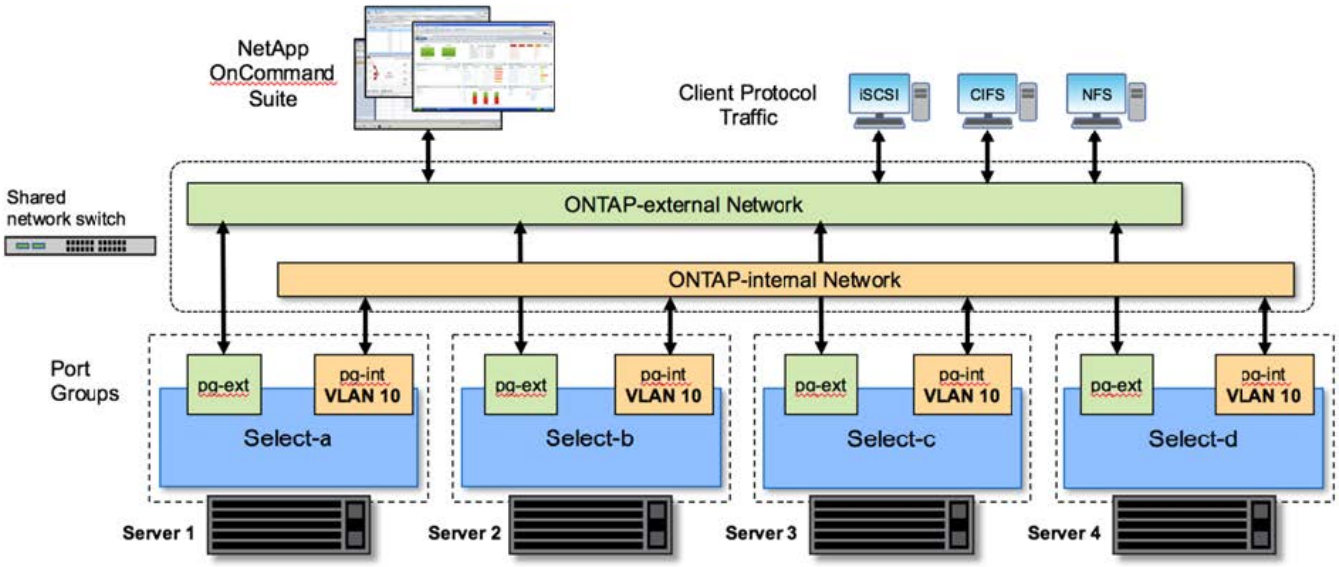
多節點ONTAP Select 支援的網路組態包含兩個網路。

這些是內部網路、負責提供叢集與內部複寫服務、以及負責提供資料存取與管理服務的外部網路。在這兩個網路中、流量的端點對端點隔離對於建立適合叢集恢復的環境來說、極為重要。

下圖顯示這些網路、顯示ONTAP Select 在VMware vSphere平台上執行的四節點VMware叢集。六節點和八節點叢集的網路配置相似。

每ONTAP Select 個實體執行個體都位於獨立的實體伺服器上。內部和外部流量會使用個別的網路連接埠群組加以隔離、這些群組會指派給每個虛擬網路介面、並允許叢集節點共用相同的實體交換器基礎架構。

- ONTAP Select 《Overview of an Singfingmultinoder cluster network configuration》（介紹一個不支持的多節點叢集



每ONTAP Select 個虛擬機器包含七個虛擬網路介面卡、以ONTAP 七個網路連接埠（e0a到e0g）的形式呈現給整個系統。雖然此類介面卡被視為實體NIC、但實際上它們是虛擬的、並透過虛擬化的網路層對應到一組實體介面。ONTAP因此、每個託管伺服器不需要六個實體網路連接埠。

不支援將虛擬網路介面卡新增ONTAP Select 至不支援的虛擬機器。

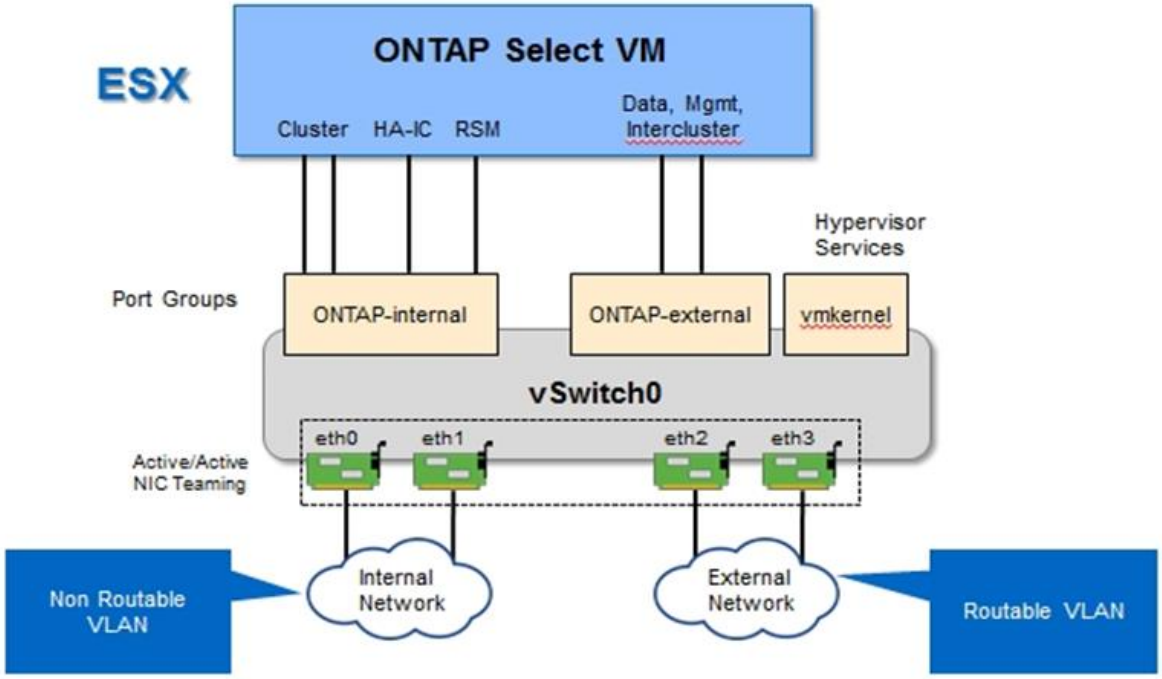
這些連接埠已預先設定為提供下列服務：

- e0a、e0b和e0g。管理與資料生命量
- e0c、e0d。叢集網路生命里數
- e0e 。RSM
- e0f 。HA互連

連接埠e0a、e0b和e0g位於外部網路上。雖然透過e0f的連接埠e0c執行數種不同的功能、但它們會共同組成內部Select網路。在制定網路設計決策時、這些連接埠應放置在單一第2層網路上。不需要將這些虛擬介面卡分隔成不同的網路。

下圖說明這些連接埠與基礎實體介面卡之間的關係、其中說明ONTAP Select ESX Hypervisor上的一個VMware叢集節點。

*多節點ONTAP Select 間單一節點的網路組態、是多節點間叢集*的一部分



將不同實體NIC之間的內部和外部流量隔離、可避免因網路資源存取不足而導致系統延遲。此外、透過NIC群組進行彙總、可確保單一網路卡故障不會妨礙ONTAP Select 到該叢集節點存取各自的網路。

請注意、外部網路和內部網路連接埠群組均以對稱的方式包含所有四個NIC介面卡。外部網路連接埠群組中的作用中連接埠是內部網路中的待命連接埠。相反地、內部網路連接埠群組中的作用中連接埠是外部網路連接埠群組中的待命連接埠。

LIF指派

隨着IPspaces的推出、ONTAP 不再使用連接埠角色。如同FAS 支援的不二陣列、ONTAP Select 叢集同時包含預設的IPspace和叢集IPspace。透過將網路連接埠e0a、e0b和e0g放入預設IPspace、然後將連接埠e0c和e0d放入叢集IPspace、這些連接埠基本上已被關閉、無法託管不屬於的生命週年。透過自動指派提供內部服務的介面、即可使用叢集中的其餘連接埠ONTAP Select 。它們不會像ONTAP 透過RSM和HA互連介面那樣透過這個解決方法暴露。

並非所有的生命都能透過ONTAP flexible命令Shell看到。HA互連和RSM介面不ONTAP 受支援、可在內部使用、以提供各自的服務。

網路連接埠和LIF會在下列各節中詳細說明。

管理和資料生命（ **e0a** 、 **e0b** 和 **e0g** ）

ONTAP 連接埠 e0a 、 e0b 和 e0g 會委派為可承載下列類型流量之生命的候選連接埠：

- SAN/NAS傳輸協定流量（CIFS、NFS和iSCSI）
- 叢集、節點和SVM管理流量
- 叢集間流量（SnapMirror和SnapVault

叢集與節點管理生命期會在ONTAP Select 執行叢集功能時自動建立。其餘的生命階段可在部署後建立。

叢集網路**lifs**（**e0c**、**e0d**）

將連接埠e0c和e0d委派為叢集介面的主連接埠。ONTAP在ONTAP Select 每個叢集節點中、ONTAP 使用連結本機IP位址（169.254.x.x）在設定過程中、會自動產生兩個叢集介面。

這些介面無法指派靜態IP位址、也不應建立其他叢集介面。

叢集網路流量必須流經低延遲、非路由的第2層網路。由於叢集處理量和延遲需求、ONTAP Select 所以不希望將此支援叢集實際放置在鄰近位置（例如、多套件、單一資料中心）。不支援跨越WAN或顯著地理距離、建立四節點、六節點或八節點的延伸叢集組態。支援使用中介器的延伸雙節點組態。

如需詳細資訊、請參閱一節 "雙節點延伸HA MetroCluster （簡稱「架構SDS」）最佳實務做法"。

為了確保叢集網路流量的最大處理量、此網路連接埠設定為使用巨型框架（7500至9000 MTU）。為確保叢集正常運作、請確認所有上游虛擬交換器和實體交換器上已啟用巨型框架、這些交換器可為ONTAP Select 叢集節點提供內部網路服務。

RAID SyncMirror 數據傳輸（**e0e**）

使用位於網路連接埠e0e的內部網路介面、在HA合作夥伴節點之間同步複寫區塊。此功能會在ONTAP 叢集設定期間使用由現象所設定的網路介面自動執行、而且管理員不需要進行任何組態設定。

連接埠e0e是ONTAP 由內部複寫流量使用的功能保留。因此、連接埠和裝載的LIF都不會顯示在ONTAP 功能區CLI或系統管理程式中。此介面設定為使用自動產生的連結本機IP位址、不支援重新指派替代IP位址。此網路連接埠需要使用巨型框架（7500至9000 MTU）。

HA互連（**e0f**）

NetApp FAS 產品組合使用專業硬體、在ONTAP 叢集中的HA配對之間傳遞資訊。不過、軟體定義環境通常不會提供這類設備（例如InfiniBand或iWARP裝置）、因此需要替代解決方案。儘管考慮到多種可能性、ONTAP 但互連傳輸上的需求卻要求在軟體中模擬此功能。因此ONTAP Select 、在一個不支援的叢集內、HA互連（傳統由硬體提供）的功能是以乙太網路作為傳輸機制、設計成作業系統。

每ONTAP Select 個節點均設定HA互連連接埠e0f。此連接埠主控HA互連網路介面、負責兩項主要功能：

- 在HA配對之間鏡射NVRAM的內容
- 在HA配對之間傳送/接收HA狀態資訊和網路活動訊息

HA互連流量會使用單一網路介面、在乙太網路封包內分層遠端直接記憶體存取（RDMA）框架、流經此網路連接埠。

以類似於RSM連接埠（e0e）的方式、使用ONTAP 者無論是從ESICLI或從System Manager、都看不到實體連接埠或代管網路介面。因此、無法修改此介面的IP位址、也無法變更連接埠的狀態。此網路連接埠需要使用巨型框架（7500至9000 MTU）。

內部和外部網路**ONTAP Select**

內部和外部網路的特性ONTAP Select 。

內部網路**ONTAP Select**

內部ONTAP Select 的不完整網路僅存在於產品的多節點機型中、負責提供ONTAP Select 含有叢集通訊、HA互連和同步複寫服務的不完整叢集。此網路包括下列連接埠和介面：

- * e0c、e0d.*託管叢集網路lifs
- * e0e.*託管RSMLIF
- 主控**HA**互連**LIF**的**e0f.**

此網路的處理量和延遲對於判斷ONTAP Select 叢集的效能和恢復能力至關重要。叢集安全性需要網路隔離、並確保系統介面與其他網路流量分開。因此ONTAP Select 、此網路必須專供該叢集使用。

不支援針對Select叢集流量以外的流量使用Select內部網路、例如應用程式或管理流量。在不支援內部VLAN的ONTAP 情況下、不可能有其他VM或主機。

流經內部網路的網路封包必須位於專屬的VLAN標記層2網路上。您可以完成下列其中一項工作來完成這項工作：

- 將VLAN標記的連接埠群組指派給內部虛擬NIC（e0c到e0f）（VST模式）
- 使用上游交換器所提供的原生VLAN、而原生VLAN不用於任何其他流量（指派沒有VLAN ID的連接埠群組、亦即、EST模式）

在任何情況下、內部網路流量的VLAN標記都是在ONTAP Select 不屬於該虛擬機器的情況下進行。

僅支援ESX標準和分散式vSwitch。不支援其他虛擬交換器或ESX主機之間的直接連線。內部網路必須完全開啟、不支援NAT或防火牆。

在一個叢集內ONTAP Select 、內部流量和外部流量會使用虛擬的第2層網路物件（稱為連接埠群組）加以分隔。正確指派這些連接埠群組的vSwitch非常重要、尤其是內部網路、負責提供叢集、HA互連和鏡像複寫服務。這些網路連接埠的網路頻寬不足、可能會導致效能降低、甚至影響叢集節點的穩定性。因此、四節點、六節點和八節點叢集需要內部ONTAP Select 的支援實體網路使用10Gb連線、不支援1Gb NIC。不過、外部網路可以做出取捨、因為將傳入資料的流程限制在ONTAP Select 某個叢集、並不會影響其可靠運作的能力。

雙節點叢集可使用四個1Gb連接埠進行內部流量、或使用單一10Gb連接埠、而非四節點叢集所需的兩個10Gb連接埠。在無法容納四張10Gb NIC卡的環境中、內部網路可使用兩張10Gb NIC卡、外部ONTAP 的支援網卡則可使用兩個1Gb NIC。

內部網路驗證與疑難排解

多節點叢集中的內部網路可使用網路連線檢查功能進行驗證。此功能可從執行的部署 CLI 中叫用 `network connectivity-check start` 命令。

執行下列命令以檢視測試的輸出：

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

此工具僅適用於對多節點選取叢集中的內部網路進行疑難排解。此工具不應用於疑難排解單節點叢集（包括vNAS組態）、ONTAP 將資源部署到ONTAP Select 非連接性、或用戶端連線問題。

叢集建立精靈（ONTAP 包含在「整合式部署GUI」中）包含內部網路檢查程式、做為建立多節點叢集期間可用的選用步驟。由於內部網路在多節點叢集中扮演重要角色、因此這項步驟成為叢集的一部分、因此建立工作流程可改善叢集建立作業的成功率。

從ONTAP 推出《2000年2月10日之際、內部網路所使用的MTU大小可設定在7、500到9、000之間。網路連線檢查工具也可用來測試7、500到9、000之間的MTU尺寸。預設的MTU值會設為虛擬網路交換器的值。如果環境中存在類似VXLAN的網路覆蓋、則必須以較小的值取代該預設值。

外部網路**ONTAP Select**

這個外部網路負責叢集的所有傳出通訊、因此存在於單節點和多節點組態上。ONTAP Select雖然此網路沒有內部網路嚴格定義的處理量需求、但系統管理員應注意、不要在用戶端ONTAP 與VMware VM之間建立網路瓶頸、因為效能問題可能會被誤認為ONTAP Select 是無法解決的問題。

> 外部流量可在vSwitch層（VST）和外部交換器層（EST）標記、方式與內部流量類似。此外、ONTAP Select 在稱為VGT的程序中、可透過不受影響的VM本身來標記外部流量。請參閱一節 "資料與管理流量分離" 以取得更多詳細資料。

下表重點說明ONTAP Select 內部和外部網路之間的主要差異。

內部與外部網路快速參考

說明	內部網路	外部網路
網路服務	叢集 HA/IC RAID SyncMirror （ RSM ）	資料管理 叢集間 （ SnapMirror 和 SnapVault ）
網路隔離	必要	選用
框架大小（MTU）	7、500至9、000	1 、 500 （預設） 9 、 000 （支援）
IP位址指派	自動產生	使用者定義
DHCP 支援	否	否

NIC群組

為了確保內部和外部網路具備必要的頻寬和恢復特性、以提供高效能和容錯能力、建議使用實體網路介面卡群組。支援具有單一10Gb連結的雙節點叢集組態。不過、NetApp建議的最佳實務做法是在ONTAP Select 內部和外部網路上使用NIC群組、

MAC位址產生

指派給ONTAP Select 所有「資訊區」網路連接埠的MAC位址會由隨附的部署公用程式自動產生。此公用程式使用NetApp專屬的平台專屬、組織專屬的識別碼（OUI）、以確保FAS 不會與任何系統發生衝突。然後、此位址的複本會儲存在ONTAP Select 內部資料庫中的安裝VM ONTAP （簡稱「整合」）中、以避免在未來的節點部署期間意外重新指派。管理員絕不能修改網路連接埠的指定MAC位址。

支援的網路組態

選取最佳的硬體、並設定網路以最佳化效能和恢復能力。

伺服器廠商瞭解客戶的需求和選擇各不相同、這點至關重要。因此、在購買實體伺服器時、在決定網路連線時、有許多選擇可供選擇。大多數的市售系統都隨附各種NIC選項、可提供單一連接埠和多連接埠選項、並可針對速度和處理量進行不同的排列。這包括支援 VMware ESX 的 25Gb/s 和 40Gb/s NIC 介面卡。

由於VMware的效能ONTAP Select 與基礎硬體的特性直接相關、因此選擇速度較快的NIC來增加VM的處理量、會導致叢集效能較高、整體使用者體驗也較佳。可使用四個10Gb NIC或兩個較高速NIC（25/40 GB/s）來實現高效能網路配置。還有許多其他組態也受到支援。對於雙節點叢集、支援4個1Gb連接埠或1個10Gb連接埠。對於單一節點叢集、支援2個1Gb連接埠。

網路最低和建議的組態

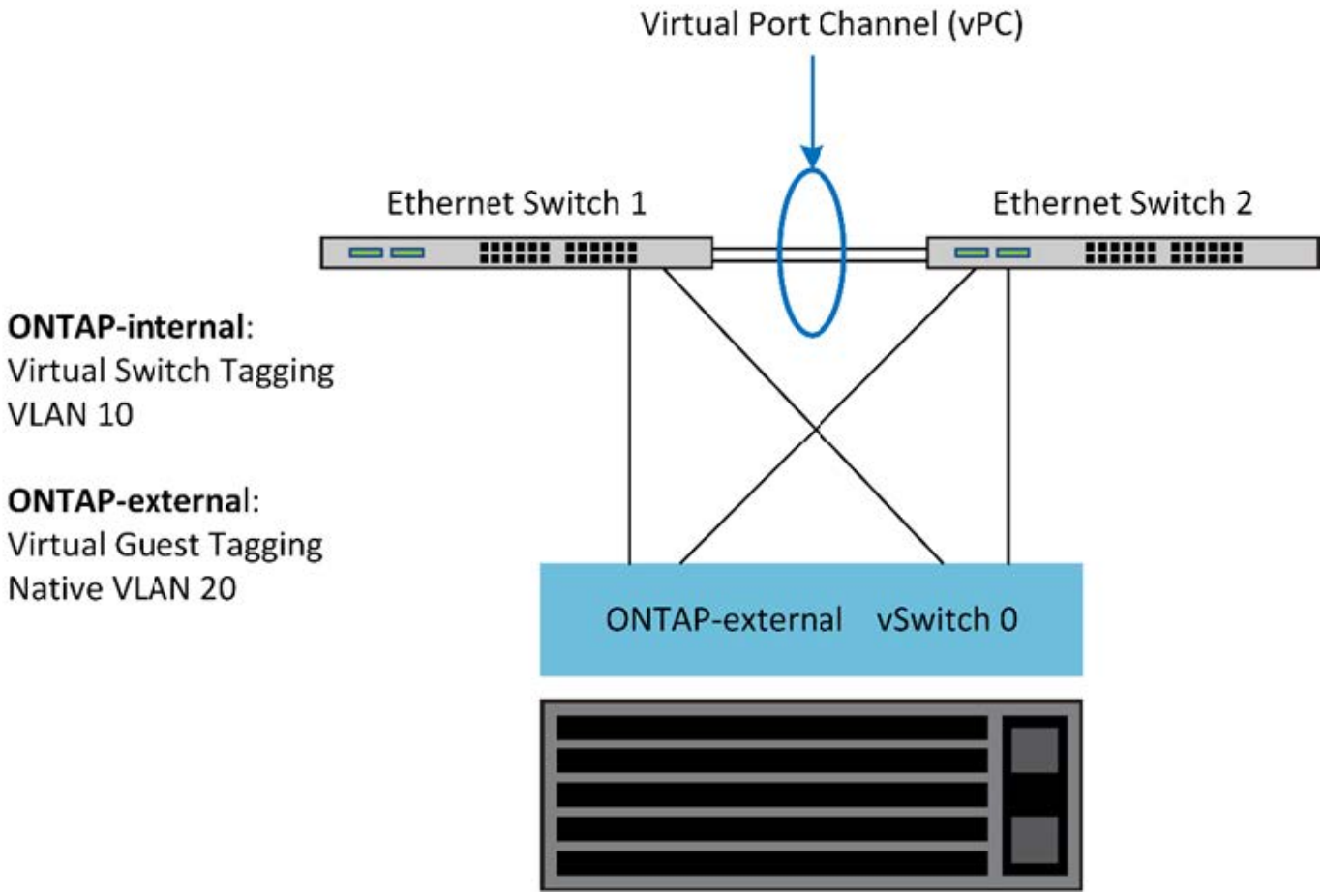
根據叢集大小、有多種支援的乙太網路組態。

叢集大小	最低要求	建議
單節點叢集	2 x 1GbE	2 個 10GbE
雙節點叢集或MetroCluster 不含任何功能的SDS	4 x 1GbE 或 1 x 10GbE	2 個 10GbE
4/6/8節點叢集	2 個 10GbE	4個10GbE或2個25/40GbE

由於可能需要在每個拓撲所需的不同 NIC 群組組態之間轉換、因此不支援在執行中叢集上的單一連結和多個連結拓撲之間進行轉換。

使用多個實體交換器進行網路組態

當有足夠的硬體可用時、NetApp建議使用下圖所示的多重交換器組態、因為新增了實體交換器故障防護功能。



VMware vSphere vSwitch 組態

適用於雙NIC和四NIC組態的VMware VSwitch組態和負載平衡原則。ONTAP Select

支援使用標準和分散式vSwitch組態。ONTAP Select分散式vSwitch支援連結集合架構（LACP）。連結集合體是一種通用的網路架構、用於彙總多個實體介面卡的頻寬。LACP是一種廠商中立的標準、可為網路端點提供開放式傳輸協定、將實體網路連接埠群組整合到單一邏輯通道中。可搭配設定為連結集合群組（LAG）的連接埠群組使用。ONTAP Select不過、NetApp建議使用個別實體連接埠作為簡單的上行（主幹）連接埠、以避免LAG組態。在這些情況下、標準和分散式VSwitches的最佳實務做法是相同的。

本節說明在雙NIC和四NIC組態中應使用的vSwitch組態和負載平衡原則。

設定ONTAP Select 要由VMware使用的連接埠群組時、應遵循下列最佳實務做法；連接埠群組層級的負載平衡原則是根據來源虛擬連接埠ID進行路由。VMware建議在連接至ESXi主機的交換器連接埠上、將STP設為PortFast。

所有vSwitch組態都需要至少兩個實體網路介面卡、並將其套裝到單一NIC群組中。支援雙節點叢集的單一10Gb連結。ONTAP Select不過、透過NIC集合來確保硬體備援是NetApp的最佳實務做法。

在vSphere伺服器上、NIC群組是用於將多個實體網路介面卡組合到單一邏輯通道的集合架構、可在所有成員連接埠之間共享網路負載。請務必記住、不需實體交換器的支援、就能建立NIC群組。負載平衡和容錯移轉原則可直接套用至NIC群組、而NIC群組並未察覺上游交換器組態。在此情況下、原則只會套用至傳出流量。

靜態連接埠通道不受ONTAP Select 支援。支援LACP的通道可支援分散式vSwitch、但使用LACP LAG可能會導致跨LAG成員的負載分配不均勻。

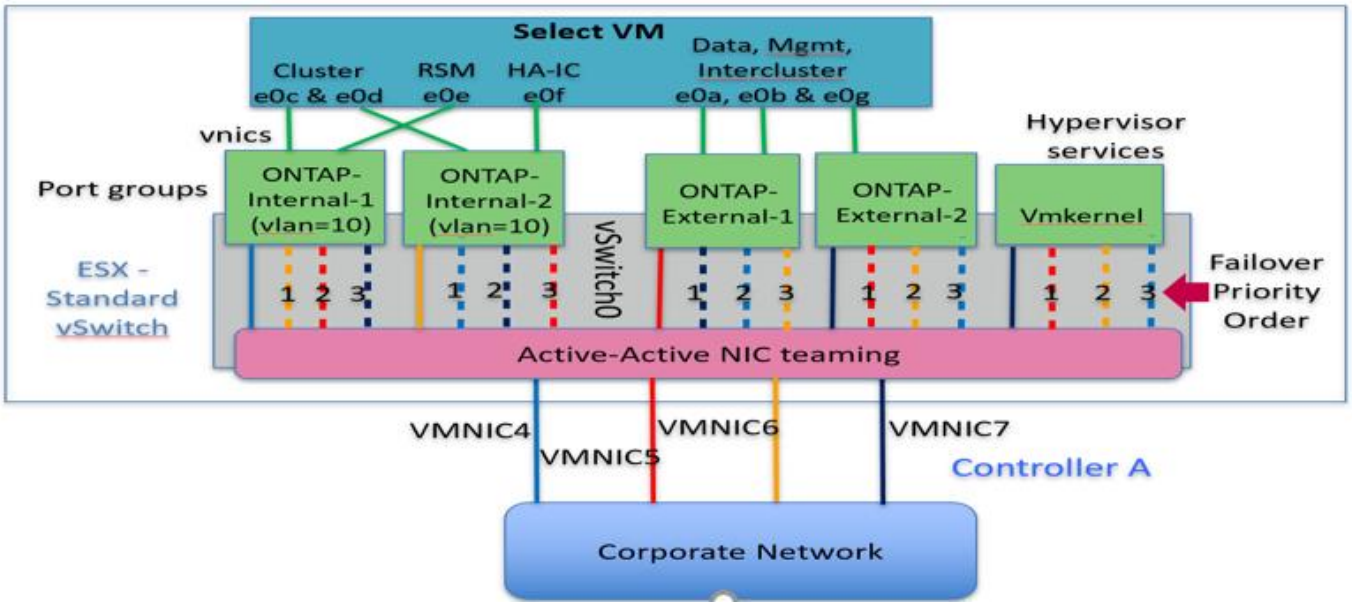
針對單一節點叢集、ONTAP 透過使用連接ONTAP Select 埠群組來設定適用於外部網路的Sfor VM、以及相同的連接埠群組、或選用不同的連接埠群組來處理叢集和節點管理流量。對於單一節點叢集、所需的實體連接埠數量可新增至外部連接埠群組作為作用中介面卡。

針對多節點叢集、ONTAP 將每ONTAP Select 個EfingVM設定為使用一或兩個連接埠群組用於內部網路、另一個或兩個連接埠群組用於外部網路。叢集和節點管理流量可以使用與外部流量相同的連接埠群組、也可以選擇使用個別的連接埠群組。叢集與節點管理流量無法與內部流量共用同一個連接埠群組。

標準或分散式**vSwitch**、以及每個節點四個實體連接埠

可以將四個連接埠群組指派給多節點叢集中的每個節點。每個連接埠群組都有一個作用中實體連接埠和三個備用實體連接埠、如下圖所示。

*每個節點有四個實體連接埠的vSwitch *



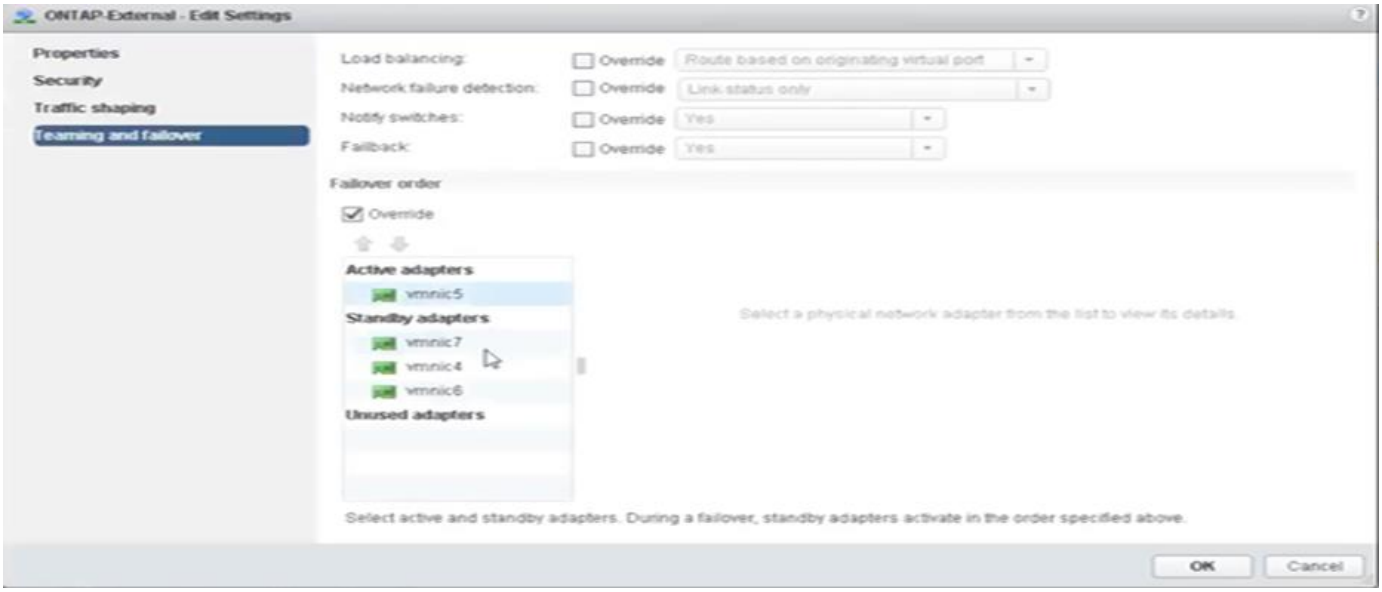
待機清單中連接埠的順序很重要。下表提供四個連接埠群組的實體連接埠發佈範例。

網路最低和建議的組態

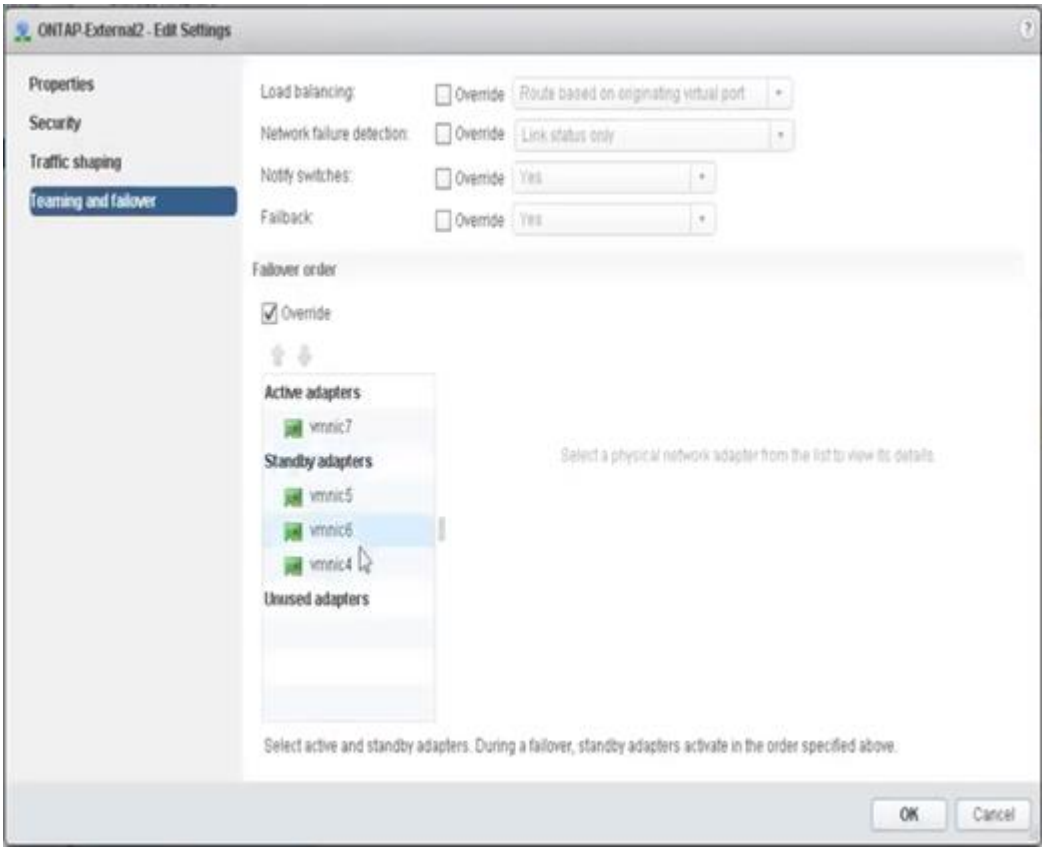
連接埠群組	外部**1.**	外部 **2.**	內部**1**	內部**2.**
使用中	vmnic0	vmnic1	vmnic2	vmnic3.
待機1	vmnic1	vmnic0	vmnic3.	vmnic2
待命2.	vmnic2	vmnic3.	vmnic0	vmnic1
待機3.	vmnic3.	vmnic2	vmnic1	vmnic0

下圖顯示來自vCenter GUI（ONTAP外部和ONTAP外部2）的外部網路連接埠群組組態。請注意、使用中的介面卡來自不同的網路卡。在此設定中、vmnic 4和vmnic 5是同一個實體NIC上的雙埠、而vmnic 6和vminc 7則是獨立NIC上的雙埠（本範例中不使用vnmics 0到3）。待命介面卡的順序可提供階層式容錯移轉、內部網路的連接埠則是最後一個。待命清單中的內部連接埠順序會在兩個外部連接埠群組之間進行類似的交換。

第**1**部分**ONTAP Select** ：外部連接埠群組組態



- 第 2 部分： ONTAP Select 外部連接埠群組組態 *



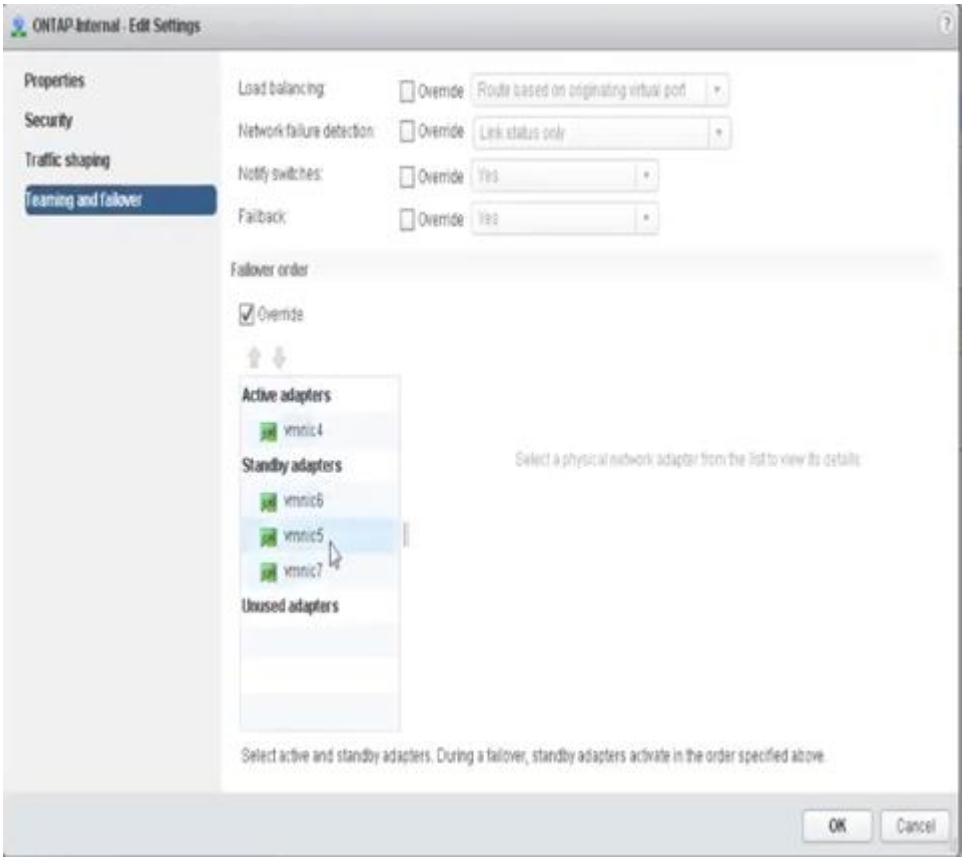
為了方便閱讀、指派內容如下：

ONTAP外部	**ONTAP-External2**
作用中介面卡： vmnic5 待命介面卡： vmnic7 、 vmnic4 、 vmnic6	作用中介面卡： vmnic7 待命介面卡： vmnic5 、 vmnic6 、 vmnic4

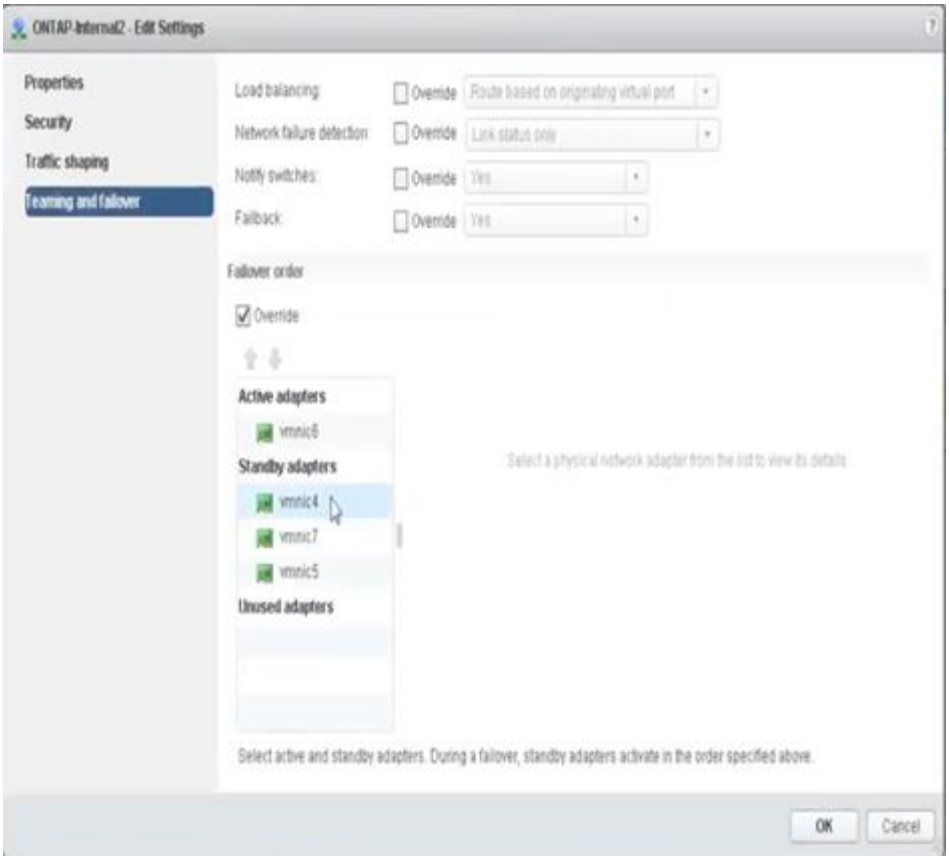
下圖顯示內部網路連接埠群組（ONTAP內部和ONTAP內部2）的組態。請注意、使用中的介面卡來自不同的網

路卡。在此設定中、vmnic 4和vmnic 5是相同實體ASIC上的雙埠、而vmnic 6和vmnic 7則是獨立ASIC上的雙埠。待命介面卡的順序可提供階層式容錯移轉、外部網路的連接埠則是最後一個。待機清單中的外部連接埠順序會在兩個內部連接埠群組之間進行類似的交換。

第**1**部分**ONTAP Select** ：內部連接埠群組組態



第**2**部分**ONTAP Select** ：內部連接埠群組



為了方便閱讀、指派內容如下：

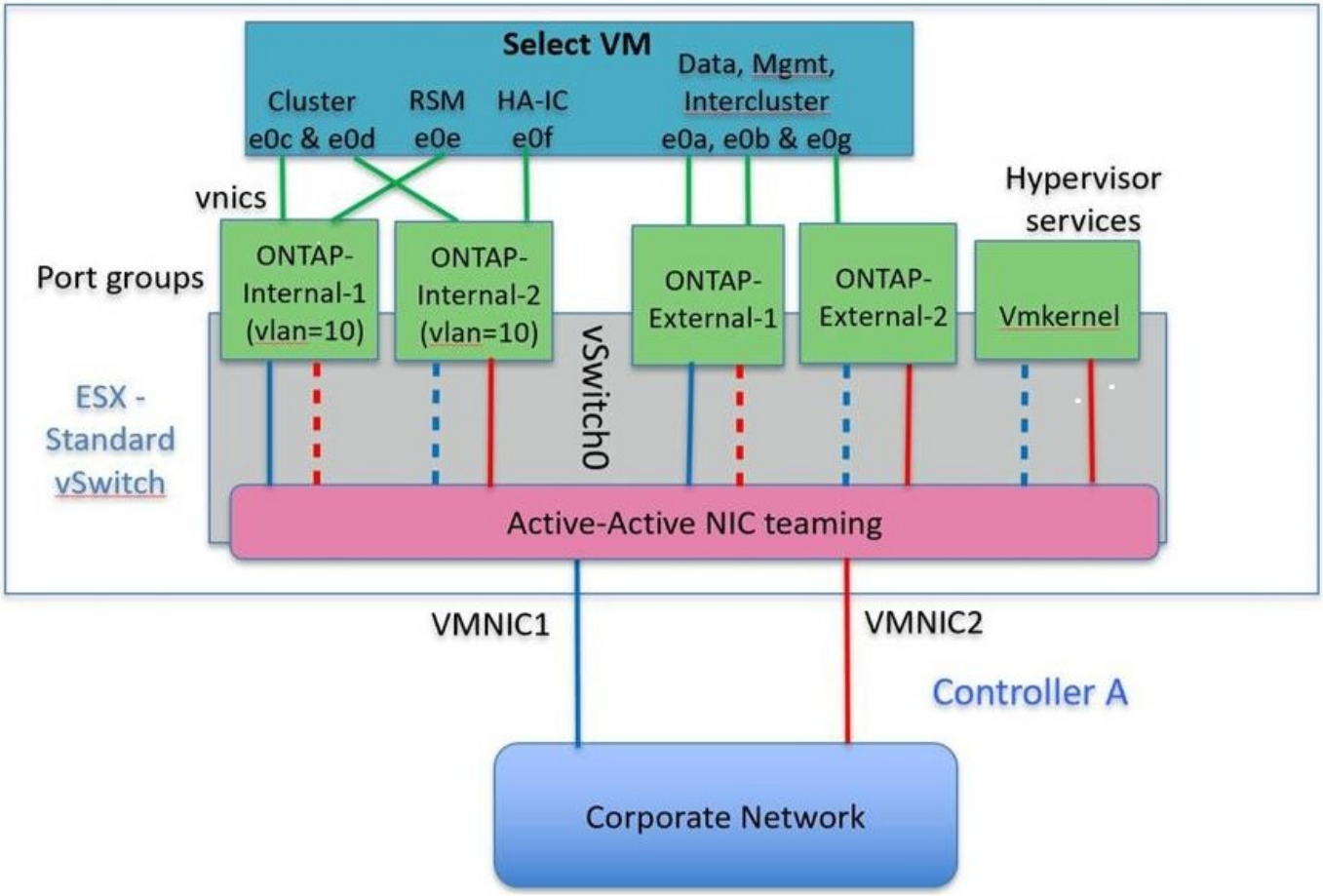
ONTAP內部	**ONTAP** 內部環境 **2.**
作用中介面卡： vmnic4 待命介面卡： vmnic6 、 vmnic5 、 vmnic7	作用中介面卡： vmnic6 待命介面卡： vmnic4 、 vmnic7 、 vmnic5

標準或分散式**vSwitch**、以及每個節點兩個實體連接埠

使用兩個高速（25/40Gb）NIC時、建議的連接埠群組組態在概念上與使用四個10Gb介面卡的組態非常類似。即使只使用兩個實體介面卡、仍應使用四個連接埠群組。連接埠群組指派如下：

連接埠群組	外部**1**（**e0a**、**e0b**）	內部**1**（**e0c**、**e0e**）	內部 **2** （ **e0d** 、 **e0f** ）	外部**2**（**e0g**）
使用中	vmnic0	vmnic0	vmnic1	vmnic1
待命	vmnic1	vmnic1	vmnic0	vmnic0

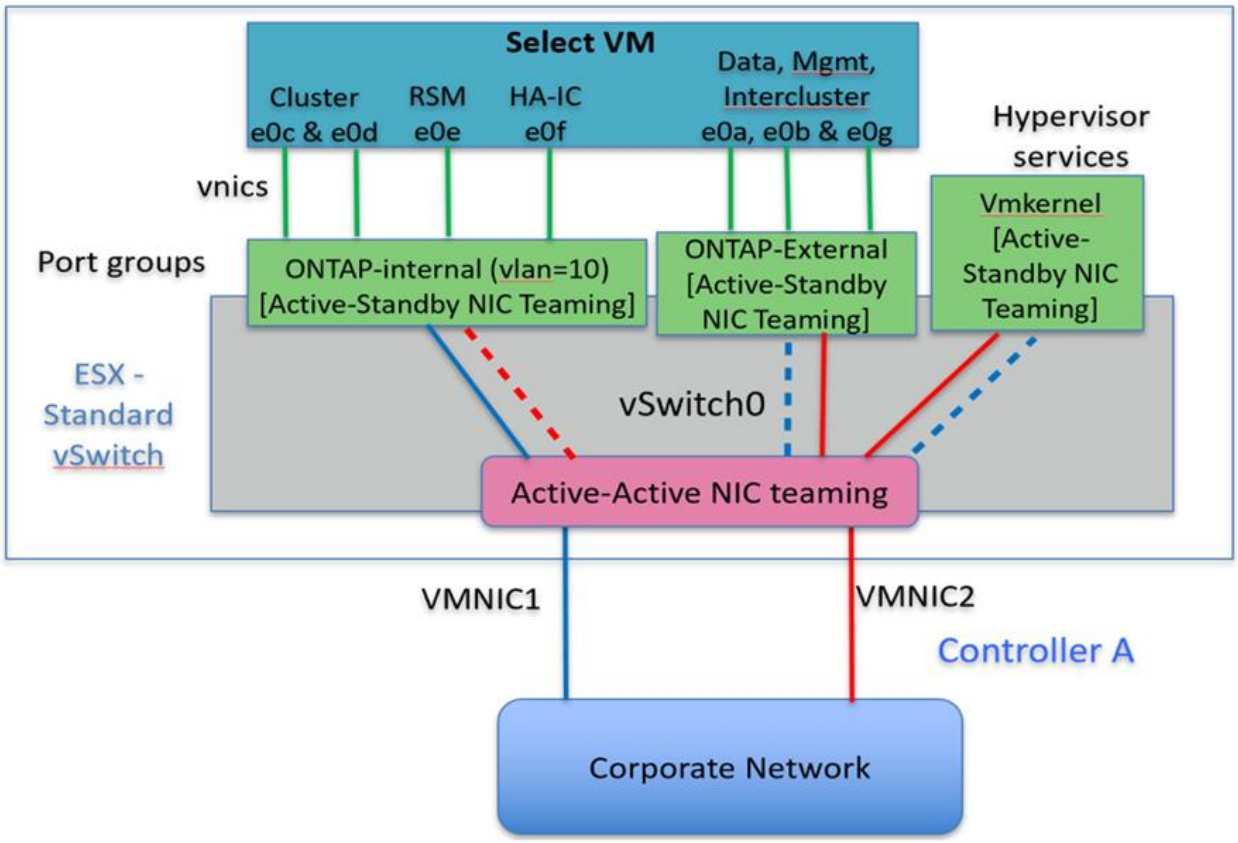
*每個節點有兩個高速（25/40Gb）實體連接埠*的vSwitch



使用兩個實體連接埠（10Gb或更少）時、每個連接埠群組都應該有一個作用中的介面卡和一個彼此相對設定的待命介面卡。內部網路僅適用於多節點ONTAP Select 的叢集。對於單節點叢集、兩個介面卡都可在外部連接埠群組中設定為作用中。

以下範例顯示vSwitch的組態、以及負責處理多節點ONTAP Select 間叢集內部與外部通訊服務的兩個連接埠群組。外部網路可以在網路中斷時使用內部網路vmnic、因為內部網路vmnics是此連接埠群組的一部分、而且設定為待命模式。外部網路的情況則相反。在兩個連接埠群組之間交換作用中和待命的vmnics、對於ONTAP Select在網路中斷期間適當地容錯移轉到各個虛擬機器來說、是非常重要的。

*每個節點有兩個實體連接埠（10Gb或更少）的vSwitch *

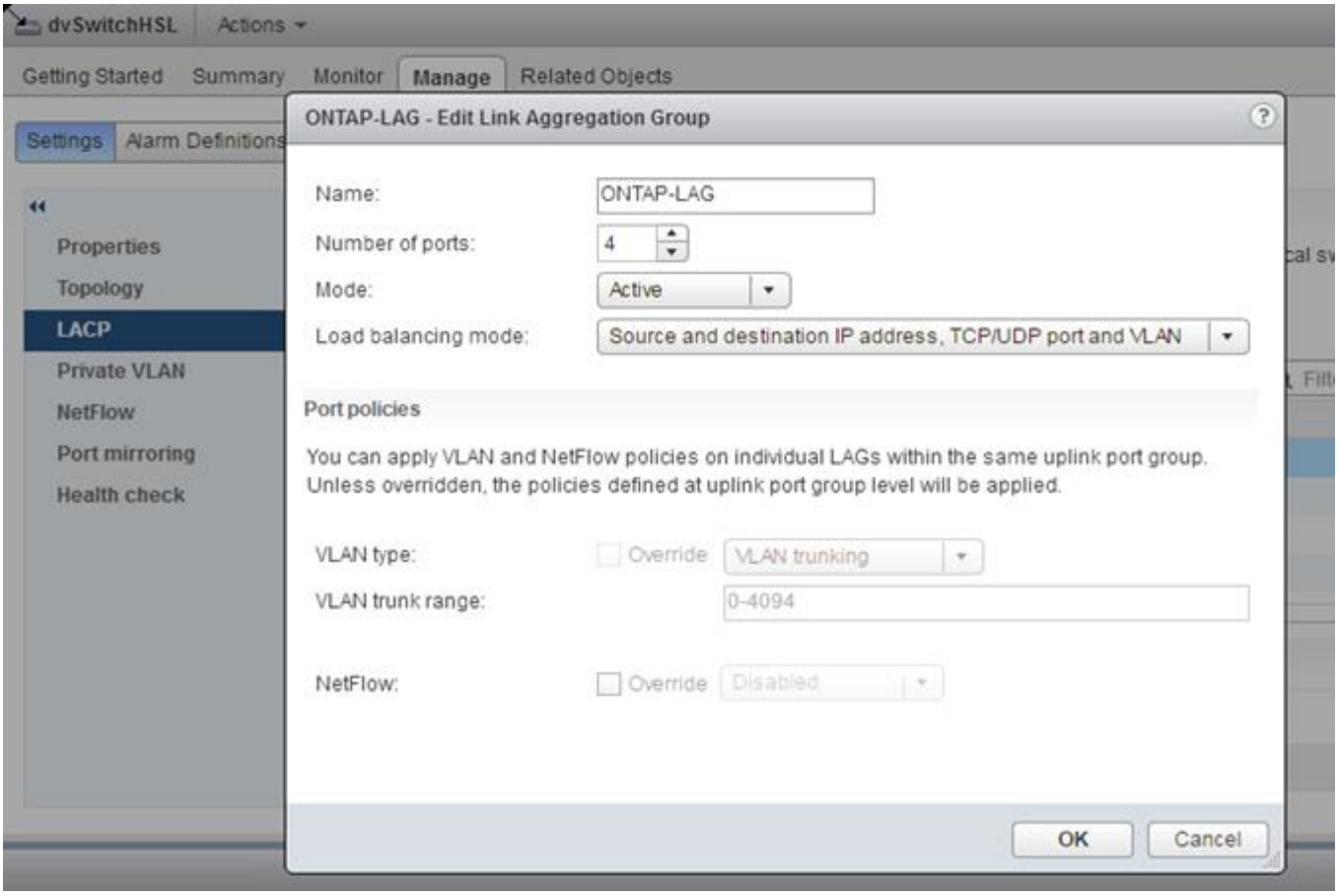


採用**LACP**的分散式**vSwitch**

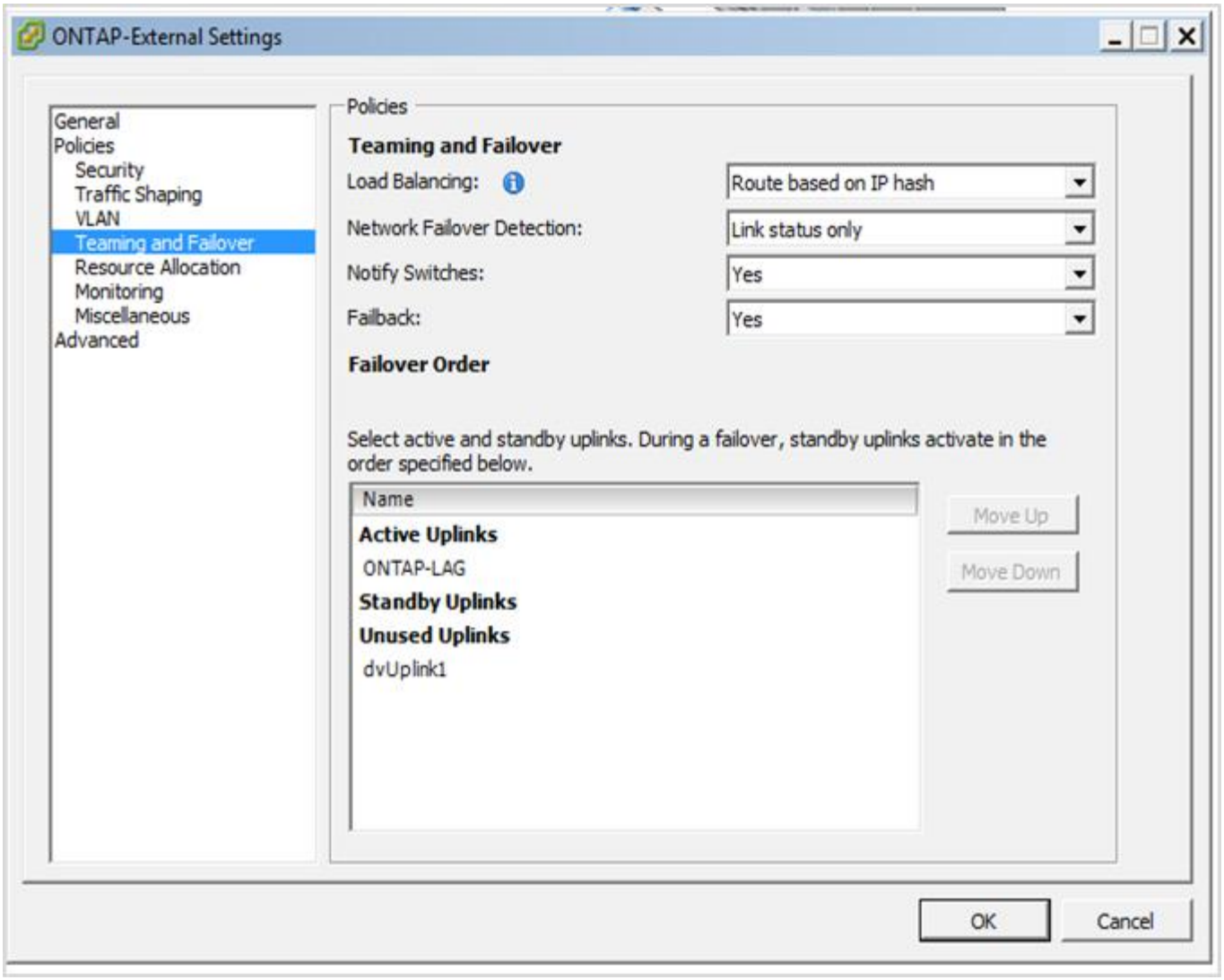
在組態中使用分散式vSwitch時、可使用LACP（雖然這不是最佳實務做法）來簡化網路組態。唯一支援的LACP組態需要所有的vmnics都位於單一LAG中。上行鏈路實體交換器必須在通道中的所有連接埠上支援7,500到9、000之間的MTU大小。內部ONTAP Select 和外部的不確定網路應該在連接埠群組層級隔離。內部網路應使用不可路由（隔離）的VLAN。外部網路可使用VST、EST或VGT。

下列範例顯示使用LACP的分散式vSwitch組態。

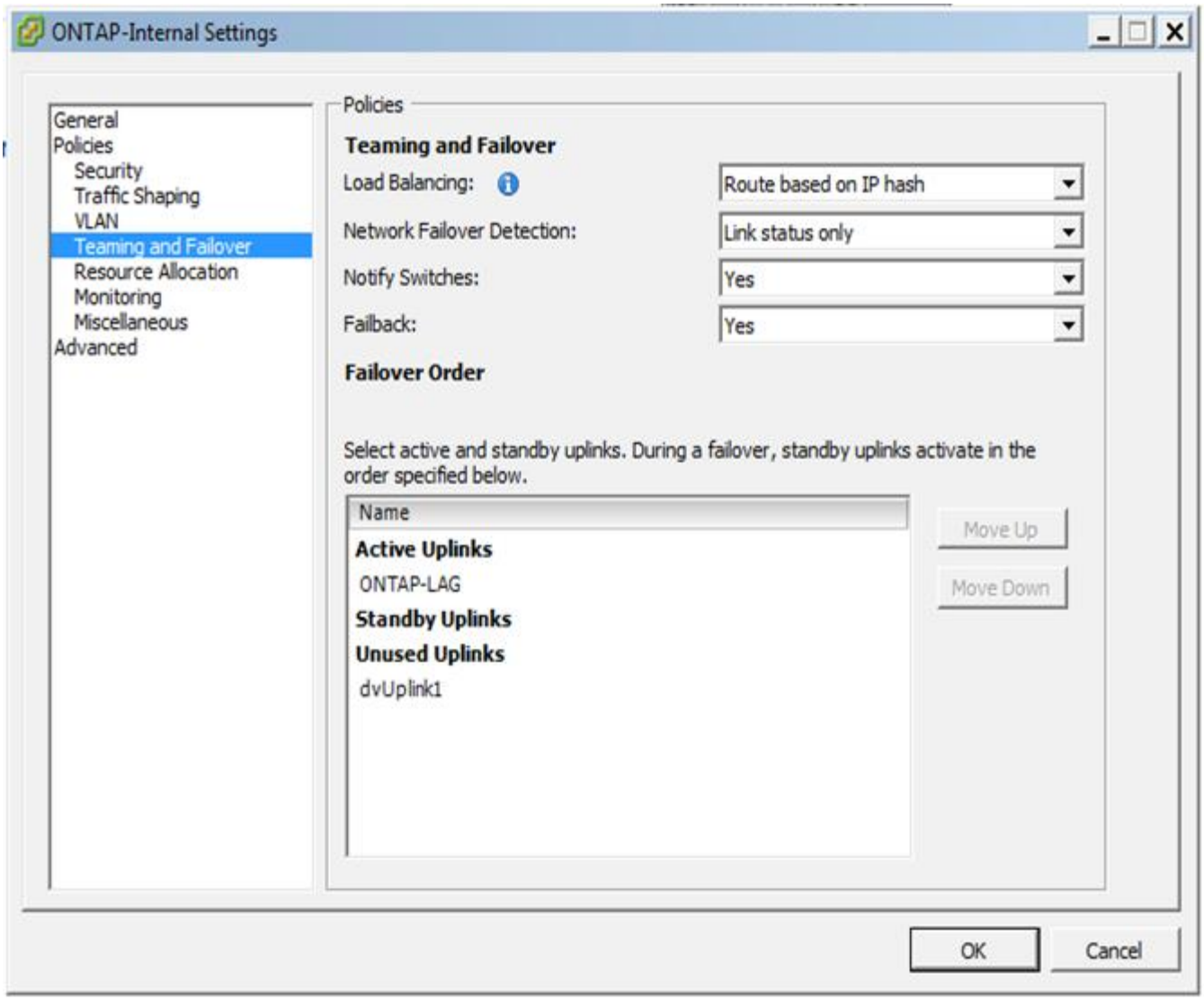
*使用LAPAC*時的LAG內容



*外部連接埠群組組態、使用已啟用LACP的分散式vSwitch *



*內部連接埠群組組態、使用已啟用LACP的分散式vSwitch *



LACP需要將上游交換器連接埠設定為連接埠通道。在分散式vSwitch上啟用此功能之前、請確定已正確設定啟用LACP的連接埠通道。

實體交換器組態

上游實體交換器組態詳細資料、以單一交換器和多重交換器環境為基礎。

在決定從虛擬交換器層到實體交換器之間的連線時、應謹慎考量。將內部叢集流量與外部資料服務區隔、應透過第2層VLAN提供的隔離功能、延伸到上游實體網路層。

實體交換器連接埠應設定為中繼連接埠。外部流量可透過兩種方式之一、在多個第2層網路之間分隔。ONTAP Select其中一種方法是使用ONTAP 含有VLAN標記的虛擬連接埠搭配單一連接埠群組。另一種方法是將VST模式中的個別連接埠群組指派給管理連接埠e0a。您也必須指派資料連接埠至e0b和e0c/e0g、視ONTAP Select 乎版本、單節點或多節點組態而定。如果外部流量是跨多個第2層網路分隔、則上行鏈路實體交換器連接埠的允許VLAN清單中應包含這些VLAN。

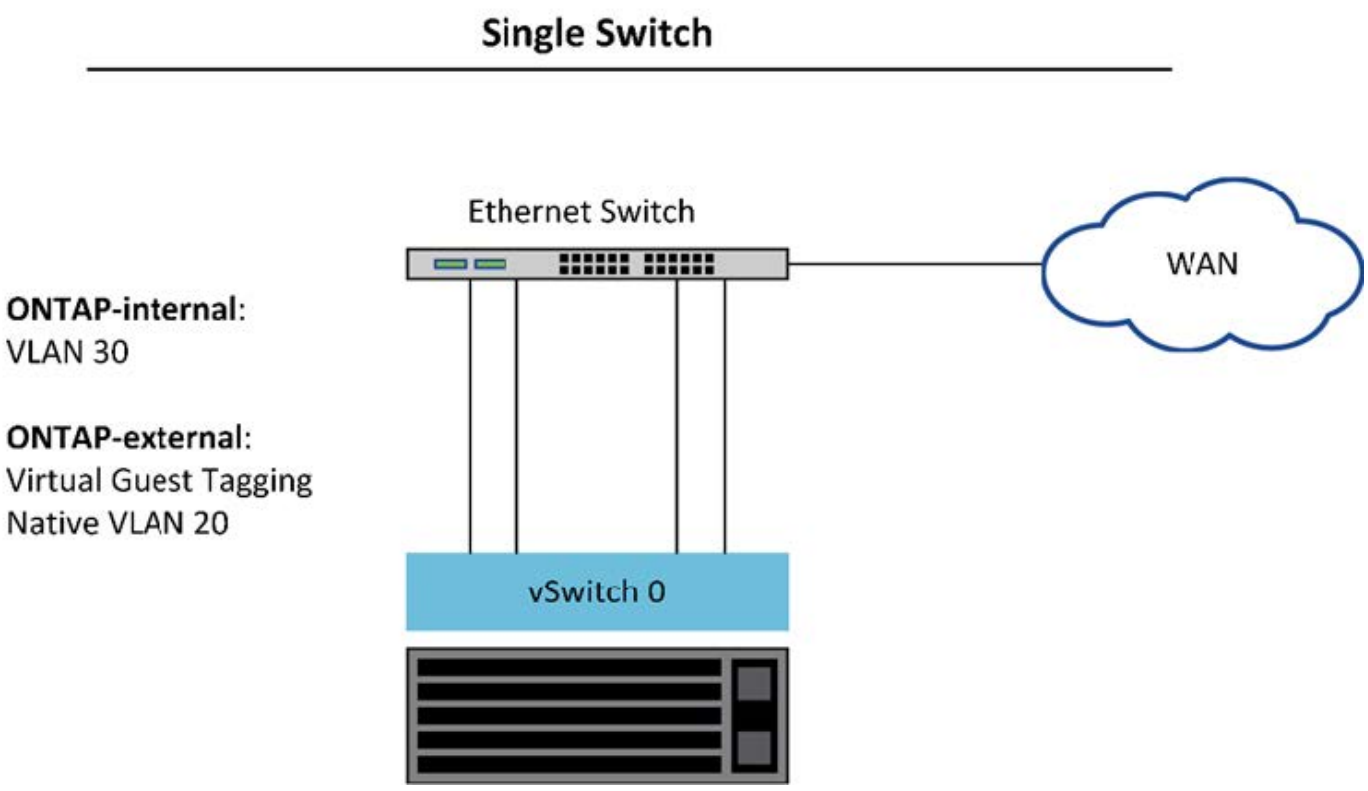
內部網路流量使用以連結本機IP位址定義的虛擬介面進行。ONTAP Select由於這些IP位址是不可路由的、因此叢集節點之間的內部流量必須流經單一第2層網路。不支援在叢集節點之間路由躍點ONTAP Select 。

共享實體交換器

下圖說明多節點ONTAP Select 間可能使用的交換器組態、在此範例中、裝載內部和外部網路連接埠群組的VSwitches所使用的實體NIC、會以纜線連接至同一個上游交換器。交換器流量會使用包含在不同VLAN中的廣播網域來隔離。

對於內部的網路、標記是在連接埠群組層級完成。ONTAP Select雖然下列範例使用VGT作為外部網路、但該連接埠群組同時支援VGT和VST。

使用共享實體交換器的網路組態

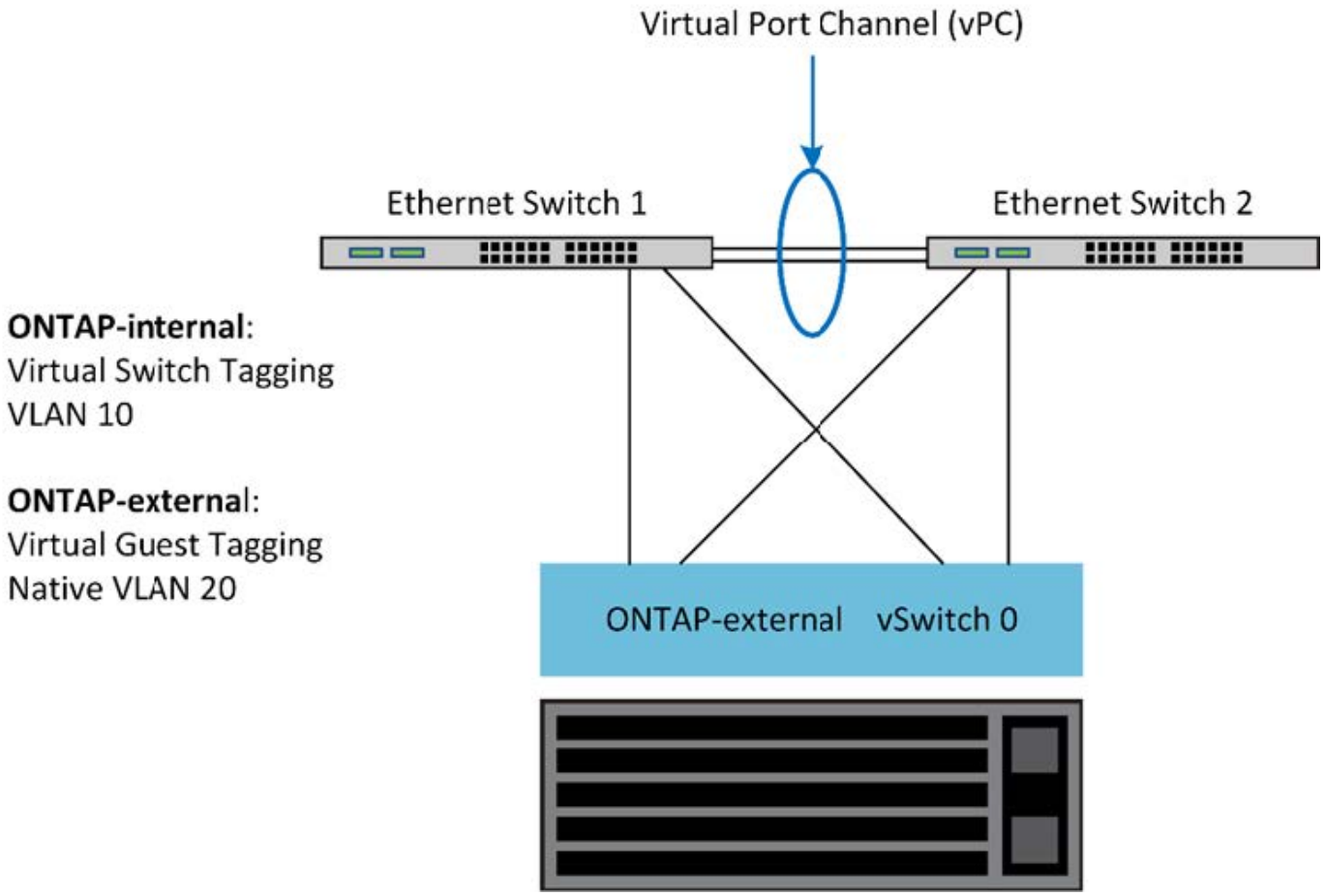


在此組態中、共享交換器會變成單點故障。如果可能、應使用多個交換器來防止實體硬體故障導致叢集網路中斷。

多個實體交換器

需要備援時、應使用多個實體網路交換器。下圖顯示多節點ONTAP Select 間的一個節點所使用的建議組態。內部和外部連接埠群組的NIC均連接至不同的實體交換器、保護使用者免於單一硬體交換器故障。在交換器之間設定虛擬連接埠通道、以避免產生跨距樹狀結構問題。

使用多個實體交換器進行網路組態



資料與管理流量分離

將資料流量和管理流量隔離到獨立的第2層網路。

外部網路流量的定義為資料（CIFS、NFS和iSCSI）、管理和複寫（SnapMirror）流量。ONTAP Select在一個叢集內ONTAP 、每種流量類型都會使用獨立的邏輯介面、而該介面必須裝載在虛擬網路連接埠上。在的多節點組態ONTAP Select 上、這些端口被指定為連接埠e0a和e0b/e0g。在單一節點組態中、這些連接埠會指定為e0a和e0b/e0c、而其餘連接埠則保留給內部叢集服務。

NetApp建議將資料流量和管理流量隔離到獨立的第2層網路。在這個環境中、這是使用VLAN標記來完成。ONTAP Select這可透過將VLAN標記的連接埠群組指派給網路介面卡1（連接埠e0a）進行管理流量來達成。然後、您可以將個別的連接埠群組指派給連接埠e0b和e0c（單一節點叢集）、以及e0b和e0g（多節點叢集）以進行資料流量。

如果本文前面所述的VST解決方案不足、則可能需要將資料與管理lifs分成一組、放在同一個虛擬連接埠上。若要這麼做、請使用稱為VGT的程序、VM會在其中執行VLAN標記。

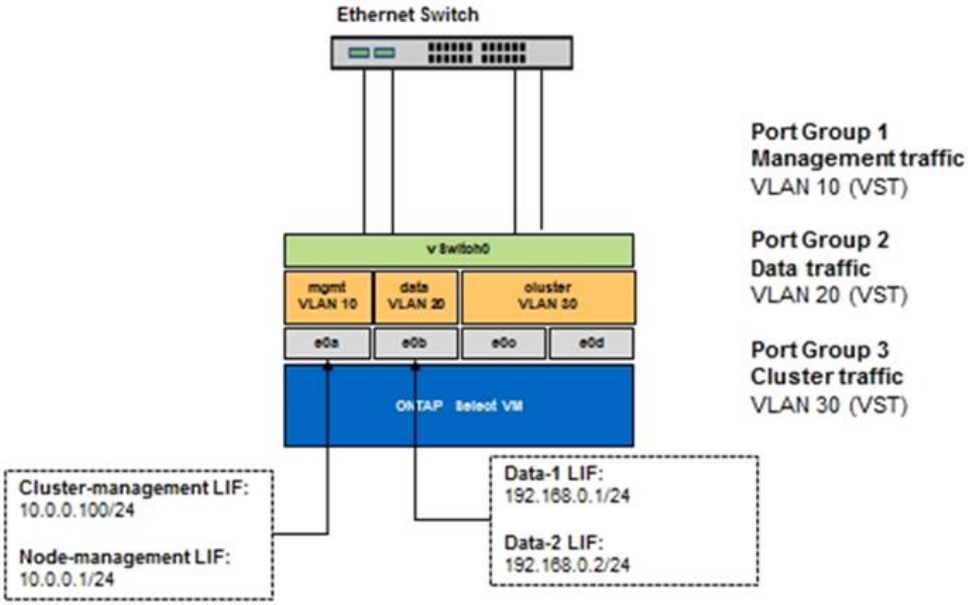
使用ONTAP 此功能時、無法使用透過VGT分隔資料與管理網路。此程序必須在叢集設定完成後執行。

使用VGT和雙節點叢集時、還有額外的警告。在雙節點叢集組態中、節點管理IP位址用於在ONTAP 完全可用之前建立與中介器的連線。因此、對應至節點管理LIF（連接埠e0a）的連接埠群組僅支援EST和VST標記。此外、如果管理和資料流量都使用相同的連接埠群組、則整個雙節點叢集僅支援EST/VST。

同時支援VST和VGT兩種組態選項。下圖顯示第一個案例VST、其中流量會透過指派的連接埠群組在vSwitch層

標記。在此組態中、叢集和節點管理生命體會指派ONTAP 至連接埠e0a、並透過指派的連接埠群組以VLAN ID 10標記。資料生命週期會指派給連接埠e0b和e0c或e0g、並使用第二個連接埠群組指派給VLAN ID 20。叢集連接埠使用第三個連接埠群組、且位於VLAN ID 30上。

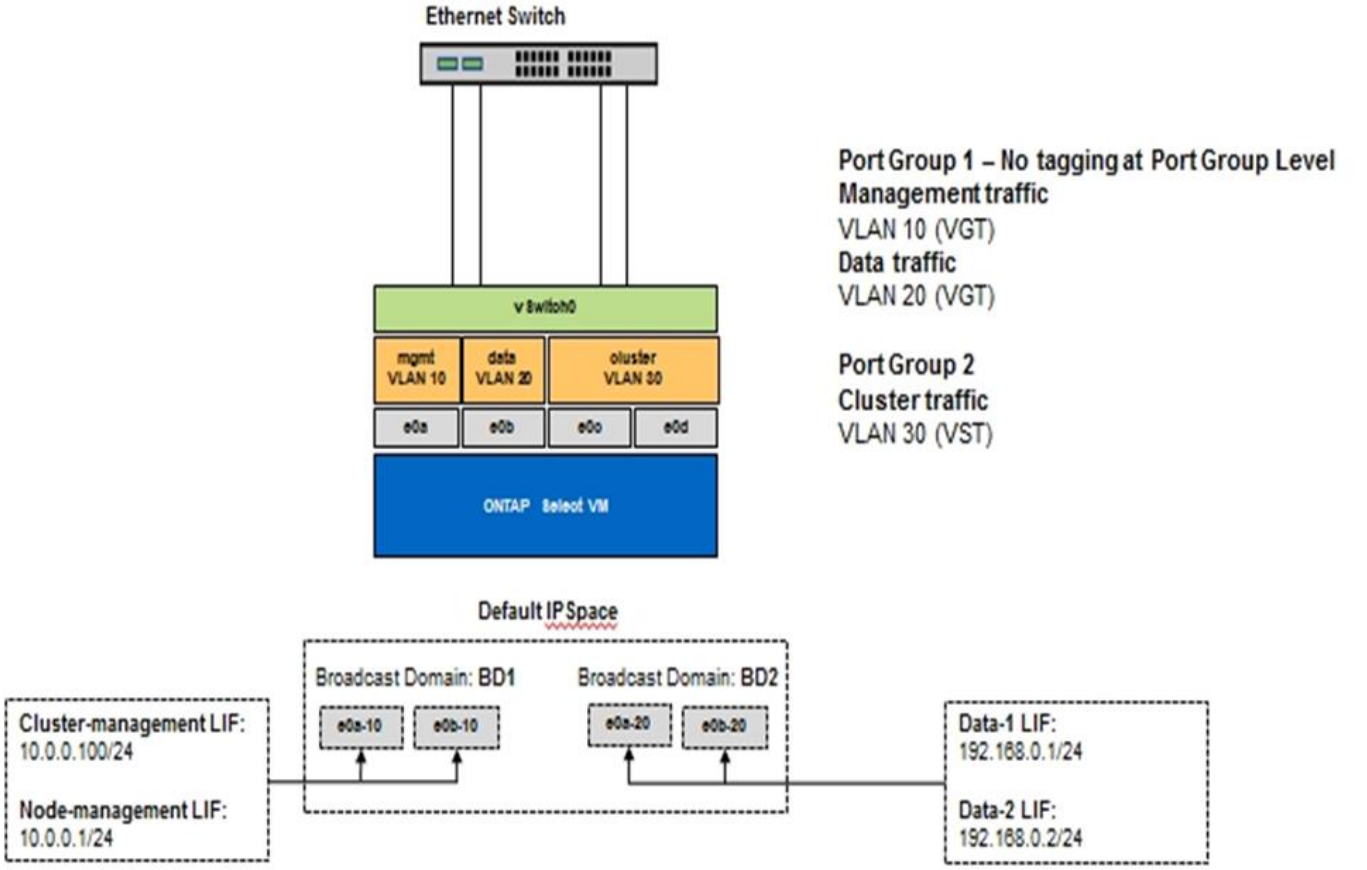
*使用VST*的資料與管理分離



下圖顯示第二個案例VGT、ONTAP 亦即使用放置在不同廣播網域中的VLAN連接埠、將流量標記於由EVM4所標記的情況。在此範例中、虛擬連接埠e0A-10/e0b-10/（e0c或e0g）10和e0a -20/e0b-20會放置在VM連接埠e0a和e0b的頂端。此組態可讓網路標記直接在ONTAP 支援區內執行、而非在vSwitch層執行。管理與資料生命期會放置在這些虛擬連接埠上、讓單一VM連接埠內的第2層子分割區得以進一步發展。叢集VLAN（VLAN ID 30）仍在連接埠群組中標記。

- 附註： *
- 當使用多個IPspace時、這種組態形式尤其理想。如果需要進一步的邏輯隔離和多租戶共享、請將VLAN連接埠分組為個別的自訂IPspaces。
- 若要支援VGT、ESX/ESX主機網路介面卡必須連接至實體交換器上的主幹連接埠。連接至虛擬交換器的連接埠群組必須將其VLAN ID設為4095、才能在連接埠群組上啟用主幹連線。

*使用VG*分隔資料與管理



高可用度架構

高可用度組態

探索高可用度選項、為您的環境選擇最佳HA組態。

雖然客戶開始將應用程式工作負載從企業級儲存設備移轉至以軟體為基礎的解決方案、並在市售硬體上執行、但對於恢復能力和容錯能力的期望和需求並未改變。HA解決方案提供零恢復點目標（RPO）、可保護客戶免受基礎架構堆疊中任何元件故障所造成的資料遺失。

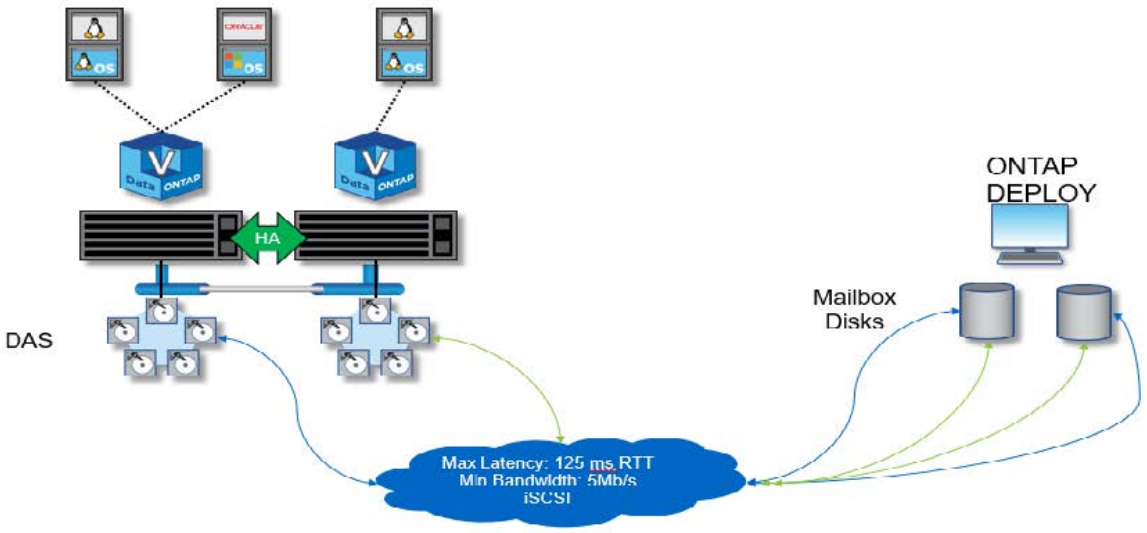
大部分的SDS市場都是以共享無儲存的概念為基礎、而軟體複寫則可在不同的儲存封閉環境中儲存多個使用者資料複本、提供資料恢復能力。利用由支援的同步複寫功能（RAID還原）、在這個前提下建置、以在叢集中儲存額外的使用者資料複本。ONTAP Select SyncMirror ONTAP這是發生在HA配對的內容中。每個HA配對都會儲存兩份使用者資料複本：一份儲存在本機節點所提供的儲存設備上、另一份儲存在HA合作夥伴所提供的儲存設備上。在一個叢集內ONTAP Select 、HA和同步複寫會相互連結、而且兩者的功能無法獨立分離或使用。因此、同步複寫功能僅適用於多節點產品。

> ⓘ 在一個不同步叢集中、同步複寫功能是HA實作的功能、而非非非同步SnapMirror或還原複寫引擎的替代功能。ONTAP Select SnapVault無法從HA獨立使用同步複寫。

共有兩ONTAP Select 種功能：多節點叢集（四個、六個或八個節點）和雙節點叢集。雙節點ONTAP Select 的不二叢集的主要特色是使用外部中介服務來解決腦分割案例。此功能可作為預設的協調器、用於其所設定的所有雙節點HA配對。ONTAP

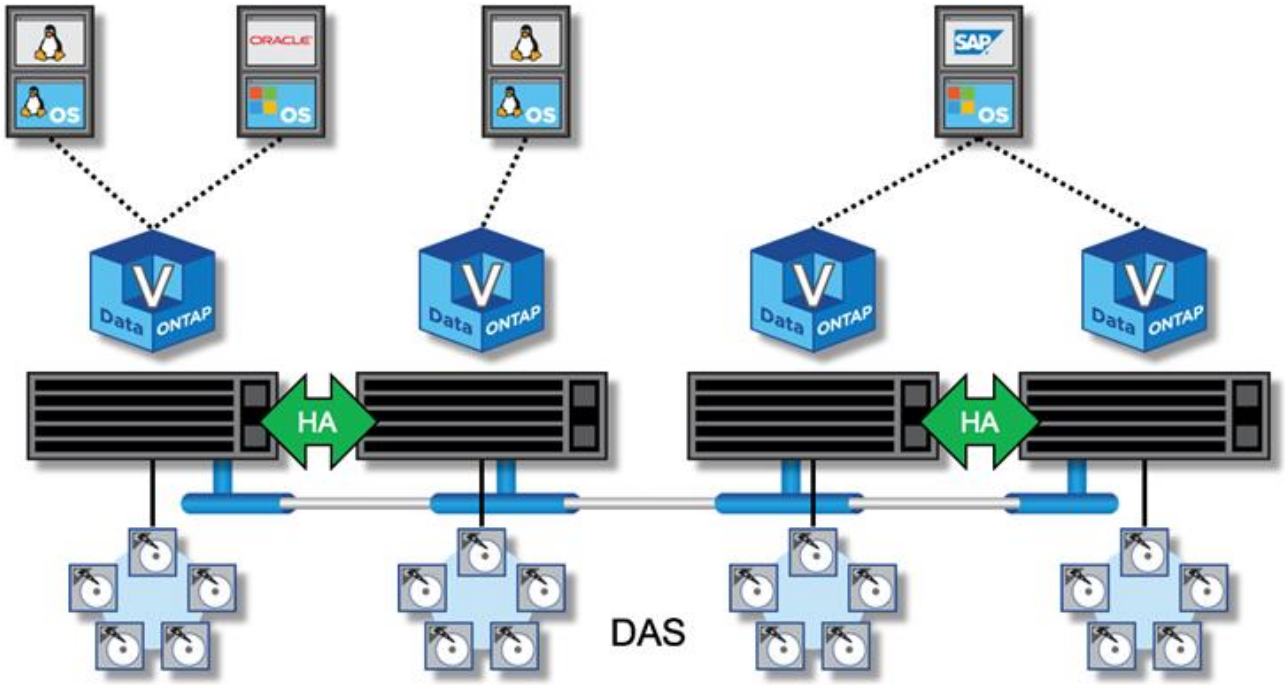
這兩種架構如下圖所示。

- ONTAP Select 使用遠端中介器和本機附加儲存設備*的雙節點版S叢 集



雙節點ONTAP Select 的叢集由一對HA配對和一位中介器組成。在HA配對中、每個叢集節點上的資料集合體都會同步鏡射、而且在發生容錯移轉時、資料不會遺失。

- ONTAP Select 使用本機附加儲存設備*的四節點效能叢集



- 四節點ONTAP Select 的叢集由兩個HA配對組成。六節點和八節點叢集分別由三對和四對HA組成。在每個HA配對中、每個叢集節點上的資料集合體都會同步鏡射、而且在發生容錯移轉時、資料不會遺失。
- 使用DAS儲存設備時、實體伺服器上只ONTAP Select 能有一個執行個體。不共享的系統本機RAID控制器存取權限、是專為管理本機附加磁碟而設計、如果沒有實體連線到儲存設備、就不可能做到這一點。ONTAP Select

雙節點**HA**與多節點**HA**的比較

與FAS 不一樣的是、ONTAP Select HA配對中的各個節點只能透過IP網路進行通訊。這表示IP網路是單點故障（SPOF）、防範網路分割區和腦分割案例、是設計的重要層面。多節點叢集可承受單一節點故障、因為叢集仲裁可由三個以上的正常節點建立。雙節點叢集仰賴ONTAP 由該解決方案代管的中介服務來達成相同的結果。

由於在各個節點與「支援不中斷」中介服務之間的活動訊號網路流量ONTAP Select 極少且彈性十分ONTAP 靈活、ONTAP 因此可將「部署虛擬機器」託管在ONTAP Select 不同於「支援不中斷」雙節點叢集的資料中心。

> ⓘ 當用作該叢集的中介器時、支援功能可將虛擬機器整合為雙節點叢集的一部分。ONTAP如果無法使用中介服務、雙節點叢集會繼續提供資料、但ONTAP Select 會停用該叢集的儲存容錯移轉功能。因此ONTAP 、使用「零位元部署中介」服務時、必須持續與ONTAP Select HA配對中的每個節點通訊。為了讓叢集仲裁正常運作、至少需要5Mbps的頻寬和125ms的最大往返時間（RTT）延遲。

如果ONTAP 扮演中介角色的原地部署VM暫時或可能永久無法使用、可使用二線ONTAP 的原地部署VM來還原雙節點叢集仲裁。如此一來、新ONTAP 的S不到 位部署VM的組態就無法管理ONTAP Select 這些節點、但它成功參與了叢集仲裁演算法。透過透過IPV4使用iSCSI傳輸協定、即可在支援節點與支援資源中心之間進行通訊ONTAP Select 。ONTAP此為啟動器的靜態節點管理IP位址、而以此為目標的是部署VM IP位址。ONTAP Select ONTAP因此、建立雙節點叢集時、無法支援節點管理IP位址的IPv6位址。在建立雙節點叢集時、會自動建立及遮罩至適當的站台管理IP位址的子網路部署託管信箱磁碟。ONTAP ONTAP Select整個組態會在設定期間自動執行、不需要進一步的管理動作。建立叢集的程式集部署執行個體是該叢集的預設中介程式。ONTAP

如果必須變更原始的中介位置、則需要執行管理動作。即使原始ONTAP 的還原部署VM遺失、也能恢復叢集仲裁。不過、NetApp建議您在ONTAP 每個雙節點叢集具現化之後、備份該還原部署資料庫。

雙節點**HA**與雙節點延伸**HA**（**MetroCluster** 不含雙節點）

您可以將雙節點主動/主動式HA叢集延伸至更遠的距離、並可能將每個節點放置在不同的資料中心。雙節點叢集與雙節點延伸叢集（也稱為MetroCluster 「效益SDS」）之間的唯一區別、是節點之間的網路連線距離。

雙節點叢集定義為叢集、兩個節點位於同一個資料中心、距離300公尺以內。一般而言、兩個節點都有上行鏈路可連至同一個網路交換器或一組交換器間連結（ISL）網路交換器。

雙節點MetroCluster 的不完整資料SDS定義為叢集、其節點實體分隔（不同的空間、不同的建築物和不同的資料中心）超過300公尺。此外、每個節點的上行鏈路連線都會連接至不同的網路交換器。不需要專用硬體即可實現此功能。MetroCluster不過、環境應符合延遲需求（RTT最長5毫秒、不穩定性最長5毫秒、總共10毫秒）和實體距離（最大10公里）。

MetroCluster SDS 是一項優質功能、需要 Premium 授權或 Premium XL 授權。Premium授權可支援建立中小型VM、以及HDD和SSD媒體。Premium XL 授權也支援 NVMe 磁碟機的建立。

> ⓘ 本地附加儲存設備（DAS）和共享儲存設備（vNAS）均支援使用支援的支援。MetroCluster請注意、VNAS組態通常會因為ONTAP Select 在整個過程中使用的是介於整個過程中的虛擬機器和共享儲存設備、而產生較高的內在延遲。包含共享儲存延遲在內的各個節點之間、均必須提供最多10毫秒的延遲。MetroCluster換句話說、只測量Select VM之間的延遲並不足夠、因為這些組態的共享儲存延遲並不微不足道。

HA RSM與鏡射Aggregate

使用RAID SyncMirror 功能（RSM2）、鏡射集合體和寫入路徑來防止資料遺失。

同步複寫

此功能是以HA合作夥伴的概念為基礎打造而成。ONTAP透過使用支援RAID功能（RSMs）、將此架構延伸至非共享的市售伺服器世界、此功能可在叢集節點之間複寫資料區塊、並在HA配對之間提供兩份使用者資料複本。ONTAP Select SyncMirror ONTAP

具有中介器的雙節點叢集可橫跨兩個資料中心。如需詳細資訊、請參閱一節 "雙節點延伸HA MetroCluster （簡稱「架構SDS」）最佳實務做法"。

鏡射**Aggregate**

一個由兩到八個節點組成的叢集。ONTAP Select每個HA配對都包含兩份使用者資料複本、透過IP網路在節點之間同步鏡射。此鏡射對使用者而言是透明的、而且是資料集合體的屬性、會在資料集合體建立程序期間自動設定。

在發生節點容錯移轉時、必須鏡射整個叢集中的所有集合體ONTAP Select 、以確保資料可用度、並避免發生硬體故障時出現SPOF。叢集中的Aggregate ONTAP Select 是從HA配對中的每個節點所提供的虛擬磁碟建置、並使用下列磁碟：

- 本機磁碟集（由目前ONTAP Select 的節點所提供）
- 鏡射磁碟集（由目前節點的HA合作夥伴提供）

用來建置鏡射Aggregate的本機磁碟和鏡射磁碟大小必須相同。這些集合體稱為plex 0和plex 1（分別表示本機和遠端鏡像配對）。實際的叢數在您的安裝中可能有所不同。

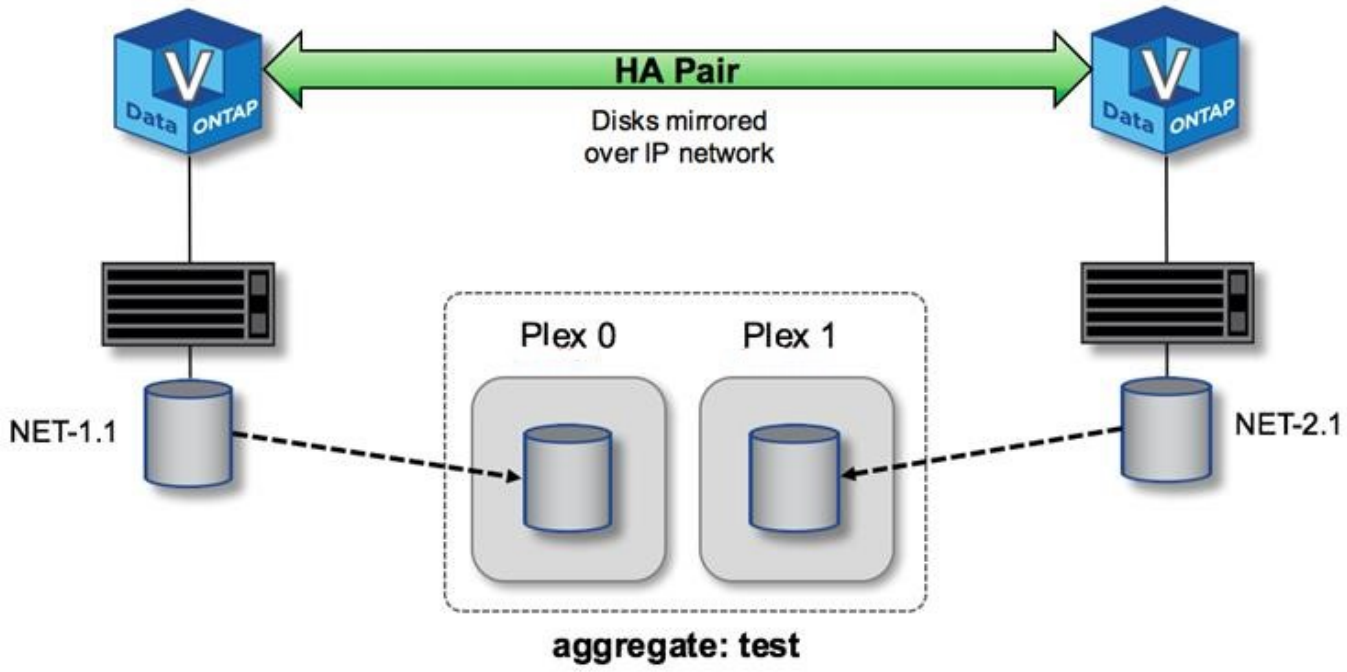
這種方法與標準ONTAP 的仰賴叢集運作方式有根本的不同。這適用於ONTAP Select 整個叢集內的所有根磁碟和資料磁碟。Aggregate同時包含本機和鏡射資料複本。因此，包含N個虛擬磁碟的集合體可提供不含2個磁碟的獨特儲存容量，因為第二個資料複本位於其專屬的磁碟上。

下圖顯示四節點ONTAP Select 的叢集內的HA配對。在此叢集中、是使用兩個HA合作夥伴儲存設備的單一集合體（測試）。此資料Aggregate由兩組虛擬磁碟組成：由ONTAP Select 故障轉移合作夥伴（Plex 1）所提供的實體叢集節點（Plex 0）所構成的本機磁碟集、以及遠端磁碟集。

Plex 0是存放所有本機磁碟的儲存區。Plex 1是存放鏡射磁碟的儲存區、或是負責儲存第二個使用者資料複本的磁碟。擁有Aggregate的節點會將磁碟貢獻給Plex 0、而該節點的HA合作夥伴則會將磁碟貢獻給Plex 1。

下圖中有一個鏡射Aggregate、其中包含兩個磁碟。此Aggregate的內容會鏡射至我們的兩個叢集節點、並將本機磁碟NET-1.1放入Plex 0儲存區、並將遠端磁碟NET-2.1放入Plex 1儲存區。在此範例中、Aggregate測試由左側的叢集節點擁有、並使用本機磁碟NET-1.1和HA合作夥伴鏡射磁碟NET-2.1。

*鏡射Aggregate * ONTAP Select



部署一個叢集時ONTAP Select 、系統上的所有虛擬磁碟都會自動指派給正確的叢集、不需要使用者就磁碟指派採取額外步驟。如此可防止意外將磁碟指派給不正確的叢、並提供最佳的鏡射磁碟組態。

寫入路徑

在叢集節點之間同步鏡射資料區塊、以及在系統故障時不需遺失資料、對於傳入寫入透過ONTAP Select 叢集傳播時所採用的路徑、會造成重大影響。此程序包含兩個階段：

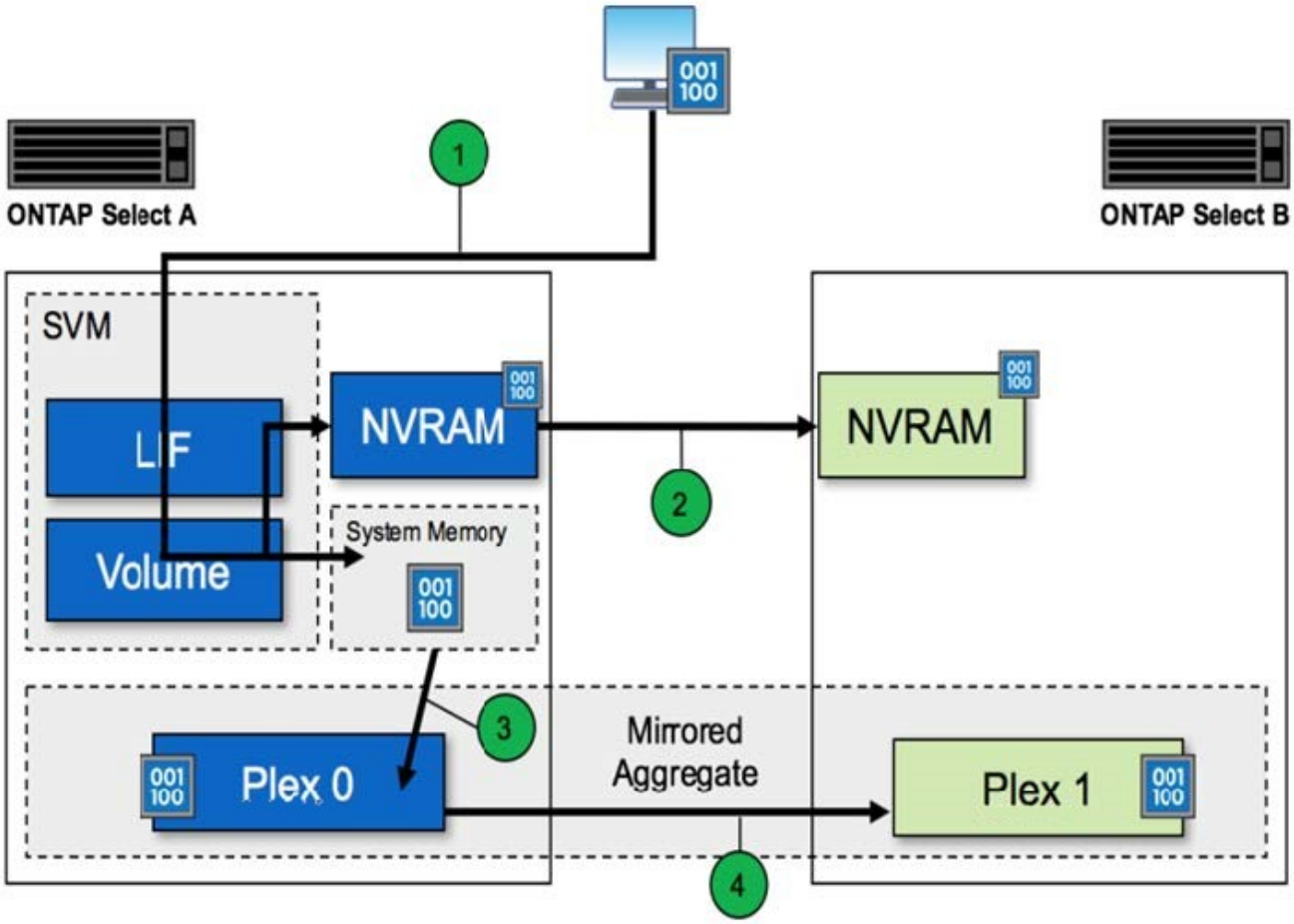
- 確認
- 減少需求

寫入目標磁碟區的作業會透過資料LIF進行、並將寫入到虛擬化的NVRAM分割區、此分割區會出現在ONTAP Select 節點的系統磁碟上、然後再確認回用戶端。在HA組態上、會執行額外的步驟、因為這些NVRAM寫入作業會在確認之前、立即鏡射到目標磁碟區擁有者的HA合作夥伴。如果原始節點發生硬體故障、此程序可確保HA合作夥伴節點上的檔案系統一致性。

將寫入作業提交至NVRAM後、ONTAP 將此分割區的內容定期移至適當的虛擬磁碟、這是稱為「減少磁碟空間」的程序。此程序只會在擁有目標Volume的叢集節點上執行一次、而不會發生在HA合作夥伴上。

下圖顯示傳入寫入要求至ONTAP Select 某個節點的寫入路徑。

不寫入路徑工作流程 ONTAP Select



傳入寫入認可包括下列步驟：

- 寫入會透過ONTAP Select 由節點A擁有的邏輯介面進入系統
- 寫入作業會提交至節點A的NVRAM、並鏡射至HA合作夥伴節點B
- 在兩個HA節點上都有I/O要求之後、該要求會被確認回用戶端。

從NVRAM降級至資料集合體（RACP）包括下列步驟：ONTAP Select ONTAP

- 寫入作業會從虛擬NVRAM移轉至虛擬資料Aggregate。
- 鏡射引擎會將區塊同步複寫到兩個叢集中。

HA其他詳細資料

HA磁碟心跳轉、HA信箱、HA心跳轉、HA容錯移轉及還原等功能、可強化資料保護。

磁碟心跳

雖然這個架構運用了許多傳統的指令碼路徑、但仍存在一些例外情況。ONTAP Select FAS其中一項例外是執行磁碟型心率、這是叢集節點用來防止網路隔離造成腦分割行為的非網路型通訊方法。分割腦部案例是叢集分割的結果、通常是網路故障所造成、因此每一方都認為另一方已停機、並嘗試接管叢集資源。

企業級HA實作必須妥善處理這類案例。透過客製化的磁碟型心跳法來達成此目標。ONTAP這是HA信箱的工

作、這是叢集節點用來傳遞活動訊號訊息的實體儲存設備位置。這有助於叢集判斷連線能力、因此可在容錯移轉時定義仲裁。

在使用共享儲存HA架構的不同時陣列上、用下列方法解決大腦分裂問題：FAS ONTAP

- SCSI持續保留
- 持續HA中繼資料
- 透過HA互連傳送HA狀態

然而ONTAP Select 、在不共用的架構中、節點只能看到自己的本機儲存設備、而無法看到HA合作夥伴的儲存設備。因此、當網路分割區隔離HA配對的每一端時、就無法使用先前的判斷叢集仲裁和容錯移轉行為的方法。

雖然無法使用現有的分割腦偵測和避免方法、但仍需要一種協調方法、這種方法符合共享無事環境的限制。利用NetApp技術、可進一步擴充現有的信箱基礎架構、在網路分割的情況下、也能做為中繼方法。ONTAP Select由於共享儲存設備無法使用、因此透過NAS存取信箱磁碟來完成中繼作業。這些磁碟會使用iSCSI傳輸協定散佈到整個叢集、包括雙節點叢集中的中介器。因此、叢集節點可根據對這些磁碟的存取來做出智慧型容錯移轉決策。如果某個節點可以存取HA合作夥伴以外其他節點的信箱磁碟、則可能正常運作。

> ⓘ 信箱架構和以磁碟為基礎的核心跳轉方法來解決叢集仲裁和腦分割問題、是ONTAP Select 導致多節點變體需要四個獨立節點、或是兩節點叢集的中介器的原因。

HA信箱張貼

HA信箱架構使用訊息POST模式。叢集節點會以重複的時間間隔、將訊息張貼到叢集內的所有其他信箱磁碟、包括中介器、說明節點已啟動且正在執行。在健全的叢集內、叢集節點上的單一信箱磁碟會在任何時間點張貼來自所有其他叢集節點的訊息。

附加至每個Select叢集節點的虛擬磁碟是專供共用信箱存取使用的磁碟。此磁碟稱為中介信箱磁碟、因為它的主要功能是在節點故障或網路分割時做為叢集中繼的方法。此信箱磁碟包含每個叢集節點的分割區、並由其他Select叢集節點透過iSCSI網路掛載。這些節點會定期將健全狀況狀態張貼到信箱磁碟的適當分割區。使用分散於整個叢集的網路存取信箱磁碟、可讓您透過可到達性對照表來推斷節點健全狀況。例如、叢集節點A和B可以POST至叢集節點D的信箱、但不能POST至節點C的信箱此外、叢集節點D無法POST至節點C的信箱、因此節點C可能已停機或是網路已隔離、因此應該接管。

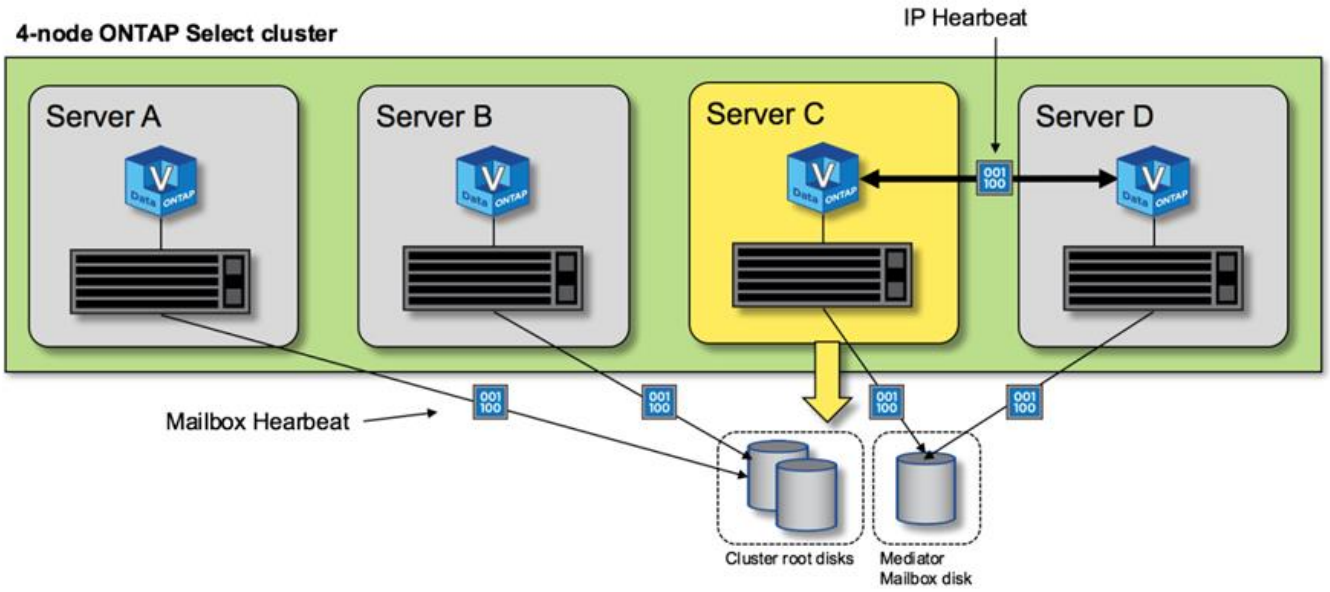
HA令人心碎

如同NetApp FAS 的支援平台、ONTAP Select VMware會定期透過HA互連傳送HA活動訊息。在這個叢集內ONTAP Select 、這是透過HA合作夥伴之間存在的TCP/IP網路連線來執行。此外、磁碟型的活動訊號訊息會傳遞至所有HA信箱磁碟、包括中介信箱磁碟。這些訊息會每隔幾秒傳遞一次、並定期讀取。發送和接收這些訊號的頻率可讓ONTAP Select 故障叢集在約15秒內偵測HA故障事件、FAS 而這個視窗可在各個平台上使用。當不再讀取活動訊息時、會觸發容錯移轉事件。

下圖顯示從單ONTAP Select 一的故障叢集節點C的觀點、透過HA互連磁碟和中介磁碟傳送和接收訊號訊息的程序

> ⓘ 網路訊號會透過HA互連傳送至HA合作夥伴節點D、而磁碟訊號則會在所有叢集節點、A、B、C和D之間使用信箱磁碟

在四節點叢集中執行**HA**資料中心：穩定狀態



HA容錯移轉與還原

在容錯移轉作業期間、正常運作的節點會使用HA合作夥伴資料的本機複本、承擔對等節點的資料服務責任。用戶端I/O可以不中斷地繼續運作、但必須先複寫此資料的變更、才能進行還原。請注意ONTAP Select 、由於這會導致儲存在存續節點上的變更遺失、所以不支援強制恢復。

重新開機的節點重新加入叢集時、會自動觸發同步處理回復作業。同步回溯所需的時間取決於多項因素。這些因素包括必須複寫的變更數目、節點之間的網路延遲、以及每個節點上磁碟子系統的速度。同步處理回饋所需的時間可能超過自動回饋10分鐘。在這種情況下、需要在同步回溯後手動恢復。您可以使用下列命令來監控同步處理回復的進度：

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

效能

效能

效能會因硬體組態而異。

由於基礎硬體和組態的特性、所以無法大幅改變VMware叢集的效能ONTAP Select 。特定的硬體組態是特定ONTAP Select 執行個體效能的最大因素。以下是影響特定ONTAP Select 執行個體效能的一些因素：

- 核心頻率。一般而言、較高的頻率較為理想。
- 單插槽與多插槽。雖然不使用多插槽功能、但支援多插槽組態的Hypervisor負荷會在整體效能方面造成一定程度的差異。ONTAP Select
- * RAID卡組態及相關Hypervisor驅動程式*。Hypervisor提供的預設驅動程式可能需要由硬體廠商驅動程式取代。
- * RAID群組中的磁碟機類型和磁碟機數目*。
- * Hypervisor版本與修補程式層級*。

效能：優質**HA**直接附加**SSD**儲存設備

參考平台的效能資訊。

參考平台

（優質XL）硬體（每個節點ONTAP Select ）

- Fujitsu PRIMERGY RX2540 M4：
  - Intel（R）Xeon（R）Gold 6142b CPU、速度為2.6 GHz
  - 32個實體核心（16 x 2個插槽）、64個邏輯核心
  - 256 GB RAM
  - 每個主機的磁碟機：24個960GB SSD
  - ESX 6.5U1.

用戶端硬體

- 5個NFSv3 IBM 3550m4用戶端

組態資訊

- 軟體RAID 1 x 9 + 2 RAID-DP（11個磁碟機）
- 22+1 RAID-5（ONTAP 位於EASEID中的RAID-0）/ RAID快取NVRAM
- 沒有使用儲存效率功能（壓縮、重複資料刪除、Snapshot複本、SnapMirror等）

下表列出ONTAP Select 使用軟體RAID和硬體RAID、在高可用度（HA）的一對邊節點上針對讀寫工作負載所測量的處理量。效能測量是使用SIO負載產生工具進行。

這些效能數據均以ONTAP Select VMware 9.6為基礎。

使用**ONTAP Select** 軟體**RAID**和硬體**RAID**、在直接附加儲存（DAS）SSD上、單一節點（四節點中型執行個體的一部分）的效能結果

說明	連續讀取**64KiB**	連續寫入**64KiB**	隨機讀取**8KiB**	隨機寫入**8KiB**	隨機**WR/ RD**（**50/50**）**8KiB**
採用DAS（SSD）軟體RAID的大型執行個體ONTAP Select	2171億次	559 MiBps	954 MiBps	394-393MBps	564 MBps
採用DAS（SSD）軟體RAID的中型執行個體ONTAP Select	2090億美元	592 次 MiBps	677 MiBps	335 個 MiBps	441 3億次

說明	連續讀取**64KiB**	連續寫入**64KiB**	隨機讀取**8KiB**	隨機寫入**8KiB**	隨機**WR/ RD**（**50/50**）**8KiB**
採用DAS（SSD）硬體RAID的中型執行個體ONTAP Select	2038 MiBps	520億分之一	578台MiBps	325迷你吧	399 MiBps

64K循序讀取

詳細資料：

- SIO直接I/O已啟用
- 2個節點
- 每個節點2個資料NIC
- 每個節點1個資料集合體（2TB硬體RAID）、（8TB軟體RAID）
- 64個SIO處理程序、每個處理程序1個執行緒
- 每個節點32個磁碟區
- 每個處理器1個檔案；每個檔案為12000MB

64K循序寫入

詳細資料：

- SIO直接I/O已啟用
- 2個節點
- 每個節點2個資料網路介面卡（NIC）
- 每個節點 1 個資料集合體（ 2TB 硬體 RAID ）、（ 4TB 軟體 RAID ）
- 128 個 SIO 處理器、每個處理器 1 個執行緒
- 每個節點的磁碟區數：32（硬體RAID）、16（軟體RAID）
- 每個處理器1個檔案；每個檔案30720MB

8K隨機讀取

詳細資料：

- SIO直接I/O已啟用
- 2個節點
- 每個節點2個資料NIC
- 每個節點 1 個資料集合體（ 2TB 硬體 RAID ）、（ 4TB 軟體 RAID ）
- 64個SIO處理程序、每個處理程序8個執行緒
- 每個節點的磁碟區數：32個
- 每個處理器 1 個檔案；每個檔案 12228MB

8K隨機寫入

詳細資料：

- SIO直接I/O已啟用
- 2個節點
- 每個節點2個資料NIC
- 每個節點 1 個資料集合體（ 2TB 硬體 RAID ）、（ 4TB 軟體 RAID ）
- 64個SIO處理程序、每個處理程序8個執行緒
- 每個節點的磁碟區數：32個
- 每個處理器有1個檔案；每個檔案為819MB

8K隨機50%寫入50%讀取

詳細資料：

- SIO直接I/O已啟用
- 2個節點
- 每個節點2個資料NIC
- 每個節點 1 個資料集合體（ 2TB 硬體 RAID ）、（ 4TB 軟體 RAID ）
- 每個處理器64個SIO pro208執行緒
- 每個節點的磁碟區數：32個
- 每個處理器 1 個檔案；每個檔案 12228MB

版權資訊

Copyright © 2025 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。 美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 http://www.netapp.com/TM 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。