



深入探究 ONTAP Select

NetApp
January 29, 2026

目錄

深入探究	1
儲存	1
ONTAP Select儲存：一般概念和特徵	1
適用於ONTAP Select本機連線儲存的硬體 RAID 服務	6
適用於本機連接儲存的ONTAP Select軟體 RAID 配置服務	12
ONTAP Select VSAN 和外部陣列配置	20
增加ONTAP Select存儲容量	23
ONTAP Select儲存效率支持	26
聯網	28
ONTAP Select網路概念與特性	28
ONTAP Select單節點與多節點網路配置	30
ONTAP Select內部與外部網路	35
支援的ONTAP Select網路配置	37
ONTAP SelectVMware vSphere vSwitch 配置	38
ONTAP Select實體交換器配置	47
ONTAP Select資料與管理流量分離	49
高可用性架構	51
ONTAP Select高可用性配置	51
ONTAP Select HA RSM 和鏡像聚合	53
ONTAP Select HA 增強資料保護	56
表現	58
ONTAP Select效能概述	58
ONTAP Select 9.6 效能：進階 HA 直連 SSD 存儲	59

深入探究

儲存

ONTAP Select儲存：一般概念和特徵

在探索特定儲存組件之前，先了解適用於ONTAP Select環境的一般儲存概念。

儲存配置階段

ONTAP Select主機儲存的主要配置階段包括以下內容：

- 部署前的先決條件
 - 確保每個虛擬機器管理程式主機都已配置並準備好進行ONTAP Select部署。
 - 配置涉及實體磁碟機、RAID 控制器和群組、LUN 以及相關的網路準備。
 - 此配置在ONTAP Select之外執行。
- 使用虛擬機器管理程式管理員實用程式進行配置
 - 您可以使用虛擬機器管理程式管理公用程式（例如，VMware 環境中的 vSphere）來設定儲存的某些方面。
 - 此配置在ONTAP Select之外執行。
- 使用ONTAP Select Deploy 管理實用程式進行配置
 - 您可以使用部署管理公用程式來設定核心邏輯儲存結構。
 - 這可以透過 CLI 命令明確執行，也可以由實用程式作為部署的一部分自動執行。
- 部署後配置
 - ONTAP Select部署完成後，您可以使用ONTAP CLI 或系統管理器設定叢集。
 - 此配置在ONTAP Select Deploy 之外執行。

託管儲存與非託管儲存

ONTAP Select存取和直接控制的儲存稱為託管儲存。同一虛擬機器管理程式主機上的任何其他儲存空間稱為非託管儲存。

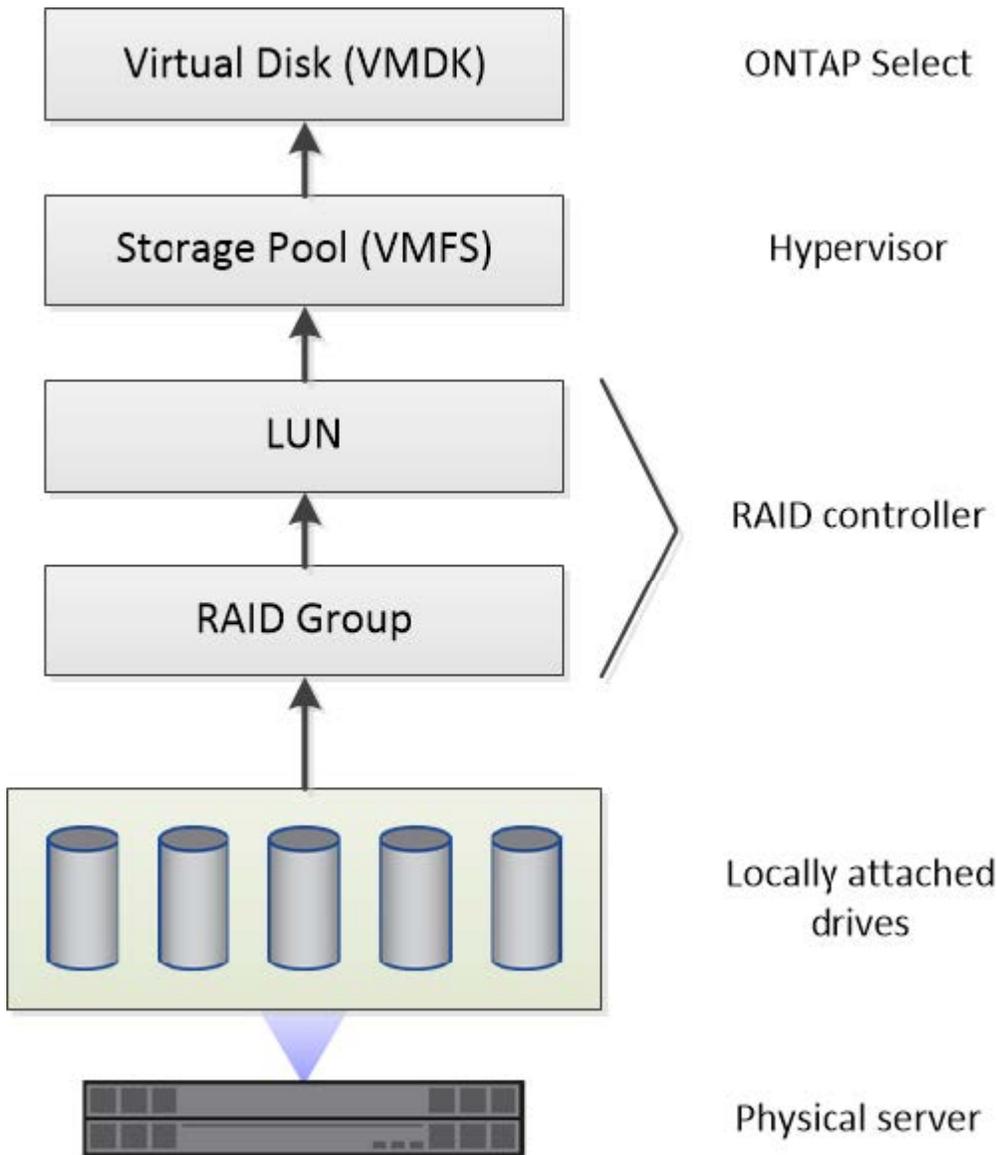
同構實體儲存

組成ONTAP Select託管儲存的所有實體磁碟機必須是同構的。也就是說，所有硬體在以下特性方面必須相同：

- 類型（SAS、NL-SAS、SATA、SSD）
- 速度（轉/分）

本地儲存環境說明

每個虛擬機器管理程式主機都包含本機磁碟和其他可供ONTAP Select使用的邏輯儲存元件。這些儲存組件按分層結構排列，從實體磁碟開始。



本機儲存組件的特性

有幾個概念適用於ONTAP Select環境中使用的本機儲存組件。在準備ONTAP Select部署之前，您應該熟悉這些概念。這些概念按類別排列：RAID 群組和 LUN、儲存池和虛擬磁碟。

將實體磁碟機分組為 RAID 群組和 LUN

可以將一個或多個實體磁碟本地連接到主機伺服器，並供ONTAP Select使用。這些實體磁碟將指派給 RAID 群組，然後作為一個或多個 LUN 呈現給虛擬機器管理程式主機作業系統。每個 LUN 都以實體硬碟呈現給虛擬機器管理程式主機作業系統。

配置ONTAP Select主機時，您應該注意以下事項：

- 所有託管儲存必須可透過單一 RAID 控制器存取
- 根據供應商的不同，每個 RAID 控制器支援每個 RAID 群組的最大磁碟機數量

一個或多個 RAID 群組

每個ONTAP Select主機必須配備單一 RAID 控制器。您應該為ONTAP Select建立單一 RAID 群組。但是，在某些情況下，您可能需要建立多個 RAID 群組。請參閱["最佳實踐摘要"](#)。

儲存池注意事項

在準備部署ONTAP Select時，您應該注意與儲存池相關的幾個問題。



在 VMware 環境中，儲存池與 VMware 資料儲存同義。

儲存池和 LUN

每個 LUN 在虛擬機器管理程式主機上被視為本機磁碟，並且可以成為一個儲存池的一部分。每個儲存池都使用虛擬機器管理程式主機作業系統可以使用的檔案系統進行格式化。

您必須確保在ONTAP Select部署過程中正確建立儲存池。您可以使用虛擬機器管理程式管理工具建立儲存池。例如，使用 VMware，您可以使用 vSphere 用戶端建立儲存池。然後，該儲存池將傳遞到ONTAP Select Deploy 管理實用程式。

管理 ESXi 上的虛擬磁碟

在準備部署ONTAP Select時，您應該注意幾個與虛擬磁碟相關的問題。

虛擬磁碟和檔案系統

ONTAP Select虛擬機器指派有多個虛擬磁碟機。每個虛擬磁碟實際上是儲存池中的一個文件，由虛擬機器管理程式維護。ONTAP Select使用多種類型的磁碟，主要是系統磁碟和資料磁碟。

您還應該了解有關虛擬磁碟的以下資訊：

- 在建立虛擬磁碟之前，儲存池必須可用。
- 在建立虛擬機器之前無法建立虛擬磁碟。
- 您必須依靠ONTAP Select Deploy 管理實用程式來建立所有虛擬磁碟（也就是說，管理員絕對不能從 Deploy 實用程式之外建立虛擬磁碟）。

配置虛擬磁碟

虛擬磁碟由ONTAP Select管理。當您使用 Deploy 管理公用程式建立叢集時，會自動建立虛擬磁碟。

ESXi 上的外部儲存環境圖示

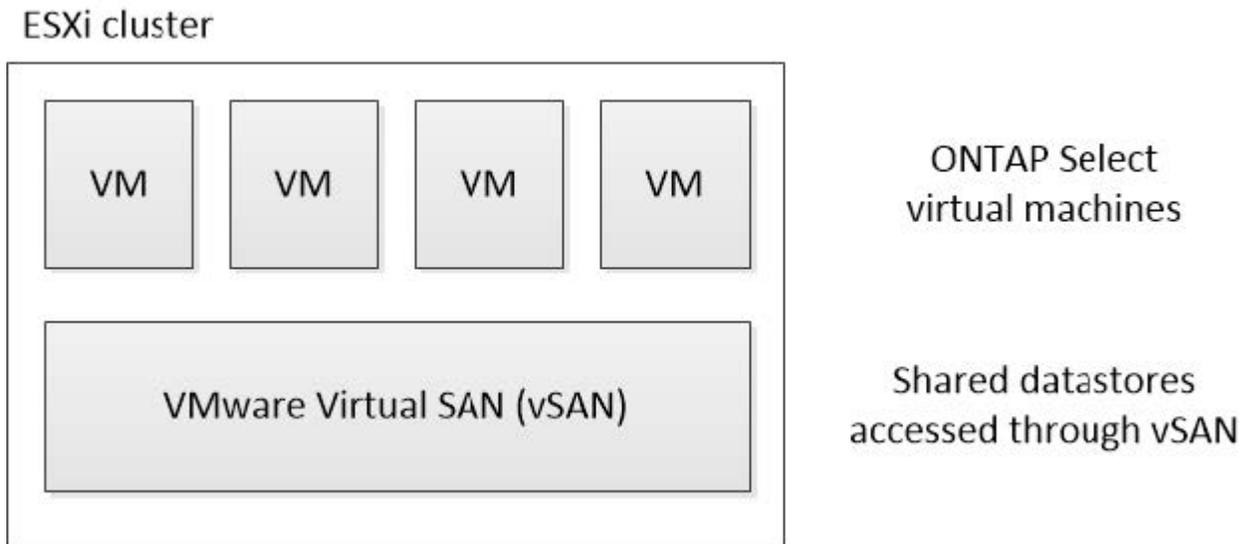
ONTAP Select vNAS 解決方案支援ONTAP Select使用位於虛擬機器管理程式主機外部儲存裝置上的資料儲存體。您可以使用 VMware vSAN 透過網路存取這些資料存儲，也可以直接在外部存儲陣列上存取。

ONTAP Select可設定為使用虛擬機器管理程式主機外部的下列類型的 VMware ESXi 網路資料儲存庫：

- vSAN (虛擬 SAN)
- VMFS
- NFS

vSAN 資料存儲

每個 ESXi 主機可以擁有一個或多個本機 VMFS 資料儲存。通常，這些資料儲存只能由本地主機存取。但是，VMware vSAN 允許 ESXi 叢集中的每個主機共用叢集中的所有資料儲存，就像它們是本機資料儲存一樣。下圖說明了 vSAN 如何建立在 ESXi 叢集中的主機之間共用的資料儲存池。

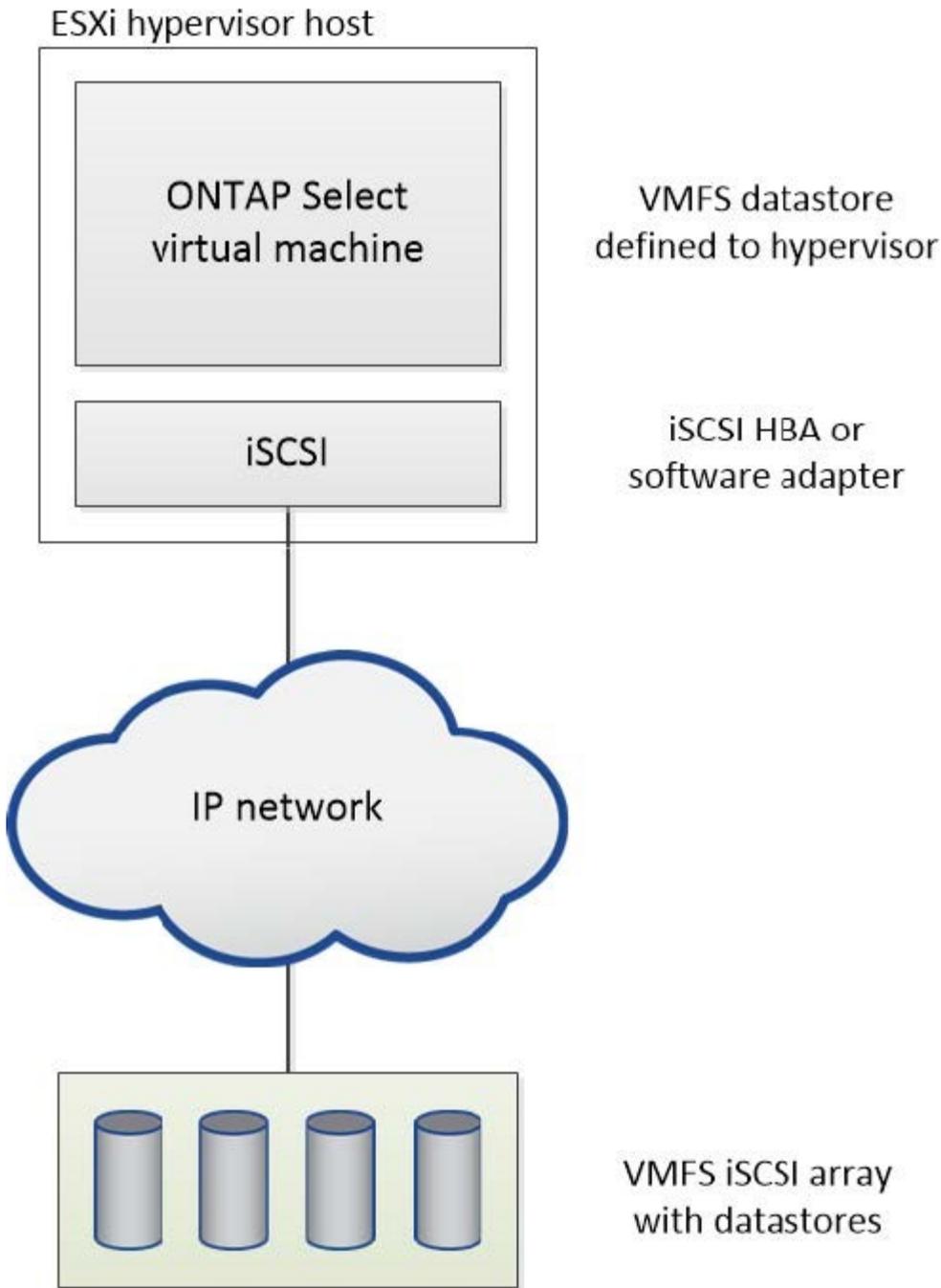


外部儲存陣列上的 VMFS 資料存儲

您可以建立位於外部儲存陣列上的 VMFS 資料儲存。可以使用多種不同的網路協定之一存取該儲存。下圖展示了使用 iSCSI 協定存取的外部儲存陣列上的 VMFS 資料儲存。

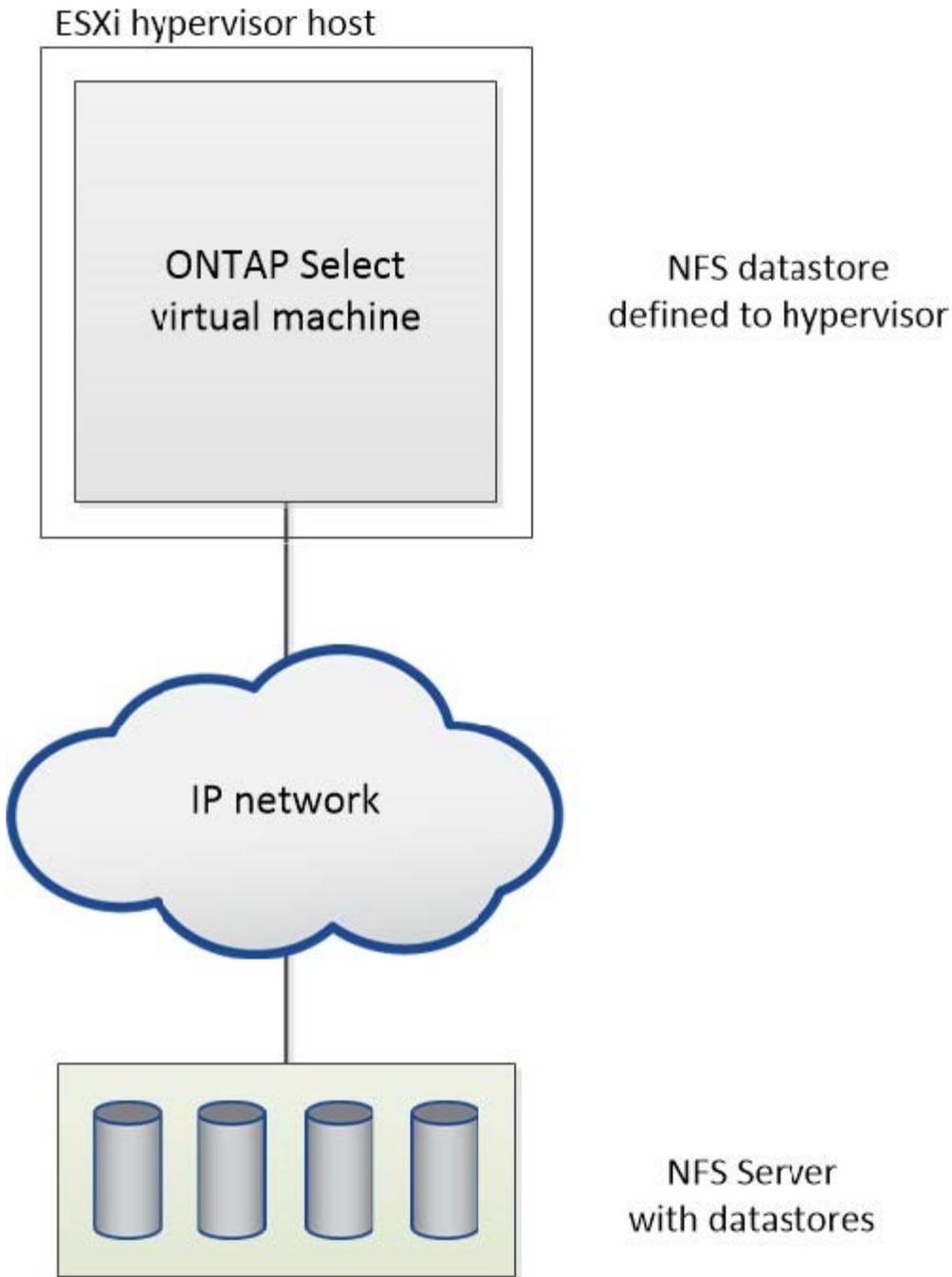


ONTAP Select 支援 VMware 儲存/SAN 相容性文件中所述的所有外部儲存陣列，包括 iSCSI、光纖通道和乙太網路光纖通道。



外部儲存陣列上的 **NFS** 資料存儲

您可以建立位於外部儲存陣列上的 NFS 資料儲存。此儲存體使用 NFS 網路協定進行存取。下圖展示了一個位於外部存儲上的 NFS 資料存儲，可透過 NFS 伺服器設備存取。



適用於ONTAP Select本機連線儲存的硬體 RAID 服務

當硬體 RAID 控制器可用時，ONTAP Select 可以將 RAID 服務移轉到硬體控制器，以提升寫入效能並防止實體磁碟機故障。因此，ONTAP Select 叢集中所有節點的 RAID 保護均由本機連接的 RAID 控制器提供，而非透過 ONTAP 軟體 RAID 提供。



ONTAP Select 資料聚合配置為使用 RAID 0，因為實體 RAID 控制器正在為底層磁碟機提供 RAID 條帶化功能。不支援其他 RAID 等級。

本機連接儲存的 RAID 控制器配置

所有為ONTAP Select提供後備儲存的本機連線磁碟都必須位於 RAID 控制器後面。大多數商用伺服器提供多種 RAID 控制器選項，涵蓋多個價位，每個選項的功能等級各不相同。我們的目標是盡可能支援這些選項，前提是它們滿足控制器的特定最低要求。



您無法從使用硬體 RAID 配置的ONTAP Select虛擬機器中分離虛擬磁碟。僅支援從使用軟體 RAID 配置的ONTAP Select虛擬機器中分離磁碟。看"[更換ONTAP Select軟體 RAID 配置中的故障磁碟機](#)"了解更多。

管理ONTAP Select磁碟的 RAID 控制器必須符合以下要求：

- 硬體 RAID 控制器必須具有電池備用單元 (BBU) 或快閃記憶體支援寫入快取 (FBWC) 並支援 12Gbps 的吞吐量。
- RAID 控制器必須支援能夠承受至少一個或兩個磁碟故障的模式 (RAID 5 和 RAID 6) 。
- 必須將驅動器快取設定為停用。
- 必須將寫入策略配置為寫回模式，並在 BBU 或快閃記憶體發生故障時回退到寫入模式。
- 必須將讀取的 I/O 策略設定為快取。

所有為ONTAP Select提供後備儲存的本機連線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 群組可使ONTAP受益於將傳入的讀取要求分散到更多磁碟上。這樣做可以顯著提升效能。對於 SAS/SSD 配置，我們針對單一 LUN 和多 LUN 配置進行了效能測試。未發現顯著差異，因此，為簡單起見，NetApp建議建立滿足配置需求所需的最少數量的 LUN。

NL-SAS 和 SATA 硬碟需要一套不同的最佳實務。出於效能考慮，磁碟的最小數量仍然為 8 個，但 RAID 群組大小不應超過 12 個磁碟機。NetAppNetApp，所有 RAID 群組都可以使用全域備用磁碟機。；



舊版 ESX 的最大範圍和資料儲存大小為 64 TB，這會影響支援這些大容量磁碟機提供的總原始容量所需的 LUN 數量。

RAID模式

許多 RAID 控制器支援最多三種操作模式，每種模式都代表寫入請求所採用的資料路徑的顯著差異。這三種模式如下：

- 直寫。所有傳入的 I/O 請求都寫入 RAID 控制器緩存，然後立即刷新到磁碟，然後再向主機確認該請求。
- 繞寫。所有傳入的 I/O 請求都直接寫入磁碟，繞過 RAID 控制器快取。
- 寫回。所有傳入的 I/O 請求都直接寫入控制器緩存，並立即返回主機確認。資料塊透過控制器異步刷新到磁碟。

回寫模式提供最短的資料路徑，資料塊進入快取後立即進行 I/O 確認。此模式為混合讀取/寫入工作負載提供最低的延遲和最高的吞吐量。然而，如果沒有 BBU 或非揮發性快閃技術，系統在此模式下運作時如果發生電源故障，使用者將面臨遺失資料的風險。

ONTAP Select需要配備備用電池或快閃記憶體單元；因此，我們可以確保在發生此類故障時，快取的區塊會刷新到磁碟。因此，要求將 RAID 控制器配置為寫回模式。

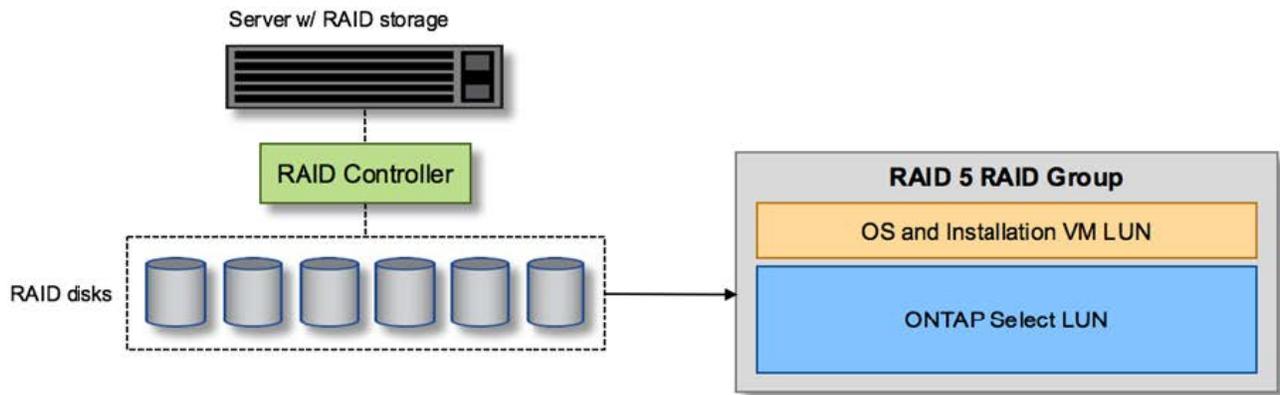
ONTAP Select與作業系統之間共用的本機磁碟

最常見的伺服器配置是所有本機連接的磁碟軸都位於單一 RAID 控制器後面。您應該至少配置兩個 LUN：一個用於虛擬機器管理程序，一個用於ONTAP Select虛擬機。

例如，假設一台 HP DL380 g8 配備六個內建硬碟和一個 Smart Array P420i RAID 控制器。所有內建硬碟均由該 RAID 控制器管理，系統中沒有其他儲存設備。

下圖顯示了這種配置方式。在此範例中，系統上沒有其他儲存；因此，虛擬機器管理程式必須與ONTAP Select節點共用儲存。

僅具有 RAID 管理主軸的伺服器 LUN 設定



透過與ONTAP Select相同的 RAID 群組配置作業系統 LUN，虛擬機器管理程式作業系統（以及同樣從該儲存配置的任何用戶端虛擬機器）可受益於 RAID 保護。此配置可防止單一磁碟機故障導致整個系統崩潰。

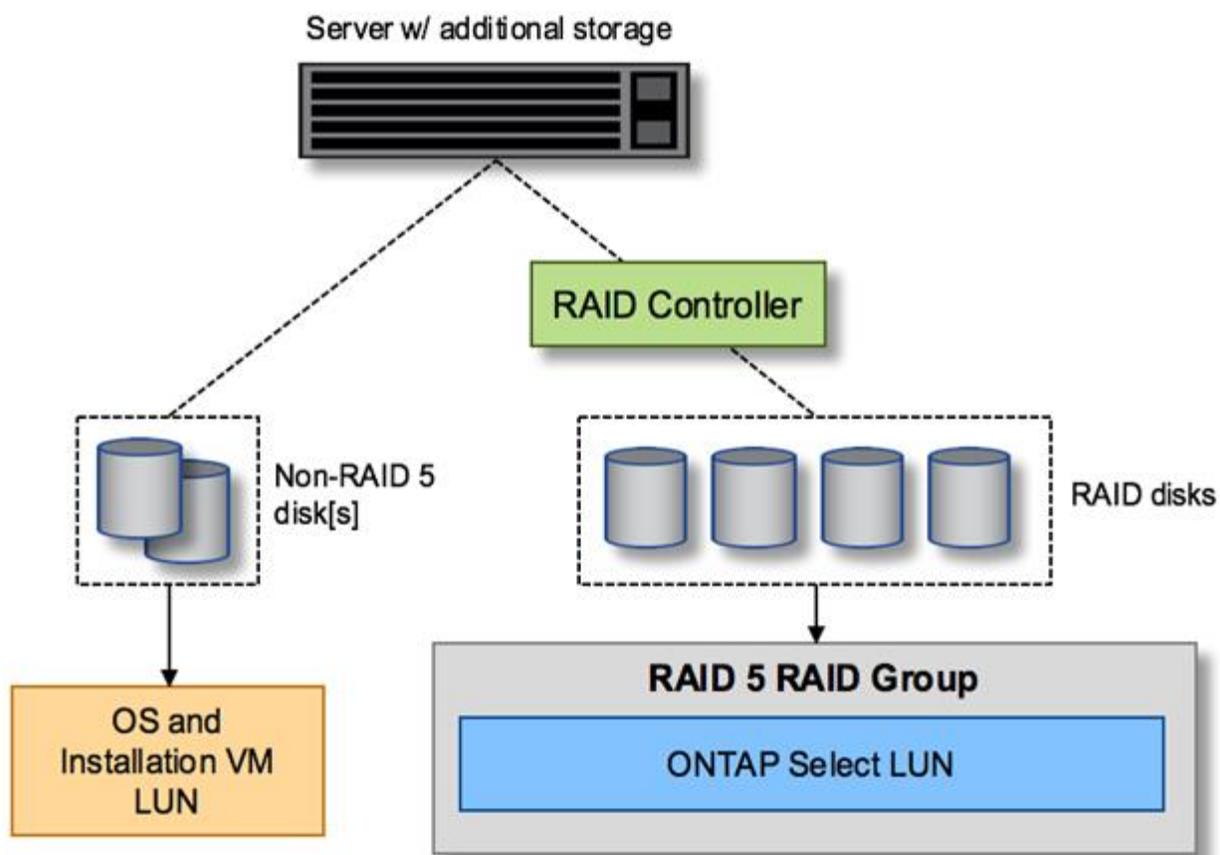
本機磁碟在ONTAP Select和 OS 之間分配

伺服器供應商提供的另一種可能的配置是使用多個 RAID 或磁碟控制器配置系統。在這種配置中，一組磁碟由磁碟控制器管理，該控制器可能提供或不提供 RAID 服務。另一組磁碟由一個硬體 RAID 控制器管理，該控制器能夠提供 RAID 5/6 服務。

採用此配置方式，位於 RAID 控制器後方、可提供 RAID 5/6 服務的磁碟軸組應由ONTAP Select虛擬機器獨佔使用。根據管理的總儲存容量，您應該將磁碟軸配置為一個或多個 RAID 群組以及一個或多個 LUN。然後，這些 LUN 將用於建立一個或多個資料儲存庫，所有資料儲存庫均受 RAID 控制器保護。

第一組磁碟保留給虛擬機管理程式作業系統和任何未使用ONTAP儲存的客戶端虛擬機，如下圖所示。

混合 RAID/非 RAID 系統上的伺服器 LUN 設定



多個 LUN

有兩種情況必須更改單一 RAID 群組/單 LUN 配置。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 的大小可能會超過底層虛擬機器管理程式的儲存限制（單一檔案系統擴充區最大大小或總儲存池最大大小）。這時，必須將底層實體儲存拆分為多個 LUN，才能成功建立檔案系統。

VMware vSphere 虛擬機器檔案系統限制

某些版本的 ESX 上資料儲存的最大大小為 64TB。

如果伺服器連接的儲存空間超過 64 TB，則可能需要設定多個 LUN，每個 LUN 的容量都小於 64 TB。建立多個 RAID 群組以縮短 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，需要考慮的重點是確保這些 LUN 具有相似且一致的效能。如果所有 LUN 都用於單一 ONTAP 聚合，這一點尤其重要。或者，如果一個或多個 LUN 的子集具有明顯不同的效能配置文件，我們強烈建議將這些 LUN 隔離到單獨的 ONTAP 聚合中。

可以使用多個檔案系統擴充區來建立單一資料儲存庫，最大可達資料儲存庫的最大大小。若要限制需要 ONTAP Select 許可證的容量，請務必在叢集安裝期間指定容量上限。此功能允許 ONTAP Select 僅使用資料儲存庫中的一部分空間（因此需要許可證）。

或者，也可以先在單一 LUN 上建立單一資料儲存庫。當需要更大容量的 ONTAP Select 授權來增加空間時，可以將該空間作為擴充區新增至相同資料儲存庫，直到達到資料儲存庫的最大大小。達到最大大小後，可以建立新的資料儲存庫並將其新增至 ONTAP Select。兩種容量擴展操作均受支持，並且可以透過使用 ONTAP Deploy 的儲存添加功能來實現。每個 ONTAP Select 節點可以配置為支援高達 400 TB 的儲存容量。從多個資料儲存庫配置

容量需要兩個步驟。

初始叢集建立可用於建立使用初始資料儲存庫中部分或全部空間的ONTAP Select叢集。第二步是使用其他資料儲存庫執行一個或多個容量添加操作，直到達到所需的總容量。此功能在本節中詳細介紹["增加儲存容量"](#)。



VMFS 開銷非零（參見["VMware 知識庫文章 1001618"](#)），並嘗試使用資料儲存區報告為可用的整個空間，導致叢集建立作業期間出現虛假錯誤。

每個資料儲存庫中都有 2% 的緩衝區未使用。此空間不需要容量許可證，因為ONTAP Select不使用它。只要未指定容量上限，ONTAP Deploy 就會自動計算緩衝區的準確 GB 數。如果指定了容量上限，則首先強制執行該大小。如果容量上限大小在緩衝區大小範圍內，則叢集建立將會失敗，並顯示錯誤訊息，其中指定了可用作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 既支援新安裝，也支援作為現有ONTAP Deploy 或ONTAP Select VM 的 Storage vMotion 作業的目標。

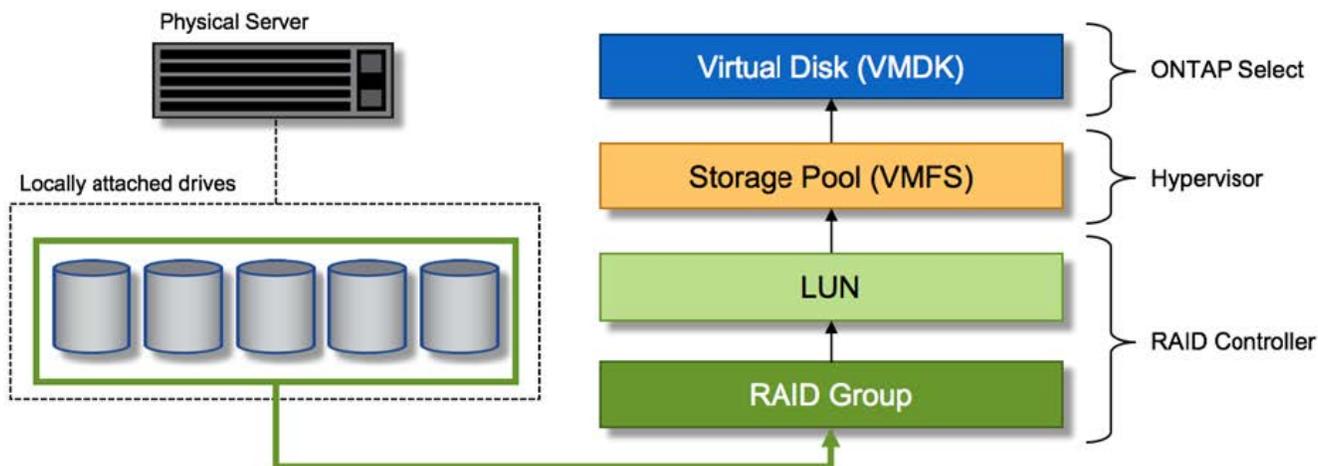
VMware 不支援從 VMFS 5 到 VMFS 6 的就地升級。因此，Storage vMotion 是唯一允許任何虛擬機器從 VMFS 5 資料儲存過渡到 VMFS 6 資料儲存的機制。但是，除了從 VMFS 5 過渡到 VMFS 6 這一特定目的之外，ONTAP Select和ONTAP Deploy 對 Storage vMotion 的支援已擴展至涵蓋其他場景。

ONTAP Select虛擬磁碟

ONTAP Select 的核心是為ONTAP提供一組從一個或多個儲存池配置的虛擬磁碟。ONTAPONTAP取得一組虛擬磁碟，並將其視為實體磁碟，而儲存堆疊的其餘部分則由虛擬機器管理程式抽象化。下圖更詳細地展示了這種關係，突顯了實體 RAID 控制器、虛擬機器管理程式和ONTAP Select虛擬機器之間的關係。

- RAID 群組和 LUN 的設定在伺服器的 RAID 控制器軟體中進行。使用 VSAN 或外部陣列時，無需進行此配置。
- 儲存池配置在虛擬機器管理程式內部進行。
- 虛擬磁碟由各個虛擬機器建立和擁有；在此範例中，由ONTAP Select建立和擁有。

虛擬磁碟到實體磁碟的映射



虛擬磁碟配置

為了提供更簡化的使用者體驗，ONTAP Select 管理工具 ONTAP Deploy 會自動從關聯的儲存池配置虛擬磁碟，並將其連接到 ONTAP Select 虛擬機器。此操作在初始設定和儲存新增操作期間都會自動執行。如果 ONTAP Select 節點屬於 HA 對，則虛擬磁碟會自動指派給本機和鏡像儲存池。

ONTAP Select 會將底層連接的儲存分割成大小相等的虛擬磁碟，每個虛擬磁碟的容量不超過 16 TB。如果 ONTAP Select 節點屬於 HA 對，則系統會在每個叢集節點上建立至少兩個虛擬磁碟，並將其指派給本機叢和鏡像叢，以便在鏡像聚合中使用。

例如，可以為 ONTAP Select 分配一個 31 TB 的資料儲存或 LUN（即部署虛擬機器並配置系統磁碟和根磁碟後剩餘的空間）。然後，建立四個約 7.75 TB 的虛擬磁碟，並將其指派給對應的 ONTAP 本機 Plex 和映像 Plex。



在 ONTAP Select 虛擬機器中新增容量可能會導致 VMDK 大小不同。如需詳細資訊，請參閱 ["增加儲存容量"](#)。與 FAS 系統不同，不同大小的 VMDK 可以存在於同一聚合中。在這些 VMDK 中使用 RAID 0 條帶，這樣就可以充分利用每個 VMDK 中的所有空間，無論其大小如何。

虛擬化 NVRAM

NetApp FAS 系統傳統上配備實體 NVRAM PCI 卡，這是一種包含非揮發性快閃記憶體的高效能卡。該卡使 ONTAP 能夠立即向客戶端確認傳入的寫入操作，從而顯著提升寫入效能。它還可以安排將已修改的資料塊移回速度較慢的儲存介質，這個過程稱為「降級暫存」。

商用系統通常不配備此類設備。因此，此 NVRAM 卡的功能已被虛擬化並放置在 ONTAP Select 系統啟動磁碟的一個分割區中。正因如此，實例的系統虛擬磁碟的放置至關重要。這也是為什麼該產品需要配備具有彈性快取的實體 RAID 控制器，以用於本地連接儲存配置。

NVRAM 位於其自己的 VMDK 上。將 NVRAM 分割到自己的 VMDK 中，可使 ONTAP Select 虛擬機器使用 vNVM 驅動程式與其 NVRAM VMDK 通訊。此外，也要求 ONTAP Select 虛擬機器使用硬體版本 13，該版本與 ESX 6.5 及更高版本相容。

資料路徑解釋：NVRAM 和 RAID 控制器

透過遍歷寫入請求進入系統時所採用的資料路徑，可以最好地突出顯示虛擬化 NVRAM 系統分區和 RAID 控制器之間的交互。

傳入ONTAP Select虛擬機器的寫入要求會以虛擬機器的NVRAM分區為目標。在虛擬化層，此分割區位於ONTAP Select系統磁碟（即連接到ONTAP Select虛擬機器的 VMDK）內。在實體層，這些請求會快取在本機 RAID 控制器中，就像所有針對底層磁碟軸的區塊變更一樣。從這裡，寫入操作會回傳給主機確認。

此時，物理上，該區塊駐留在 RAID 控制器快取中，等待刷新到磁碟。邏輯上，該區塊駐留在NVRAM中，等待降級到對應的用戶資料磁碟。

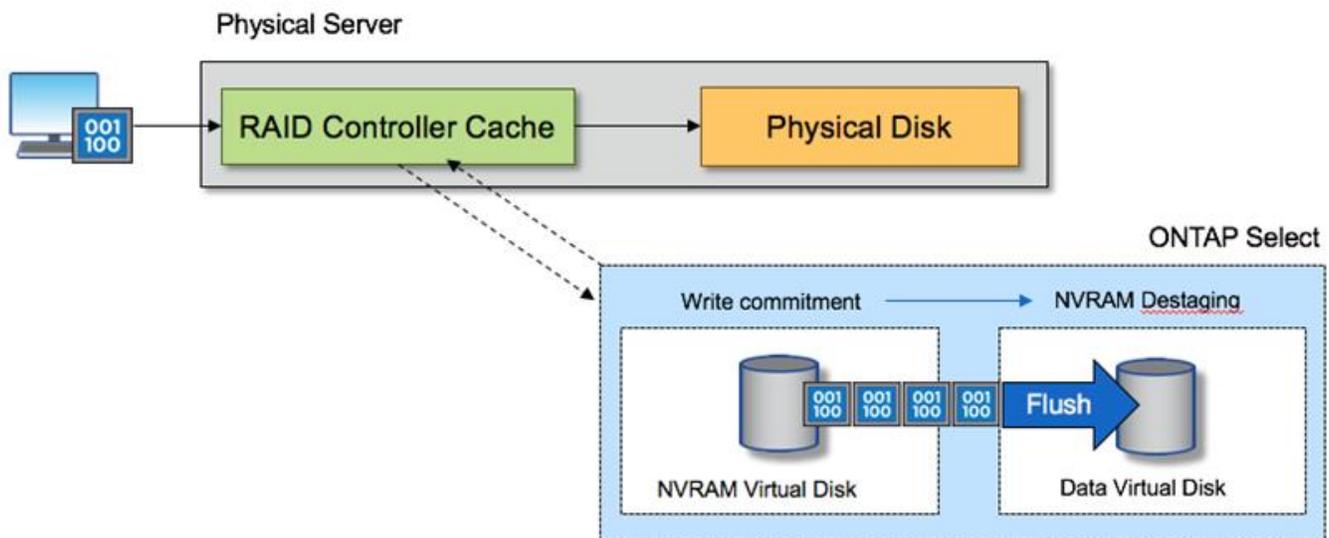
由於變更的區塊會自動儲存在 RAID 控制器的本機快取中，因此傳入NVRAM分割區的寫入作業也會自動快取並定期刷新到實體儲存媒體。請勿將此與定期將NVRAM內容刷新回ONTAP資料磁碟的操作混淆。這兩個事件互不相關，發生的時間和頻率也不同。

下圖展示了傳入寫入所採用的 I/O 路徑。它突出顯示了實體層（由 RAID 控制器快取和磁碟表示）與虛擬層（由虛擬機器的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在本機 RAID 控制器快取中，但該快取無法感知虛擬機器結構或其虛擬磁碟。它會儲存系統上所有已更改的區塊，而NVRAM只是其中的一部分。這包括發送到虛擬機器管理程式的寫入請求（前提是虛擬機器管理程式是從相同的備用主軸配置的）。

傳入ONTAP Select VM 的寫入



NVRAM分區在其自己的 VMDK 上獨立存在。此 VMDK 使用 ESX 6.5 或更高版本中提供的 vNVME 驅動程式進行連線。此變更對於使用軟體 RAID 的ONTAP Select安裝最為重要，因為此類安裝無法從 RAID 控制器快取中獲益。

適用於本機連接儲存的ONTAP Select軟體 RAID 配置服務

軟體 RAID 是在ONTAP軟體堆疊內實作的 RAID 抽象層。它提供的功能與傳統ONTAP平台（例如FAS）中的 RAID 層相同。RAID層執行磁碟機奇偶校驗計算，並針對ONTAP Select節點中的單一磁碟機故障提供保護。

除了硬體 RAID 配置之外，ONTAP Select還提供了軟體 RAID 選項。在某些環境中，例如在小型商用硬體上部署ONTAP Select時，硬體 RAID 控制器可能無法使用或不合適。軟體 RAID 擴展了可用的部署選項，使其涵蓋此類環境。要在您的環境中啟用軟體 RAID，請記住以下幾點：

- 它可透過 Premium 或 Premium XL 許可證獲得。
- 它僅支援ONTAP根磁碟和資料磁碟的 SSD 或 NVMe（需要 Premium XL 許可證）磁碟機。
- 它需要一個單獨的系統磁碟作為ONTAP Select VM 啟動分割區。
 - 選擇一個單獨的磁碟（SSD 或 NVMe 磁碟機），為系統磁碟（多節點設定中的NVRAM、Boot/CF 卡、Coredump 和 Mediator）建立資料儲存。

筆記

- 服務磁碟和系統磁碟這兩個術語可以互換使用。
 - 服務磁碟是ONTAP Select VM 內使用的 VMDK，用於服務各種項目，例如叢集、啟動等。
 - 從主機的角度來看，服務磁碟物理上位於單一實體磁碟上（統稱為服務/系統實體磁碟）。此實體磁碟必須包含 DAS 資料儲存。ONTAPDeploy 會在叢集部署期間為ONTAP ONTAP Select虛擬機器建立這些服務磁碟。
- 無法進一步跨多個資料儲存庫或跨多個實體磁碟機分離ONTAP Select系統磁碟。
- 硬體 RAID 並未被棄用。

本機連接儲存的軟體 RAID 配置

使用軟體 RAID 時，理想情況下沒有硬體 RAID 控制器，但是，如果系統確實有現有的 RAID 控制器，則必須遵守以下要求：

- 必須停用硬體 RAID 控制器，以便磁碟可以直接呈現給系統（JBOD）。此變更通常可以在 RAID 控制器 BIOS 中進行。
- 或者，硬體 RAID 控制器應處於 SAS HBA 模式。例如，某些 BIOS 配置除了 RAID 之外還允許「AHCI」模式，可以選擇該模式來啟用 JBOD 模式。這將啟用直通功能，以便實體磁碟機在主機上可以按原樣顯示。

根據控制器支援的最大驅動器數量，可能需要額外的控制器。在 SAS HBA 模式下，請確保 IO 控制器 (SAS HBA) 至少支援 6Gb/s 的速度。不過，NetApp建議使用 12Gbps 的速度。

不支援任何其他硬體 RAID 控制器模式或配置。例如，某些控制器允許 RAID 0 支持，這可以人為地啟用磁碟直通功能，但其後果可能不堪設想。支援的實體磁碟（僅限 SSD）大小在 200GB 至 16TB 之間。



管理員需要追蹤ONTAP Select VM 正在使用的驅動器，並防止在主機上無意中使用這些驅動器。

ONTAP Select虛擬磁碟和實體磁碟

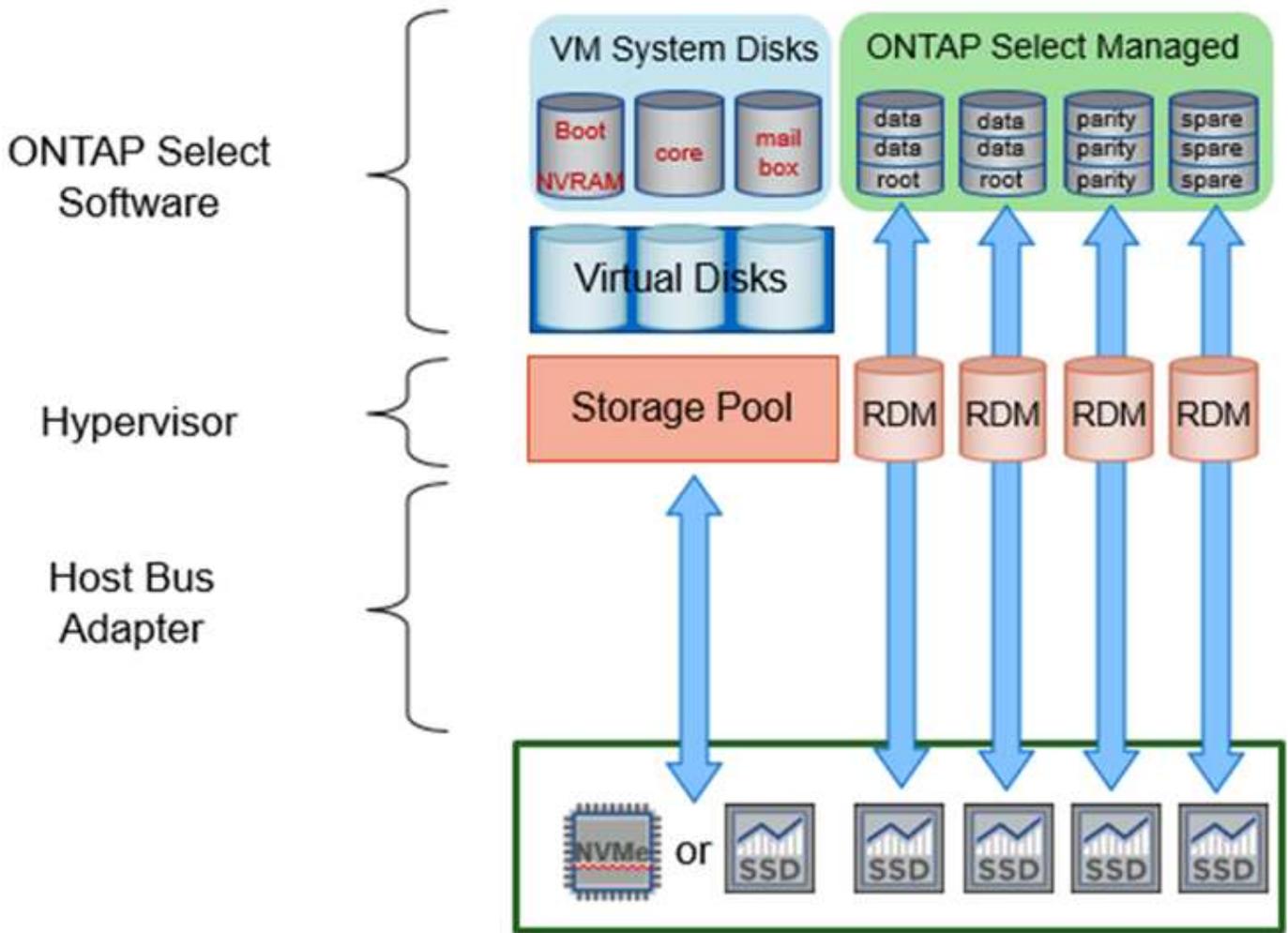
對於使用硬體 RAID 控制器的配置，實體磁碟冗餘由 RAID 控制器提供。ONTAPONTAP Select會提供一個或多個 VMDK，ONTAP管理員可以從中配置資料聚合。這些 VMDK 以 RAID 0 格式進行條帶化，因為ONTAP軟體 RAID 冗餘、低效且低效，這是由於硬體層級提供的彈性所致。此外，用於系統磁碟的 VMDK 與用於儲存使用者資料的 VMDK 位於相同資料儲存中。

使用軟體 RAID 時，ONTAP Deploy 會向ONTAP Select提供一組虛擬磁碟 (VMDK) 和實體磁碟原始設備對應 [RDM]（用於 SSD）以及直通或 DirectPath IO 設備（用於 NVMe）。

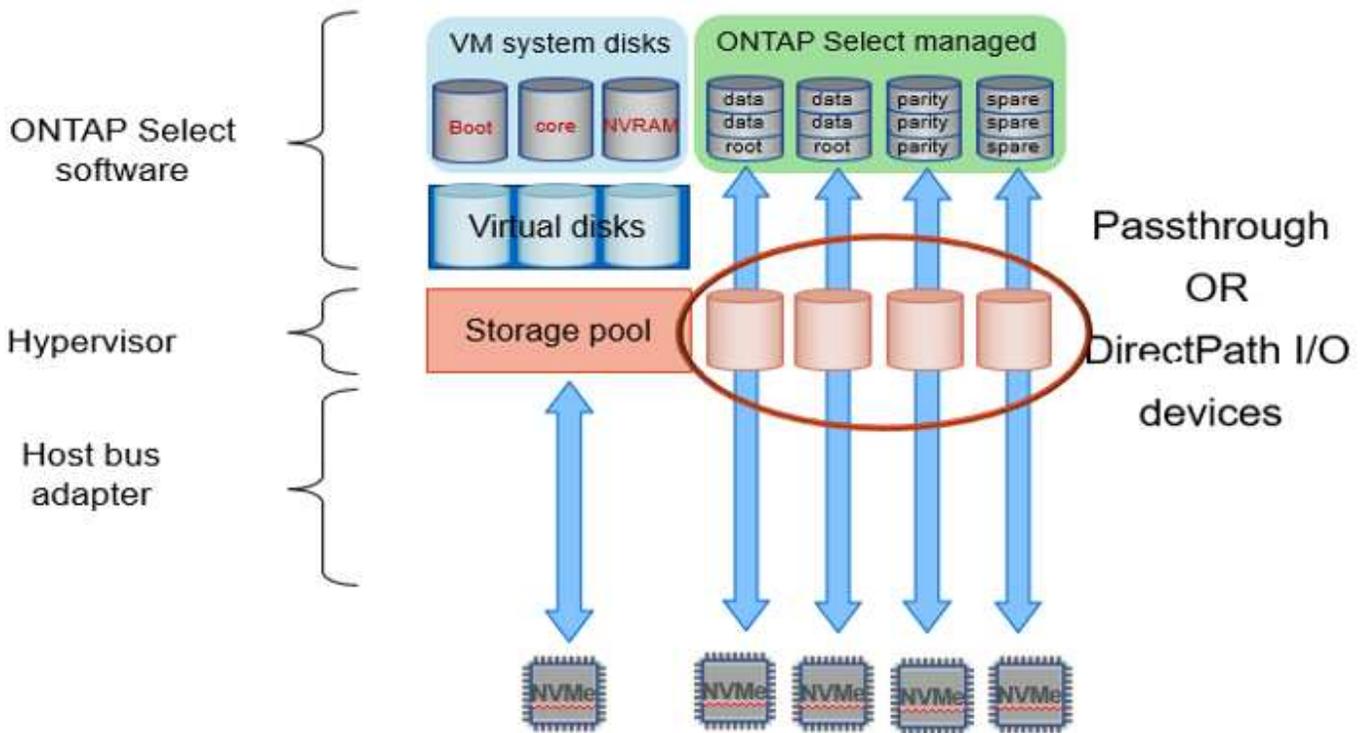
下圖更詳細地展示了這種關係，突顯了用於ONTAP Select VM 內部的虛擬化磁碟與用於儲存使用者資料的實體磁碟之間的差異。

- ONTAP Select軟體 RAID：使用虛擬化磁碟和 RDM*

ONTAP Select with Software RAID



系統磁碟 (VMDK) 位於相同資料儲存區和相同實體磁碟上。虛擬NVRAM磁碟需要快速且耐用的媒體。因此，僅支援 NVMe 和 SSD 類型的資料儲存區。



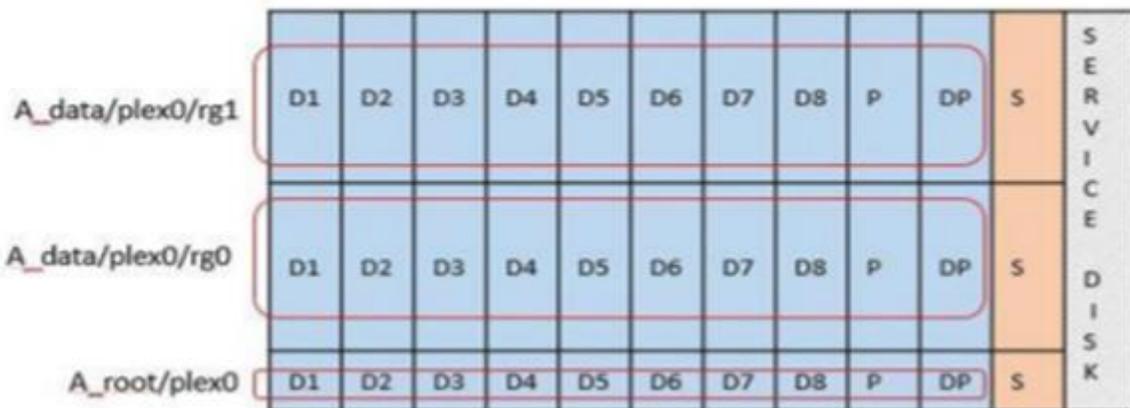
系統磁碟 (VMDK) 位於相同資料儲存區和相同實體磁碟上。虛擬NVRAM磁碟需要快速且耐用的媒體。因此，僅支援 NVMe 和 SSD 類型的資料儲存區。使用 NVMe 磁碟機儲存資料時，出於效能考慮，系統磁碟也應為 NVMe 裝置。在全 NVMe 配置中，英特爾傲騰卡是系統磁碟的理想選擇。

i 在目前版本中，無法進一步跨多個資料儲存庫或多個實體磁碟機分離ONTAP Select系統磁碟。

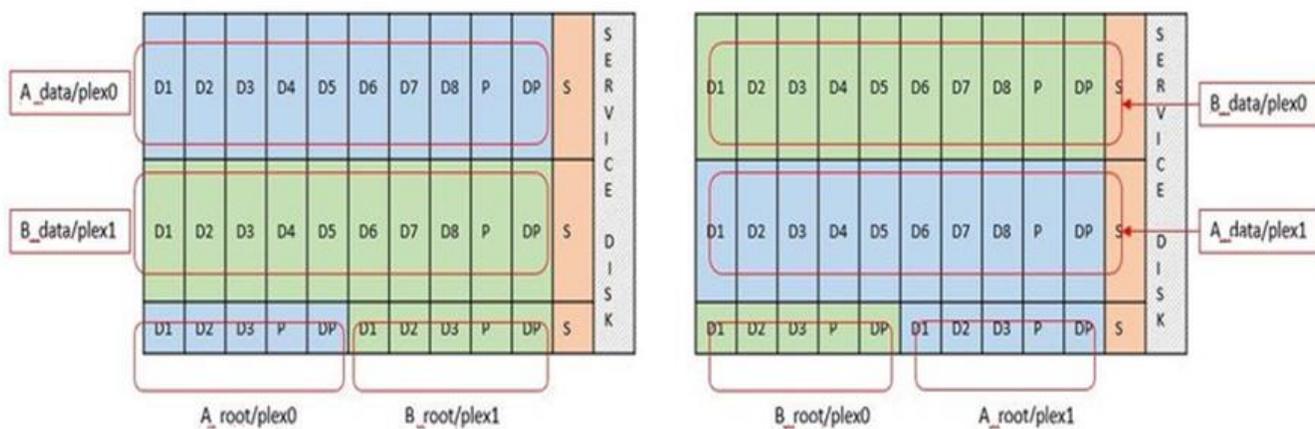
每個資料磁碟分為三個部分：一個小的根分割區（條帶）和兩個大小相等的分割區，從而建立在ONTAP Select 虛擬機器中可見的兩個資料磁碟。分區使用根資料 (RD2) 模式，如下圖所示，針對單節點叢集和 HA 對中的節點。

`P` 表示奇偶校驗磁碟機。 `DP` 表示雙奇偶校驗驅動器，且 `S` 表示備用驅動器。

單節點叢集的 RDD 磁碟分割區



多節點叢集 (HA 對) 的 RDD 磁碟分割區



ONTAP軟體 RAID 支援下列 RAID 類型：RAID 4、RAID-DP 和RAID-TEC。這些是FAS和AFF平台使用的相同 RAID 結構。對於根配置，ONTAP Select僅支援 RAID 4 和 RAID-DP。當使用RAID-TEC進行資料聚合時，整體保護是 RAID-DP。ONTAP Select HA 使用無共享架構，將每個節點的配置複製到另一個節點。這意味著每個節點必須儲存其根分區及其對等節點根分區的副本。由於資料磁碟只有一個根分割區，因此最低資料磁碟數量將取決於ONTAP Select節點是否屬於 HA 對。

對於單節點集群，所有資料分區用於儲存本地（活動）資料。對於屬於 HA 對的節點，一個資料分區用於儲存該節點的本地（活動）數據，另一個資料分區用於鏡像來自 HA 對等節點的活動資料。

直通 (DirectPath IO) 設備與原始設備對映 (RDM)

VMware ESX 目前不支援將 NVMe 磁碟用作原始裝置對應。若要讓ONTAP Select直接控制 NVMe 磁碟，必須在 ESX 中將 NVMe 磁碟機配置為直通裝置。請注意，將 NVMe 設備配置為直通設備需要伺服器 BIOS 的支持，並且這是一個中斷過程，需要重新啟動 ESX 主機。此外，每個 ESX 主機的最大直通裝置數量為 16 個。但是，ONTAP Deploy 將其限制為 14 個。每個ONTAP Select節點 14 個 NVMe 設備的限制意味著全 NVMe 配置將提供非常高的 IOP 密度 (IOP/TB)，但總容量會降低。或者，如果需要具有更大儲存容量的高效能配置，建議的配置是較大的ONTAP Select VM 大小、用於系統磁碟的 INTEL Optane 卡以及用於資料儲存的額定數量的 SSD 磁碟機。



為了充分利用 NVMe 效能，請考慮較大的ONTAP Select VM 大小。

直通設備和 RDM 之間還有一個區別。RDM可以對應到正在執行的虛擬機器。直通設備需要重新啟動虛擬機器。這意味著任何 NVMe 驅動器更換或容量擴展（驅動器添加）過程都需要重新啟動ONTAP Select虛擬機。驅動器更換和容量擴展（驅動器添加）操作由ONTAP Deploy 中的工作流程驅動。ONTAP管理單節點叢集的ONTAP Select重新啟動以及 HA 對的故障轉移/故障復原。但是，必須注意使用 SSD 資料磁碟機（不需要ONTAP Select重新啟動/故障轉移）和使用 NVMe 資料磁碟機（需要ONTAP Select重新啟動/故障轉移）之間的差異。

實體和虛擬磁碟配置

為了提供更簡化的使用者體驗，ONTAP Deploy 會自動從指定的資料儲存庫（實體系統磁碟）配置系統（虛擬）磁碟，並將其連接到ONTAP Select虛擬機器。此操作會在初始設定期間自動執行，以便ONTAP Select虛擬機器能夠啟動。RDM會進行分割區，並且會自動建置根聚合。如果ONTAP Select節點屬於 HA 對，則資料分區會自動指派給本機儲存池和鏡像儲存池。此分配操作會在叢集建立操作和儲存新增操作期間自動進行。

由於ONTAP Select VM 上的資料磁碟與底層實體磁碟相關聯，因此建立具有大量實體磁碟的配置會對效能產生影響。



根聚合的 RAID 群組類型取決於可用的磁碟數量。ONTAP 會選擇適當的 RAID 群組類型。如果指派給節點的磁碟數量足夠，則使用 RAID-DP；否則，則建立 RAID-4 根聚合。

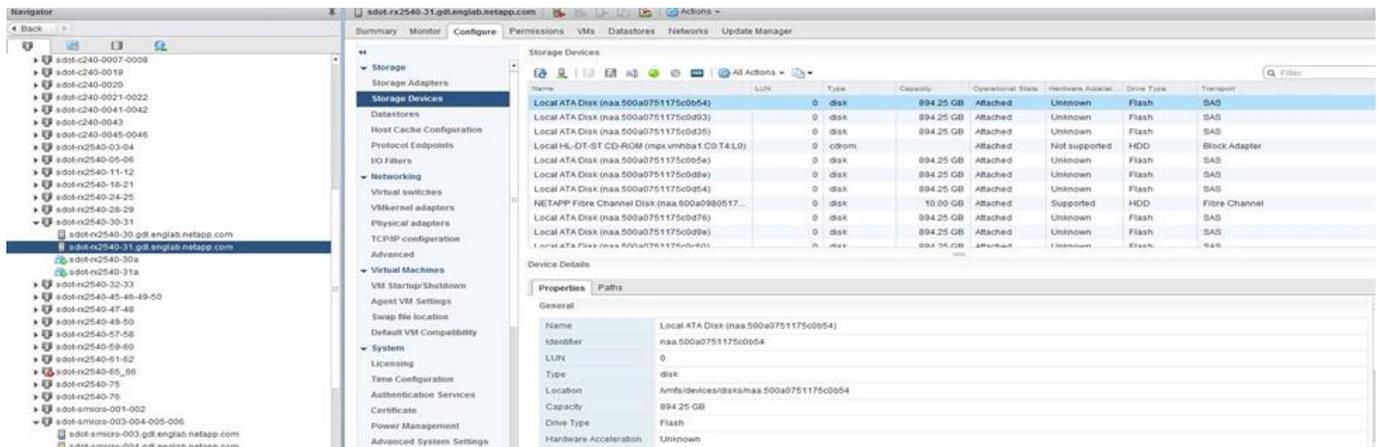
使用軟體 RAID 為 ONTAP Select 虛擬機器新增容量時，管理員必須考慮實體磁碟機的大小以及所需的磁碟機數量。有關詳細信息，請參閱“增加儲存容量”。

與 FAS 和 AFF 系統類似，只有容量相等或更大的驅動器才能新增到現有 RAID 群組。容量更大的驅動器大小合適。如果您要建立新的 RAID 群組，則新 RAID 群組的大小應與現有 RAID 群組的大小相匹配，以確保整體聚合效能不會下降。

將 ONTAP Select 磁碟與對應的 ESX 磁碟相匹配

ONTAP Select 磁碟通常標記為 NET xy 您可以使用下列 ONTAP 指令來取得磁碟 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，您可以輸入以下命令來使給定實體磁碟（由其 naa.unique-id 標識）的 LED 閃爍。

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

使用軟體 RAID 時出現多個磁碟機故障

系統可能會遇到多個驅動器同時處於故障狀態的情況。系統的行為取決於聚合 RAID 保護和故障磁碟機的數量。

RAID4 聚合可承受一個磁碟故障，RAID-DP 聚合可承受兩個磁碟故障，而RAID-TEC聚合可承受三個磁碟故障。

如果故障磁碟數量小於該 RAID 類型支援的最大故障數量，並且有備用磁碟可用，則重建過程將自動啟動。如果沒有備用磁碟可用，則聚合將以降級狀態提供數據，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量，則本機叢會被標記為故障，聚合狀態也會降級。數據由位於 HA 配對節點上的第二個叢提供。這意味著，節點 1 的任何 I/O 請求都會透過叢集互連連接埠 e0e (iSCSI) 傳送到實體上位於節點 2 上的磁碟。如果第二個叢也發生故障，則聚合會被標記為故障，資料不可用。

必須刪除並重新建立發生故障的 plex，才能恢復正確的資料鏡像。請注意，多磁碟故障導致資料聚合降級，也會導致根聚合降級。ONTAPONTAP Select使用根-資料-資料 (RDD) 分區模式將每個實體磁碟機分割為一個根分割區和兩個資料分割區。因此，遺失一個或多個磁碟可能會影響多個聚合，包括本地根聚合或遠端根聚合的副本，以及本機資料聚合和遠端資料聚合的副本。

```
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
Size      Position  Disk                               Type      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..



為了測試或模擬一個或多個驅動器故障，請使用 `storage disk fail -disk NET-x.y -immediate`` 命令。如果系統中有備用磁碟，聚合將開始重建。您可以使用命令檢查重建的狀態 ``storage aggregate show`。您可以使用ONTAP Deploy 移除類比故障磁碟機。請注意 ONTAP 已將磁碟機標示為 Broken。驅動器實際上並未損壞，可以使用ONTAP Deploy 重新添加。若要清除「Broken」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該是空的。

虛擬化NVRAM

NetApp FAS系統傳統上配備實體NVRAM PCI 卡。該卡是一款高效能卡，包含非揮發性閃存，可大幅提升寫入效能。它透過授予ONTAP立即向客戶端確認傳入寫入的能力來實現這一點。它還可以安排將已修改的資料塊移回速度較慢的儲存介質，這個過程稱為「降級暫存」。

商用系統通常不配備此類設備。因此，NVRAM卡的功能已虛擬化，並放置在ONTAP Select系統啟動磁碟的一個分割區中。正因如此，實例的系統虛擬磁碟的放置位置至關重要。

ONTAP Select VSAN 和外部陣列配置

虛擬 NAS (vNAS) 部署支援虛擬 SAN (VSAN)、部分 HCI 產品以及外部陣列類型的資料儲存上的ONTAP Select叢集。這些配置的底層基礎架構提供了資料儲存的彈性。

最低要求是底層配置受 VMware 支持，並且應在相應的 VMware HCL 上列出。

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集群，這包括同一 HA 對中的兩個ONTAP Select節點共享單一資料儲存庫（包括 vSAN 資料儲存庫）的架構。這些節點也可以安裝在同一共用外部陣列的不同資料儲存庫上。這樣可以提高陣列端儲存效率，從而減少整個ONTAP Select HA 對的整體佔用空間。ONTAPONTAP Select vNAS 解決方案的架構與具有本機 RAID 控制器的 DAS 上的ONTAP Select架構非常相似。也就是說，每個ONTAP Select節點都會繼續擁有其 HA 夥伴節點資料的副本。ONTAP儲存效率策略是節點範圍的。因此，陣列端儲存效率是可取的，因為它們可以應用於來自兩個ONTAP Select節點的資料集。

HA 對中的每個ONTAP Select節點也可能使用單獨的外部陣列。將ONTAP Select Metrocluster SDS 與外部儲存結合使用時，這是常見的選擇。

當為每個ONTAP Select節點使用單獨的外部陣列時，兩個陣列為ONTAP Select VM 提供相似的效能特徵非常重要。

vNAS 架構與配備硬體 RAID 控制器的本機 DAS

從邏輯上講，vNAS 架構與配備 DAS 和 RAID 控制器的伺服器架構最相似。在這兩種情況下，ONTAP Select都會佔用資料儲存空間。此資料儲存空間被劃分為 VMDK，這些 VMDK 構成傳統的ONTAP資料聚合。在叢集建立和儲存新增作業期間，ONTAP Deploy 會確保 VMDK 的大小正確，並指派給正確的叢（對於 HA 對）。

vNAS 與配備 RAID 控制器的 DAS 之間有兩個主要區別。最直接的區別是 vNAS 不需要 RAID 控制器。vNAS 假設底層外部陣列能夠提供配備 RAID 控制器的 DAS 所提供的資料持久性和彈性。第二個更細微的差異與NVRAM效能有關。

vNAS NVRAM

ONTAP Select NVRAM是一種 VMDK。換句話說，ONTAP Select在區塊尋址設備 (VMDK) 之上模擬了一個位元組可尋址空間 (傳統NVRAM)。然而，NVRAM的效能對於ONTAP Select節點的整體效能至關重要。

對於具有硬體 RAID 控制器的 DAS 設置，硬體 RAID 控制器快取充當事實上的NVRAM緩存，因為對NVRAM VMDK 的所有寫入都首先託管在 RAID 控制器快取中。

對於 vNAS 架構，ONTAP Deploy 會自動使用名為「單一實例資料日誌記錄 (SIDL)」的啟動參數來設定ONTAP Select節點。當此啟動參數存在時，ONTAP Select會繞過NVRAM，將資料負載直接寫入資料聚合。NVRAM用於記錄 WRITE 操作變更的區塊的位址。此功能的優點在於避免了雙重寫入：一次寫入NVRAM，另一次寫入NVRAM降級後。此功能僅適用於 vNAS，因為本地寫入 RAID 控制器快取的額外延遲可以忽略不計。

SIDL 功能並非與所有ONTAP Select儲存效率功能相容。可以使用以下命令在聚合層級停用 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。停用該聚合中所有磁碟區上的所有儲存效率策略後，可以重新啟用 SIDL 功能：

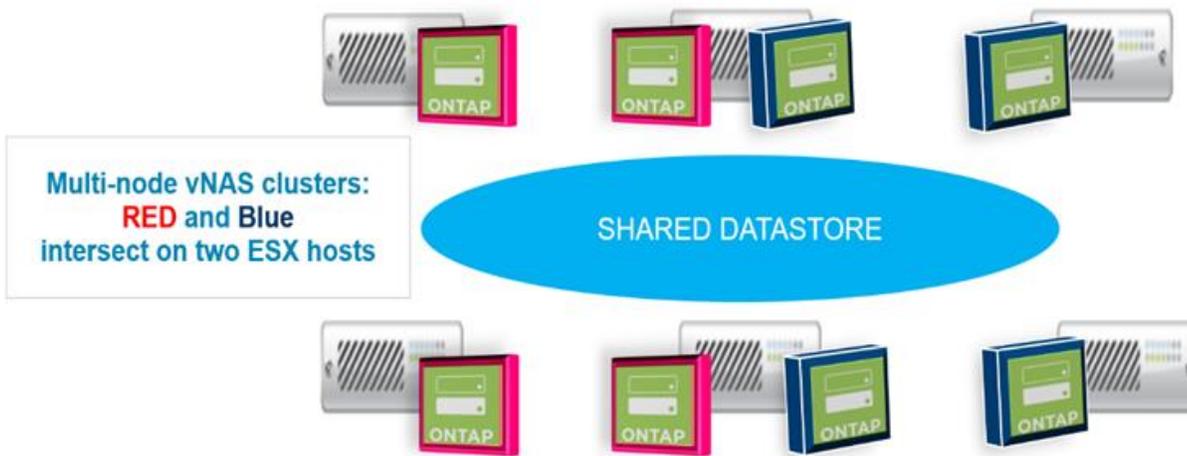
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

在 ESXi 上使用 vNAS 時並置ONTAP Select節點

ONTAP Select支援在共用儲存上部署多節點ONTAP Select叢集。ONTAP支援在同一 ESX 主機上配置多個ONTAP Select節點，前提是這些節點不屬於同一叢集。請注意，此配置僅適用於 vNAS 環境 (共用資料儲存)。使用 DAS 儲存時，不支援每個主機部署多個ONTAP Select實例，因為這些實例會爭用相同硬體 RAID 控制器。

ONTAP Deploy 確保多節點 vNAS 叢集的初始部署不會將來自相同叢集的多個ONTAP Select實例放置在同一台主機上。下圖顯示了兩個在兩台主機上相交的四節點叢集的正確部署範例。

多節點 vNAS 叢集的初始部署

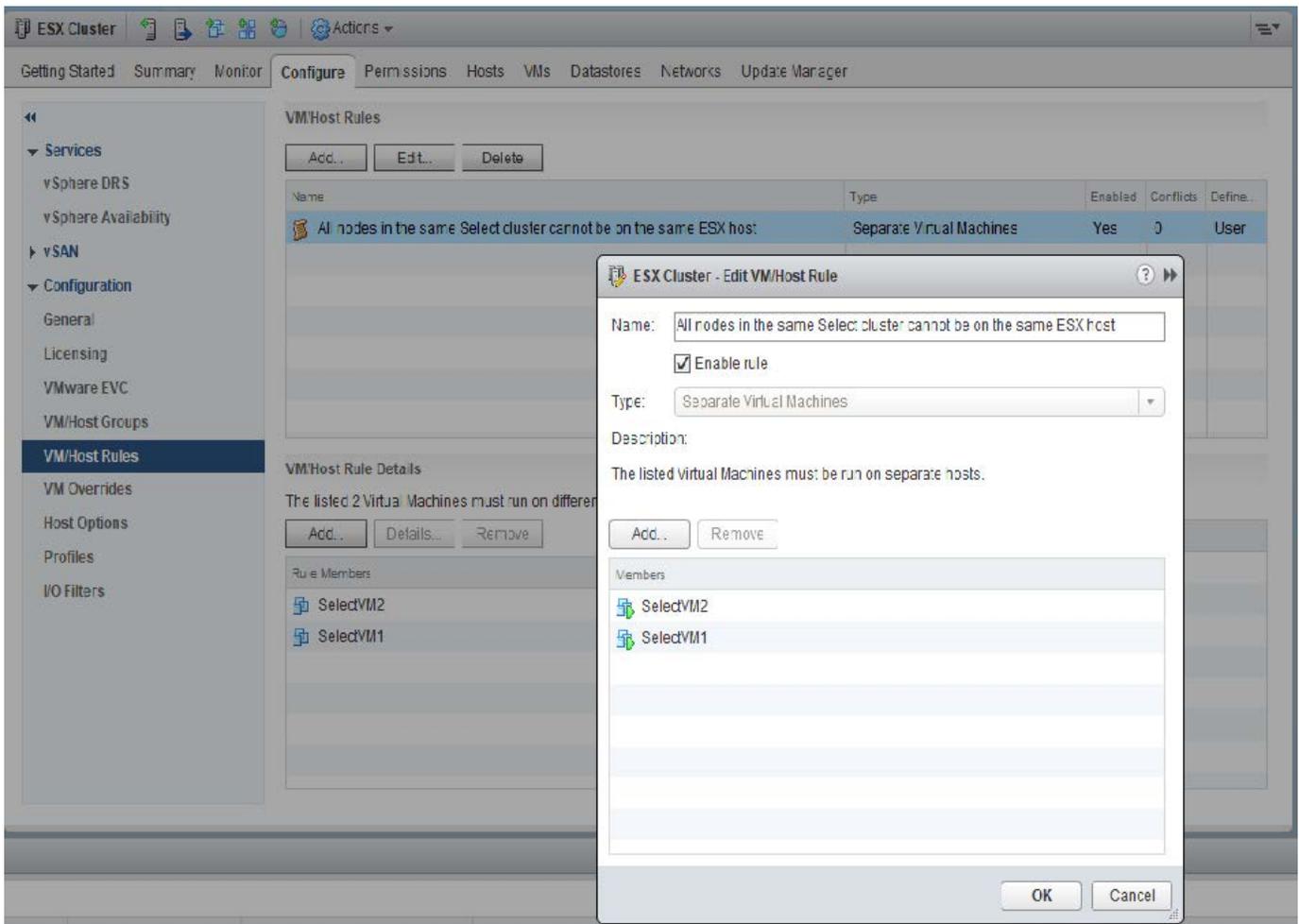


部署後，ONTAP Select節點可以在主機之間遷移。這可能會導致配置不理想且不受支持，因為同一叢集中的兩個或多個ONTAP Select節點共用同一底層主機。NetApp建議手動建立虛擬機器反關聯性規則，以便VMware自動維護同一叢集中節點之間的實體隔離，而不僅僅是同一HA對中的節點之間的實體隔離。

 反親和性規則要求在ESX叢集上啟用DRS。

請參閱以下範例，以了解如何為ONTAP Select虛擬機器建立反關聯性規則。如果ONTAP Select叢集包含多個HA對，則叢集中的所有節點都必須包含在此規則中。

The screenshot shows the vSphere configuration interface for VM/Host Rules. The left-hand navigation pane is expanded to show 'VM/Host Rules' under the 'Configuration' section. The main content area has tabs for 'Configure', 'Permissions', 'Hosts', 'VMs', 'Datastores', 'Networks', and 'Update Manager'. The 'Configure' tab is active, showing a table for VM/Host Rules. The table has columns for Name, Type, Enabled, Conflicts, and Defined By. The table is currently empty, with the text 'This list is empty.' displayed in the center. Below the table, it says 'No VM/Host rule selected'.



由於以下原因之一，同一ONTAP Select叢集中的兩個或多個ONTAP Select節點可能會位於同一 ESX 主機上：

- 由於 VMware vSphere 授權限制或未啟用 DRS，因此不存在 DRS。
- 由於 VMware HA 操作或管理員啟動的 VM 遷移優先，因此 DRS 反親和性規則被繞過。

請注意，ONTAP Deploy 不會主動監控ONTAP Select虛擬機器的位置。但是，叢集刷新操作會在ONTAP Deploy 日誌中反映此不受支援的配置：



增加ONTAP Select存儲容量

ONTAP Deploy 可用於為ONTAP Select叢集中的每個節點新增和授權額外的儲存。

ONTAP Deploy 中的儲存新增功能是增加管理儲存量的唯一方法，不支援直接修改ONTAP Select虛擬機器。下圖顯示了用於啟動儲存新增精靈的「+」圖示。



以下注意事項對於容量擴展操作的成功至關重要。增加容量需要現有許可證覆蓋總空間量（現有空間量加上新空間量）。如果儲存新增操作導致節點超出其授權容量，則操作將會失敗。應先安裝具有足夠容量的新許可證。

如果將額外容量新增至現有的ONTAP Select聚合，則新的儲存池（資料儲存）的效能設定檔應與現有儲存池（資料儲存）的效能設定檔相似。請注意，無法將非 SSD 儲存新增至安裝了類似AFF特性（啟用快閃記憶體）的ONTAP Select節點。也不支援混合使用 DAS 和外部儲存。

如果將本機連線的儲存空間新增至系統以提供額外的本機 (DAS) 儲存池，則必須建置額外的 RAID 群組和 LUN（或多個 LUN）。與FAS系統一樣，如果要為相同聚合新增空間，應注意確保新 RAID 群組的效能與原始 RAID 群組的效能相似。如果要建立新的聚合，並且充分了解新聚合的效能影響，則新的 RAID 群組佈局可能會有所不同。

如果資料儲存的總大小不超過支援的最大資料儲存大小，則可以將新空間作為擴充區新增至相同資料儲存。將資料儲存擴充區新增至已安裝ONTAP Select 的資料儲存體中可以動態完成，且不會影響ONTAP Select節點的運作。

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

在 HA 對中，每個節點都包含來自其配對節點的資料鏡像副本。在節點 1 中新增空間需要在其配對節點（節點 2）中新增相同大小的空間，以便將節點 1 的所有資料複製到節點 2。換句話說，在節點 1 的容量添加操作中加入到節點 2 的空間在節點 2 上不可見或不可存取。將該空間加入節點 2 是為了確保節點 1 的資料在 HA 事件期間得到全面保護。

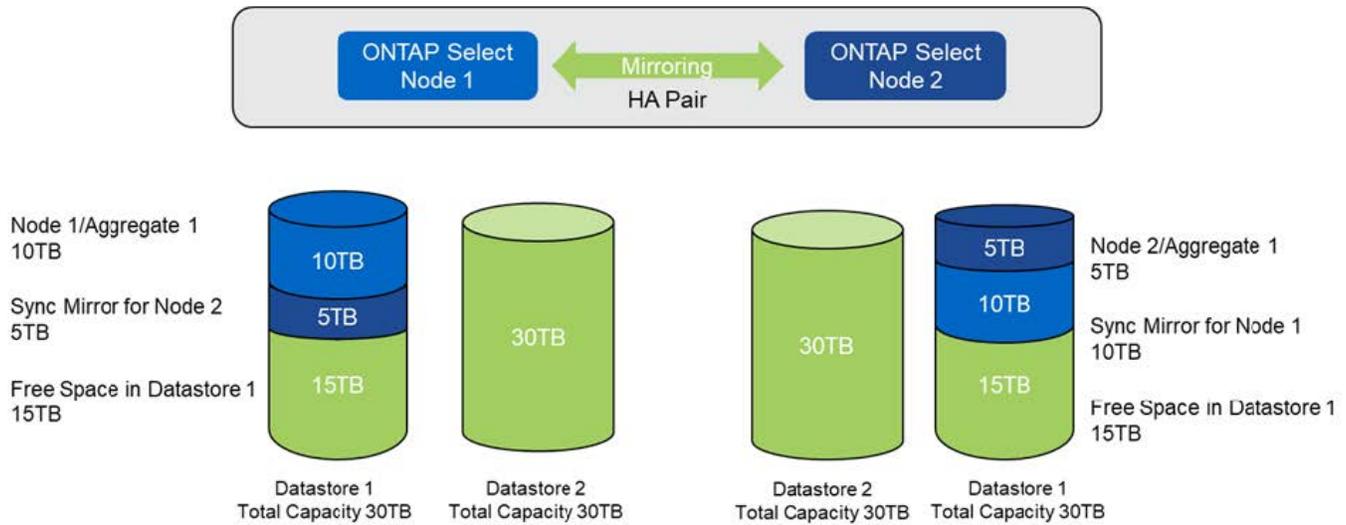
關於性能，還有一個額外的考慮因素。節點 1 上的資料會同步複製到節點 2。因此，節點 1 上新空間（資料儲存）的效能必須與節點 2 上新空間（資料儲存）的效能相符。換句話說，在兩個節點上添加空間，但使用不同的磁碟機技術或不同的 RAID 群組大小，可能會導致效能問題。這是由於 RAID SyncMirror操作作用於在配對節點上維護資料副本。

若要增加 HA 對中兩個節點上使用者可存取的容量，必須執行兩次儲存新增操作，每個節點一次。每次儲存新增操作都需要在兩個節點上增加額外的空間。每個節點所需的總空間等於節點 1 所需空間加上節點 2 所需空間。

初始設定包含兩個節點，每個節點包含兩個資料存儲，每個資料存儲有 30 TB 的空間。ONTAPDeploy 會建立一個雙節點集群，每個節點佔用資料儲存ONTAP中的 10 TB 空間。ONTAPDeploy 為每個節點配置 5 TB 的活動空間。

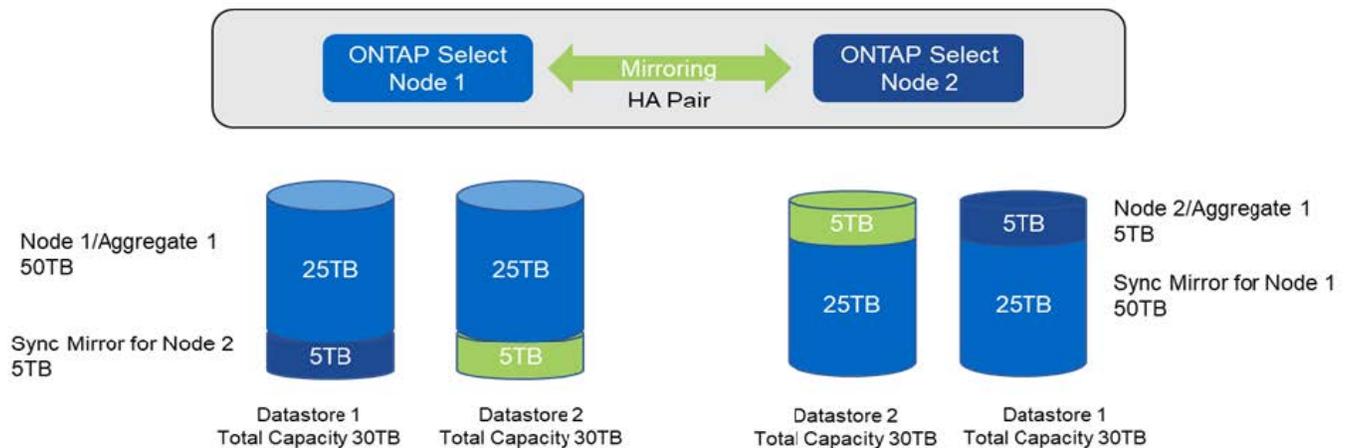
下圖顯示了節點 1 的單次儲存新增操作的結果。ONTAP ONTAP Select在每個節點上仍使用相同的儲存量 (15 TB)。但是，節點 1 的活動儲存量 (10 TB) 大於節點 2 的活動儲存量 (5 TB)。由於每個節點都託管對方節點資料的副本，因此兩個節點都受到完全保護。資料儲存庫 1 中剩餘額外的可用空間，而資料儲存庫 2 仍然完全可用。

容量分配：單次儲存新增操作後的分配與釋放空間



節點 1 上的另外兩個儲存新增作業佔用了資料儲存 1 的剩餘空間以及資料儲存 2 的一部分空間（使用容量上限）。第一次儲存新增作業佔用了資料儲存 1 中剩餘的 15 TB 可用空間。下圖顯示了第二次儲存新增操作的結果。此時，節點 1 管理 50 TB 的活躍數據，而節點 2 則擁有原始的 5 TB。

容量分佈：節點 1 執行兩次額外儲存新增作業後的分配和可用空間



容量添加操作期間使用的最大 VMDK 大小為 16 TB。叢集建立操作期間使用的最大 VMDK 大小仍為 8 TB。ONTAP 會根據您的配置（單節點或多節點叢集）和要新增的容量建立正確大小的 VMDK。但是，叢集建立作業期間每個 VMDK 的最大大小不得超過 8 TB，儲存新增作業期間每個 VMDK 的最大大小不得超過 16 TB。

使用軟體 RAID 增加 ONTAP Select 的容量

類似地，儲存新增精靈也可用於增加使用軟體 RAID 的 ONTAP Select 節點的管理容量。此精靈僅顯示可用的 DAS SSD 驅動器，這些驅動器可以作為 RDM 對應到 ONTAP Select 虛擬機器。

雖然容量許可證可以增加 1 TB，但在使用軟體 RAID 時，無法以實體方式將容量增加 1 TB。與在 FAS 或 AFF 陣列新增磁碟類似，某些因素決定了單次操作中可新增的最小儲存量。

請注意，在 HA 對中，為節點 1 新增儲存需要節點的 HA 對（節點 2）上也擁有相同數量的磁碟機。本機磁碟機

和遠端磁碟均由節點 1 上的一次儲存新增作業使用。也就是說，遠端驅動器用於確保節點 1 上的新儲存空間在節點 2 上得到複製和保護。為了在節點 2 上添加本地可用的存儲，兩個節點都必須執行單獨的存儲添加操作，並且擁有相同數量的驅動器。

ONTAP Select會將所有新磁碟機分割為與現有磁碟機相同的根分割區、資料分割區和資料分割區。分區操作在建立新聚合或在現有聚合上擴充期間進行。每個磁碟上根分割區條帶的大小設定為與現有磁碟上的現有根分割區大小相符。因此，兩個相等的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再除以二。根分區條帶大小是可變的，在初始叢集設定期間如下計算。所需的總根空間（單節點叢集為 68 GB，HA 對為 136 GB）在初始磁碟數量減去任何備用磁碟機和奇偶校驗磁碟機後分配。根分割區條帶大小在所有新增至系統的磁碟機上保持不變。

如果要建立新的聚合，則所需的最小磁碟機數量取決於 RAID 類型以及ONTAP Select節點是否屬於 HA 對的一部分。

如果將儲存空間新增至現有聚合，則需要考慮一些其他事項。假設 RAID 群組尚未達到最大限制，則可以將磁碟機新增至現有 RAID 群組。將磁碟軸新增至現有 RAID 群組的傳統FAS和AFF最佳實務也適用於此，並且在新磁碟軸上建立熱點是一個潛在的問題。此外，只有資料分割區大小相等或更大的磁碟機才能新增至現有 RAID 群組。如上所述，資料分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分割區，則新磁碟機的大小是適當的。換句話說，每個新驅動器的一部分容量仍未被利用。

您也可以使用新磁碟機建立新的 RAID 群組，作為現有聚合的一部分。在這種情況下，RAID 群組大小應與現有 RAID 群組大小相符。

ONTAP Select儲存效率支持

ONTAP Select提供的儲存效率選項與FAS和AFF陣列上的儲存效率選項類似。

使用全快閃 VSAN 或通用快閃記憶體陣列的ONTAP Select虛擬 NAS (vNAS) 部署應遵循具有非 SSD 直連式儲存 (DAS) 的ONTAP Select的最佳實務。

只要您擁有帶有 SSD 驅動器和高級許可證的 DAS 存儲，新安裝中就會自動啟用類似AFF的特性。

類似AFF的特性，以下內嵌 SE 功能會在安裝過程中自動啟用：

- 內聯零模式偵測
- 卷內嵌重複資料刪除
- 卷後台重複資料刪除
- 自適應內聯壓縮
- 內聯資料壓縮
- 聚合內嵌重複資料刪除
- 聚合後台重複資料刪除

若要驗證ONTAP Select是否已啟用所有預設儲存效率策略，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

```



要從 9.6 及更高版本升級 ONTAP Select，您必須使用高級許可證在 DAS SSD 儲存上安裝 ONTAP Select。此外，在使用 ONTAP Deploy 進行初始叢集安裝時，您必須選取「啟用儲存效率」複選框。如果先前的條件尚未滿足，則在 ONTAP 升級後啟用類似 AFF 的個人化功能需要手動建立啟動參數並重新啟動節點。請聯絡技術支援以了解更多詳細資訊。

ONTAP Select 儲存效率配置

下表總結了可用的各種儲存效率選項，這些選項預設為啟用，或預設未啟用但建議啟用，具體取決於媒體類型和軟體許可證。

ONTAP Select 功能	DAS SSD (高級版或高級 XL ¹)	DAS HDD (所有許可證)	vNAS (所有授權)
線上零點偵測	是 (預設)	是 由使用者按卷啟用	是 由使用者按卷啟用
卷內嵌重複資料刪除	是 (預設)	無法使用	不支援
32K 內聯壓縮 (二次壓縮)	是 由使用者根據每個卷啟用。	是 由使用者按卷啟用	不支援
8K 內聯壓縮 (自適應壓縮)	是 (預設)	是 由使用者根據磁碟區啟用	不支援
後台壓縮	不支援	是 由使用者根據磁碟區啟用	是 由使用者按卷啟用
壓縮掃描儀	是的	是的	是 由使用者按卷啟用
內聯資料壓縮	是 (預設)	是 由使用者根據磁碟區啟用	不支援
壓實掃描儀	是的	是的	不支援
聚合內嵌重複資料刪除	是 (預設)	不適用	不支援

ONTAP Select功能	DAS SSD (高級版或高級XL ¹)	DAS HDD (所有許可證)	vNAS (所有授權)
卷後台重複資料刪除	是 (預設)	是 由使用者根據磁碟區啟用	是 由使用者按卷啟用
聚合後台重複資料刪除	是 (預設)	不適用	不支援

[small]#¹ ONTAP Select 9.6 支援新的許可證 (Premium XL) 和新的虛擬機器大小 (Large)。但是，大型虛擬機器僅支援使用軟體 RAID 的 DAS 配置。9.6版本中的大型ONTAP Select虛擬機器不支援硬體 RAID 和 vNAS 配置。#

DAS SSD 設定升級行為注意事項

升級到ONTAP Select 9.6 或更高版本後，等待 `system node upgrade-revert show` 指令指示升級已完成，然後再驗證現有磁碟區的儲存效率值。

在升級到ONTAP Select 9.6 或更高版本的系統上，在現有聚合或新建立的聚合上建立的新磁碟區的行為與在全新部署上建立的磁碟區的行為相同。進行ONTAP Select程式碼升級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策略，但有一些差異：

場景 1

如果升級之前磁碟區上未啟用任何儲存效率策略，則：

- 卷與 `space guarantee = volume` 未啟用內聯資料壓縮、聚合內聯重複資料刪除和聚合後台重複資料刪除。這些選項可在升級後啟用。
- 卷與 `space guarantee = none` 未啟用後台壓縮。此選項可在升級後啟用。
- 升級後，現有磁碟區上的儲存效率策略設定為自動。

場景 2

如果在升級之前磁碟區上已經啟用了某些儲存效率，則：

- 卷與 `space guarantee = volume` 升級後沒有看到任何差異。
- 卷與 `space guarantee = none` 已啟用聚合後台重複資料刪除。
- 卷與 `storage policy inline-only` 將其策略設為自動。
- 具有使用者定義儲存效率策略的磁碟區在策略上沒有變化，但具有 `space guarantee = none`。這些磁碟區已啟用聚合後台重複資料刪除。

聯網

ONTAP Select網路概念與特性

首先熟悉適用於ONTAP Select環境的常規網路概念。然後，探索單節點和多節點叢集的具體特性和選項。

實體網路

實體網路主要透過提供底層第二層交換基礎架構來支援ONTAP Select叢集部署。與實體網路相關的配置包括虛擬機器管理程式主機和更廣泛的交換網路環境。

主機 NIC 選項

每個ONTAP Select虛擬機器管理程式主機必須配置兩個或四個實體連接埠。具體配置取決於以下幾個因素：

- 叢集是否包含一個或多個ONTAP Select主機
- 使用哪種虛擬機器管理程式作業系統
- 如何配置虛擬交換機
- 連結是否使用 LACP

實體交換器配置

您必須確保實體交換器的配置支援ONTAP Select部署。實體交換器與基於虛擬機器管理程式的虛擬交換器整合在一起。您選擇的具體配置取決於多種因素。主要考慮因素包括：

- 您將如何保持內部和外部網路之間的分離？
- 您會保持資料和管理網路之間的分離嗎？
- 二層 VLAN 將如何配置？

邏輯網路

ONTAP Select使用兩個不同的邏輯網路，根據類型分離流量。具體來說，流量可以在叢集內的主機之間流動，也可以流向叢集外的儲存用戶端和其他電腦。虛擬機器管理程式管理的虛擬交換器有助於支援邏輯網路。

內部網路

在多節點叢集部署中，各個ONTAP Select節點使用獨立的「內部」網路進行通訊。此網路不會暴露給ONTAP Select叢集中的節點，也不會在節點外部使用。



內部網路僅存在於多節點叢集中。

內部網路具有以下特點：

- 用於處理ONTAP叢集內流量，包括：
 - 簇
 - 高可用性互連 (HA-IC)
 - RAID 同步鏡像 (RSM)
- 基於VLAN的單二層網路
- 靜態 IP 位址由ONTAP Select分配：
 - 僅限 IPv4
 - 未使用 DHCP
 - 鏈路本地地址
- MTU 大小預設為 9000 字節，可在 7500-9000 範圍內調整（含）

外部網路

外部網路處理ONTAP Select叢集節點與外部儲存客戶端以及其他電腦之間的流量。外部網路是每個叢集部署的一部分，具有以下特點：

- 用於處理ONTAP流量，包括：
 - 資料（NFS、CIFS、iSCSI）
 - 管理（叢集和節點；可選 SVM）
 - 集群間（可選）
- 選用支援 VLAN：
 - 資料連接埠組
 - 管理連接埠組
- 根據管理員的配置選擇指派的 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集群。

虛擬機器網路環境

虛擬機器管理程式主機提供了多種網路功能。

ONTAP Select依賴透過虛擬機器公開的以下功能：

虛擬機器連接埠

ONTAP Select有多個連接埠可供使用。這些連接埠的分配和使用取決於多種因素，包括群集的大小。

虛擬交換機

虛擬機器管理程式環境中的虛擬交換器軟體（無論是 vSwitch (VMware) 或 Open vSwitch (KVM)）都會將虛擬機器公開的連接埠與實體乙太網路 NIC 連接埠連接起來。您必須根據環境為每個ONTAP Select主機配置一個 vSwitch。

ONTAP Select單節點與多節點網路配置

ONTAP Select支援單節點和多節點網路配置。

單節點網路配置

單節點ONTAP Select配置不需要ONTAP內部網路，因為沒有叢集、HA 或鏡像流量。

與ONTAP Select產品的多節點版本不同，每個ONTAP Select VM 包含三個虛擬網路適配器，分別提供給ONTAP網路連接埠 e0a、e0b 和 e0c。

這些連接埠用於提供以下服務：管理、資料和叢集間 LIF。

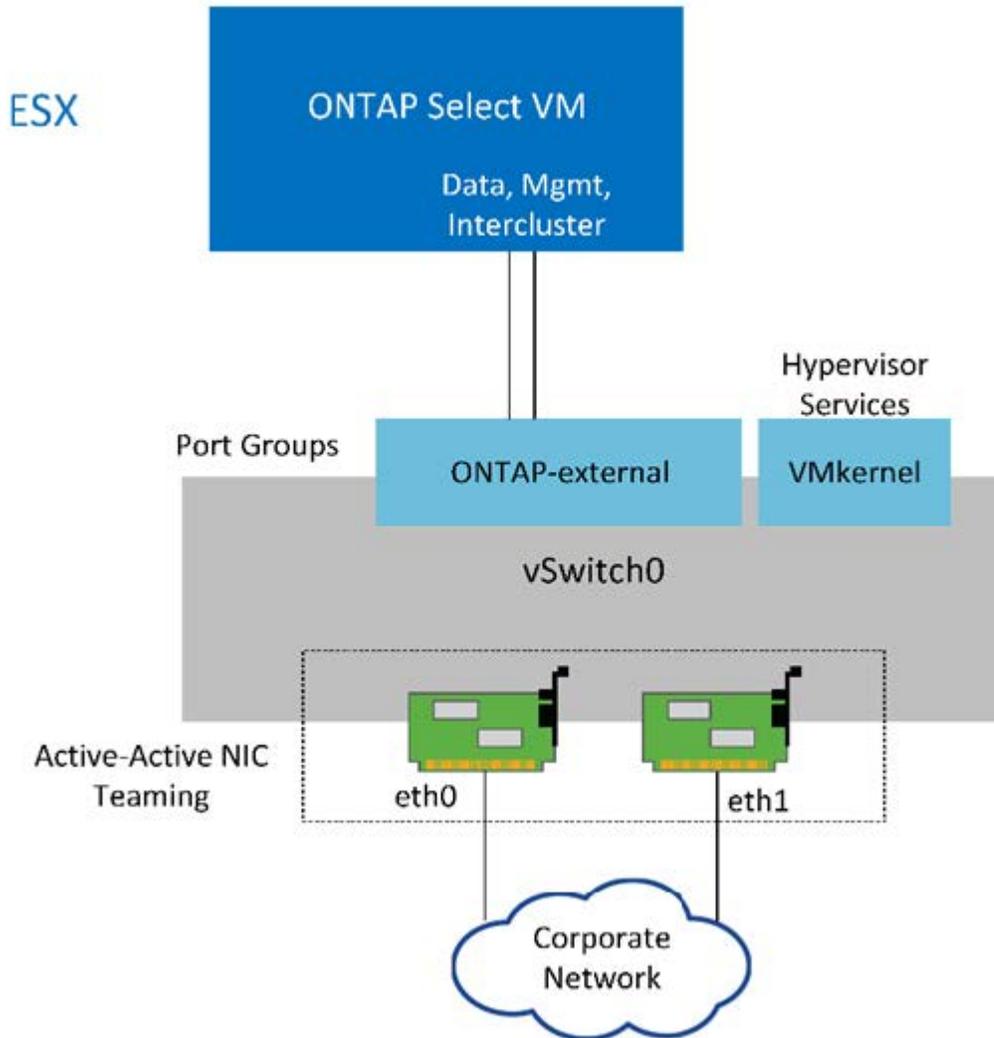
虛擬機

ONTAP Select可以部署為單節點叢集。虛擬機管理程式主機包含一個虛擬交換機，用於提供對外部網路的存取。

ESXi

下圖顯示了這些連接埠與底層實體適配器之間的關係，該圖描繪了 ESX 虛擬機器管理程式上的一個ONTAP Select叢集節點。

單節點ONTAP Select叢集的網路配置



儘管兩個適配器對於單節點叢集來說已經足夠，但仍需要 NIC 組合。

LIF 分配

如本文檔的多節點 LIF 分配部分所述，ONTAP Select使用 IP 空間將叢集網路流量與資料流量和管理流量分開。此平台的單節點版本不包含叢集網路。因此，在叢集 IP 空間中不存在任何連接埠。



叢集和節點管理 LIF 會在ONTAP Select叢集設定期間自動建立。其餘 LIF 可在部署後建立。

管理和數據 LIF (e0a、e0b 和 e0c)

ONTAP連接埠 e0a、e0b 和 e0c 被委派為承載以下類型流量的 LIF 的候選連接埠：

- SAN/NAS 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)

多節點網路配置

多節點ONTAP Select網路配置由兩個網路組成。

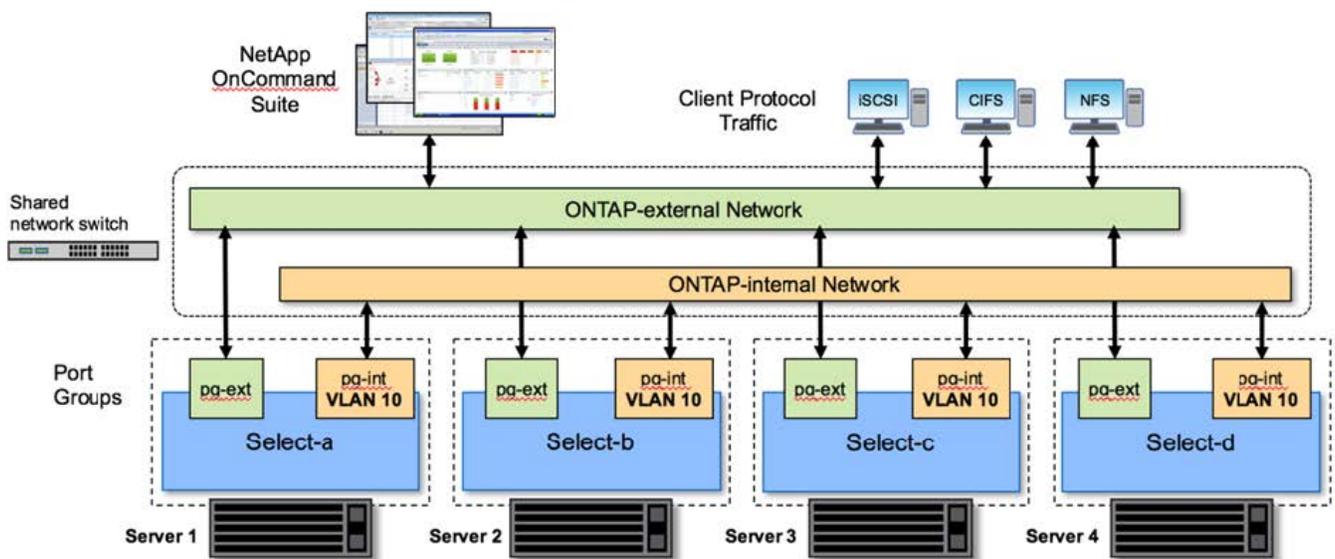
這兩個網路分別是：內部網路（負責提供叢集和內部複製服務）和外部網路（負責提供資料存取和管理服務）。對這兩個網路內部的流量進行端到端隔離，對於建立適合叢集彈性的環境至關重要。

下圖顯示了這些網路，該圖展示了在 VMware vSphere 平台上運行的四節點ONTAP Select叢集。六節點叢集和八節點叢集的網路佈局類似。



每個ONTAP Select實例都駐留在單獨的實體伺服器上。內部和外部流量使用單獨的網路連接埠群組進行隔離，這些連接埠群組分配給每個虛擬網路接口，並允許叢集節點共享相同的實體交換器基礎架構。

ONTAP Select多節點叢集網路配置概述



每個ONTAP Select虛擬機包含七個虛擬網路適配器，這些適配器以一組七個網路連接埠 (e0a 至 e0g) 的形式呈現給ONTAP。儘管ONTAP將這些適配器視為實體 NIC，但它們實際上是虛擬的，並透過虛擬化網路層映射到一組實體介面。因此，每個託管伺服器不需要六個實體網路連接埠。



不支援向ONTAP Select VM 新增虛擬網路適配器。

這些連接埠已預先配置以提供以下服務：

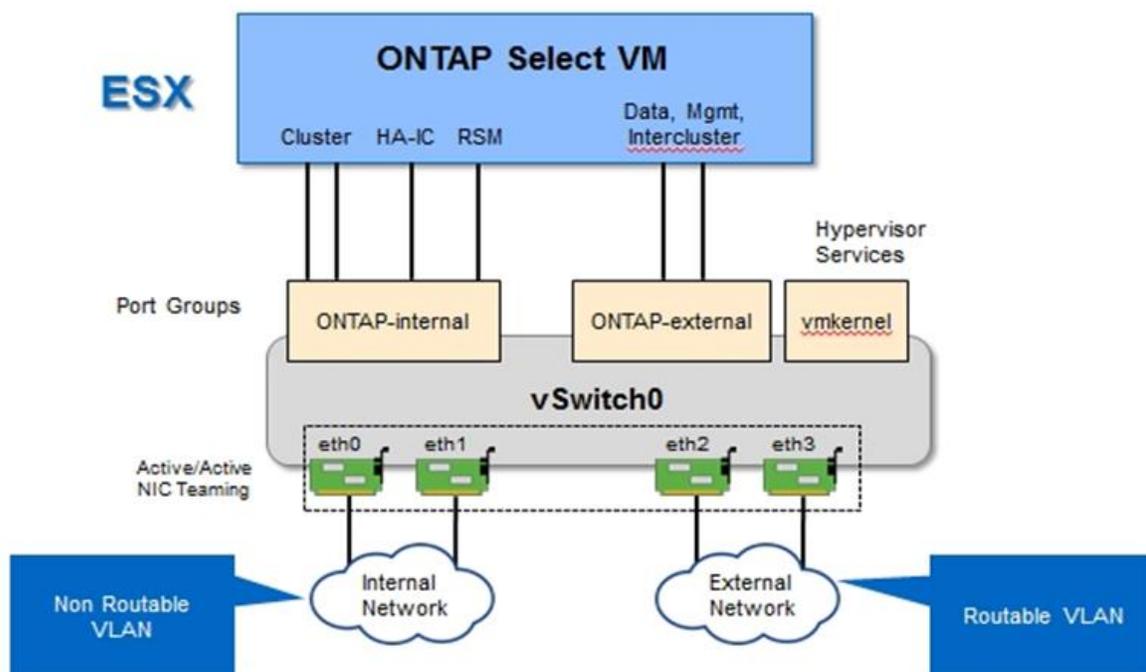
- e0a、e0b 和 e0g。管理和數據 LIF

- e0c、e0d。集群網路 LIF
- e0e.RSMRWS
- e0f.HA 互連

連接埠 e0a、e0b 和 e0g 位於外部網路上。雖然連接埠 e0c 到 e0f 執行多種不同功能，但它們共同構成了內部 Select 網路。在進行網路設計時，應將這些連接埠放置在單一二層網路上。無需將這些虛擬適配器分散在不同的網路之間。

下圖說明了這些連接埠與底層實體適配器之間的關係，該圖描繪了 ESX 虛擬機器管理程式上的一個ONTAP Select叢集節點。

多節點ONTAP Select叢集中單一節點的網路配置



跨不同實體網路卡隔離內部和外部流量，可防止網路資源存取不足而導致系統延遲。此外，透過網路卡綁定進行聚合，可確保單一網路介面卡發生故障時，ONTAP Select叢集節點仍可存取對應的網路。

請注意，外部網路和內部網路連接埠群組均以對稱方式包含所有四個網路卡適配器。外部網路連接埠群組中的活動連接埠是內部網路中的備用連接埠。相反，內部網路連接埠群組中的活動連接埠是外部網路連接埠群組中的備用連接埠。

LIF 分配

隨著 IP 空間的引入，ONTAP 連接埠角色已被棄用。與 FAS 陣列一樣，ONTAP Select 叢集包含預設 IP 空間和叢集 IP 空間。透過將網頁連接埠 e0a、e0b 和 e0g 放入預設 IP 空間，並將連接埠 e0c 和 e0d 放入叢集 IP 空間，這些連接埠實際上被隔離，無法託管不屬於該 IP 空間的 LIF。ONTAP Select 叢集中的其餘連接埠透過自動指派提供內部服務的介面來使用。它們不會像 RSM 和 HA 互連接口那樣透過 ONTAP Shell 公開。



並非所有 LIF 都可透過 ONTAP 指令 Shell 看到。HA 互連接口和 RSM 介面在 ONTAP 中處於隱藏狀態，僅供內部使用，以提供各自的服務。

以下章節將詳細介紹網路連接埠和 LIF。

管理和數據 LIF (e0a、e0b 和 e0g)

ONTAP 連接埠 e0a、e0b 和 e0g 被委派為承載以下類型流量的 LIF 的候選連接埠：

- SAN/NAS 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)



叢集和節點管理 LIF 會在 ONTAP Select 叢集設定期間自動建立。其餘 LIF 可在部署後建立。

集群網路 LIF (e0c、e0d)

ONTAP 連接埠 e0c 和 e0d 被指定為叢集介面的主連接埠。在每個 ONTAP Select 叢集節點中，ONTAP 設定期間會使用連結本機 IP 位址 (169.254.xx) 自動產生兩個叢集介面。



這些介面不能指派靜態 IP 位址，且不應建立額外的叢集介面。

叢集網路流量必須流經低延遲、非路由的第 2 層網路。由於叢集吞吐量和延遲要求，ONTAP Select 叢集應位於實體位置較近的位置 (例如，多機櫃、單一資料中心)。不支援透過跨 WAN 或跨較遠地理距離分離 HA 節點來建立四節點、六節點或八節點延伸叢集配置。支援使用中介器的延伸雙節點配置。

有關詳細信息，請參閱"[雙節點擴展 HA \(MetroCluster SDS\) 最佳實踐](#)"。



為確保叢集網路流量的最大吞吐量，此網路連接埠配置為使用巨型幀 (7500 到 9000 MTU)。為確保叢集正常運行，請驗證所有為 ONTAP Select 叢集節點提供內部網路服務的上游虛擬交換器和實體交換器上是否都啟用了巨型訊框。

RAID SyncMirror 流量 (e0e)

使用位於網路連接埠 e0e 上的內部網路接口，跨 HA 合作夥伴節點同步複製區塊。此功能使用 ONTAP 在叢集設定期間配置的網路介面自動運行，無需管理員進行任何配置。



ONTAP 保留埠 e0e 用於內部複製流量。因此，此連接埠和託管 LIF 在 ONTAP 命令列介面 (CLI) 或系統管理員 (System Manager) 中均不可見。此介面配置為使用自動產生的連結本機 IP 位址，不支援重新指派備用 IP 位址。此網路連接埠需要使用巨型幀 (7500 到 9000 MTU)。

HA 互連 (e0f)

NetApp FAS 陣列使用專用硬體在 ONTAP 叢集中的 HA 對之間傳遞訊息。然而，軟體定義環境往往缺乏此類設備 (例如 InfiniBand 或 iWARP 設備)，因此需要一種替代解決方案。儘管考慮了多種可能性，但 ONTAP 對互連傳輸的要求要求在軟體中模擬此功能。因此，在 ONTAP Select 叢集中，HA 互連的功能 (傳統上由硬體提供) 已設計到作業系統中，並使用乙太網路作為傳輸機制。

每個 ONTAP Select 節點都配置有一個 HA 互連埠 e0f。此連接埠託管 HA 互連網路接口，該接口負責兩個主要功能：

- 在 HA 對之間鏡像 NVRAM 的內容

- 在 HA 對之間傳送/接收 HA 狀態訊息和網路心跳訊息

HA 互連流量透過在乙太網路封包內分層遠端直接記憶體存取 (RDMA) 幀，使用單一網路介面流經此網路連接埠。



與 RSM 連接埠 (e0e) 類似，使用者無法從 ONTAP CLI 或 System Manager 中看到實體連接埠和託管網路介面。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幀 (MTU 為 7500 到 9000)。

ONTAP Select 內部與外部網路

ONTAP Select 內部和外部網路的特性。

ONTAP Select 內部網路

內部 ONTAP Select 網路僅存在於產品的多節點版本中，負責為 ONTAP Select 叢集提供叢集通訊、HA 互連和同步複製服務。此網路包含以下連接埠和介面：

- *e0c、e0d。*託管叢集網路 LIF
- *e0e.*託管 RSM LIF
- *e0f.*託管 HA 互連 LIF

此網路的吞吐量和延遲對於確定 ONTAP Select 叢集的效能和彈性至關重要。為了確保叢集安全並確保系統介面與其他網路流量隔離，需要進行網路隔離。因此，此網路必須由 ONTAP Select 叢集專用。



不支援將 Select 內部網路用於 Select 叢集流量以外的其他流量（例如應用程式流量或管理流量）。ONTAP ONTAPVLAN 上不能有其他虛擬機器或主機。

穿越內部網路的網路封包必須位於帶有 VLAN 標記的專用第 2 層網路上。這可以透過完成以下任務之一來實現：

- 將 VLAN 標記的連接埠群組指派給內部虛擬 NIC (e0c 到 e0f) (VST 模式)
- 使用上游交換器提供的本機 VLAN，其中本機 VLAN 不用於任何其他流量（指派沒有 VLAN ID 的連接埠群組，即 EST 模式）

在所有情況下，內部網路流量的 VLAN 標記都是在 ONTAP Select VM 之外完成的。



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

在 ONTAP Select 叢集中，內部流量和外部流量使用稱為連接埠群組的虛擬二層網路物件進行隔離。正確分配這些連接埠群組的 vSwitch 至關重要，尤其是對於負責提供叢集、高可用性互連和鏡像複製服務的內部網路而言。這些網路連接埠的網路頻寬不足會導致效能下降，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要求內部 ONTAP Select 網路使用 10Gb 連線；不支援 1Gb 網路卡。但是，可以對外部網路進行權衡，因為限制流入 ONTAP Select 叢集的資料流不會影響其可靠運作的能力。

雙節點叢集可以使用四個 1Gb 端口進行內部流量傳輸，也可以使用 10Gb 端口，而無需像四節點叢集那樣使用兩個 10Gb 端口。如果伺服器環境不適合安裝四個 10Gb NIC 卡，則可以將兩個 10Gb NIC 卡用於內部網路，將兩個 1Gb NIC 卡用於外部 ONTAP 網路。

可以使用網路連接檢查器功能來驗證多節點叢集中的內部網路。此功能可以從執行以下命令的 Deploy CLI 呼叫：``network connectivity-check start`` 命令。

執行以下命令以查看測試的輸出：

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

此工具僅適用於對多節點 Select 叢集中的內部網路進行故障排除。不應用於排除單節點叢集（包括 vNAS 配置）、ONTAP Deploy 到ONTAP Select 的連線問題或客戶端連線問題。

叢集建立精靈（ONTAP Deploy GUI 的一部分）包含內部網路檢查器，作為建立多節點叢集期間可用的選用步驟。鑑於內部網路在多節點叢集中發揮的重要作用，將此步驟作為叢集建立工作流程的一部分可以提高叢集創建操作的成功率。

從ONTAP Deploy 2.10 開始，內部網路使用的 MTU 大小可以設定在 7,500 到 9,000 之間。網路連線檢查器也可用於測試 7,500 到 9,000 之間的 MTU 大小。預設 MTU 值設定為虛擬網路交換器的值。如果環境中存在 VXLAN 等網路覆蓋，則必須將該預設值替換為較小的值。

ONTAP Select外部網路

ONTAP Select外部網路負責叢集的所有出站通信，因此在單節點和多節點配置中均存在。雖然此網路不像內部網路那樣具有嚴格定義的吞吐量要求，但管理員應注意不要在客戶端和ONTAP虛擬機器之間造成網路瓶頸，因為效能問題可能會被誤認為是ONTAP Select問題。



與內部流量類似，外部流量可以在 vSwitch 層 (VST) 和外部交換器層 (EST) 進行標記。此外，外部流量也可以由ONTAP Select虛擬機本身透過稱為 VGT 的過程進行標記。請參閱["數據和管理流量分離"](#)了解更多詳情。

下表重點介紹了ONTAP Select內部和外部網路之間的主要差異。

內部網路與外部網路快速參考

描述	內部網路	外部網路
網路服務	群集 HA/IC RAID SyncMirror (RSM)	叢集間資料管理 (SnapMirror 和 SnapVault)
網路隔離	必需的	選修的
幀大小 (MTU)	7,500 至 9,000	1,500 (預設) 9,000 (支援)
IP位址分配	自動生成	使用者定義
DHCP 支援	不	不

NIC 組合

為了確保內部和外部網路均具備提供高效能和容錯能力所需的頻寬和彈性特性，建議使用實體網路適配器綁定。支援具有單一 10Gb 連結的雙節點叢集配置。但是，NetApp建議的最佳實務是在ONTAP Select叢集的內部和外部網路上都使用 NIC 綁定。

MAC 位址生成

指派給所有ONTAP Select網路連接埠的 MAC 位址均由隨附的部署公用程式自動產生。此實用程式使用NetApp特有的平台特定組織唯一識別碼 (OUI)，以確保不與FAS系統衝突。之後，此位址的副本將儲存在ONTAP Select安裝虛擬機器 (ONTAP Deploy) 的內部資料庫中，以防止在未來的節點部署過程中意外重新分配。管理員在任何情況下都不應修改已指派的網路連接埠 MAC 位址。

支援的ONTAP Select網路配置

選擇最佳硬體並配置您的網路以優化效能和彈性。

伺服器供應商深知客戶的需求各不相同，選擇至關重要。因此，在購買實體伺服器時，網路連線決策中存在著許多選擇。大多數商用系統都隨附各種網路卡 (NIC)，提供單一連接埠和多連接埠選項，速度和吞吐量各不相同。其中包括支援 VMware ESX 的 25Gb/s 和 40Gb/s 網路卡適配器。

由於ONTAP Select虛擬機器的效能與底層硬體的特性直接相關，因此透過選擇更高速的 NIC 來提高虛擬機器的吞吐量，可以提高叢集效能並改善整體使用者體驗。可以使用四個 10 Gb NIC 或兩個高速 NIC (25/40 Gb/s) 來實現高效能網路佈局。此外，還支援許多其他配置。對於雙節點集群，支援 4 個 1 Gb 連接埠或 1 個 10 Gb 連接埠。對於單節點集群，支援 2 個 1 Gb 連接埠。

網路最低配置和推薦配置

根據叢集大小，有幾種支援的乙太網路配置。

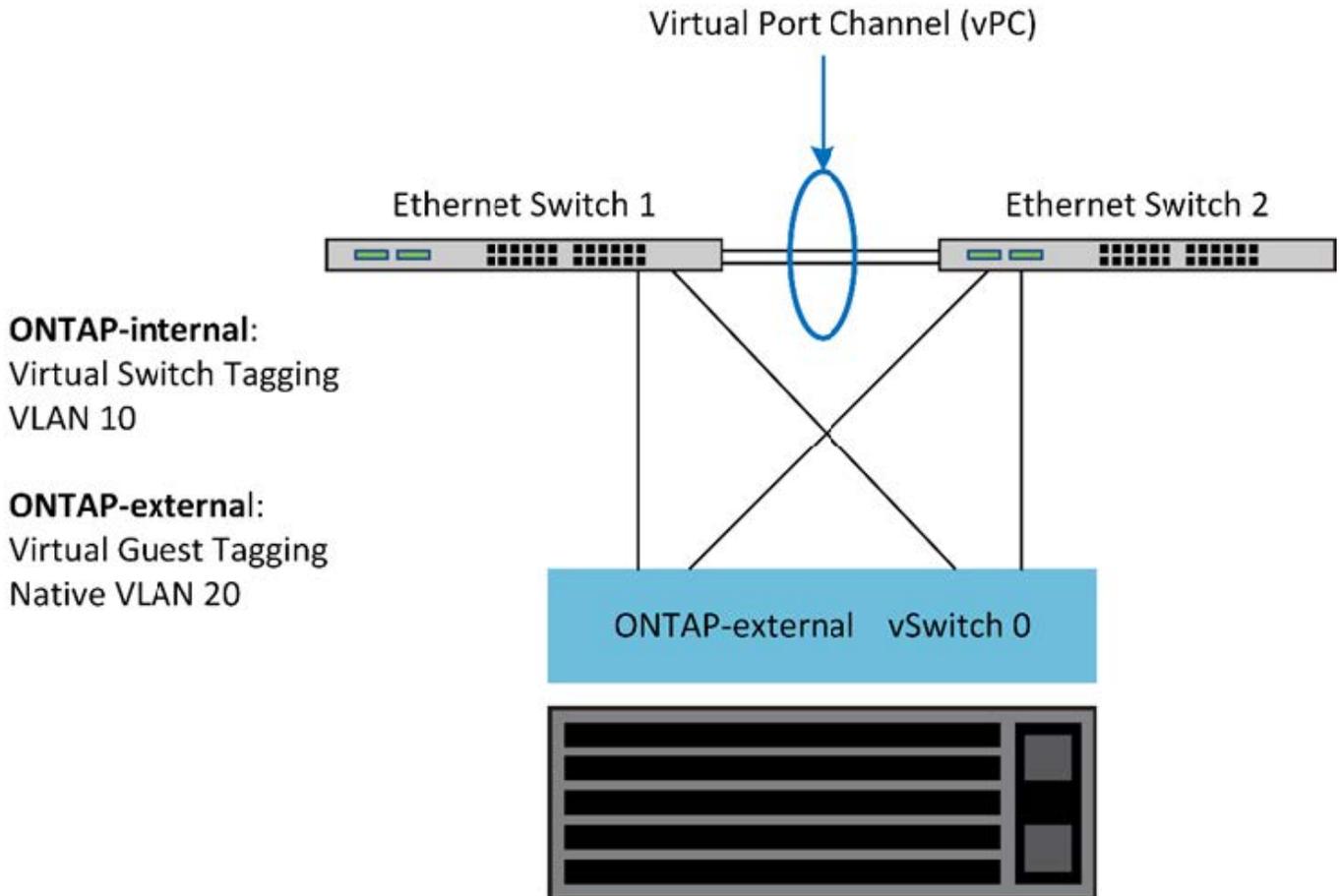
簇大小	最低要求	推薦
單節點集群	2 個 1GbE	2 個 10GbE
雙節點集群或MetroCluster SDS	4 個 1GbE 或 1 個 10GbE	2 個 10GbE
4/6/8節點集群	2 個 10GbE	4 個 10GbE 或 2 個 25/40GbE



不支援在正在運行的群集上進行單鏈路和多鏈路拓撲之間的轉換，因為可能需要在每個拓撲所需的不同的 NIC 組合配置之間進行轉換。

使用多個實體交換器的網路配置

當有足夠的硬體可用時，NetApp建議使用下圖所示的多交換器配置，因為它可以增加對實體交換器故障的保護。



ONTAP Select VMware vSphere vSwitch 配置

ONTAP Select vSwitch 配置和雙 NIC 和四 NIC 配置的負載平衡策略。

ONTAP Select 支援使用標準和分散式 vSwitch 配置。分散式 vSwitch 支援鏈路聚合結構 (LACP)。鏈路聚合是一種常見的網路結構，用於聚合跨多個實體适配器的頻寬。LACP 是一種與供應商無關的標準，它為網路端點提供了一種開放協議，可將多組實體網路連接埠捆綁到單一邏輯通道中。ONTAP Select 可以與配置為鏈路聚合組 (LAG) 的連接埠組配合使用。但是，NetApp 建議將各個實體連接埠用作簡單的上行鏈路 (中繼) 端口，以避免 LAG 配置。在這些情況下，標準和分散式 vSwitch 的最佳實踐是相同的。

本節介紹雙 NIC 和四 NIC 配置中應使用的 vSwitch 配置和負載平衡策略。

配置 ONTAP Select 所使用的連接埠群組時，應遵循以下最佳實務；連接埠群組層級的負載平衡策略是基於來源虛擬連接埠 ID 的路由。VMware 建議在連接到 ESXi 主機交換器連接埠上將 STP 設定為 Portfast。

所有 vSwitch 配置都需要至少兩個實體網路介面卡捆綁到一個 NIC 群組中。ONTAP Select 支援雙節點叢集使用單一 10Gb 連結。但是，NetApp 的最佳實踐是透過 NIC 聚合來確保硬體冗餘。

在 vSphere 伺服器上，NIC 群組是一種聚合結構，用於將多個實體網路适配器捆綁到單一邏輯通道中，從而允許所有成員連接埠分擔網路負載。需要注意的是，NIC 群組可以在沒有實體交換器支援的情況下建立。負載平衡和故障轉移策略可以直接套用於 NIC 群組，而 NIC 組則無需感知上游交換器的配置。在這種情況下，策略僅適用於出站流量。



ONTAP Select 不支援靜態連接埠通道。分散式 vSwitch 支援啟用 LACP 的通道，但使用 LACP LAG 可能會導致 LAG 成員之間的負載分配不均。

對於單節點集群，ONTAP Deploy 會將ONTAP Select虛擬機配置為使用一個連接埠組作為外部網路，並使用相同連接埠組（或可選地使用不同的連接埠組）作為集群和節點管理流量。對於單節點集群，可以將所需數量的實體連接埠作為活動適配器新增至外部連接埠組。

對於多節點集群，ONTAP Deploy 會將每個ONTAP Select虛擬機配置為使用一個或兩個連接埠組用於內部網路，並單獨使用一個或兩個連接埠組用於外部網路。叢集和節點管理流量可以與外部流量使用同一個連接埠群組，也可以選擇使用單獨的連接埠群組。叢集和節點管理流量不能與內部流量共用同一個連接埠群組。

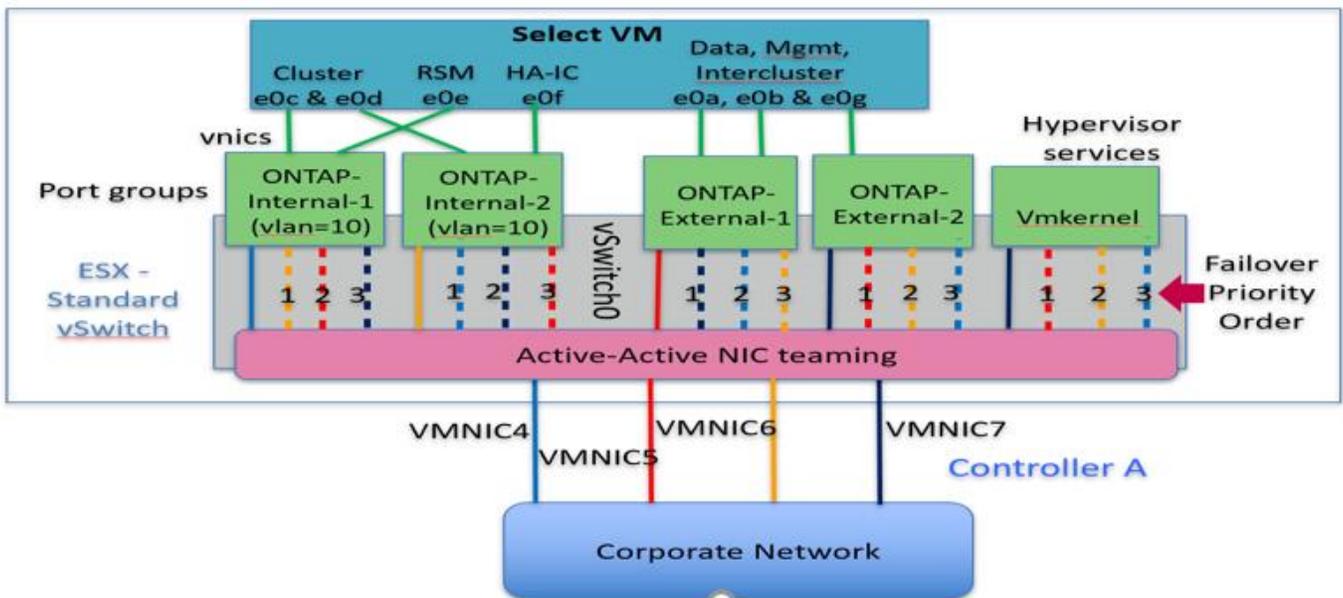


ONTAP Select最多支援四個 VMNIC。

標準或分佈式 vSwitch 和每個節點四個實體端口

多節點叢集中的每個節點可以分配四個連接埠群組。每個端口組都有一個活動實體端口和三個備用實體端口，如下圖所示。

每個節點具有四個實體連接埠的 vSwitch



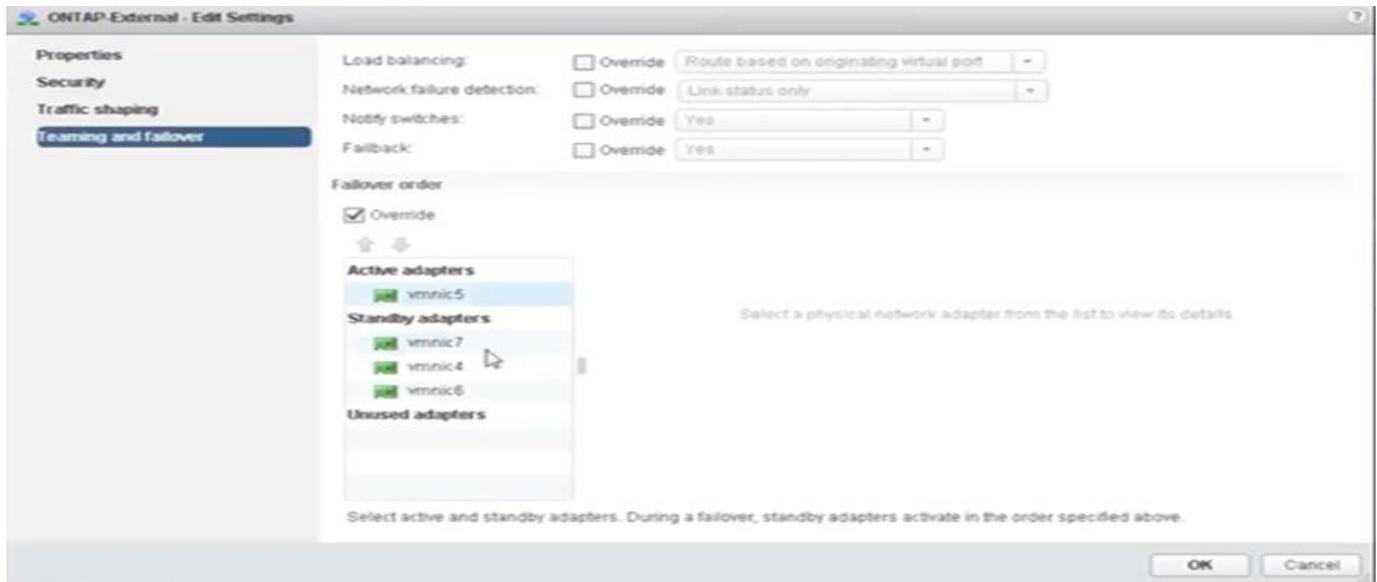
備用清單中連接埠的順序很重要。下表提供了四個連接埠組中實體連接埠分佈的範例。

網路最低配置和建議配置

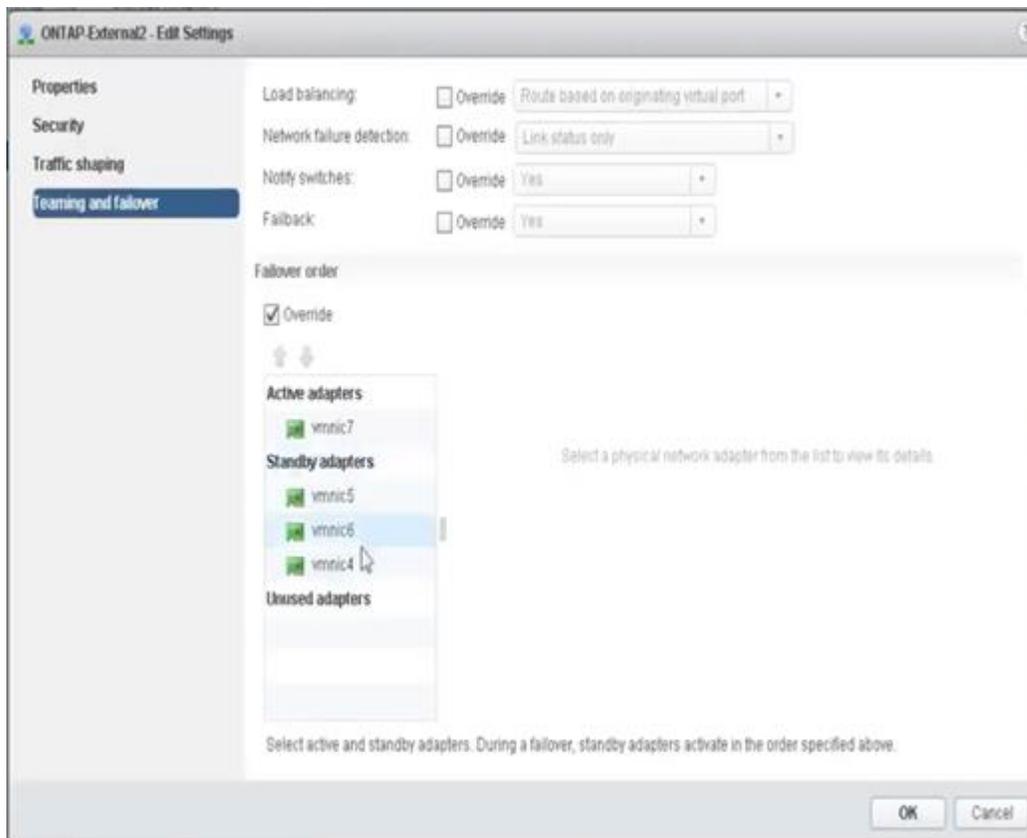
港口集團	外部 1	外部 2	內部 1	內部 2
積極的	vmnic0	vmnic1	vmnic2	vmnic3
待機1	vmnic1	vmnic0	vmnic3	vmnic2
待機2	vmnic2	vmnic3	vmnic0	vmnic1
待機 3	vmnic3	vmnic2	vmnic1	vmnic0

下圖顯示了 vCenter GUI 中外部網路連接埠群組的配置（ONTAP-External 和ONTAP-External2）。請注意，活動適配器來自不同的網路卡。在此設定中，vmnic 4 和 vmnic 5 是同一實體網路卡上的雙端口，而 vmnic 6 和 vmnic 7 同樣是單獨網路卡上的雙端口（本例中未使用 vmnics 0 至 3）。備用適配器的順序提供了分層故障轉移，其中來自內部網路的連接埠位於最後。備用清單中的內部連接埠順序在兩個外部連接埠組之間進行類似交換。

第 1 部分：ONTAP Select外部連接埠群組配置



第 2 部分：ONTAP Select外部連接埠群組配置

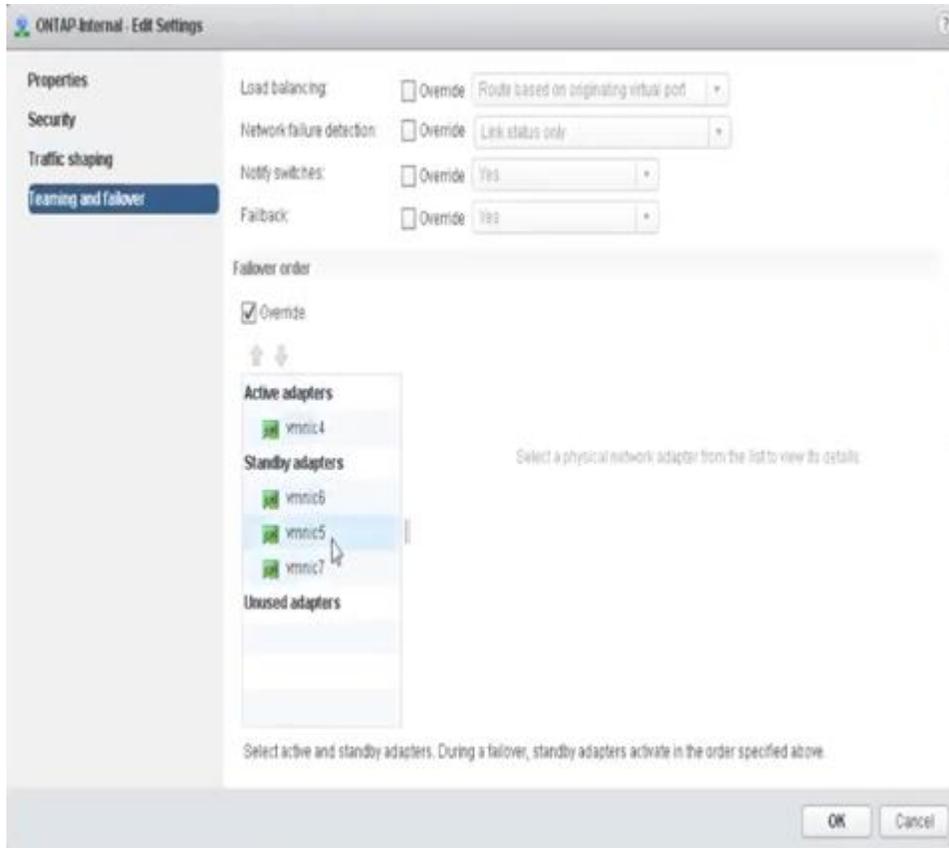


為了方便閱讀，任務分配如下：

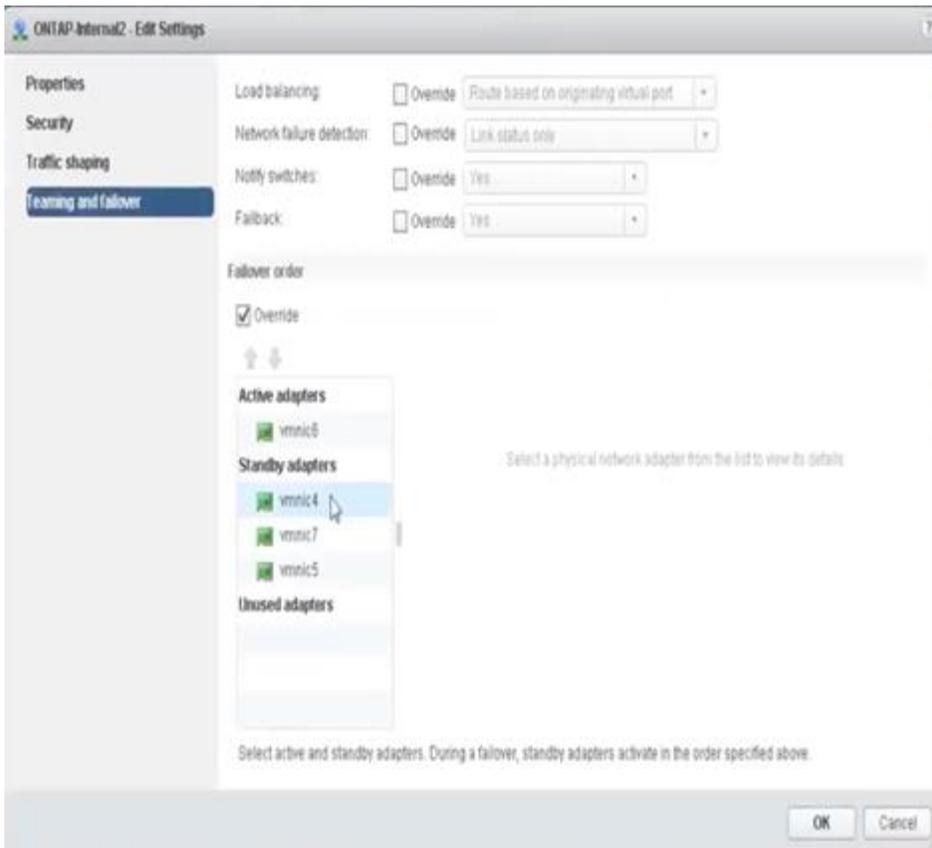
ONTAP- 外部	ONTAP外部 2
活動適配器：vmnic5 備用適配器：vmnic7、vmnic4、vmnic6	活動適配器：vmnic7 備用適配器：vmnic5、vmnic6、vmnic4

下圖顯示了內部網路連接埠群組（ONTAP-Internal 和ONTAP-Internal2）的配置。請注意，活動適配器來自不同的網路卡。在此設定中，vmnic 4 和 vmnic 5 是同一實體 ASIC 上的雙端口，而 vmnic 6 和 vmnic 7 同樣是單獨 ASIC 上的雙端口。備用適配器的順序提供了分層故障轉移，來自外部網路的連接埠位於最後。備用清單中外部連接埠的順序在兩個內部連接埠組之間進行類似交換。

第 1 部分：ONTAP Select內部連接埠群組配置



第 2 部分：ONTAP Select內部連接埠群組



為了方便閱讀，任務分配如下：

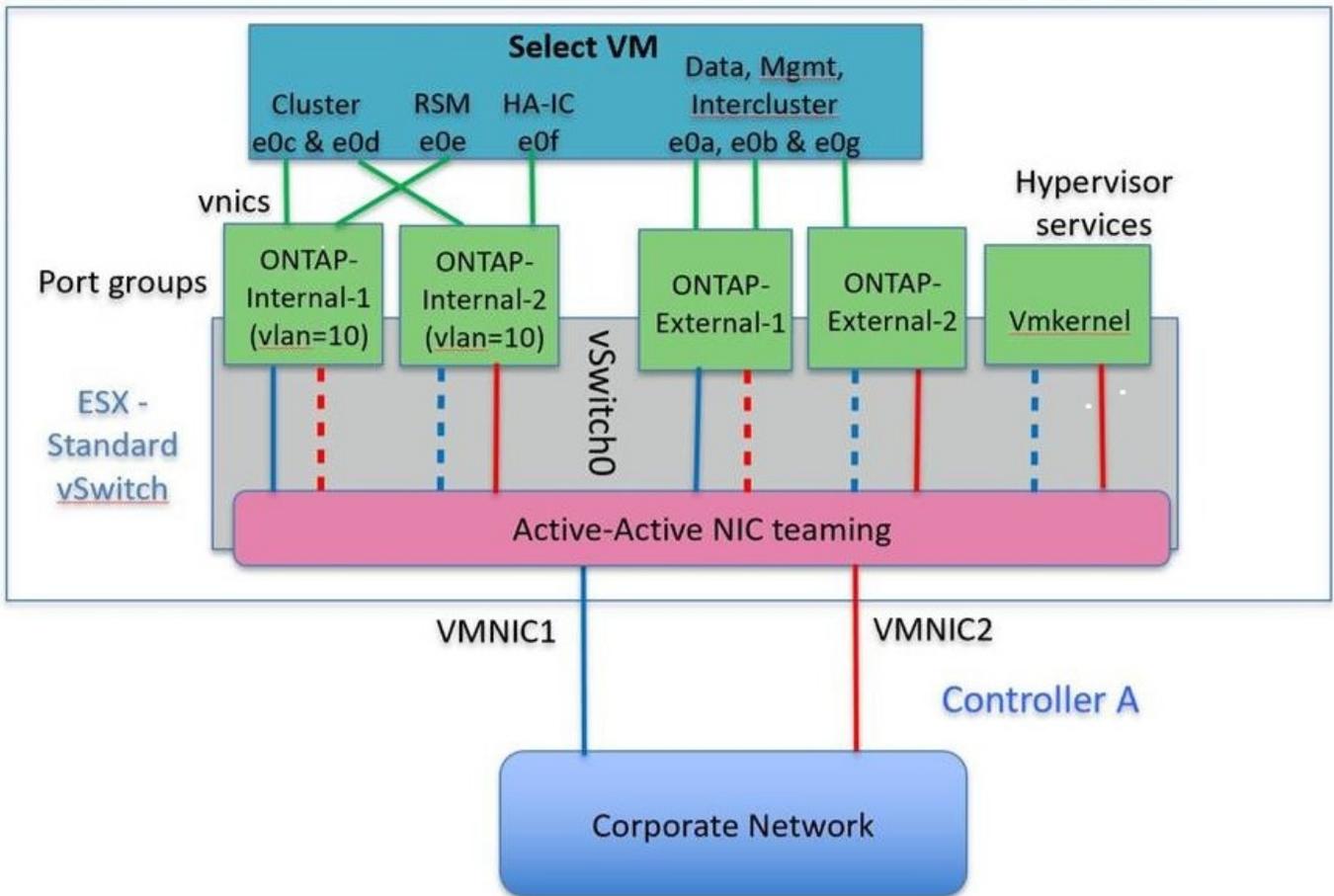
ONTAP- 內部	ONTAP內部 2
活動適配器：vmnic4 備用適配器：vmnic6、vmnic5、vmnic7	活動適配器：vmnic6 備用適配器：vmnic4、vmnic7、vmnic5

標準或分佈式 **vSwitch** 和每個節點兩個實體端口

使用兩個高速 (25/40 Gb) 網路卡時，建議的連接埠組配置在概念上與使用四個 10 Gb 適配器的配置非常相似。即使僅使用兩個實體適配器，也應使用四個連接埠組。連接埠群組分配如下：

港口集團	外部 1 (e0a、e0b)	內部 1 (e0c、e0e)	內部 2 (e0d、e0f)	外部 2 (e0g)
積極的	vmnic0	vmnic0	vmnic1	vmnic1
支援	vmnic1	vmnic1	vmnic0	vmnic0

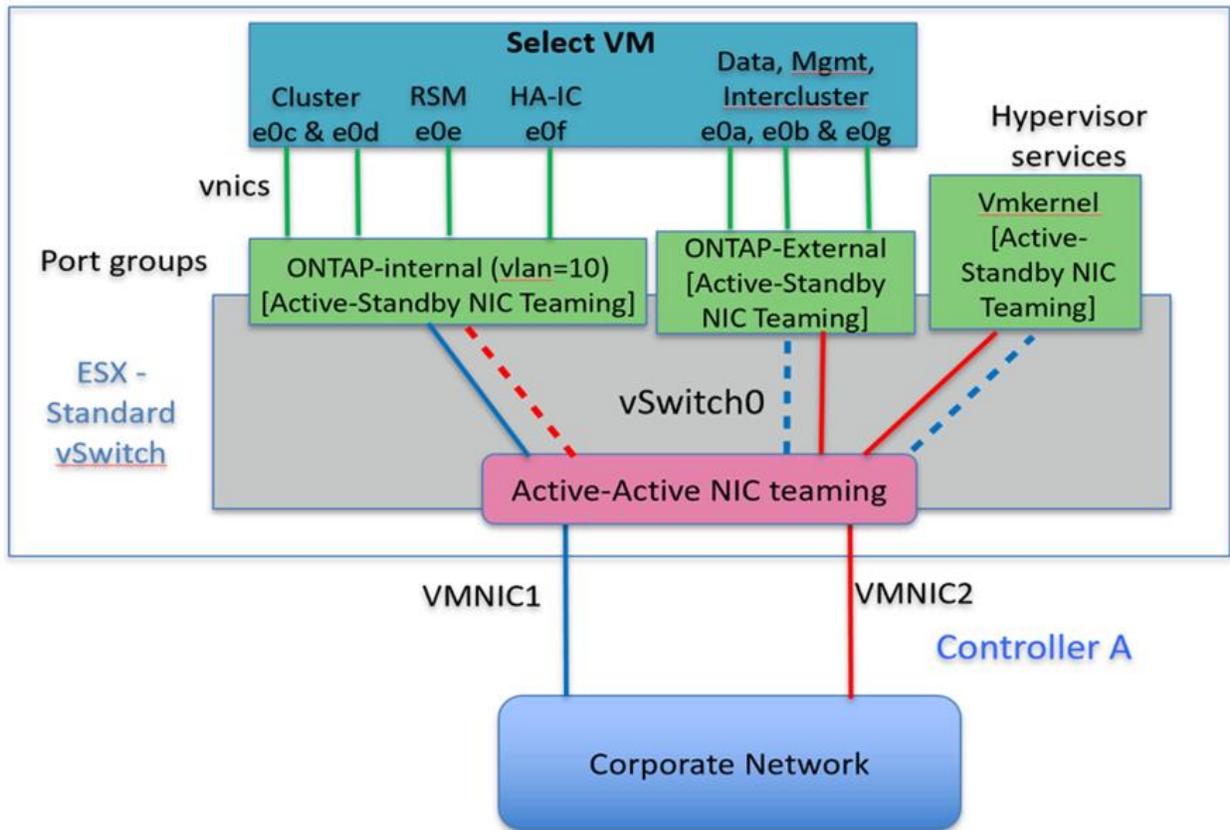
每個節點配備兩個高速 (25/40Gb) 實體連接埠的 **vSwitch**



使用兩個實體連接埠（10 Gb 或更小）時，每個連接埠組應配置一個活動適配器和一個備用適配器，且兩個適配器的連接埠方向相反。內部網路僅適用於多節點ONTAP Select叢集。對於單節點叢集，兩個適配器都可以在外部連接埠組中配置為活動狀態。

以下範例展示了一個 vSwitch 的配置以及兩個連接埠組，這兩個連接埠組負責處理多節點ONTAP Select叢集的內部和外部通訊服務。由於內部網路 VMNIC 屬於此連接埠群組，且配置為待機模式，因此在發生網路中斷時，外部網路可以使用內部網路 VMNIC。外部網路的情況則相反。在兩個連接埠群組之間交替使用活動和待機 VMNIC 對於ONTAP Select虛擬機器在網路中斷期間正確進行故障轉移至關重要。

每個節點具有兩個實體連接埠（10Gb 或更少）的 vSwitch

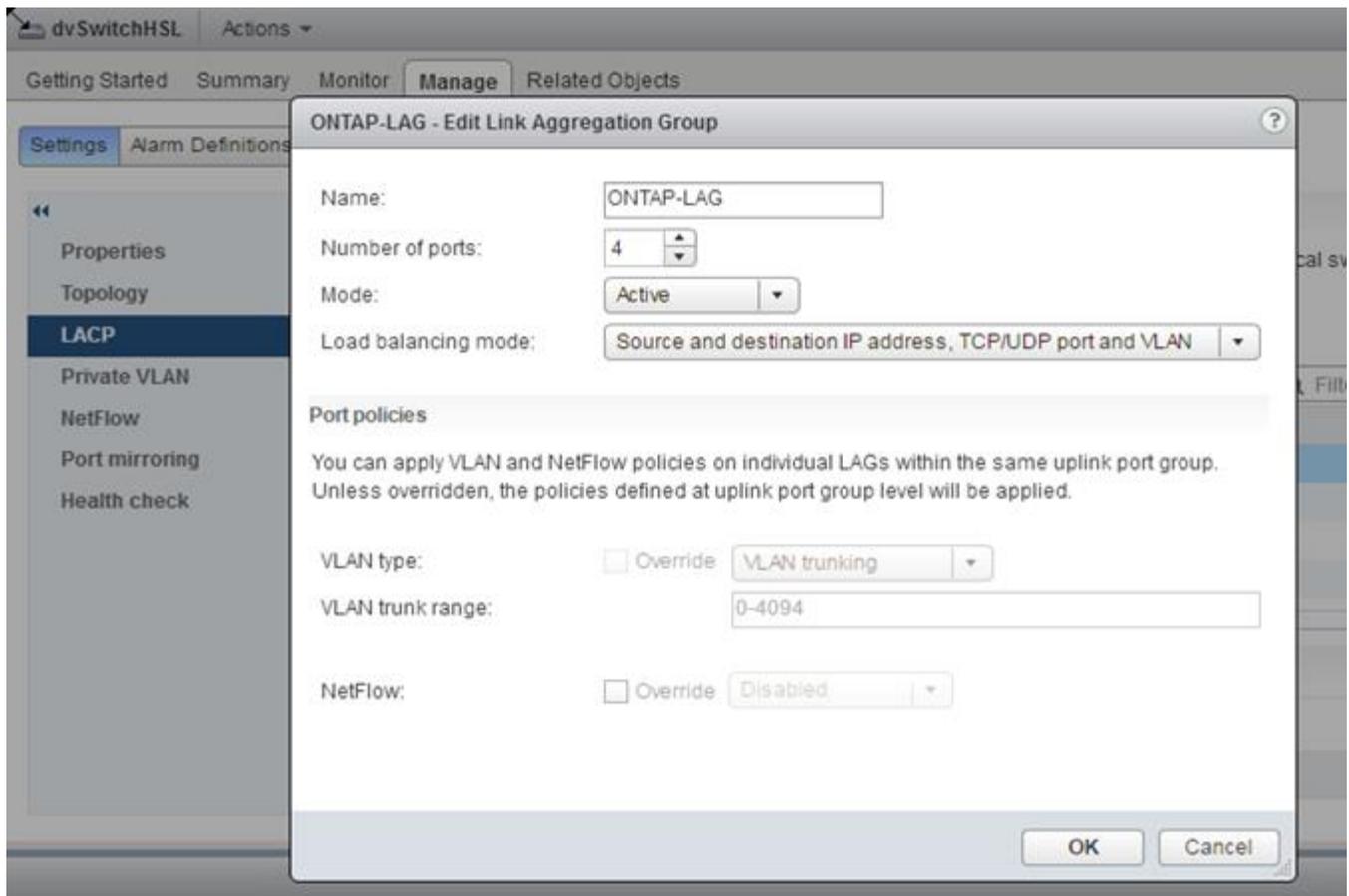


具有 LACP 的分散式 vSwitch

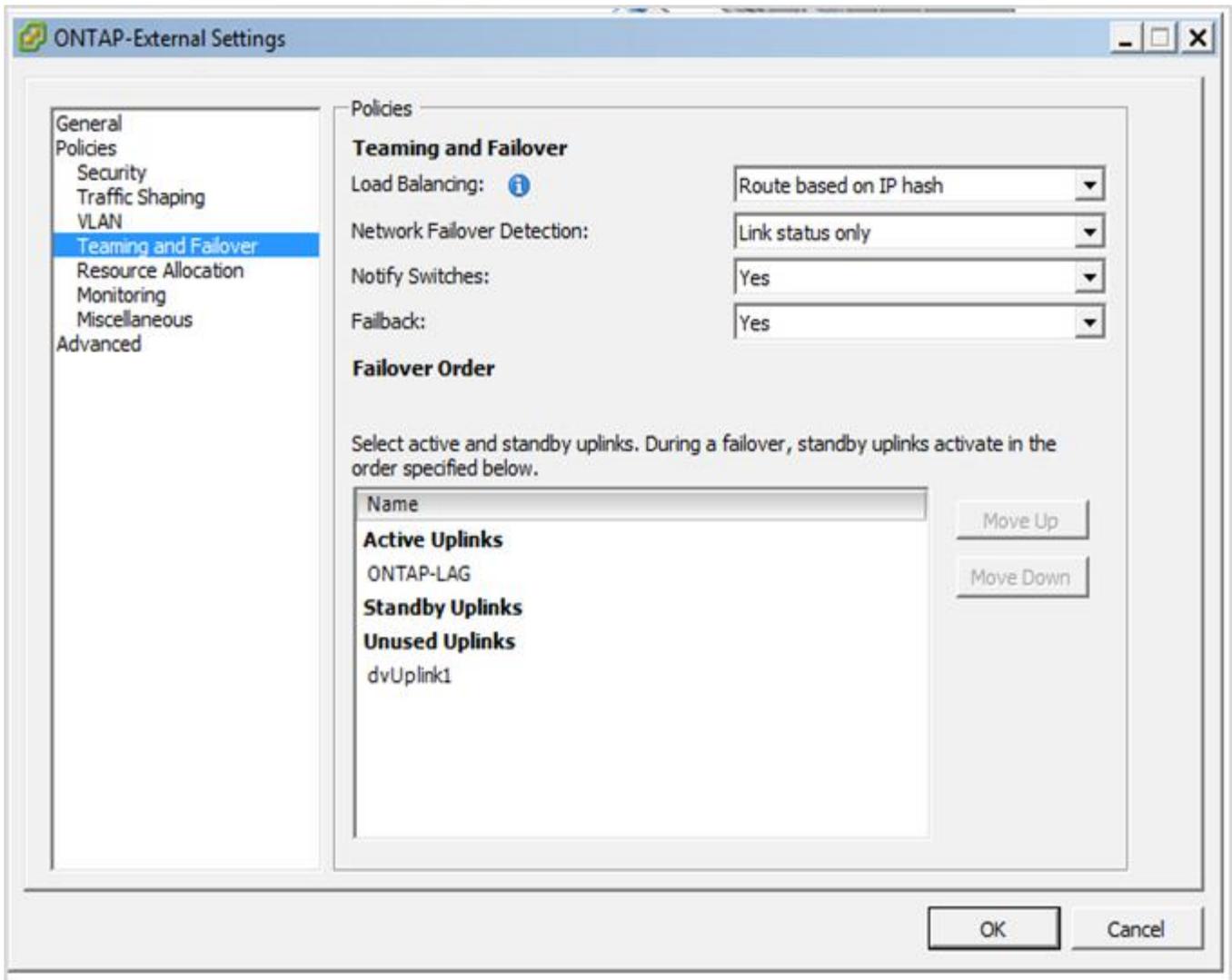
在組態中使用分散式 vSwitch 時，可以使用 LACP（儘管這並非最佳實踐）來簡化網路配置。唯一受支援的 LACP 配置要求所有 VMNIC 位於單一 LAG 中。上行鏈路實體交換器必須在通道中的所有連接埠上支援 7,500 到 9,000 之間的 MTU 大小。內部和外部 ONTAP Select 網路應在連接埠群組層級隔離。內部網路應使用不可路由（隔離）的 VLAN。外部網路可以使用 VST、EST 或 VGT。

以下範例顯示了使用 LACP 的分散式 vSwitch 配置。

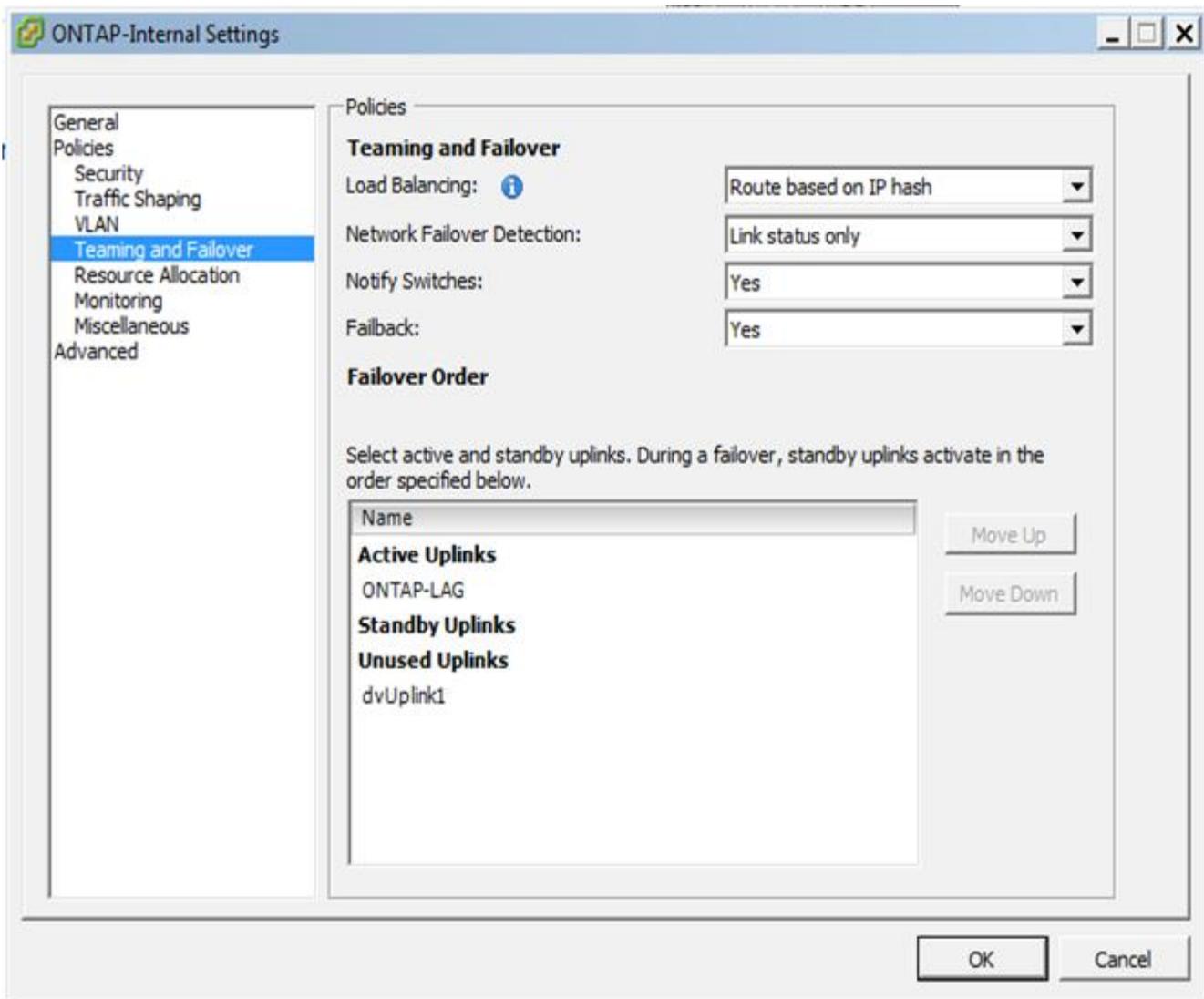
使用 LACP 時的 LAG 屬性



使用啟用了 **LACP** 的分散式 **vSwitch** 的外部連接埠群組配置



使用啟用 LACP 的分散式 vSwitch 的內部連接埠群組配置



LACP 要求您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，請確保已正確配置啟用 LACP 的連接埠通道。

ONTAP Select 實體交換器配置

基於單交換器和多交換器環境的上游實體交換器配置詳細資訊。

從虛擬交換器層到實體交換器的連接決策時，應仔細考慮。叢集內部流量與外部資料服務的分離應透過二層 VLAN 提供的隔離機制延伸至上游實體網路層。

實體交換器連接埠應配置為中繼埠。ONTAP Select 外部流量可以透過兩種方式在多個二層網路中分離。一種方法是使用單一連接埠群組的 ONTAP VLAN 標記虛擬連接埠。另一種方法是在 VST 模式下將單獨的連接埠群組指派給管理連接埠 e0a。您還必須根據 ONTAP Select 版本以及單節點或多節點配置，將資料連接埠指派給 e0b 和 e0c/e0g。如果外部流量在多個二層網路中分離，則上行鏈路實體交換器連接埠應將這些 VLAN 新增至其允許的 VLAN 清單中。

ONTAP Select 內部網路流量使用透過連結本機 IP 位址定義的虛擬介面進行傳輸。由於這些 IP 位址不可路由，因此叢集節點之間的內部流量必須透過單一第 2 層網路傳輸。ONTAP Select 叢集節點之間的路由躍點不受支援。

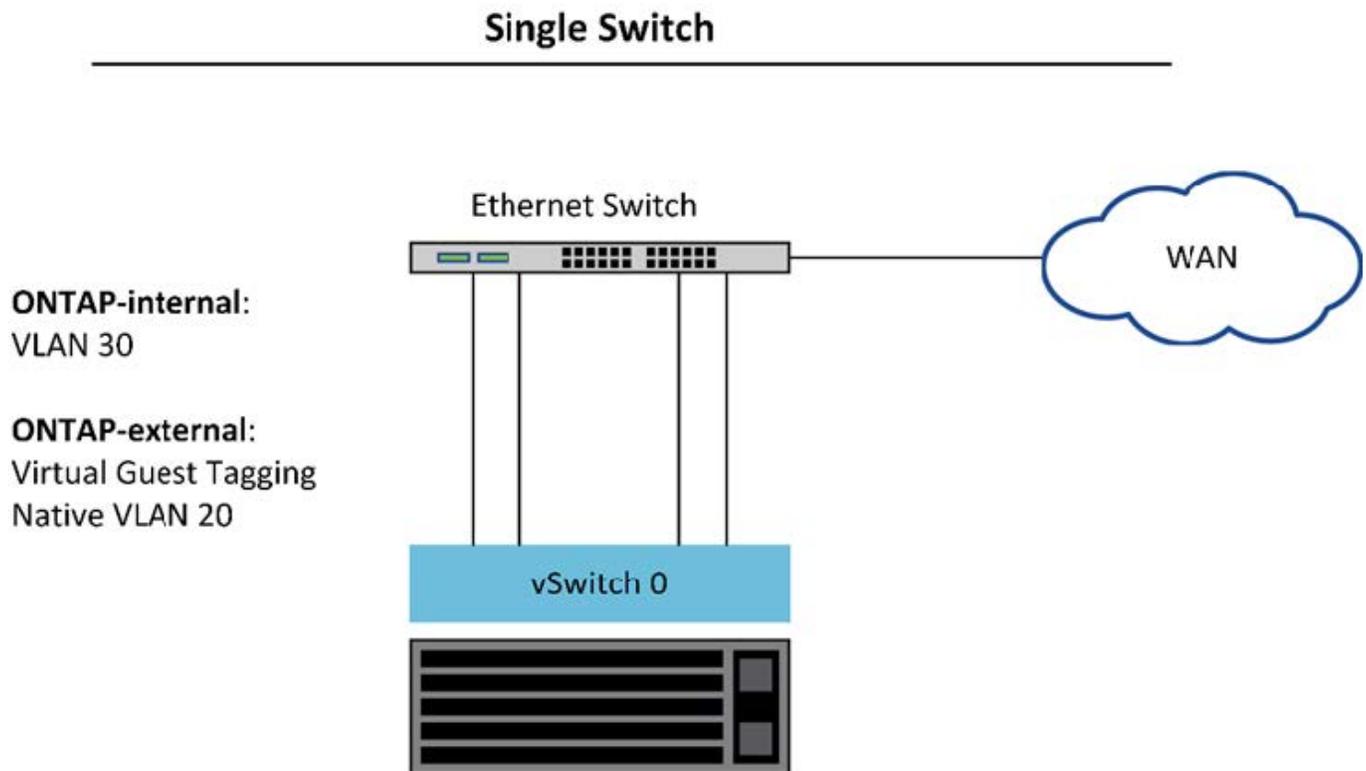
共享實體交換機

下圖展示了多節點ONTAP Select叢集中一個節點可能使用的交換器配置。在此範例中，託管內部和外部網路連接埠群組的 vSwitch 使用的實體網路卡透過纜線連接到同一上游交換器。交換器流量透過包含在不同 VLAN 中的廣播域進行隔離。



對於ONTAP Select內部網路，標記是在連接埠群組層級完成的。雖然以下範例使用 VGT 作為外部網路，但此連接埠組同時支援 VGT 和 VST。

使用共用實體交換器的網路設定

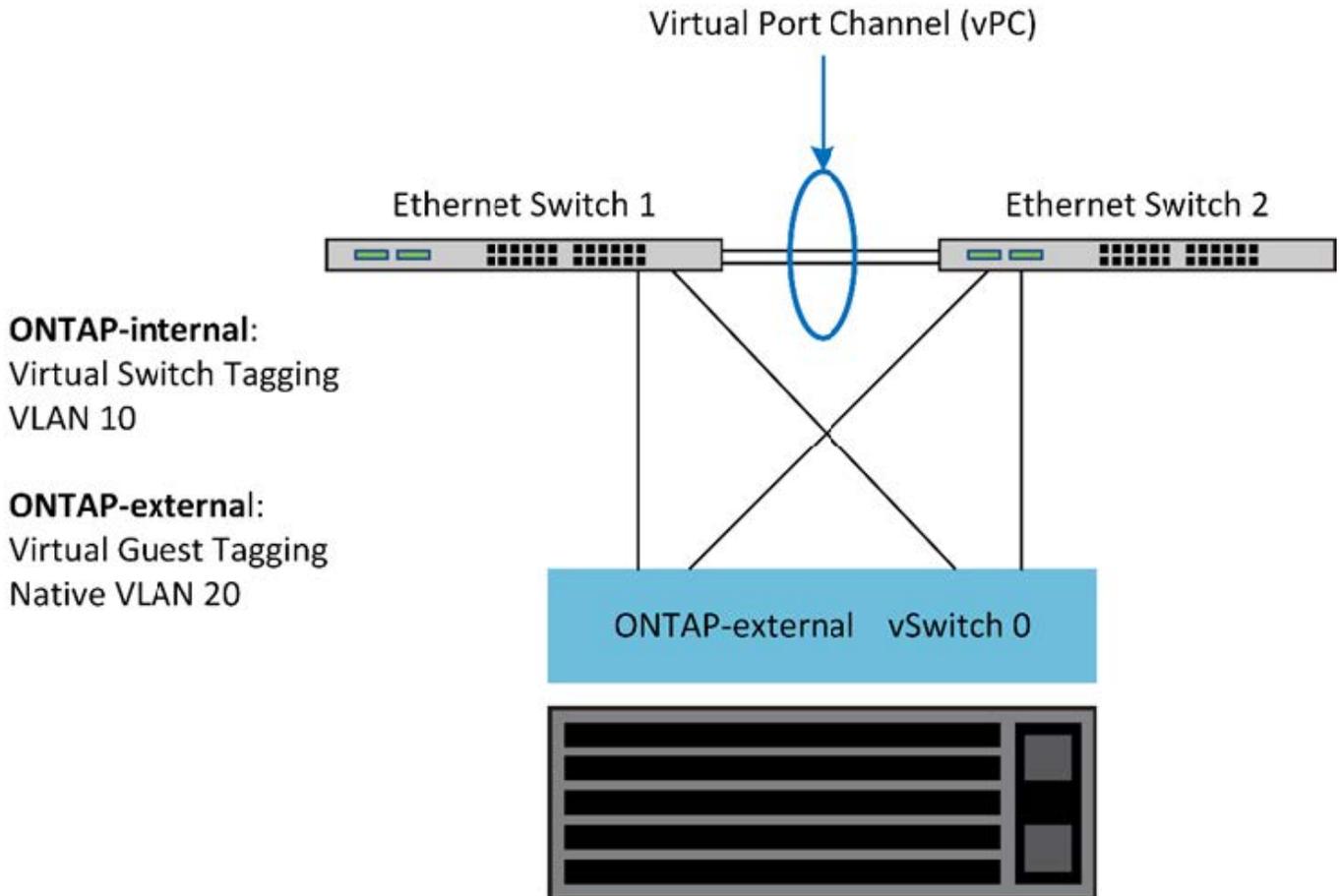


在此配置中，共用交換器將成為單點故障。如果可能，應使用多個交換機，以防止實體硬體故障導致叢集網路中斷。

多個物理交換機

如果需要冗餘，應使用多個實體網路交換器。下圖顯示了多節點ONTAP Select叢集中一個節點的建議配置。內部和外部連接埠群組的網路卡都連接到不同的實體交換機，以保護使用者免受單一硬體交換器故障的影響。交換器之間配置了虛擬連接埠通道，以防止生成樹問題。

使用多個實體交換器的網路設定



ONTAP Select 資料與管理流量分離

將資料流量和管理流量隔離到單獨的第 2 層網路。

ONTAP Select 外部網路流量定義為資料 (CIFS、NFS 和 iSCSI)、管理和複製 (SnapMirror) 流量。在 ONTAP 叢集中，每種類型的流量都使用單獨的邏輯接口，該接口必須託管在虛擬網路連接埠上。在 ONTAP Select 的多節點配置中，這些連接埠被指定為連接埠 e0a 和 e0b/e0g。在單節點配置中，這些連接埠被指定為 e0a 和 e0b/e0c，其餘連接埠則保留用於內部叢集服務。

NetApp 建議將資料流量和管理流量隔離到單獨的第 2 層網路。在 ONTAP Select 環境中，這可以透過使用 VLAN 標記來實現。您可以透過將帶有 VLAN 標記的連接埠群組指派給網路介面卡 1 (連接埠 e0a) 來實現管理流量。然後，您可以為連接埠 e0b 和 e0c (單節點叢集) 以及 e0b 和 e0g (多節點叢集) 指派單獨的連接埠群組來傳輸資料流量。

如果本文檔前面介紹的 VST 解決方案不夠用，則可能需要在同一個虛擬連接埠上共置資料 LIF 和管理 LIF。為此，請使用稱為 VGT 的流程，其中 VLAN 標記由虛擬機器執行。



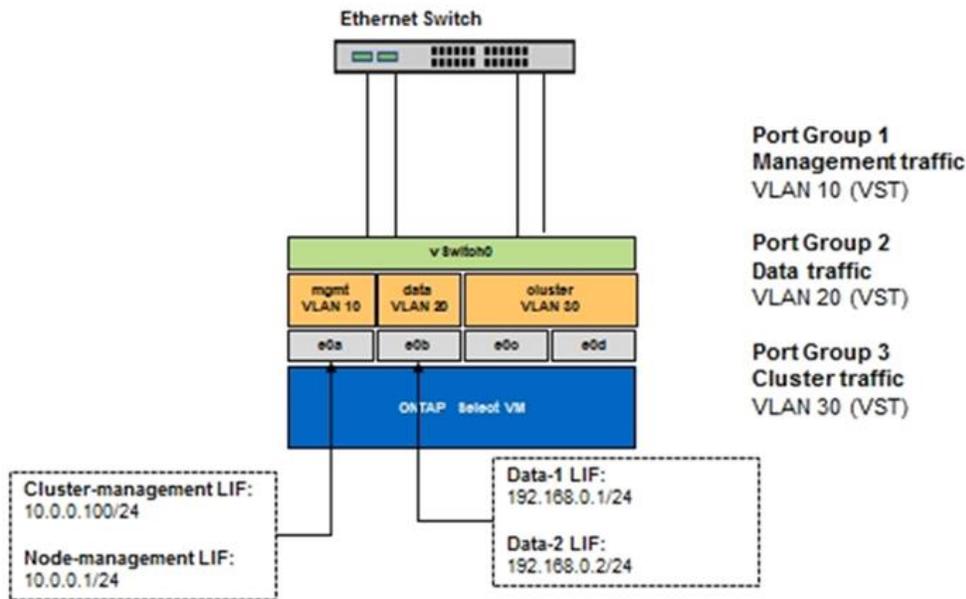
使用 ONTAP Deploy 公用程式時，無法透過 VGT 分離資料和管理網路。此過程必須在叢集設定完成後執行。

使用 VGT 和雙節點叢集時，還有一個需要注意的事項。在雙節點叢集配置中，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，對應到節點管理 LIF (連接埠 e0a) 的連接埠組僅支援 EST 和 VST 標記。此外，如果管理和資料流量都使用同一個連接埠群組，則整個雙節點叢集僅支援 EST/VST。

VST 和 VGT 兩種設定選項均受支援。下圖顯示了第一種方案 VST，其中流量透過分配的連接埠群組在 vSwitch

層進行標記。在此配置中，叢集和節點管理 LIF 指派給ONTAP連接埠 e0a，並透過指派的連接埠群組標記為 VLAN ID 10。資料 LIF 指派給連接埠 e0b 以及 e0c 或 e0g，並使用第二個連接埠群組指定 VLAN ID 20。叢集連接埠使用第三個連接埠群組，位於 VLAN ID 30 上。

使用 VST 分離資料和管理

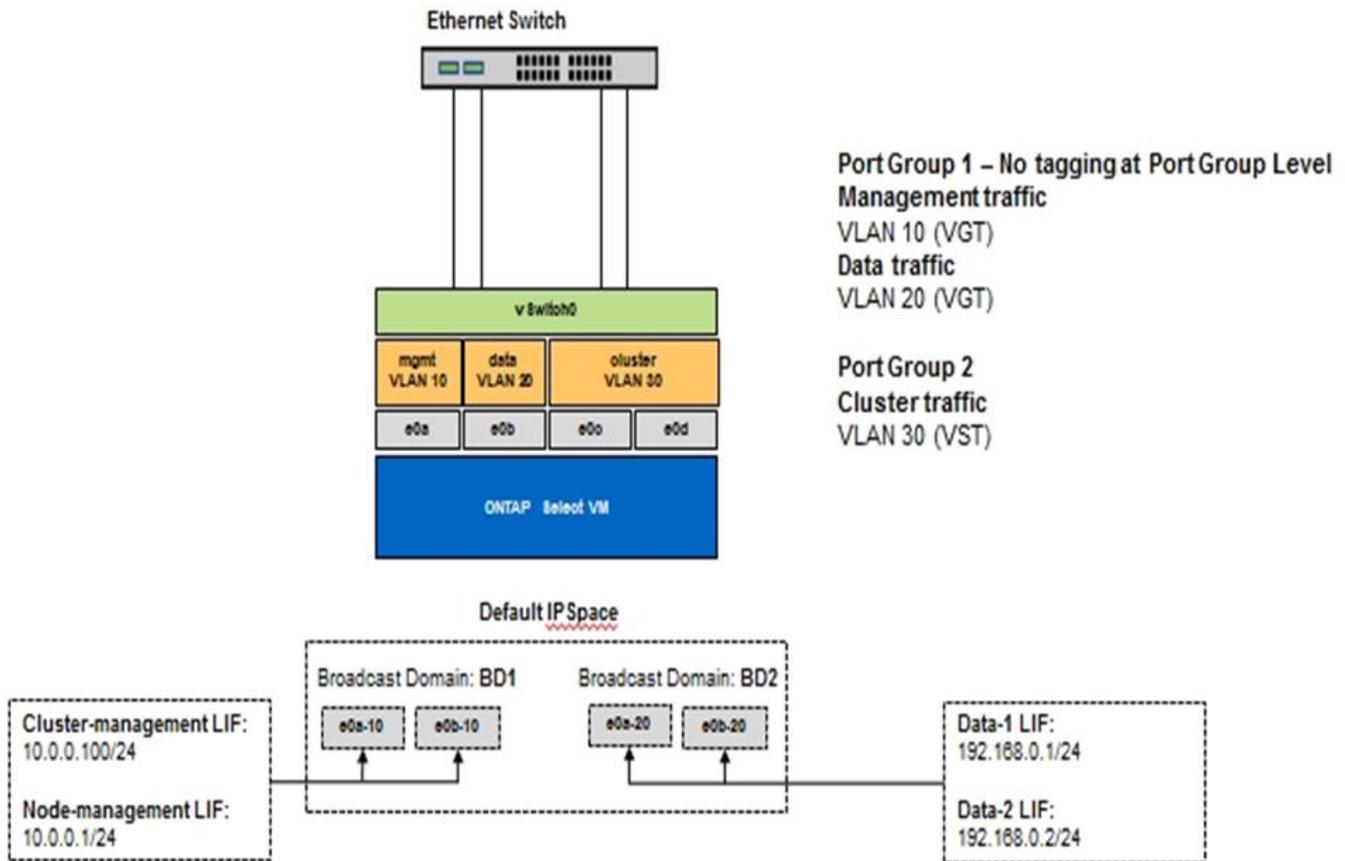


下圖顯示了第二種場景 VGT，其中ONTAP虛擬機器使用放置在單獨廣播域中的 VLAN 連接埠標記流量。在此範例中，虛擬連接埠 e0a-10/e0b-10/(e0c 或 e0g)-10 和 e0a-20/e0b-20 位於虛擬機器連接埠 e0a 和 e0b 之上。此配置允許直接在ONTAP內部執行網路標記，而不是在 vSwitch 層執行。管理和資料 LIF 位於這些虛擬連接埠上，從而允許在單一虛擬機器連接埠內進一步進行第 2 層細分。集群 VLAN (VLAN ID 30) 仍在連接埠群組中標記。

筆記：

- 這種配置方式在使用多個 IP 空間時尤其適用。如果需要進一步的邏輯隔離和多租用戶功能，可以將 VLAN 連接埠分組到單獨的自訂 IP 空間。
- 若要支援 VGT，ESXi/ESX 主機網路適配器必須連接到實體交換器上的中繼連接埠。連接到虛擬交換器的連接埠群組必須將其 VLAN ID 設定為 4095，以便在連接埠群組上啟用中繼功能。

使用 VGT 進行資料和管理分離



高可用性架構

ONTAP Select高可用性配置

探索高可用性選項，為您的環境選擇最佳的 HA 配置。

儘管客戶開始將應用程式工作負載從企業級儲存設備遷移到運行在商用硬體上的基於軟體的解決方案，但對彈性和容錯能力的期望和需求並未改變。提供零復原點目標 (RPO) 的高可用性 (HA) 解決方案可保護客戶免受基礎架構堆疊中任何元件故障導致的資料遺失。

很大一部分 SDS 市場建立在無共享儲存的概念之上，軟體複製透過在不同的儲存孤島中儲存多個用戶資料副本來提供資料彈性。ONTAP Select基於此前提構建，使用ONTAP提供的同步複製功能 (RAID SyncMirror) 在叢集內儲存額外的使用者資料副本。這發生在 HA 對的環境中。每個 HA 對都儲存兩個使用者資料副本：一個在本地節點提供的儲存上，另一個在 HA 合作夥伴提供的儲存上。在ONTAP Select叢集中，HA 和同步複製綁定在一起，兩者的功能不能分離或獨立使用。因此，同步複製功能僅在多節點產品中可用。

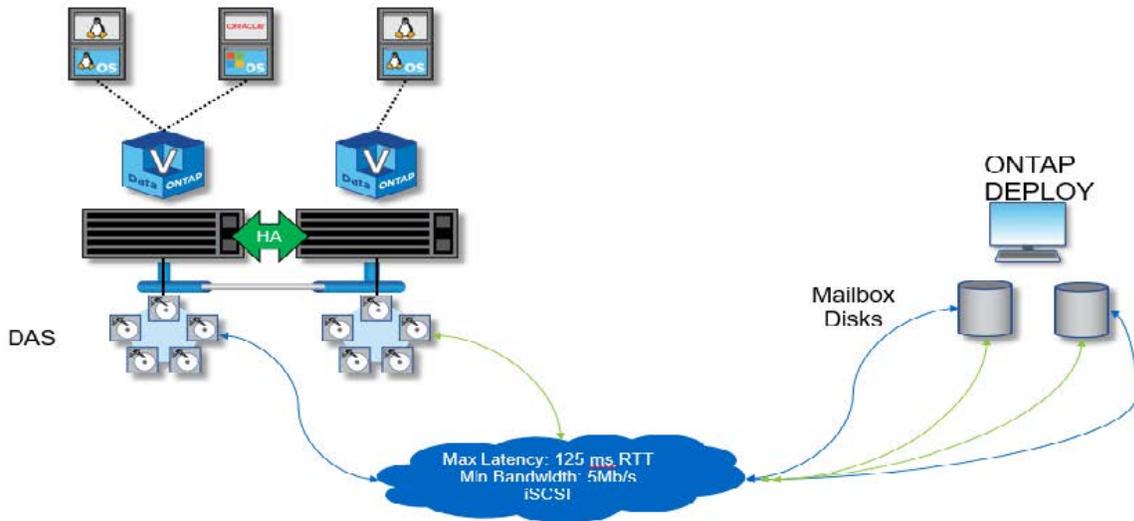


在ONTAP Select叢集中，同步複製功能是 HA 實現的功能，而不是非同步SnapMirror或SnapVault複製引擎的替代品。同步複製不能獨立於 HA 使用。

ONTAP Select HA 部署模型有兩種：多節點叢集（四節點、六節點或八節點）和雙節點叢集。雙節點ONTAP Select叢集的顯著特徵是使用外部調解服務來解決腦裂問題。ONTAPDeployONTAP機器可作為其配置的所有雙節點 HA 對的預設調解器。

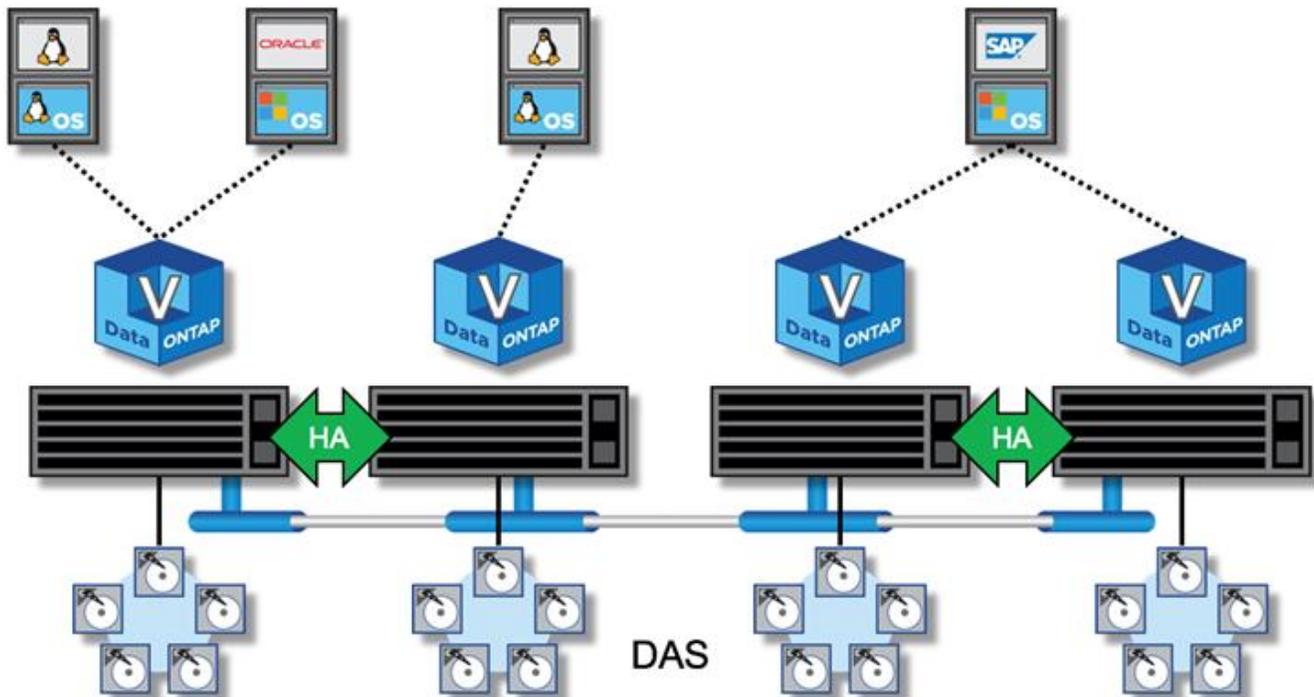
下圖表示了這兩種架構。

帶有遠端調解器並使用本地連接存儲的雙節點ONTAP Select集群



雙節點ONTAP Select叢集由一個 HA 對和一個調解器組成。在 HA 對中，每個叢集節點上的資料聚合都會同步鏡像，因此即使發生故障轉移，也不會遺失資料。

使用本地連接儲存的四節點ONTAP Select叢集



- 四節點ONTAP Select叢集由兩個 HA 對組成。六節點和八節點叢集分別由三個和四個 HA 對組成。在每個 HA 對中，每個叢集節點上的資料聚合都會同步鏡像，因此即使發生故障轉移，也不會遺失資料。
- 使用 DAS 儲存時，實體伺服器上只能存在一個ONTAP Select實例。ONTAP Select需要對系統的本機 RAID 控制器進行非共用訪問，並且旨在管理本機連接的磁碟，而如果没有與儲存裝置的實體連接，則無法實現這一點。

雙節點 HA 與多節點 HA

與FAS陣列不同，HA 對中的ONTAP Select節點僅透過 IP 網路進行通訊。這意味著 IP 網路存在單點故障 (SPOF)，因此，防止網路分區和腦裂情況成為設計的重要方面。多節點叢集可以承受單節點故障，因為叢集仲裁可以由三個或更多倖存節點建立。雙節點叢集依賴ONTAP Deploy 虛擬機器所託管的調解器服務來實現相同的結果。

ONTAP Select節點和ONTAP Deploy 中介服務之間的心跳網路流量極小且具有彈性，因此ONTAP Deploy VM 可以託管在與ONTAP Select雙節點叢集不同的資料中心。



ONTAP Deploy 虛擬機器在充當雙節點叢集的調解器時，將成為該叢集不可或缺的一部分。如果調解器服務不可用，雙節點叢集將繼續提供數據，但ONTAP Select叢集的儲存故障轉移功能將被停用。因此，ONTAP Deploy 調解器服務必須與 HA 對中的每個ONTAP Select節點保持持續通訊。為了確保叢集仲裁正常運行，最低頻寬要求為 5Mbps，最大往返時間 (RTT) 延遲要求為 125 毫秒。

如果充當調解器的ONTAP Deploy 虛擬機器暫時或可能永久無法使用，則可以使用輔助ONTAP Deploy 虛擬機器來還原雙節點叢集仲裁。這會導致新的ONTAP Deploy 虛擬機器無法管理ONTAP Select節點，但可以成功參與叢集仲裁演算法。ONTAP Select節點與ONTAP Deploy 虛擬機器之間的通訊是透過 IPv4 上的 iSCSI 協定完成的。ONTAP Select節點管理 IP 位址是啟動器，ONTAP Deploy 虛擬機器 IP 位址是目標。因此，在建立雙節點叢集時，節點管理 IP 位址無法支援 IPv6 位址。在建立雙節點叢集時，系統會自動建立ONTAP Deploy 託管的郵件磁碟，並將其封鎖為正確的ONTAP Select節點管理 IP 位址。整個配置在設定過程中自動執行，無需進一步的管理操作。建立叢集的ONTAP Deploy 實例是該叢集的預設調解器。

如果必須變更原始調解器位置，則需要執行管理操作。即使原始ONTAP Deploy 虛擬機器遺失，也可以復原叢集仲裁。但是，NetApp建議您在每個雙節點叢集實例化後備份ONTAP Deploy 資料庫。

雙節點 HA 與雙節點延伸 HA (MetroCluster SDS)

可以將雙節點主動/主動 HA 叢集擴展到更遠的距離，並可能將每個節點放置在不同的資料中心。雙節點群集和雙節點延伸群集（也稱為MetroCluster SDS）之間的唯一區別在於節點之間的網路連接距離。

雙節點叢集定義為兩個節點位於相同資料中心且相距 300 公尺以內的叢集。通常，兩個節點都具有到同一網路交換器或一組交換器間連結 (ISL) 網路交換器的上行鏈路。

雙節點MetroCluster SDS 是指節點物理上相距超過 300 公尺（不同房間、不同建築物和不同資料中心）的群集。此外，每個節點的上行鏈路連接分別連接到單獨的網路交換器。MetroCluster不需要專用硬體。但是，該環境應滿足延遲（RTT 最大 5 毫秒，抖動最大 5 毫秒，總計 10 毫秒）和物理距離（最大 10 公里）的要求。

MetroCluster SDS 是一項進階功能，需要 Premium 授權或 Premium XL 授權。Premium授權支援建立中小型虛擬機器以及 HDD 和 SSD 媒體。PremiumXL 許可證還支援建立 NVMe 驅動器。



MetroCluster SDS 支援本地連接儲存 (DAS) 和共用儲存 (vNAS)。請注意，由於ONTAP Select虛擬機器與共用儲存之間的網路，vNAS 配置通常具有較高的固有延遲。MetroClusterSDS 配置必須在節點之間提供最多 10 毫秒的延遲，其中包括共用儲存延遲。換句話說，僅測量 Select 虛擬機器之間的延遲是不夠的，因為對於這些配置來說，共用儲存延遲不可忽略。

ONTAP Select HA RSM 和鏡像聚合

使用 RAID SyncMirror (RSM)、鏡像聚合和寫入路徑防止資料遺失。

同步複製

ONTAP高可用性 (HA) 模型建立在高可用性合作夥伴的概念之上。ONTAPONTAP Select將此架構擴展到非共用商用伺服器領域，利用ONTAP中提供的 RAID SyncMirror (RSM) 功能在叢集節點之間複製資料區塊，從而在高可用性 (HA) 對中提供兩個使用者資料副本。

具有調解器的雙節點叢集可以跨越兩個資料中心。有關更多信息，請參閱["雙節點擴展 HA \(MetroCluster SDS\) 最佳實踐"](#)。

鏡像聚合

ONTAP Select叢集由 2 到 8 個節點組成。每個 HA 對包含兩個使用者資料副本，透過 IP 網路跨節點同步鏡像。此鏡像對使用者透明，並且是資料聚合的屬性，會在資料聚合建立過程中自動配置。

ONTAP Select叢集中的所有聚合都必須進行鏡像，以便在發生節點故障轉移時確保資料可用性，並在發生硬體故障時避免出現單點故障 (SPOF)。ONTAPONTAP Select叢集中的聚合基於 HA 對中每個節點提供的虛擬磁碟構建，並使用以下磁碟：

- 一組本機磁碟（由目前ONTAP Select節點提供）
- 一組鏡像磁碟（由目前節點的 HA 夥伴提供）



用於建構鏡像聚合的本機磁碟和鏡像磁碟的大小必須相同。這些聚合稱為叢 0 和叢 1（分別表示本地鏡像對和遠端鏡像對）。實際的叢編號在您的安裝中可能有所不同。

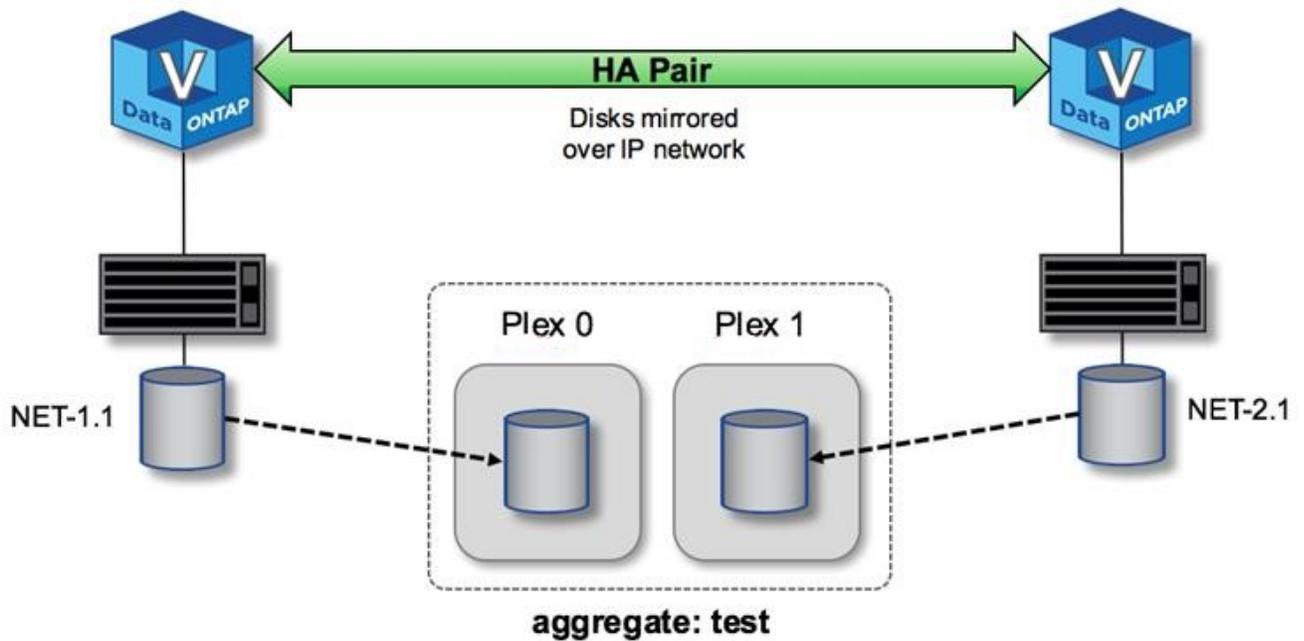
這種方法與標準ONTAP集群的工作方式有著根本的不同。這適用於ONTAP Select叢集中的所有根磁碟和資料磁碟。聚合包含資料的本機副本和鏡像副本。因此，包含 N 個虛擬磁碟的聚合可提供相當於 N/2 個磁碟的唯一存儲，因為第二個資料副本位於其自己的唯一磁碟上。

下圖顯示了四節點ONTAP Select叢集中的 HA 對。此叢集中有一個聚合（測試），它使用來自兩個 HA 配對節點的儲存。此資料聚合由兩組虛擬磁碟組成：一組本機磁碟，由ONTAP Select所屬叢集節點 (Plex 0) 提供；一組遠端磁碟，由故障轉移配對節點 (Plex 1) 提供。

Plex 0 是用於存放所有本機磁碟的儲存桶。Plex1 是用於存放鏡像磁碟（即負責儲存使用者資料第二個複製副本的磁碟）的儲存桶。擁有聚合的節點會將磁碟提供給 Plex 0，而該節點的 HA 配對節點會將磁碟提供給 Plex 1。

下圖中有一個包含兩個磁碟的鏡像聚合。此聚合的內容在兩個叢集節點之間進行鏡像，本地磁碟 NET-1.1 放置在 Plex 0 儲存桶中，遠端磁碟 NET-2.1 放置在 Plex 1 儲存桶中。在此範例中，聚合測試歸左側的叢集節點所有，並使用本機磁碟 NET-1.1 和 HA 夥伴鏡像磁碟 NET-2.1。

- ONTAP Select鏡像聚合*



部署ONTAP Select叢集時，系統上的所有虛擬磁碟都會自動指派給正確的 Plex，而無需使用者執行任何與磁碟分配相關的額外步驟。這可以防止磁碟意外分配給錯誤的 Plex，並提供最佳的鏡像磁碟配置。

寫入路徑

叢集節點之間的資料區塊同步鏡像以及系統故障時不遺失資料的要求對傳入寫入作業在ONTAP Select叢集中傳播時所採用的路徑有重大影響。此過程包含兩個階段：

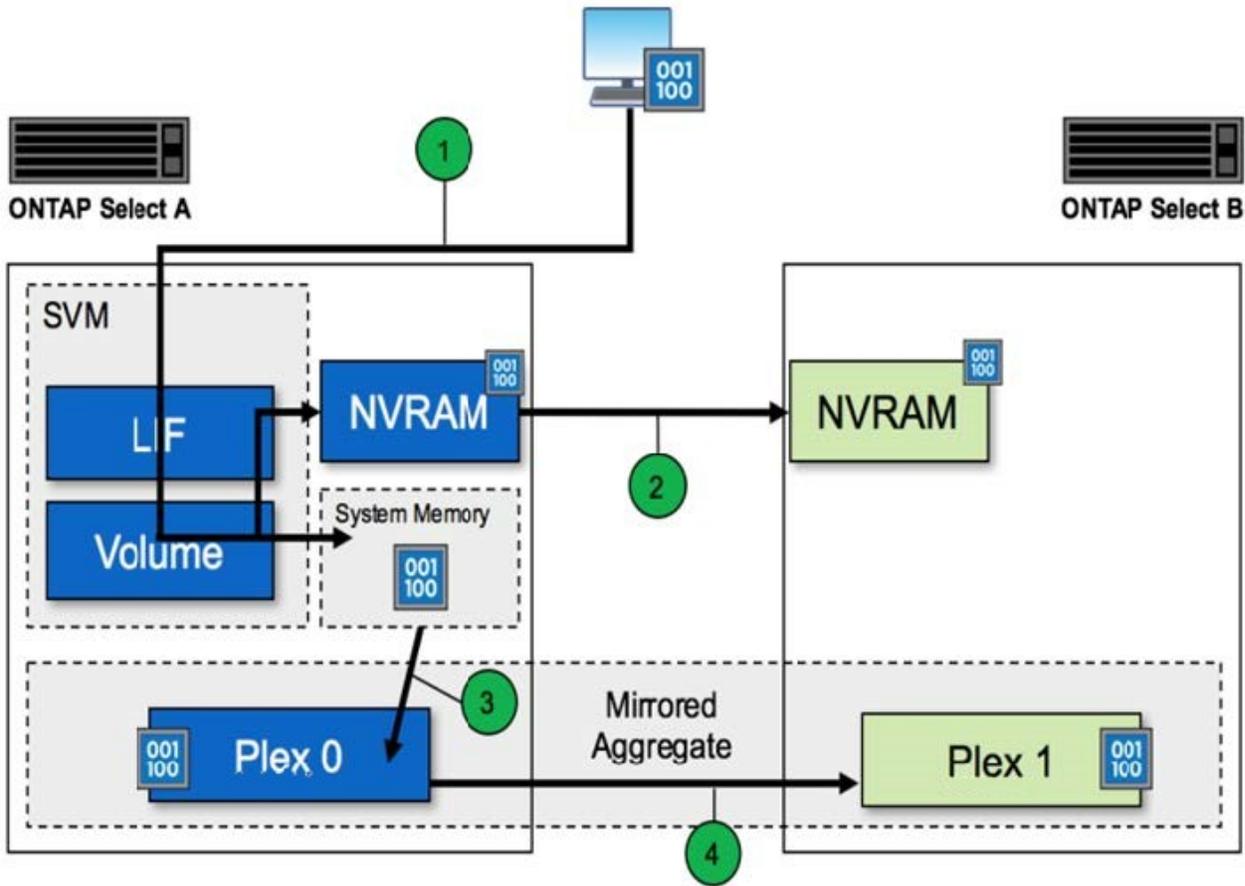
- 致謝
- 降級

對目標磁碟區的寫入作業透過資料 LIF 進行，並提交到ONTAP Select節點系統磁碟上的虛擬化NVRAM分割區，然後再向客戶端確認。在 HA 設定中，也會執行一個額外的步驟，因為這些NVRAM寫入作業會在確認之前立即鏡像到目標磁碟區擁有者的 HA 配對節點。此流程可確保在原始節點發生硬體故障時，HA 配對節點上的檔案系統保持一致。

將寫入內容提交至NVRAM後，ONTAP會定期將此分割區的內容移至對應的虛擬磁碟，此程序稱為降級轉儲。此過程僅在擁有目標磁碟區的叢集節點上發生一次，不會在 HA 配對節點上發生。

下圖顯示了傳入ONTAP Select節點的寫入請求的寫入路徑。

- ONTAP Select寫入路徑工作流程*



傳入寫入確認包括以下步驟：

- 寫入透過ONTAP Select節點 A 擁有的邏輯介面進入系統。
- 寫入作業提交到節點 A 的 NVRAM 並鏡像到 HA 夥伴節點 B。
- 當兩個 HA 節點都出現 I/O 請求後，該請求就會被確認回客戶端。

ONTAP Select 從 NVRAM 降級到資料聚合 (ONTAP CP) 包括以下步驟：

- 寫入操作從虛擬 NVRAM 轉入虛擬資料聚合。
- 鏡像引擎同步將區塊複製到兩個 plex。

ONTAP Select HA 增強資料保護

高可用性 (HA) 磁碟心跳、HA 信箱、HA 心跳、HA 故障轉移和回饋功能可增強資料保護。

磁碟心跳

儘管 ONTAP Select HA 架構充分利用了傳統 FAS 陣列的許多程式碼路徑，但也存在一些例外。其中一個例外是基於磁碟的心跳機制的實現，這是一種非基於網路的通訊方法，叢集節點使用這種方法來防止網路隔離導致裂腦行為。裂腦場景是群集分區的結果，通常由網路故障引起，導致群集兩端都認為對方已宕機並試圖接管群集資源。

企業級 HA 實施必須妥善處理此類情況。ONTAP 透過一種基於磁碟的客製化心跳機制來實現這一點。ONTAP

郵箱負責處理此任務，它是叢集節點用於傳遞心跳訊息的實體儲存位置。這有助於集群確定連接性，從而在發生故障轉移時確定仲裁。

在使用共享儲存 HA 架構的FAS陣列上，ONTAP透過以下方式解決裂腦問題：

- SCSI 持久預留
- 持久 HA 元數據
- 透過 HA 互連發送 HA 狀態

然而，在ONTAP Select叢集的無共用架構中，節點只能看到自己的本機存儲，而看不到 HA 配對節點的本機儲存。因此，當網路分區隔離 HA 對的每一側時，上述確定集群仲裁和故障轉移行為的方法將無法使用。

雖然現有的裂腦檢測和避免方法無法使用，但仍需要一種能夠適應無共享環境約束的調解方法。ONTAP Select進一步擴展了現有的郵箱基礎架構，使其能夠在網路分區時充當調解方法。由於共用儲存不可用，因此調解是透過 NAS 存取郵箱磁碟來完成的。這些磁碟分佈在整個叢集中，包括雙節點叢集中的調解器，並使用 iSCSI 協定。因此，叢集節點可以根據對這些磁碟的存取做出智慧故障轉移決策。如果一個節點可以存取其 HA 夥伴節點之外的其他節點的郵箱磁碟，則該節點很可能已啟動且運作正常。



解決叢集仲裁和裂腦問題的郵箱架構和基於磁碟的心跳方法是ONTAP Select多節點變體需要四個獨立節點或雙節點叢集的調解器的原因。

HA 郵箱發帖

HA 郵箱架構採用訊息發布模型。叢集節點會定期向叢集中所有其他郵件磁碟（包括中介節點）發布訊息，表示該節點已啟動並正在執行。在正常運作的叢集中，任何時間點，叢集節點上的單一郵箱磁碟都會收到來自所有其他叢集節點的訊息。

每個 Select 叢集節點都附加一個虛擬磁碟，專門用於共用郵箱存取。此磁碟被稱為中介郵箱磁碟，因為其主要功能是在發生節點故障或網路分割時充當叢集中介。此郵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。這些節點會定期將運作狀況發佈到郵件匣磁碟的對應分割區。使用分佈在整個叢集中的網路可存取郵箱磁碟，您可以透過可達性矩陣推斷節點的運作狀況。例如，叢集節點 A 和 B 可以向叢集節點 D 的郵箱傳送郵件，但不能傳送郵件給節點 C 的郵件。此外，叢集節點 D 無法向節點 C 的郵箱發送郵件，因此節點 C 很可能已關閉或網路隔離，應該被接管。

HA 心跳

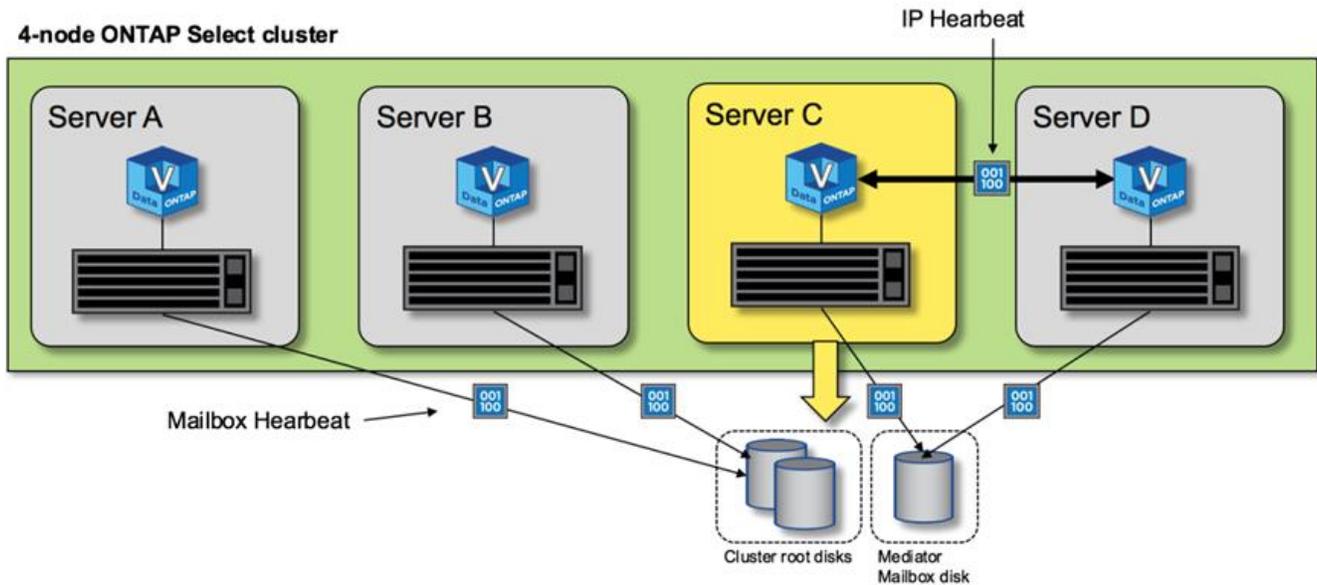
與NetApp FAS平台一樣，ONTAP Select會定期透過 HA 互連發送 HA 心跳訊息。在ONTAP Select叢集中，此操作透過 HA 夥伴節點之間的 TCP/IP 網路連線執行。此外，基於磁碟的心跳訊息會傳遞到所有 HA 郵箱磁碟，包括中介郵箱磁碟。這些訊息每隔幾秒鐘傳遞一次，並定期讀取。如此高的發送和接收頻率使ONTAP Select叢集能夠在大約 15 秒內檢測到 HA 故障事件，這與FAS平台上的可用時間視窗相同。當不再讀取心跳訊息時，將觸發故障轉移事件。

下圖從單一ONTAP Select叢集節點（節點 C）的角度顯示了透過 HA 互連和調解磁碟發送和接收心跳訊息的過程。



網路心跳透過 HA 互連傳送到 HA 夥伴節點 D，而磁碟心跳使用跨所有叢集節點 A、B、C 和 D 的郵件磁碟。

四節點叢集中的 HA 心跳：穩定狀態



HA 故障轉移與恢復

在故障轉移作業期間，倖存節點將使用其 HA 夥伴節點資料的本機副本承擔其對等節點的資料服務責任。客戶端 I/O 可以繼續不間斷地運行，但必須先複製對此資料的更改，然後才能進行交還。請注意，ONTAP Select 不支援強制交還，因為這會導致儲存在倖存節點上的變更遺失。

重啟的節點重新加入叢集時，會自動觸發同步恢復原操作。同步恢復所需的時間取決於多種因素。這些因素包括必須複製的更改數量、節點之間的網路延遲以及每個節點上磁碟子系統的速度。同步恢復所需的時間可能會超過 10 分鐘的自動交還視窗。在這種情況下，需要在同步恢復後進行手動交還。您可以使用以下命令監控同步恢復的進度：

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

表現

ONTAP Select 效能概述

由於底層硬體和配置的特性，ONTAP Select 叢集的效能可能會有很大差異。具體的硬體配置是影響特定 ONTAP Select 實例效能的最大因素。以下是一些影響特定 ONTAP Select 實例效能的因素：

- 核心頻率。一般來說，頻率越高越好。
- 單插槽與多插槽。ONTAP Select 不使用多插槽功能，但支援多插槽配置的虛擬機器管理程式開銷會導致整體效能出現一定程度的偏差。
- RAID 卡配置和相關的虛擬機器管理程式驅動程式。虛擬機器管理程式提供的預設驅動程式可能需要替換為硬體供應商的驅動程式。
- RAID 群組中的磁碟機類型和磁碟機數量。
- 虛擬機器管理程式版本和修補程式等級。

ONTAP Select 9.6 效能：進階 HA 直連 SSD 存儲

參考平台的效能資訊。

參考平台

ONTAP Select (Premium XL) 硬體 (每個節點)

- 富士通 PRIMERGY RX2540 M4：
 - 英特爾® 至強® 金牌 6142b CPU，2.6 GHz
 - 32 個實體核心 (16 x 2 個插槽)，64 個邏輯核心
 - 256 GB 內存
 - 每台主機的驅動器：24 960GB SSD
 - ESX 6.5U1

用戶端硬體

- 5 個 NFSv3 IBM 3550m4 用戶端

配置資訊

- SW RAID 1 x 9 + 2 RAID-DP (11 個磁碟機)
- 22+1 RAID-5 (ONTAP中為 RAID-0) /RAID 快取NVRAM
- 未使用任何儲存效率功能 (壓縮、重複資料刪除、Snapshot 副本、SnapMirror等)

下表列出了使用軟體 RAID 和硬體 RAID 在ONTAP Select高可用性 (HA) 節點對上針對讀取/寫入工作負載測得的吞吐量。效能測量是使用 SIO 負載產生工具進行的。



這些性能數字基於ONTAP Select 9.6。

單節點 (四節點中型實例的一部分) **ONTAP Select**叢集在直接連接儲存 (DAS) SSD 上的效能結果，具有軟體 RAID 和硬體 RAID

描述	順序讀取 64KiB	順序寫入 64KiB	隨機讀取 8KiB	隨機寫入 8KiB	隨機寫/讀 (50/50) 8KiB
具有 DAS (SSD) 軟體 RAID 的ONTAP Select 大型實例	2171 MiBps	559 MiBps	954 MiBps	394 MiBps	564 MiBps
帶有 DAS (SSD) 軟體 RAID 的ONTAP Select 中型實例	2090 MiBps	592 MiBps	677 MiBps	335 MiBps	441 3MiBps

描述	順序讀取 64KiB	順序寫入 64KiB	隨機讀取 8KiB	隨機寫入 8KiB	隨機寫/讀 (50/50) 8KiB
具有 DAS (SSD) 硬體 RAID 的ONTAP Select 中型實例	2038 MiBps	520 MiBps	578 MiBps	325 MiBps	399 MiBps

64K順序讀取

細節：

- 啟用 SIO 直接 I/O
- 2個節點
- 每個節點 2 個資料 NIC
- 每個節點 1 個資料聚合 (2TB 硬體 RAID) ， (8TB 軟體 RAID)
- 64 個 SIO 行程，每個行程 1 個執行緒
- 每個節點 32 個卷
- 每個進程 1 個檔案；每個檔案 12000MB

64K順序寫入

細節：

- 啟用 SIO 直接 I/O
- 2個節點
- 每個節點 2 個資料網路介面卡 (NIC)
- 每個節點 1 個資料聚合 (2TB 硬體 RAID) ， (4TB 軟體 RAID)
- 128 個 SIO 進程，每個進程 1 個執行緒
- 每個節點的捲數：32 (硬體 RAID) ， 16 (軟體 RAID)
- 每個進程 1 個檔案；每個檔案 30720MB

8K隨機讀取

細節：

- 啟用 SIO 直接 I/O
- 2個節點
- 每個節點 2 個資料 NIC
- 每個節點 1 個資料聚合 (2TB 硬體 RAID) ， (4TB 軟體 RAID)
- 64 個 SIO 行程，每個行程 8 個執行緒
- 每個節點的捲數：32
- 每個進程 1 個檔案；每個檔案 12228MB

8K隨機寫入

細節：

- 啟用 SIO 直接 I/O
- 2個節點
- 每個節點 2 個資料 NIC
- 每個節點 1 個資料聚合 (2TB 硬體 RAID) ， (4TB 軟體 RAID)
- 64 個 SIO 行程，每個行程 8 個執行緒
- 每個節點的捲數：32
- 每個進程 1 個檔案；每個檔案 8192MB

8K 隨機 50% 寫入 50% 讀取

細節：

- 啟用 SIO 直接 I/O
- 2個節點
- 每個節點 2 個資料 NIC
- 每個節點 1 個資料聚合 (2TB 硬體 RAID) ， (4TB 軟體 RAID)
- 64 個 SIO 行程，每個行程 208 個執行緒
- 每個節點的捲數：32
- 每個進程 1 個檔案；每個檔案 12228MB

版權資訊

Copyright © 2026 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。