深入探討
ONTAP Select

NetApp
November 08, 2026

This PDF was generated from https://docs.netapp.com/zh-tw/ontap-select-9191/concept_stor_concepts_chars.html on November 08, 2026. Always check docs.netapp.com for the latest.

目錄

深入探討 . . . 1
儲存 . . . 1
了解 ONTAP Select 的一般儲存設備特性 . . . 1
適用於具有本機附加儲存設備的 ONTAP Select 的硬體 RAID 服務 . . . 6
ONTAP Select 軟體 RAID 組態服務，適用於本機附加儲存設備 . . . 12
ONTAP Select vSAN 和外部陣列組態 . . . 20
關於增加 ONTAP Select 儲存容量的考慮因素 . . . 24
ONTAP Select 儲存效率支援 . . . 27
網路 . . . 29
ONTAP Select 網路概念和特性 . . . 29
ONTAP Select 單節點和多節點網路組態 . . . 31
ONTAP Select 內部和外部網路 . . . 36
支援的 ONTAP Select 網路組態 . . . 38
ESXi 上的 ONTAP Select VMware vSphere vSwitch 組態 . . . 39
ONTAP Select 實體交換器組態 . . . 48
ONTAP Select 資料和管理流量分離 . . . 50
高可用性架構 . . . 52
ONTAP Select 高可用性組態 . . . 52
ONTAP Select HA RSM 和鏡像 Aggregate . . . 54
ONTAP Select HA 的資料保護增強功能 . . . 57
了解 ONTAP Select 的效能 . . . 59

深入探討

儲存

了解 **ONTAP Select** 的一般儲存設備特性

在深入了解特定的儲存組件之前，請先了解適用於 ONTAP Select 環境的一般儲存概念。

儲存組態階段

ONTAP Select 主機儲存的主要組態階段包括以下幾個方面：

- 部署前準備工作
  - 確保每個 Hypervisor 主機都已設定完成，並準備好進行 ONTAP Select 部署。
  - 組態涉及實體磁碟機、RAID 控制器和群組、LUN 以及相關的網路準備。
  - 此配置是在 ONTAP Select 之外執行的。
- 使用 hypervisor 管理程式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）來設定儲存資料的某些方面。
  - 此配置是在 ONTAP Select 之外執行的。
- 使用 ONTAP Select Deploy 管理公用程式進行組態設定
  - 您可以使用 Deploy 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，也可以由該實用程式在部署過程中自動執行。
- 部署後組態
  - ONTAP Select 部署完成後，您可以使用 ONTAP CLI 或 System Manager 設定叢集。
  - 此組態是在 ONTAP Select Deploy 之外執行的。

託管儲存與非託管儲存

由 ONTAP Select 直接存取和控制的儲存設備是託管儲存設備。同一 Hypervisor 主機上的任何其他儲存設備都是非託管儲存設備。

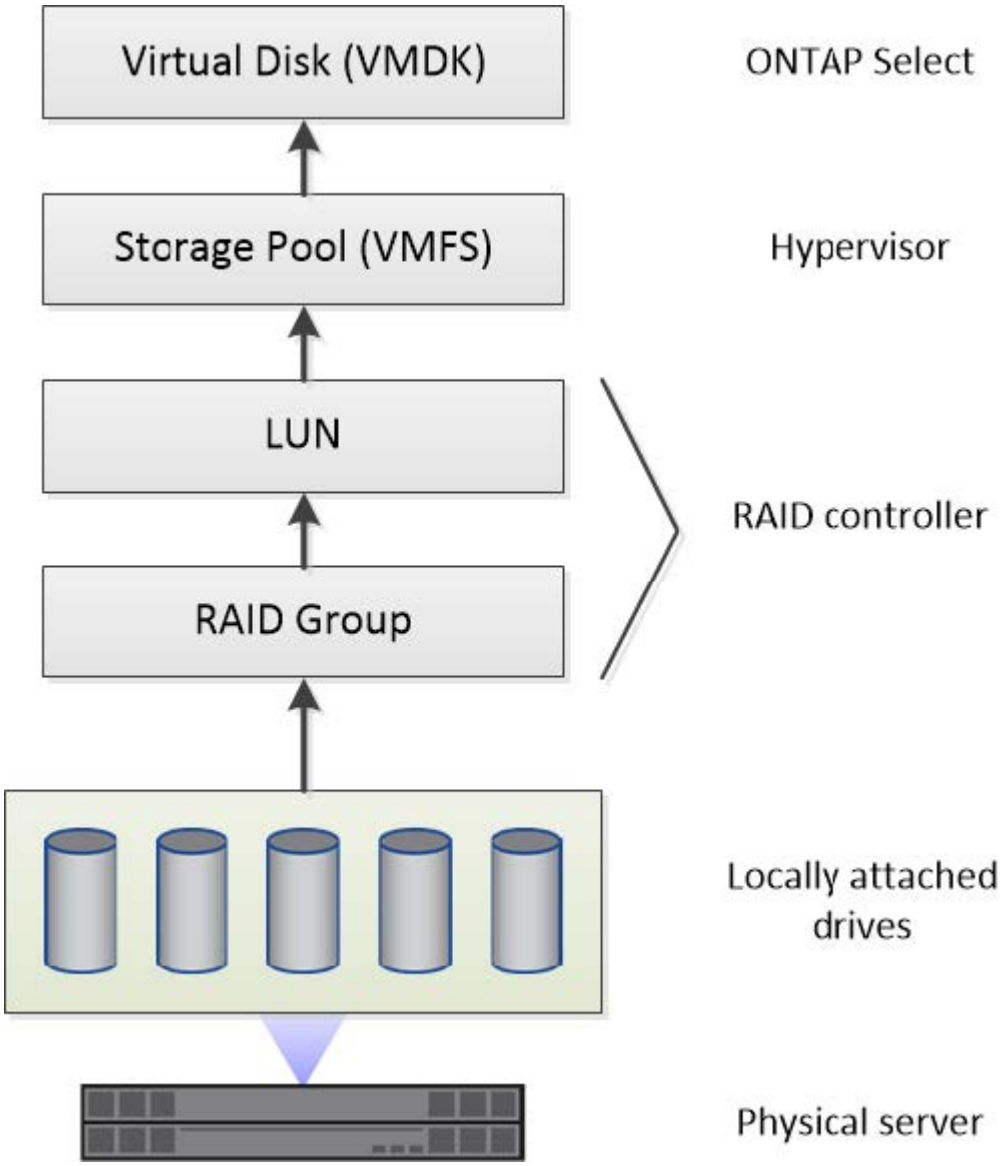
同質實體儲存設備

組成 ONTAP Select 管理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）
- 轉速 (RPM)

本機儲存環境示意圖

每個 Hypervisor 主機都包含可供 ONTAP Select 使用的本機磁碟和其他邏輯儲存元件。這些儲存元件以分層結構排列，從實體磁碟開始。



本機儲存元件的特性

ONTAP Select 環境中使用的本機儲存元件涉及多個概念。在準備部署 ONTAP Select 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、儲存資源池以及虛擬磁碟。

將實體磁碟機分組到 **RAID** 群組和 **LUN** 中

一個或多個實體磁碟可以本地連接到主機伺服器，並供 ONTAP Select 使用。這些實體磁碟被指派到 RAID 群組，然後以一個或多個 LUN 的形式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：

- 所有託管儲存設備都必須可透過單一 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也不同

一個或多個 **RAID** 群組

每個 ONTAP Select 主機必須只有一個 RAID 控制器。您應該為 ONTAP Select 建立一個 RAID 群組。但是，在某些情況下，您可以考慮建立多個 RAID 群組。請參閱 "最佳實務做法摘要"。

儲存資源池考量事項

在準備部署 ONTAP Select 時，您應該了解與儲存資源池相關的幾個問題。

在 VMware 環境中，儲存資源池與 VMware 資料存放區是同義詞。

儲存資源池和 **LUN**

每個 LUN 在虛擬機器管理程式主機上都被視為本機磁碟，並且可以是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資源池。您可以使用 Hypervisor 管理工具建立儲存資源池。例如，在 VMware 中,您可以使用 vSphere 用戶端建立儲存資源池。然後,該儲存資源池將傳遞給 ONTAP Select Deploy 管理公用程式。

管理 **ESXi** 上的虛擬磁碟

在準備部署 ONTAP Select 時，您應該了解與虛擬磁碟相關的幾個問題。

虛擬磁碟和檔案系統

ONTAP Select 虛擬機器被指派了多個虛擬磁碟機。每個虛擬磁碟實際上都是儲存池中的一個檔案，由虛擬機器管理程式維護。ONTAP Select 使用幾種類型的磁碟，主要包括系統磁碟和資料磁碟。

您還應該注意以下有關虛擬磁碟的事項：

- 必須先有可用的儲存資源池，才能建立虛擬磁碟。
- 在建立虛擬機器之前，無法建立虛擬磁碟。
- 您必須依賴 ONTAP Select Deploy 管理公用程式來建立所有虛擬磁碟（也就是說，管理員絕不能透過 Deploy 公用程式以外的方式建立虛擬磁碟）。

設定虛擬磁碟

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，它們會自動建立。

ESXi 上的外部儲存環境示意圖

ONTAP Select vNAS 解決方案可讓 ONTAP Select 能夠使用位於 Hypervisor 主機外部儲存設備上的資料存放區。這些資料存放區可以透過網路使用 VMware vSAN 存取，也可以直接在外部儲存陣列存取。

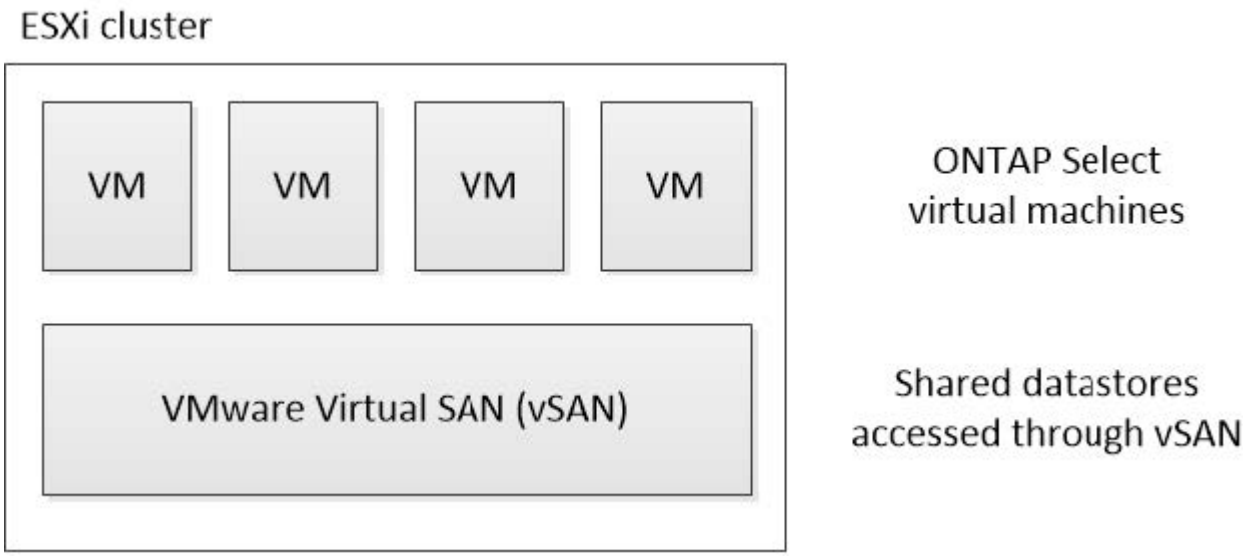
ONTAP Select 可以設定為使用下列類型的 VMware ESXi 網路資料存放區，這些資料存放區位於 Hypervisor 主機外部：

- vSAN（Virtual SAN）
- VMFS

- NFS

vSAN 資料存放區

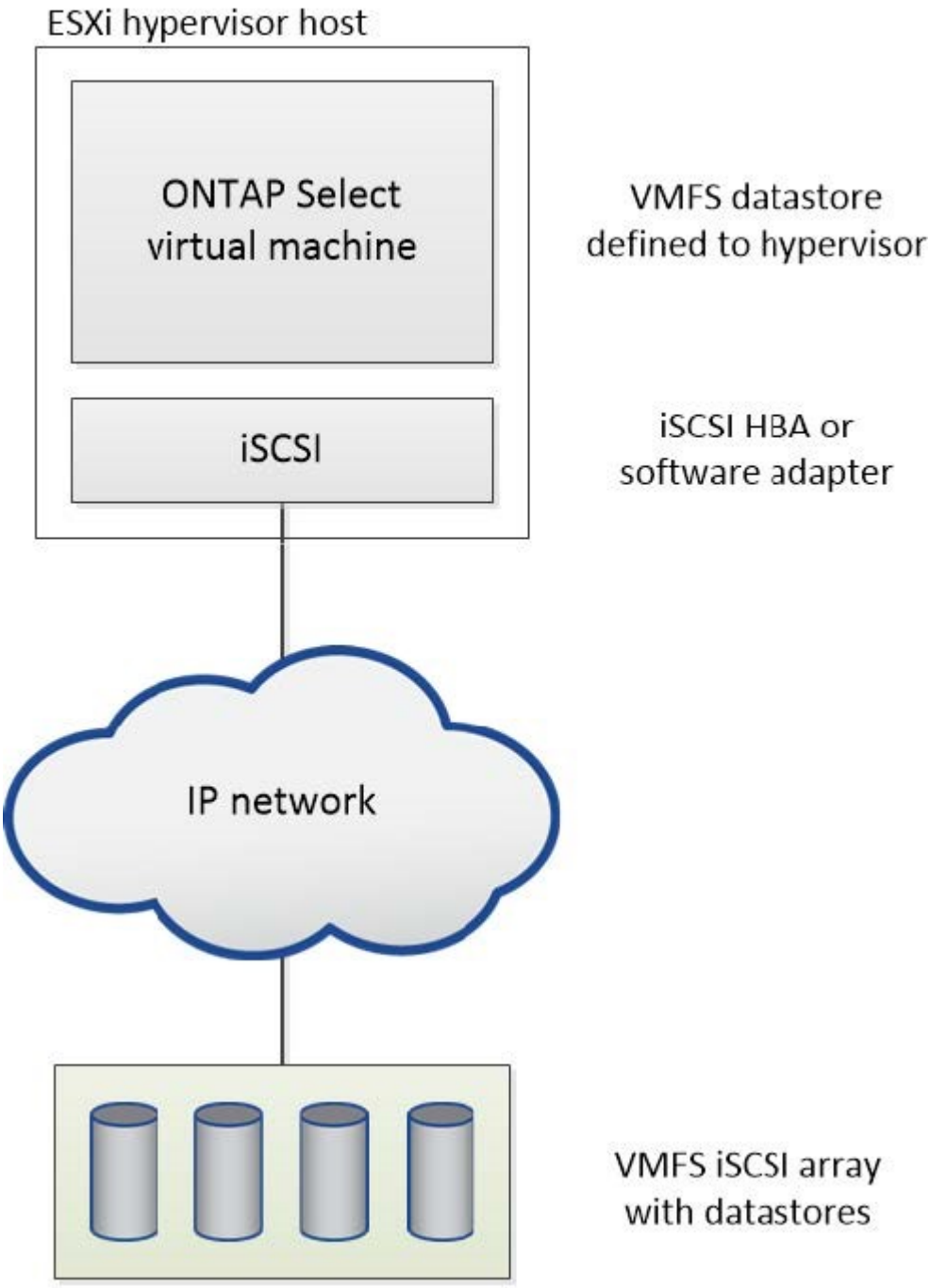
每個 ESXi 主機可以擁有一個或多個本機 VMFS 資料儲存區。通常情況下，這些資料儲存區只能由本機主機存取。但是，VMware vSAN 允許 ESXi 叢集中的每個主機共用叢集中的所有資料儲存區，就像這些資料儲存區是本機儲存區一樣。下圖展示了 vSAN 如何建立一個資料儲存區集區，供 ESXi 叢集中的主機共用。



外部儲存陣列上的 **VMFS** 資料存放區

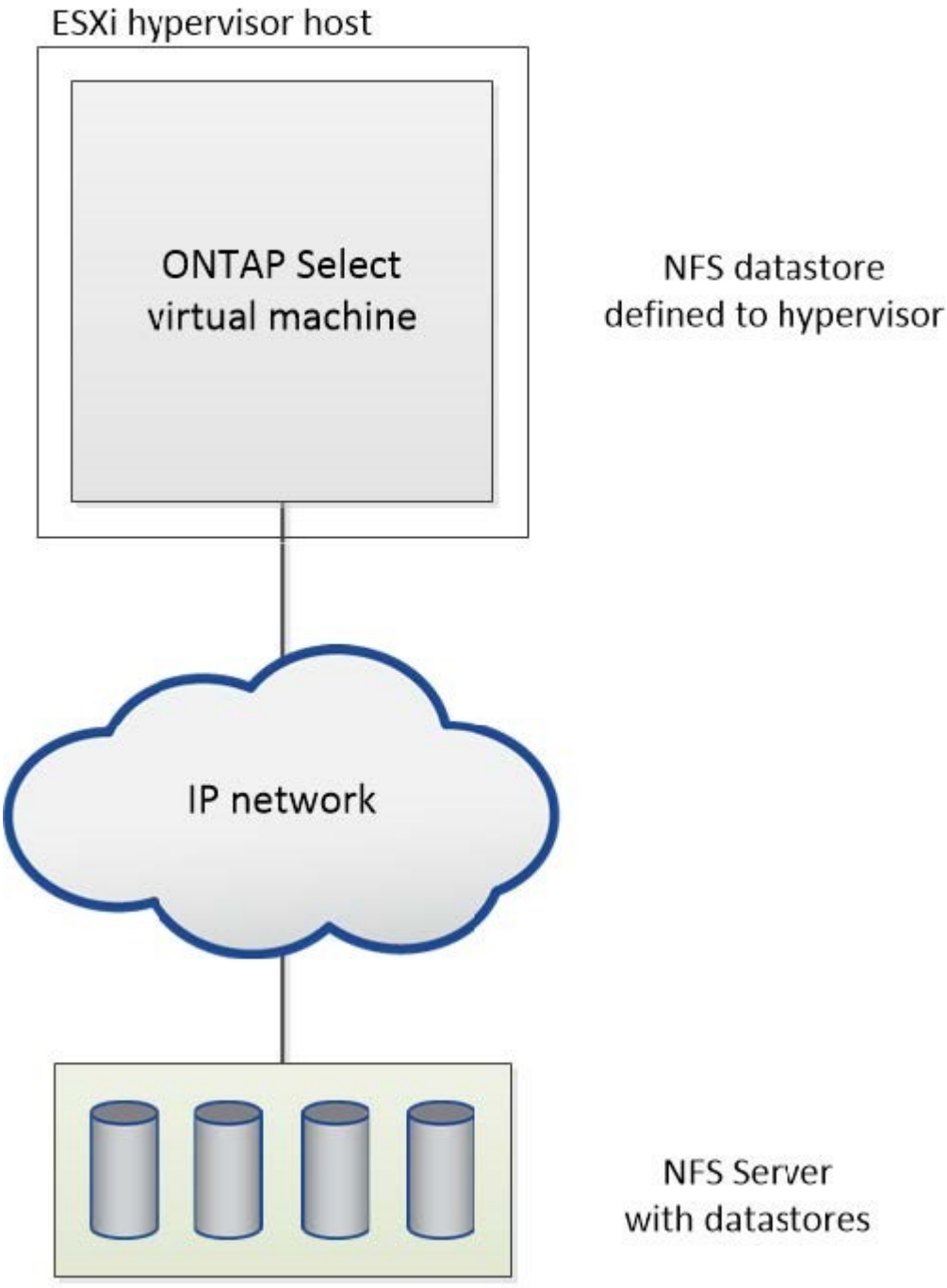
您可以建立位於外部儲存陣列上的 VMFS 資料儲存區。該儲存可透過多種不同的網路協定進行存取。下圖展示了使用 iSCSI 協定存取外部儲存陣列上的 VMFS 資料儲存區。

ONTAP Select 支援 VMware Storage/SAN Compatibility 文件中所述的所有外部儲存陣列，包括 iSCSI、Fiber Channel 和 Fiber Channel over Ethernet。



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 資料儲存區。此儲存空間透過 NFS 網路傳輸協定進行存取。下圖展示了透過 NFS 伺服器應用裝置存取的位於外部儲存上的 NFS 資料儲存區。



適用於具有本機附加儲存設備的 **ONTAP Select** 的硬體 **RAID** 服務

當硬體 RAID 控制器可用時，ONTAP Select 可以將 RAID 服務移轉到此硬體控制器，從而提升寫入效能並防止實體硬碟故障。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。

ONTAP Select 資料 Aggregate 設定為使用 RAID 0，因為實體 RAID 控制器為基礎磁碟機提供 RAID 等量分割。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 控制器之後。大多數通用伺服器都提供多種價位的 RAID 控制器選項，每種選項的功能等級各不相同。我們的目標是盡可能支援所有這些選項，前提是它們滿足控制器的最低要求。

> 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。分離磁碟僅適用於使用軟體 RAID 配置的 ONTAP Select 虛擬機器。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 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份單元（BBU）或快閃記憶體備份寫入快取（FBWC），並支援 12Gbps 的處理量。
- RAID 控制器必須支援能夠承受至少一次或兩次磁碟故障的模式（RAID 5 和 RAID 6）。
- 必須將磁碟機快取設為停用。
- 寫入原則必須設定為回寫模式，並在 BBU 或 Flash 故障時回退到直寫模式。
- 讀取的 I/O 原則必須設為快取。

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 硬碟和 SSD，使用最多 24 個硬碟的 RAID 群組可以讓 ONTAP 充分利用將傳入讀取請求分散到更多磁碟上的優勢。這樣做可以顯著提升效能。對於 SAS/SSD 組態，我們針對單一 LUN 與多 LUN 組態進行了效能測試。結果顯示兩者之間沒有顯著差異,因此,為了簡化起見,NetApp 建議建立滿足您組態需求的最少數量 LUN。

NL-SAS 和 SATA 磁碟機需要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 群組大小不應超過 12 個磁碟機。NetApp 也建議每個 RAID 群組使用一個備用磁碟；不過，也可以使用所有 RAID 群組的全域備用磁碟。例如，每三個 RAID 群組可以使用兩個備用磁碟，每個 RAID 群組包含 8 到 12 個磁碟機。

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

RAID 模式

許多 RAID 控制器支援多達三種作業模式，每種模式都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：

- 直寫式。所有傳入的 I/O 要求都會寫入 RAID 控制器快取、然後立即清除至磁碟、再將要求確認回主機。
- Writearound。所有傳入的 I/O 請求都直接寫入磁碟，繞過 RAID 控制器快取。
- 回寫模式。所有傳入的 I/O 要求都會直接寫入控制器快取，並立即確認回傳給主機。資料區塊會使用控制器以非同步方式排清至磁碟。

寫回模式提供最短的資料路徑，I/O 確認在資料塊進入快取後立即發生。此模式為混合讀寫工作負載提供最低的延遲和最高的吞吐量。但是，如果沒有 BBU 或非揮發性快閃技術，使用者在該模式下運作時，如果系統發生斷電，則存在資料遺失的風險。

ONTAP Select 需要配備電池備援或快閃記憶體單元；因此，我們可以確保在發生此類故障時，快取區塊會刷新到磁碟。正因如此，RAID 控制器必須配置為回寫模式。

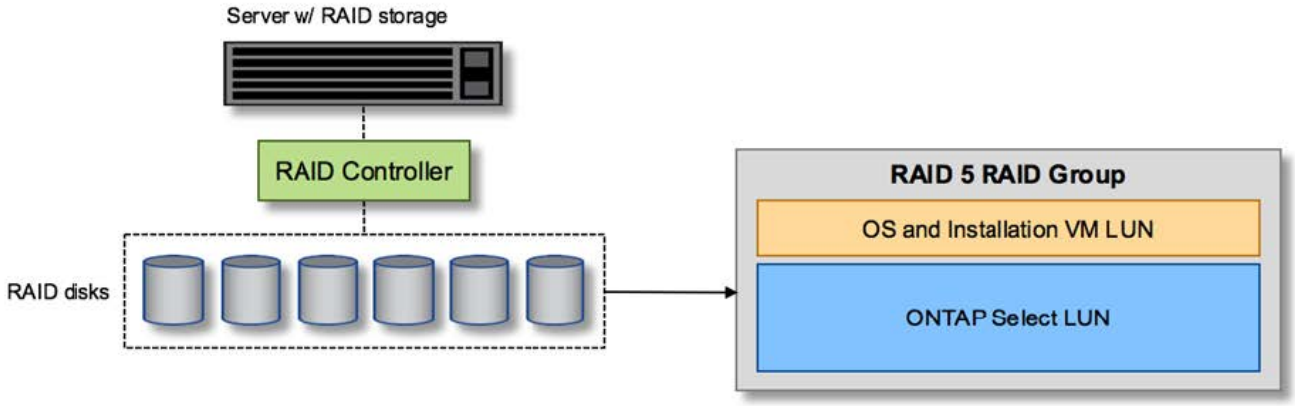
ONTAP Select 與作業系統之間共用的本機磁碟

最常見的伺服器組態是所有本機連接的主軸都位於單一 RAID 控制器之後。您至少應該配置兩個 LUN：一個用於 Hypervisor、一個用於 ONTAP Select VM。

例如，考慮一台配備六個內部磁碟機和一個 Smart Array P420i RAID 控制器的 HP DL380 g8 伺服器。所有內部磁碟機都由該 RAID 控制器管理，系統中沒有其他儲存設備。

下圖顯示此配置樣式。在此範例中，系統上沒有其他儲存設備；因此，Hypervisor 必須與 ONTAP Select 節點共用儲存設備。

僅使用 **RAID** 管理磁碟的伺服器 **LUN** 組態



將作業系統 LUN 設定到與 ONTAP Select 相同的 RAID 群組中，可使虛擬機器管理程式作業系統（以及任何也從該儲存配置的用戶端虛擬機器）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。

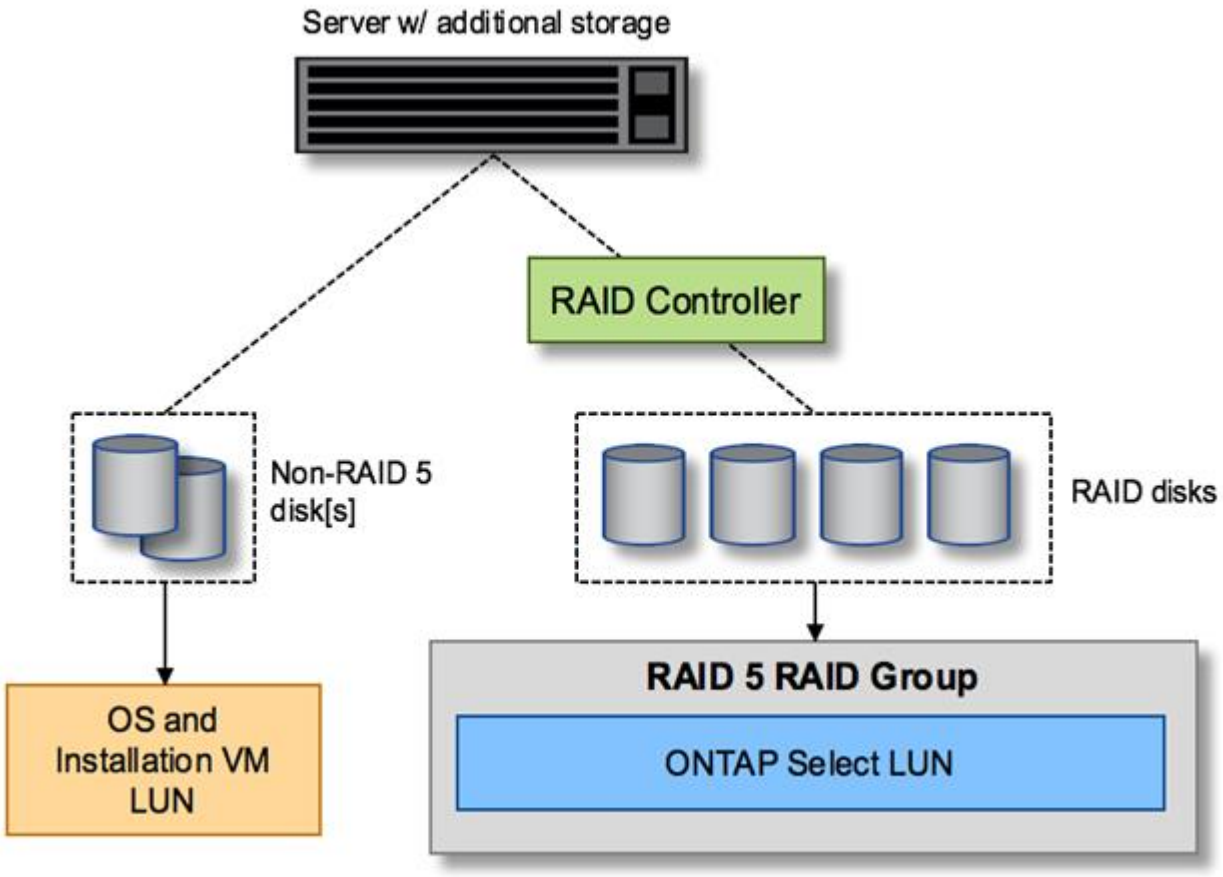
本機磁碟在 **ONTAP Select** 和 **OS** 之間分配

伺服器廠商提供的另一種可能組態是使用多個 RAID 或磁碟控制器來設定系統。在此組態中,一組磁碟由一個磁碟控制器管理,該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。

在這種配置方式下，位於 RAID 控制器後方、可提供 RAID 5/6 服務的磁碟組應僅供 ONTAP Select VM 使用。根據所管理的總儲存容量，您應將磁碟組配置為一個或多個 RAID 群組以及一個或多個 LUN。這些 LUN 將用於建立一個或多個資料存放區，所有資料存放區均受 RAID 控制器保護。

第一組磁碟保留給 Hypervisor 作業系統和任何未使用 ONTAP 儲存設備的用戶端 VM，如下圖所示。

混合 **RAID**/非 **RAID** 系統上的伺服器 **LUN** 組態



多個 **LUN**

在兩種情況下，單一 RAID 群組 / 單一 LUN 組態必須變更。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（個別檔案系統範圍大小上限或儲存資源池大小總計上限）。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

某些版本的 ESXi 上資料存放區的最大大小為 64TB 。

如果伺服器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，一個重要的考慮因素是確保這些 LUN 具有相似且一致的效能。如果所有 LUN 都用於單一 ONTAP Aggregate 中，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特徵，我們強烈建議將這些 LUN 隔離到單獨的 ONTAP Aggregate 中。

可以使用多個檔案系統擴充區來建立單一資料存放區，其大小不超過資料存放區的最大容量。為了限制需要 ONTAP Select 授權的容量，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（因此也僅需要授權）資料存放區中的部分空間。

或者，您可以先在單一 LUN 上建立一個資料儲存區。當需要更多空間（需要更大的 ONTAP Select 容量授權）時，可以將該空間作為範圍新增至同一個資料儲存區，直到達到資料儲存區的最大容量。達到最大容量後，可以建立新的資料儲存區並將其新增至 ONTAP Select。兩種容量擴充作業均受支援，並且可以透過 ONTAP Deploy 的 storage-add 功能實現。每個 ONTAP Select 節點最多可設定為支援 400TB 的儲存設備。從多個資料儲存區

配置容量需要分兩個步驟進行。

初始叢集建立操作可用於建立 ONTAP Select 叢集，該叢集可使用初始資料存放區中的部分或全部空間。第二步是使用其他資料存放區執行一次或多次容量擴充操作，直到達到所需的總容量。此功能在相關章節中有詳細介紹"增加儲存容量"。

> VMFS 開銷不為零（請參閱 VMware KB 1001618），嘗試使用資料存放區報告的整個可用空間會導致叢集建立作業期間出現虛假錯誤。

每個資料儲存中都留有 2% 的緩衝區未使用。由於 ONTAP Select 不使用這部分空間，因此無需容量授權。只要未指定容量上限，ONTAP Deploy 就會自動計算緩衝區的確切大小（以 GB 為單位）。如果指定了容量上限，則首先強制執行該大小。如果容量上限的大小小於緩衝區大小，則叢集建立將會失敗，並顯示錯誤訊息，其中會指定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 的儲存 vMotion 操作的目標。

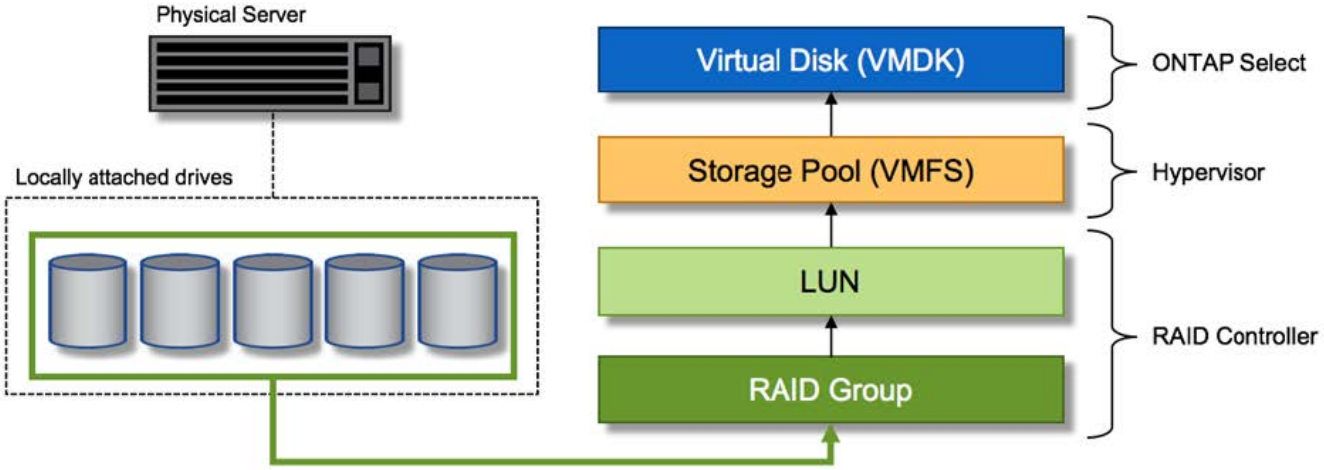
VMware 不支援從 VMFS 5 直接升級到 VMFS 6。因此，儲存 vMotion 是唯一允許虛擬機器從 VMFS 5 資料儲存遷移到 VMFS 6 資料儲存的機制。不過，ONTAP Select 和 ONTAP Deploy 對儲存 vMotion 的支援已擴展至涵蓋 VMFS 5 到 VMFS 6 遷移之外的其他場景。

ONTAP Select 虛擬磁碟

ONTAP Select 的核心在於向 ONTAP 提供一組從一個或多個儲存池配置的虛擬磁碟。ONTAP 將這些虛擬磁碟視為實體磁碟，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，重點突出了實體 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 組態是在伺服器的 RAID 控制器軟體中進行。使用 VSAN 或外部陣列時不需要此組態。
- 儲存資源池組態是在 Hypervisor 內進行。
- 虛擬磁碟由各個 VM 建立和擁有；在本例中，由 ONTAP Select 建立和擁有。

虛擬磁碟到實體磁碟對應



虛擬磁碟資源配置

為了提供更流暢的使用者體驗，ONTAP Select 管理工具 ONTAP Deploy 會自動從關聯的儲存池配置虛擬磁碟，並將其附加到 ONTAP Select VM。此操作在初始設定和新增儲存操作期間都會自動執行。如果 ONTAP Select 節點是 HA 配對的一部分，則虛擬磁碟會自動指派給本機儲存池和鏡像儲存池。

ONTAP Select 將底層附加儲存分割成大小相等的虛擬磁碟，每個虛擬磁碟不超過 16TB。如果 ONTAP Select 節點是 HA 對的一部分，則每個叢集節點上至少建立兩個虛擬磁碟，並將其指派給本機 plex 和鏡像 plex，以便在鏡像 Aggregate 中使用。

例如，ONTAP Select 可以指派一個 31TB 的資料儲存區或 LUN（虛擬機器部署完畢並配置好系統磁碟和根磁碟後剩餘的空間）。然後，系統會建立四個約 7.75TB 的虛擬磁碟，並將其指派給對應的 ONTAP 本機 plex 和鏡射 plex。

> 在 ONTAP Select 虛擬機器中新增容量可能會導致產生大小不同的 VMDK。詳情請參閱相關章節"增加儲存容量"。與 FAS 系統不同，不同大小的 VMDK 可以存在於同一個集合體中。ONTAP Select 在這些 VMDK 之間使用 RAID 0 條帶化，因此無論每個 VMDK 的大小如何，都能充分利用其所有空間。

虛擬化 **NVRAM**

NetApp FAS 系統通常配備一塊實體 NVRAM PCI 卡，這是一種高效能卡，內建非揮發性快閃記憶體。該卡使 ONTAP 能夠立即向用戶端確認收到的寫入操作，從而顯著提升寫入效能。它還可以安排將修改後的資料區塊移回速度較慢的儲存媒體，這一過程稱為 destaging。

普通系統通常不配備此類設備。因此，此 NVRAM 卡的功能已虛擬化，並放置在 ONTAP Select 系統開機磁碟的一個分割區中。正因如此，執行個體的系統虛擬磁碟的放置位置至關重要。這也是為什麼該產品要求本地附加儲存組態必須配備具有彈性快取的實體 RAID 控制器。

NVRAM 被放置在獨立的 VMDK 中。將 NVRAM 分割到單獨的 VMDK 中，使得 ONTAP Select VM 能夠使用 vNVMe 驅動程式與其 NVRAM VMDK 進行通訊。此外，這也要求 ONTAP Select VM 使用硬體版本 13，該版本與 ESXi 8.0 及更高版本相容。

資料路徑詳解：**NVRAM** 與 **RAID** 控制器

要了解虛擬化 NVRAM 系統分區與 RAID 控制器之間的交互，最好的方法是分析寫入要求進入系統時所經過的

資料路徑。

傳送到 ONTAP Select VM 的寫入請求指向 VM 的 NVRAM 分割區。在虛擬化層，此分割區位於 ONTAP Select 系統磁碟（一個附加到 ONTAP Select VM 的 VMDK）中。在實體層，這些請求會像所有指向底層磁碟的區塊變更一樣，快取在本機 RAID 控制器中。然後，寫入操作會從這裡確認並傳回給主機。

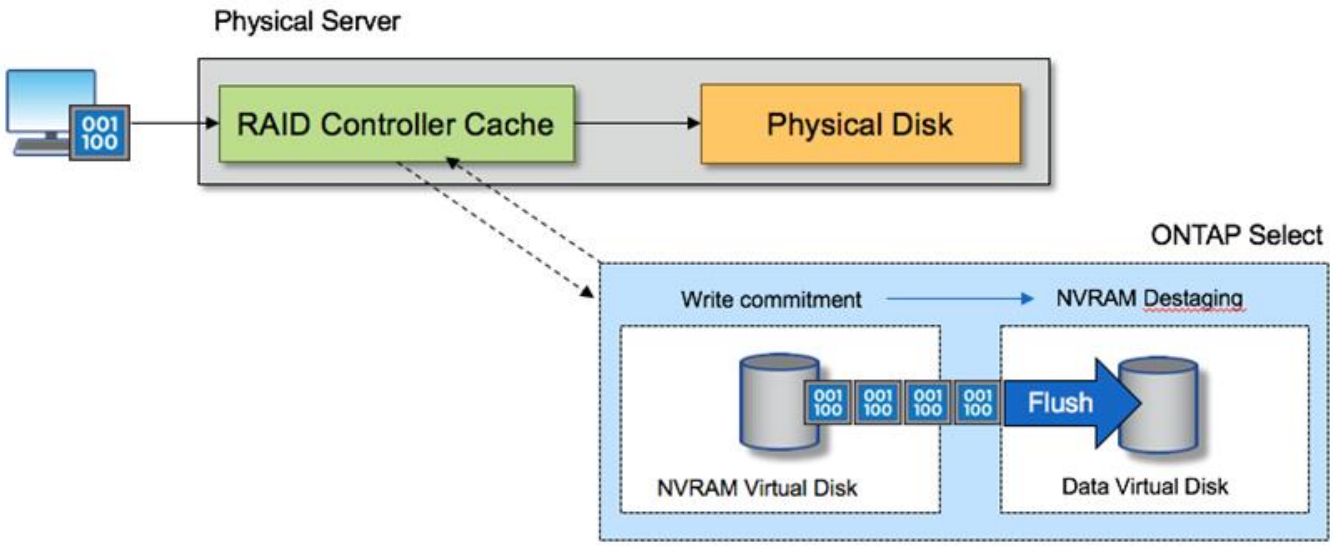
此時，從物理層面來說，該資料塊位於 RAID 控制器快取中，等待刷新到磁碟。從邏輯層面來說，該資料塊位於 NVRAM 中，等待移出到對應的使用者資料磁碟。

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 NVRAM 分割區的寫入作業會自動快取，並定期清除至實體儲存媒體。這不應與定期將 NVRAM 內容清除回 ONTAP 資料磁碟混淆。這兩個事件互不相關，發生的時間和頻率也不同。

下圖顯示了傳入寫入操作的 I/O 路徑。它突出了實體層（由 RAID 控制器快取和磁碟表示）和虛擬層（由虛擬機器的 NVRAM 和資料虛擬磁碟表示）之間的差異。

儘管 NVRAM VMDK 上更改的區塊會被快取到本機 RAID 控制器快取中、但該快取並不了解 VM 架構或其虛擬磁碟。它儲存系統上所有已更改的區塊、而 NVRAM 只是其中的一部分。這包括繫結至 Hypervisor 的寫入要求、如果是從相同的備用主軸配置。

傳入 **ONTAP Select VM** 的寫入操作



NVRAM 分割區位於獨立的 VMDK 檔案中。此 VMDK 檔案透過 ESXi 8.0 或更高版本中提供的 vNVME 驅動程式進行掛載。對於使用軟體 RAID 的 ONTAP Select 安裝而言，此變更尤其重要，因為此類安裝無法利用 RAID 控制器快取。

ONTAP Select 軟體 RAID 組態服務，適用於本機附加儲存設備

軟體 RAID 是 ONTAP 軟體堆疊中實作的 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。此 RAID 層執行磁碟機同位校驗計算，並為 ONTAP Select 節點提供單一磁碟機故障的保護。

除了硬體 RAID 組態之外，ONTAP Select 也提供軟體 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適用。軟體 RAID 擴充了可用的部署選項，以納入此類環境。若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 授權取得。
- 它僅支援 SSD 或 NVMe（需要 Premium XL 授權）磁碟機用於 ONTAP 根磁碟和資料磁碟。
- ONTAP Select VM 開機分割區需要單獨的系統磁碟。
  - 選擇單獨的磁碟（SSD 或 NVMe 磁碟機）來建立系統磁碟（多節點設定中的 NVRAM、Boot/CF 卡、Coredump 和 Mediator）的資料存放區。



- 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 中用於服務各種項目（例如叢集、開機等）的虛擬磁碟（VMDK）。
  - 從主機角度來看，服務磁碟位於單一實體磁碟（統稱為服務/系統實體磁碟）上。此實體磁碟必須包含 DAS 資料儲存區。ONTAP Deploy 會在叢集部署期間為 ONTAP Select VM 建立這些服務磁碟。
- 無法將 ONTAP Select 系統磁碟進一步分散到多個資料儲存區或多個實體磁碟機。
- 硬體 RAID 並未被棄用。

本機附加儲存設備的軟體 **RAID** 組態

使用軟體 RAID 時，理想情況下不需要硬體 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：

- 您必須停用硬體 RAID 控制器，以便磁碟可以直接呈現給系統（JBOD）。通常可以在 RAID 控制器 BIOS 中進行此變更。
- 或者，硬體 RAID 控制器應設定為 SAS HBA 模式。例如，某些 BIOS 設定除了 RAID 模式外，還允許使用「AHCI」模式，您可以選擇啟用該模式來啟用 JBOD 模式。這樣可以實現直通，使實體磁碟機能夠直接被主機識別。

根據控制器支援的最大磁碟機數量，可能需要額外的控制器。使用 SAS HBA 模式時，請確保 I/O 控制器（SAS HBA）的最低速度為 6Gbps。然而，NetApp 建議使用 12Gbps 的速度。

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，可以人為地啟用磁碟直通，但影響可能不理想。支援的實體磁碟（僅限 SSD）大小介於 200GB 至 16TB 之間。



管理員需要追蹤 ONTAP Select VM 正在使用的磁碟機,並防止在主機上意外使用這些磁碟機。

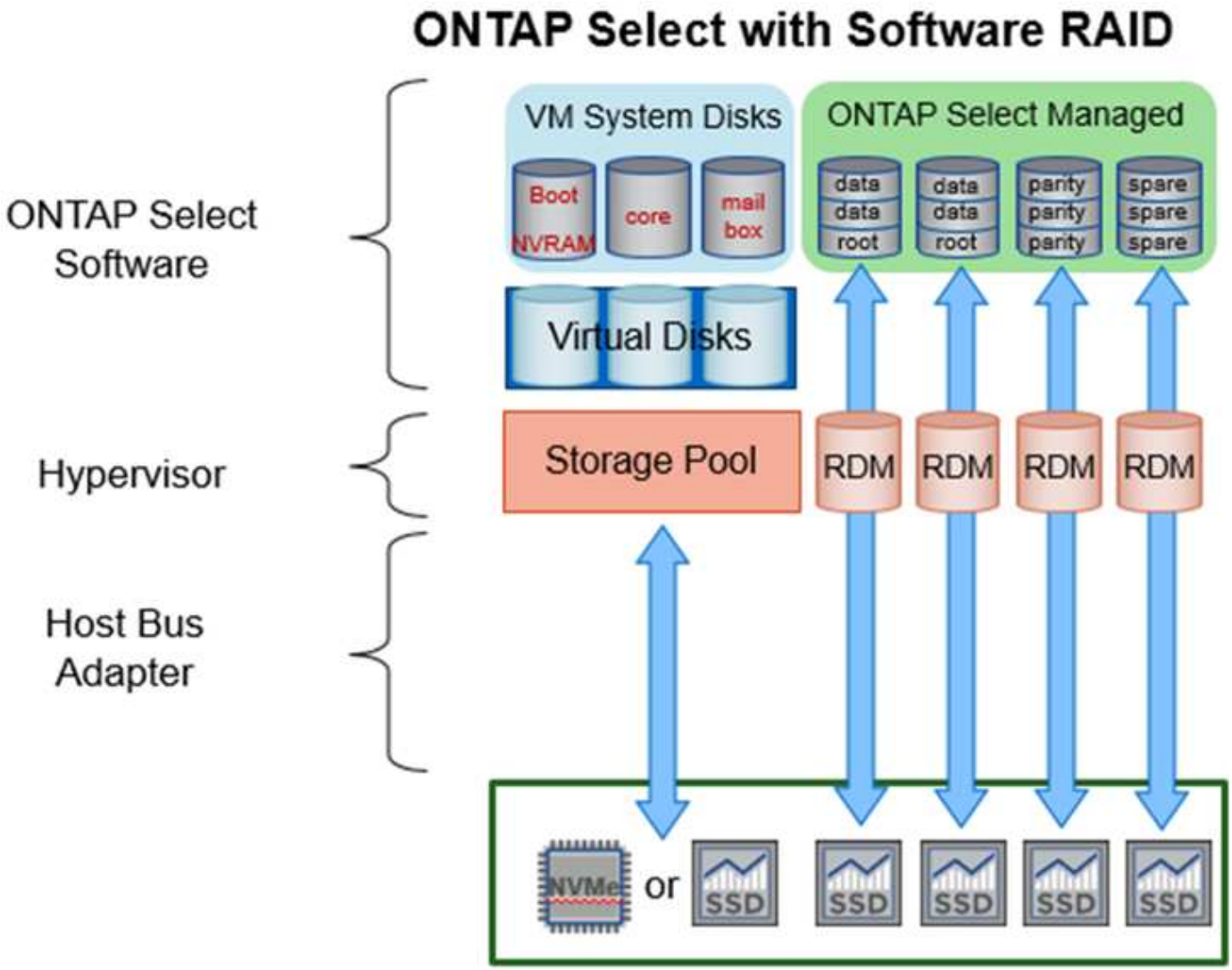
ONTAP Select 虛擬和實體磁碟

對於採用硬體 RAID 控制器的組態、實體磁碟備援由 RAID 控制器提供。ONTAP Select 會提供一個或多個 VMDK、ONTAP 管理員可從中設定資料集合體。這些 VMDK 會以 RAID 0 格式進行等量分割、因為使用 ONTAP 軟體 RAID 是多餘的、效率不佳、而且由於硬體層級提供的恢復能力而無效。此外、用於系統磁碟的 VMDK 與用於儲存使用者資料的 VMDK 位於相同的資料存放區中。

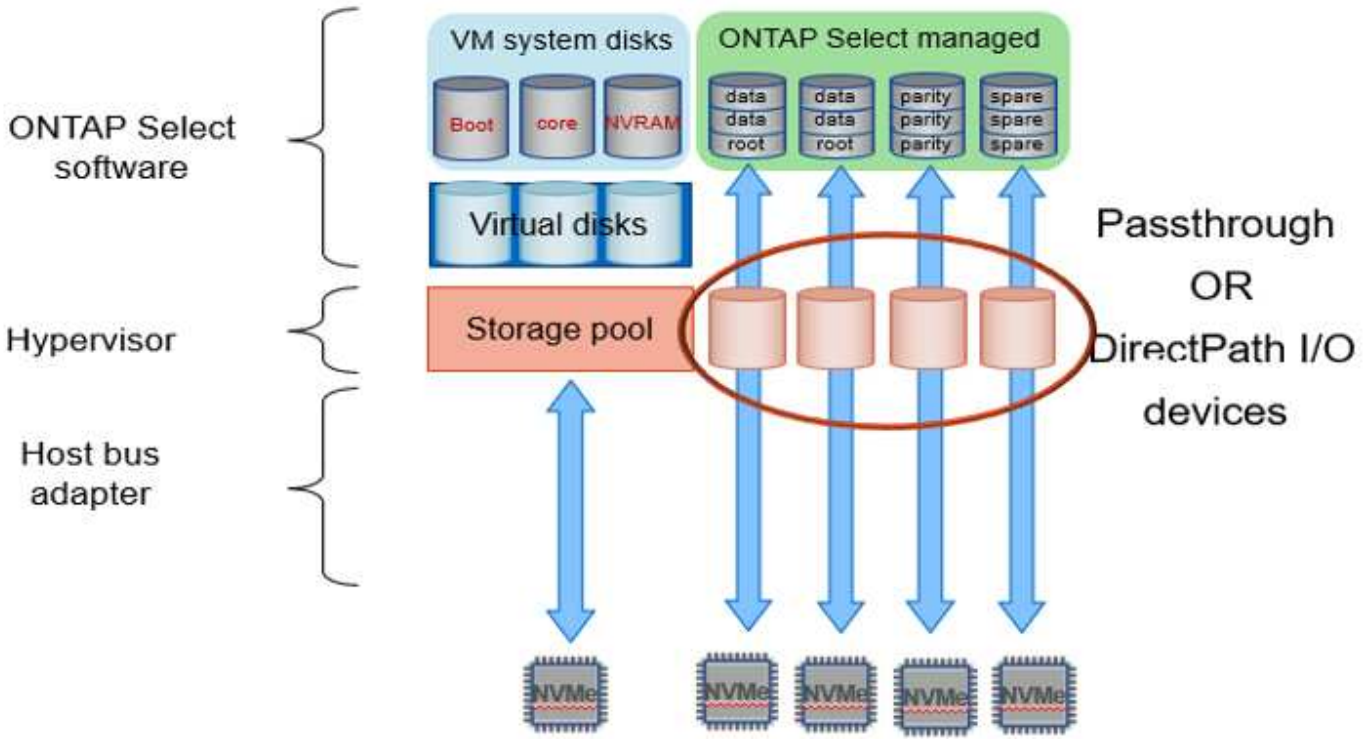
使用軟體 RAID 時，ONTAP Deploy 會向 ONTAP Select 提供一組 VMDK 和實體磁碟，以及 SSD 的裸設備對映 [RDM] 和 NVMe 的直通或 DirectPath IO 設備。

以下圖表更詳細地顯示了這種關係，突顯了 ONTAP Select VM 內部使用的虛擬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**：使用虛擬磁碟和 **RDM**



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。



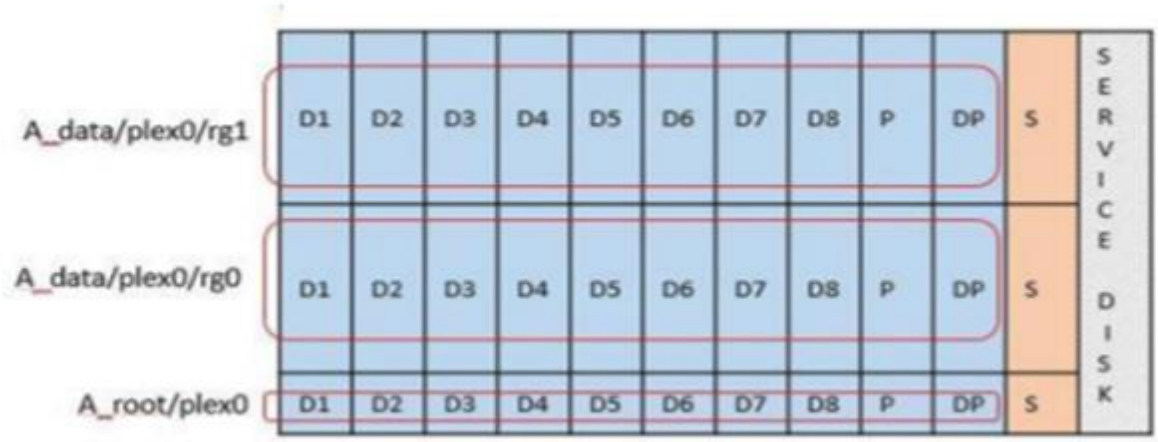
系統磁碟（VMDK）位於相同資料儲存區和相同實體磁碟上。虛擬 NVRAM 磁碟需要快速且耐用的媒體。因此，僅支援 NVMe 和 SSD 類型的資料儲存區。如果使用 NVMe 驅動器儲存資料，出於效能考慮，系統磁碟也應該是 NVMe 裝置。在全 NVMe 組態中，INTEL Optane 卡是系統磁碟的理想選擇。

目前版本無法將 ONTAP Select 系統磁碟進一步分散到多個資料儲存或多個實體磁碟機。

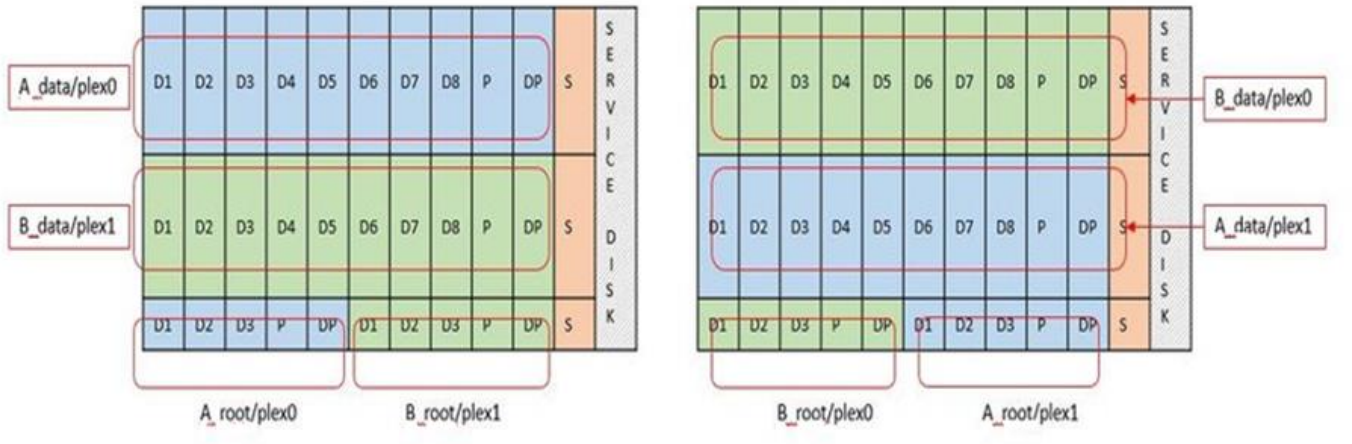
每個資料磁碟被分成三個部分：一個小的根分割區（條帶）和兩個大小相等的分割區，從而在 ONTAP Select VM 中建立兩個資料磁碟。分割區採用根資料資料（RD2）架構，如下圖所示，分別展示了單節點叢集和高可用性（HA）配對中單一節點的分割區情況。

`P` 表示奇偶校驗磁碟機、 `DP` 表示雙奇偶校驗磁碟機、 `S` 表示備用磁碟機。

單節點叢集的 **RDD** 磁碟分割



多節點叢集（**HA** 配對）的 **RDD** 磁碟分割



ONTAP 軟體 RAID 支援下列 RAID 類型：RAID 4、RAID-DP 和 RAID-TEC。這些 RAID 結構與 FAS 和 AFF 平台使用的 RAID 結構相同。對於根分割區配置，ONTAP Select 僅支援 RAID 4 和 RAID-DP。當使用 RAID-TEC 進行資料聚合時，整體保護等級為 RAID-DP。ONTAP Select HA 採用無共享架構，將每個節點的組態複製到其他節點。這意味著每個節點必須儲存其根分割區以及其對等節點根分割區的副本。資料磁碟只有一個根分割區。這意味著資料磁碟的最小數量取決於 ONTAP Select 節點是否屬於 HA 配對。

對於單節點叢集，所有資料分割區都用於儲存本機（作用中）資料。對於屬於 HA 配對一部分的節點，一個資料分割區用於儲存該節點的本機（作用中）資料，第二個資料分割區用於鏡射來自 HA 對等節點的作用中資料。

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

ESXi 和 KVM 虛擬機器管理程式不支援將 NVMe 磁碟作為裸裝置對應（RDM）使用。若要讓 ONTAP Select 直接控制 NVMe 磁碟、您必須在 ESXi 或 KVM 中將這些磁碟機配置為直通裝置。將 NVMe 裝置配置為直通裝置需要伺服器 BIOS 的支援、而且您可能需要重新啟動主機。此外、每個主機可指派的直通裝置數量有限制、具體數量可能因平台而異。不過、ONTAP Deploy 將每個 ONTAP Select 節點的 NVMe 裝置數量限制為 14 個。這表示 NVMe 組態以犧牲總容量為代價、提供了非常高的 IOPS 密度（IOPS/TB）。或者、如果您需要具有更大儲存容量的高效能組態、建議的組態是使用較大的 ONTAP Select VM 大小、Intel Optane 卡作為系統磁碟、以及適量的 SSD 磁碟機進行資料儲存。

若要充分發揮 NVMe 的效能，請考慮使用較大的 ONTAP Select VM 容量。

直通設備和 RDM 之間還有另一個區別。RDM 可以對應到正在執行的虛擬機器 (VM)。而直通設備則需要重新啟動 VM。這表示任何 NVMe 驅動器更換或容量擴充（新增驅動器）操作都需要重新啟動 ONTAP Select VM。驅動器更換和容量擴充（新增驅動器）操作由 ONTAP Deploy 中的工作流程驅動。ONTAP Deploy 管理單節點叢集的 ONTAP Select 重新啟動以及高可用性 (HA) 配對的容錯移轉/容錯回復。但是，需要注意的是，使用 SSD 資料驅動器（無需 ONTAP Select 重新啟動/容錯移轉）和使用 NVMe 資料驅動器（需要 ONTAP Select 重新啟動/容錯移轉）之間存在差異。

實體和虛擬磁碟資源配置

為了提供更流暢的使用者體驗，ONTAP Deploy 會自動從指定的資料儲存區（實體系統磁碟）配置系統（虛擬）磁碟，並將其附加到 ONTAP Select VM。此作業會在初始設定期間自動執行，以便 ONTAP Select VM 能夠開機。RDM 會被分割，並自動建置根 Aggregate。如果 ONTAP Select 節點是 HA 配對的一部分，則資料分割區會自動指派給本機儲存資源池和鏡射儲存資源池。此指派會在叢集建立作業和儲存設備新增作業期間自動執行。

由於 ONTAP Select VM 上的資料磁碟與底層實體磁碟相關聯，因此建立具有更多實體磁碟的組態會對效能產生影響。

根 Aggregate 的 RAID 群組類型取決於可用磁碟的數量。ONTAP Deploy 會選擇適當的 RAID 群組類型。如果節點已指派足夠的磁碟，則使用 RAID-DP；否則，將建立 RAID-4 根 Aggregate。

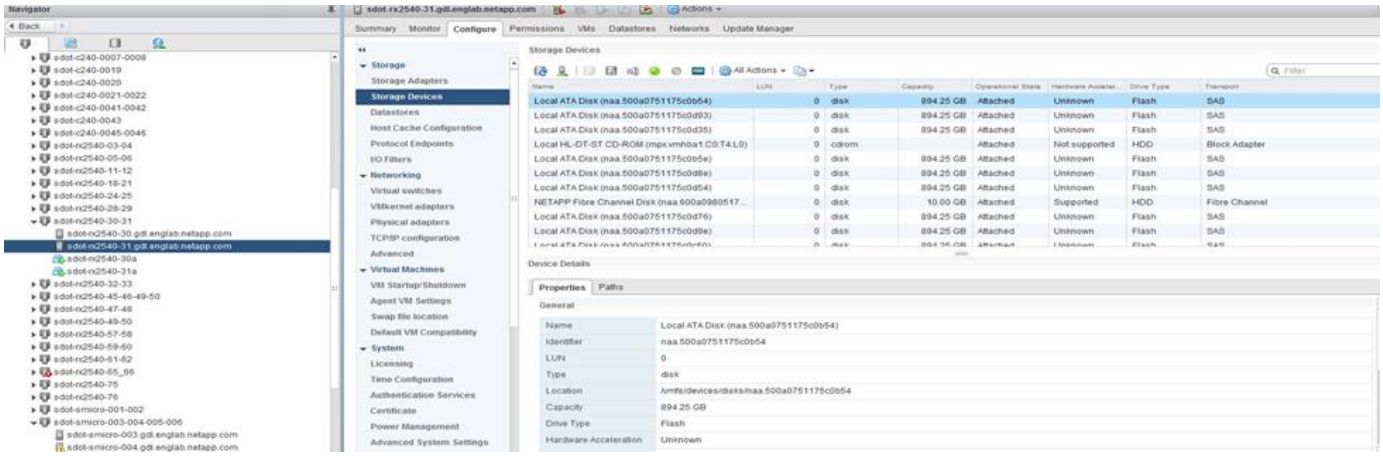
使用軟體 RAID 為 ONTAP Select VM 新增容量時，管理員必須考慮實體磁碟機的大小和所需的磁碟機數量。如需詳細資訊，請參閱 "增加儲存容量"。

與 FAS 和 AFF 系統類似，您只能在現有 RAID 群組中新增容量等於或大於現有 RAID 群組的磁碟機。容量較大的磁碟機會調整為適當大小。如果您要建立新的 RAID 群組，則新 RAID 群組的大小應與現有 RAID 群組的大小相符，以確保整體 Aggregate 效能不會降低。

將 **ONTAP Select** 磁碟與對應的 **ESXi** 或 **KVM** 磁碟相匹配

ONTAP Select 磁碟通常標示為 NET x.y。您可以使用下列 ONTAP 指令取得磁碟 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 或 KVM Shell 中、您可以輸入下列命令來閃爍指定實體磁碟（由其 naa.unique-id 識別）的 LED。

ESXi

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

KVM

```
cat /sys/block/<block_device_id>/device/wwid
```

使用軟體 **RAID** 時出現多個磁碟機故障

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 可以承受一個磁碟故障、RAID-DP Aggregate 可以承受兩個磁碟故障，而 RAID-TEC Aggregate 可以承受三個磁碟故障。

如果故障磁碟數量少於 RAID 類型支援的最大故障數量，且有備用磁碟可用，則重建程序會自動啟動。如果沒有備用磁碟可用，則 Aggregate 會以降級狀態提供資料服務，直到新增備用磁碟為止。

如果故障磁碟的數量超過 RAID 類型支援的最大故障數量、則本機 plex 會標示為故障、且 Aggregate 狀態會降級。資料是從位於 HA 合作夥伴上的第二個 plex 提供。這表示節點 1 的任何 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 plex 也故障、則 Aggregate 會標示為故障、且資料無法使用。

若要還原正確的資料鏡像，必須刪除並重新建立故障的 plex。請注意，當多磁碟故障導致資料聚合降級時，根聚合也會降級。ONTAP Select 使用根-資料-資料（RDD）分區方案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。

在以下範例輸出中，失敗的 plex 會被刪除並重新建立：

```
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..
```

若要測試或模擬一個或多個磁碟機故障,請使用 `storage disk fail -disk NET-x.y -immediate`命令。如果系統中有備援磁碟機,`Aggregate` 將開始重建。您可以使用命令 `storage aggregate show`檢查重建狀態。您可以使用 `ONTAP Deploy` 移除模擬故障的磁碟機。請注意,`ONTAP` 已將該磁碟機標記為 `Broken`。該磁碟機實際上並未損壞,可以使用 ONTAP Deploy 將其重新新增。若要清除 Broken 標籤,請在 ONTAP Select CLI 中輸入以下命令:

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

最後一條命令的輸出應該為空。

虛擬化 **NVRAM**

NetApp FAS 系統通常配備一塊實體 NVRAM PCI 卡。這塊卡片效能卓越，內建非揮發性快閃記憶體，可大幅提升寫入效能。其原理是使 ONTAP 能夠立即向客戶端確認收到的寫入資料。此外，它還可以安排將修改後的資料區塊移回速度較慢的儲存媒體，這一過程稱為 destaging。

普通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 系統開機磁碟的一個分割區中。正因如此，執行個體的系統虛擬磁碟的放置位置至關重要。

ONTAP Select vSAN 和外部陣列組態

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儲存區。這些組態的基礎基礎架構可提供資料儲存區恢復能力。

最低要求是您使用的 Hypervisor（在支援的 Linux 主機上執行的 VMware ESXi 或 KVM）支援底層組態。如果 Hypervisor 是 ESXi，則它應該列在對應的 VMware HCL 中。

vNAS 架構

vNAS 命名規則適用於所有不使用 DAS 的設定。對於多節點 ONTAP Select 叢集，這包括同一 HA 配對中的兩個 ONTAP Select 節點共享單一資料存放區（包括 vSAN 資料存放區）的架構。節點也可以安裝在來自相同共用外部陣列的不同資料存放區上。這可以提高陣列端的儲存效率，從而減少整個 ONTAP Select HA 配對的整體佔用空間。ONTAP Select vNAS 解決方案的架構與使用本機 RAID 控制器的 ONTAP Select on DAS 非常相似。也就是說，每個 ONTAP Select 節點都會繼續保留其 HA 合作夥伴的資料副本。ONTAP 儲存效率原則的範圍為節點。因此，陣列端的儲存效率更為理想，因為它們可能會套用至來自兩個 ONTAP Select 節點的資料集。

HA 配對中的每個 ONTAP Select 節點也可以使用個別的外部陣列。當使用 ONTAP Select MetroCluster SDS 搭配外部儲存設備時，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，兩個陣列必須為 ONTAP Select VM 提供類似的效能特徵，這一點非常重要。

vNAS 架構與採用硬體 RAID 控制器的本機 DAS

vNAS 架構在邏輯上與配備 DAS 和 RAID 控制器的伺服器架構最為相似。在這兩種情況下、ONTAP Select 都會佔用資料儲存區空間。該資料儲存區空間會被分割成 VMDK、而這些 VMDK 會形成傳統的 ONTAP 資料 Aggregate。ONTAP Deploy 可確保在叢集建立和儲存新增作業期間、VMDK 的大小適當、並指派給正確的 Plex（在 HA 配對的情況下）。

vNAS 和配備 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 假定底層外部陣列能夠提供與配備 RAID 控制器的 DAS 相同的資料持續性和恢復能力。第二個差別，也是更細微的差別，與 NVRAM 的效能有關。

vNAS NVRAM

ONTAP Select NVRAM 是一個 VMDK。這意味著 ONTAP Select 在區塊尋址裝置（VMDK）之上模擬了一個位元組尋址空間（傳統 NVRAM）。然而，NVRAM 的效能對 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 設定，硬體 RAID 控制器快取會做為 NVRAM 快取，因為所有寫入 NVRAM VMDK 的作業都會先託管在 RAID 控制器快取中。

對於 VNAS 架構， ONTAP Deploy 會自動為 ONTAP Select 節點配置一個名為「單一實例資料記錄 (SIDL)」的啟動參數。啟用此啟動參數後， ONTAP Select 將繞過 NVRAM ，直接將資料有效負載寫入資料聚合。NVRAM 僅用於記錄 WRITE 操作變更的資料區塊的位址。此功能的優點在於避免了雙重寫入：一次寫入 NVRAM ，另一次寫入是在 NVRAM 被移除時進行的。此功能僅針對 vNAS 啟用，因為對 RAID 控制器快取的本機寫入所帶來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 Aggregate 層級停用 SIDL 功能：

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

如果關閉 SIDL 功能，寫入效能會受到影響。在停用該 Aggregate 中所有 Volume 上的所有儲存效率原則後，可以重新啟用 SIDL 功能：

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

```
affected aggregate)
```

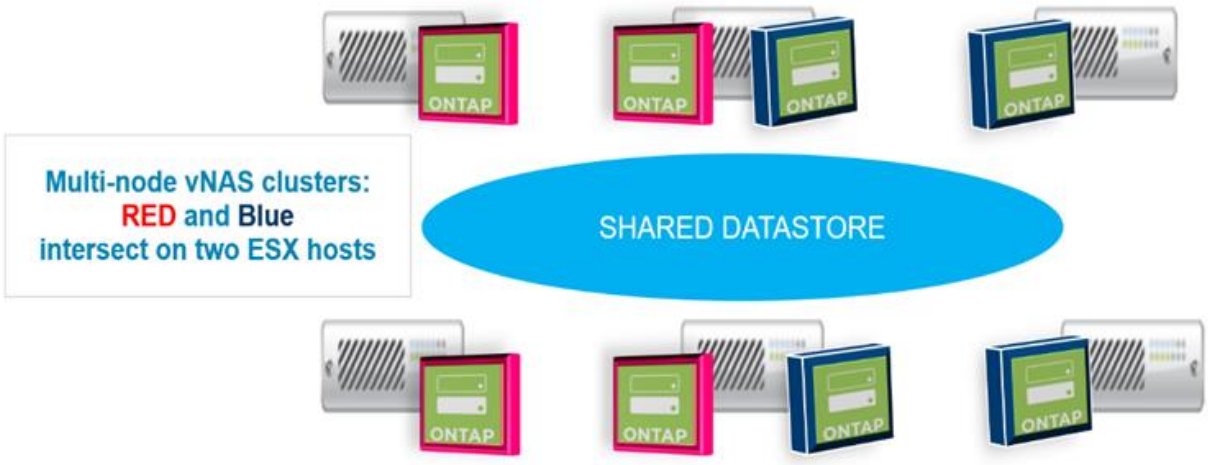
在 **ESXi** 上使用 **vNAS** 時、共置 **ONTAP Select** 節點

ONTAP Select 支援在共用儲存設備上建立多節點 ONTAP Select 叢集。ONTAP Deploy 可在同一個 ESXi 主機上設定多個 ONTAP Select 節點，前提是這些節點不屬於同一個叢集。

> 此組態僅適用於 VNAS 環境（共用資料存放區）。使用 DAS 儲存設備時，不支援每個主機有多個 ONTAP Select 執行個體，因為這些執行個體會競爭相同的硬體 RAID 控制器。

ONTAP Deploy 可確保多節點 VNAS 叢集的初始部署不會將相同叢集中的多個 ONTAP Select 執行個體放置在相同主機上。下圖展示了兩個四節點叢集在兩台主機上交叉部署的正確範例。

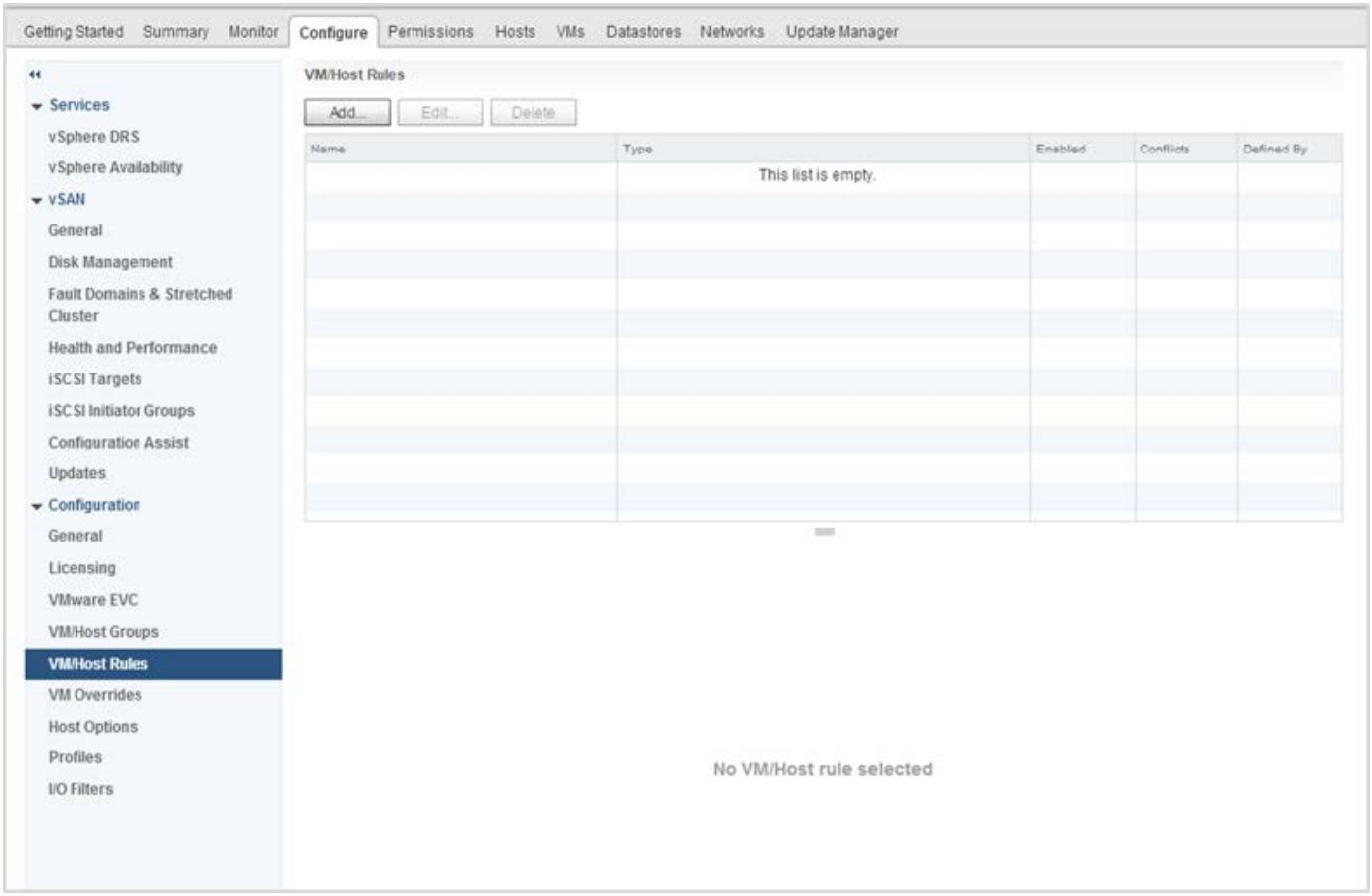
多節點 **VNAS** 叢集的初始部署

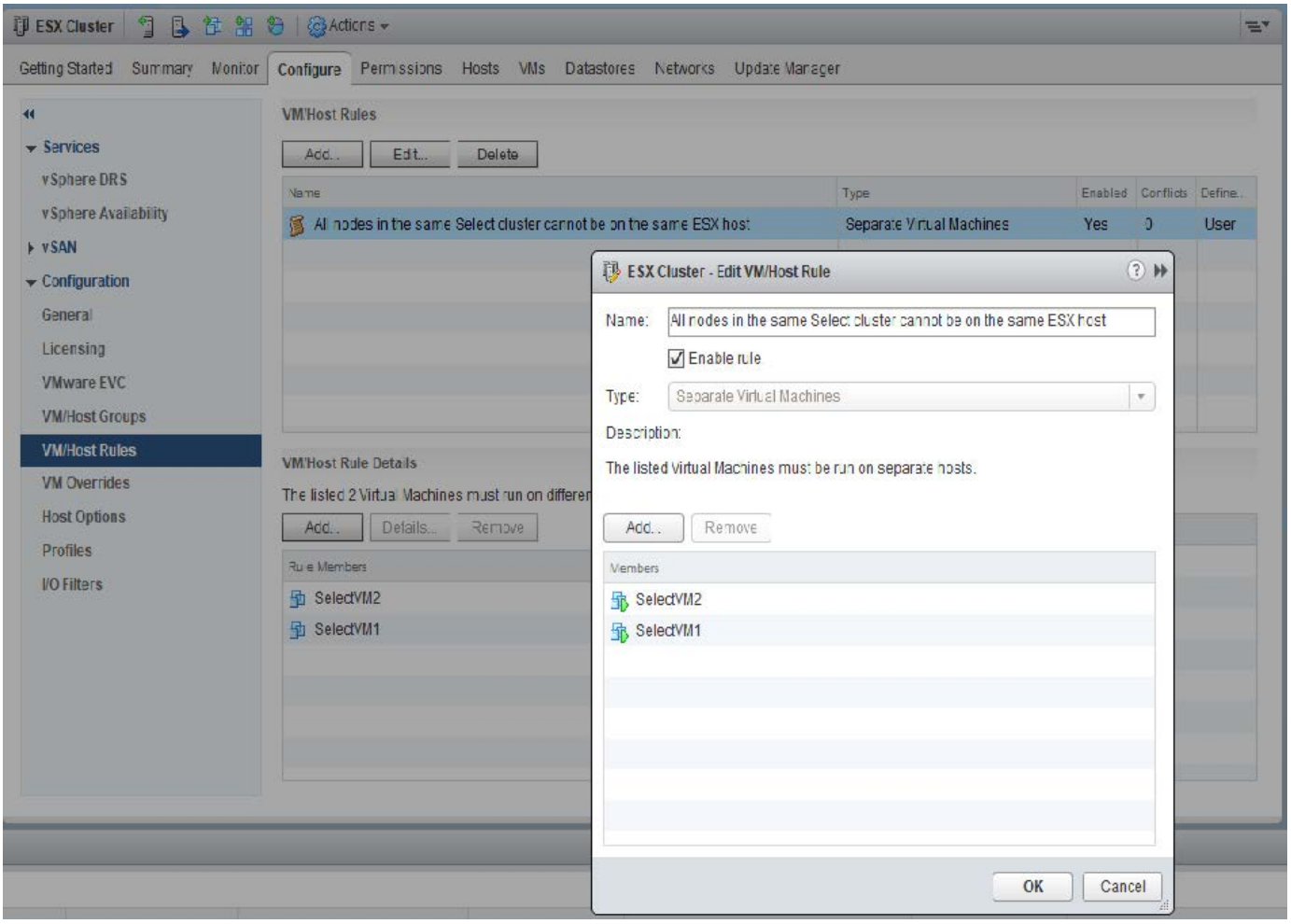


部署後、ONTAP Select 節點可以在主機之間移轉。這可能會導致非最佳且不受支援的組態、使同一叢集中的兩個或多個 ONTAP Select 節點共用相同的基礎主機。NetApp 建議手動建立 VM 反關聯性規則、以便 VMware 自動維護同一叢集節點之間的實體分隔、而不僅僅是同一 HA 配對的節點。

> 反親和性規則要求在 ESXi 叢集上啟用 DRS。

請參考以下範例，了解如何為 ONTAP Select VM 建立反親和性規則。如果 ONTAP Select 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。


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



由於以下原因之一，來自同一 ONTAP Select 叢集的兩個或多個 ONTAP Select 節點可能位於同一 ESXi 主機上：

- 由於 VMware vSphere 授權限制或 DRS 未啟用，因此不會顯示 DRS。
- DRS 反關聯性規則會被略過，因為 VMware HA 作業或系統管理員啟動的 VM 移轉具有優先權。

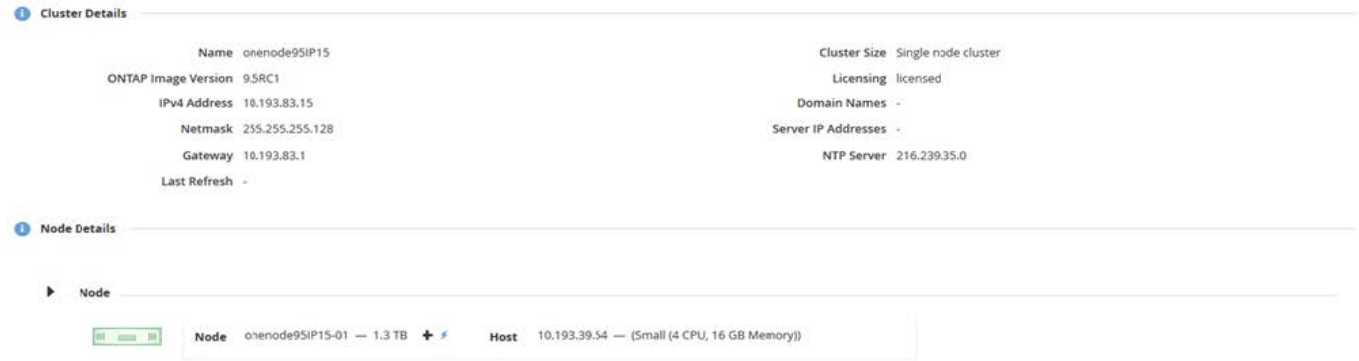
ONTAP Deploy 不會主動監控 ONTAP Select VM 位置。但是、叢集重新整理作業會在 ONTAP Deploy 記錄中反映此不受支援的組態：



關於增加 **ONTAP Select** 儲存容量的考慮因素

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。

在 ONTAP Deploy 中，儲存新增功能是增加管理儲存空間的唯一方法，不支援直接修改 ONTAP Select VM。下圖顯示了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。增加容量需要現有授權能夠涵蓋全部空間（現有空間加新增空間）。如果儲存新增操作導致節點超出其授權容量,則該操作將失敗。應先安裝具有足夠容量的新授權。

如果將額外容量新增至現有的 ONTAP Select Aggregate 中，則新的儲存資源池（資料存放區）的效能設定檔應與現有儲存資源池（資料存放區）的效能設定檔相似。請注意，無法為安裝了類似 AFF 特性（啟用 Flash）的 ONTAP Select 節點新增非 SSD 儲存設備。此外，也不支援混合使用 DAS 和外部儲存設備。

如果在系統中新增本機附加儲存以提供額外的本機（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，如果在相同聚合中新增空間，則應注意確保新 RAID 群組的效能與原始 RAID 群組的效能相似。如果建立新的聚合，並且充分了解新聚合的效能影響，則可以採用不同的 RAID 群組佈局。

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功能添加到相同資料儲存中。在已安裝 ONTAP Select 的資料儲存中新增資料儲存擴充功能是動態的，並且不會影響 ONTAP Select 節點的運作。

如果 ONTAP Select 節點是 HA 配對的一部分，則需要考慮一些其他問題。

在 HA 配對中，每個節點都包含其合作夥伴資料的鏡像複本。將空間新增至節點 1 時，需要將相同數量的空間新增至其合作夥伴節點 2，以便將節點 1 的所有資料複寫至節點 2。換句話說，作為節點 1 容量新增作業的一部分而新增至節點 2 的空間，在節點 2 上是不可見或無法存取的。將空間新增至節點 2 是為了在 HA 事件期間充分保護節點 1 資料。

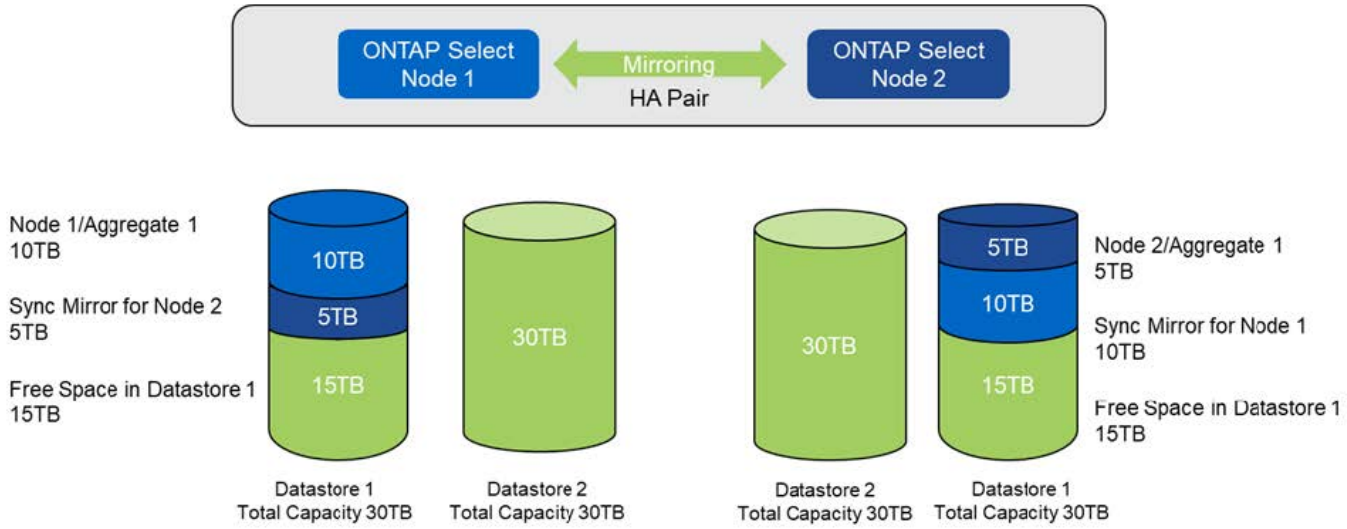
關於效能，還有一點要考慮。節點 1 上的資料會同步複製到節點 2。因此，節點 1 上新增空間（資料存放區）的效能必須與節點 2 上新增空間（資料存放區）的效能相符。換句話說，如果在兩個節點上都增加了空間，但使用了不同的磁碟機技術或不同的 RAID 群組大小，則可能會導致效能問題。這是因為 RAID SyncMirror 操作用於在合作夥伴節點上維護資料副本。

若要增加 HA 配對中兩個節點的使用者可存取容量，必須執行兩次儲存新增作業，每個節點一次。每次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，每個節點有兩個資料存放區，每個資料存放區有 30TB 的空間。ONTAP Deploy 建立一個雙節點叢集，每個節點從資料存放區 1 佔用 10TB 的空間。ONTAP Deploy 為每個節點配置 5TB 的作用中空間。

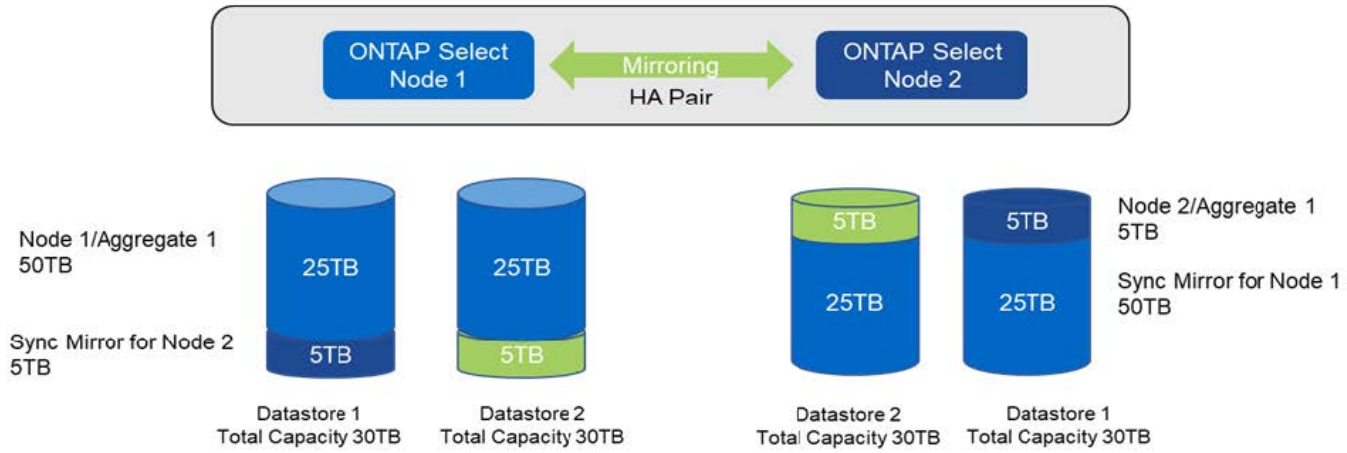
下圖顯示了對節點 1 執行一次儲存新增操作的結果。ONTAP Select 在每個節點上仍使用相同的儲存空間 (15TB)。但是，節點 1 的活動儲存空間 (10TB) 比節點 2 (5TB) 多。由於每個節點都託管著另一個節點的資料副本，因此兩個節點都受到完全保護。資料儲存區 1 中還有剩餘可用空間，資料儲存區 2 仍然完全可用。

容量分配：單次儲存新增作業後的分配空間與可用空間



在節點 1 上執行的兩個額外的儲存新增操作會佔用資料儲存區 1 的剩餘空間和資料儲存區 2 的一部分空間（使用容量上限）。第一次儲存新增作業會佔用資料儲存區 1 中剩餘的 15TB 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 50TB 的作用中資料，而節點 2 則保留了最初的 5TB 資料。

容量分配：節點 **1** 執行兩次額外儲存新增作業後的分配空間和可用空間



在容量新增作業期間，VMDK 的最大大小為 16TB。在叢集建立操作期間，VMDK 的最大大小仍為 8TB。ONTAP Deploy 會根據您的配置（單節點或多節點叢集）和新增的容量大小建立大小適當的 VMDK。但是，在叢集建立操作期間，每個 VMDK 的最大大小不應超過 8TB；在儲存新增作業期間，每個 VMDK 的最大大小不得超過 16TB。

使用軟體 **RAID** 增加 **ONTAP Select** 的容量

同樣，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。此精靈僅顯示可用且可對應為 RDM 到 ONTAP Select 虛擬機器的 DAS SSD 磁碟機。

雖然可以增加單一 TB 的容量授權，但在使用軟體 RAID 時，實際上無法增加單一 TB 的容量。與在 FAS 或 AFF 陣列新增磁碟類似，某些因素決定了單次操作可新增的最小儲存容量。

在 HA 配對中、將儲存設備新增至節點 1 時、節點的 HA 配對（節點 2 ）上也必須有相同數量的磁碟機可用。節點 1 上的單一儲存設備新增作業會同時使用本機磁碟機和遠端磁碟。也就是說、遠端磁碟機可確保節點 1 上的新儲存設備在節點 2 上進行複寫和保護。若要在節點 2 上新增本機

> 可用的儲存設備、則必須在兩個節點上執行個別的儲存設備新增作業、且兩個節點上都必須有個別且相同數量的磁碟機可用。

ONTAP Select 會將所有新磁碟機分割到與現有磁碟機相同的根分割區、資料分割區和資料分割區中。分割區操作在建立新 Aggregate 或擴充現有 Aggregate 時進行。每個磁碟上的根分割區條帶大小設定為與現有磁碟上的根分割區大小相符。因此，兩個相等的資料分割區大小均可計算為磁碟總容量減去根分割區大小再除以二。根分割區條帶大小是可變的，它在初始叢集設定期間按如下方式計算。所需的根分割區總空間（單節點叢集為 68GB、HA 配對為 136GB）分配到初始磁碟數量（減去任何備用磁碟和同位元檢查磁碟）。新增至系統中的所有磁碟機的根分割區條帶大小保持不變。

如果您要建立新的 Aggregate，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 配對。

如果要為現有 Aggregate 新增儲存設備，則需要考慮一些額外因素。可為現有 RAID 群組新增磁碟機，前提是該 RAID 群組尚未達到最大容量限制。傳統的 FAS 和 AFF 最佳實務同樣適用於向現有 RAID 群組新增磁碟，但需要注意的是，新磁碟可能會產生熱點。此外，只有資料分割區大小等於或大於現有分割區大小的磁碟機才能新增至現有 RAID 群組中。如上所述，資料分割區大小與磁碟機原始大小不同。如果新增的資料分割區大於現有分割區，則新磁碟機將進行「正確大小」調整。換句話說，每個新磁碟機的一部分容量將未被使用。

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 儲存效率支援

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）部署應遵循使用非 SSD 直接連接儲存設備（DAS）的 ONTAP Select 最佳實務做法。

只要您擁有配備 SSD 硬碟的 DAS 儲存和進階授權，新安裝的系統就會自動啟用類似 AFF 的特性。

具有類似 AFF 的特性，以下內嵌 SE 功能會在安裝過程中自動啟用：

- 內嵌零模式偵測
- Volume 內嵌重複資料刪除
- Volume 背景重複資料刪除
- 自適應內嵌壓縮
- 內嵌資料壓縮
- Aggregate 內嵌重複資料刪除
- Aggregate 背景重複資料刪除

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
```

對於從 9.6 及更高版本升級的 ONTAP Select，您必須將 ONTAP Select 安裝在具有高級授權的 DAS SSD 儲存設備上。此外，您必須在使用 ONTAP Deploy 進行初始叢集安裝時勾選 **Enable Storage Efficiencies** 核取方塊。如果在 ONTAP 升級後未滿足先前條件，則啟用類似 AFF 的特性需要手動建立開機引數並重新啟動節點。請聯絡技術支援以了解更多詳情。

ONTAP Select 儲存效率組態

下表總結了各種可用的儲存效率選項，這些選項預設為啟用，或預設不啟用但建議啟用，具體取決於媒體類型和軟體授權。

ONTAP Select 功能	**DAS SSD**（**premium** 或 **premium XL**[1]）	**DAS HDD** （所有授權）	**vNAS** （所有授權）
內嵌零偵測	是 (預設)	是 使用者可按磁碟區啟用此功能	是 使用者可按磁碟區啟用此功能
Volume 內嵌重複資料刪除	是 (預設)	無法使用	不支援
32K 內嵌壓縮（次要壓縮）	是的，使用者可按磁碟區啟用此功能。	是 使用者可按磁碟區啟用此功能	不支援
8K 內嵌壓縮（自適應壓縮）	是 (預設)	是 使用者可按磁碟區啟用	不支援
背景壓縮	不支援	是 使用者可按磁碟區啟用	是 使用者可按磁碟區啟用此功能
壓縮掃描器	是的	是的	是 使用者可按磁碟區啟用此功能
內嵌資料壓縮	是 (預設)	是 使用者可按磁碟區啟用	不支援
壓實掃描器	是的	是的	不支援
Aggregate 內嵌重複資料刪除	是 (預設)	N/A	不支援
Volume 背景重複資料刪除	是 (預設)	是 使用者可按磁碟區啟用	是 使用者可按磁碟區啟用此功能
Aggregate 背景重複資料刪除	是 (預設)	N/A	不支援

[1]ONTAP Select 9.6 支援新的授權（premium XL）和新的 VM 大小（large）。但是，large VM 僅支援使用軟體 RAID 的 DAS 組態。9.6 版本中的 large ONTAP Select VM 不支援硬體 RAID 和 vNAS 組態。

關於 **DAS SSD** 組態升級行為的注意事項

升級到 ONTAP Select 9.6 或更高版本後，請等待 `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 或更高版本的系統中，在現有 aggregate 或新建立的 aggregate 上建立的新 volume 與在全新部署上建立的 volume 具有相同的行為。經歷 ONTAP Select 程式碼升級的現有 volume 與新建立的 volume 具有大部分相同的儲存效率原則，但也存在一些差異：

情境 **1**

如果在升級之前磁碟區上未啟用任何儲存效率原則，則：

- 具有 `space guarantee = volume`的磁碟區未啟用即時資料壓縮、Aggregate 即時重複資料刪除和 Aggregate 背景重複資料刪除。這些選項可在升級後啟用。
- 具有 `space guarantee = none`的磁碟區未啟用背景壓縮。此選項可在升級後啟用。
- 升級後，現有磁碟區的儲存效率原則設定為 auto。

情境 **2**

如果在升級之前已在磁碟區上啟用某些儲存效率功能，則：

- 具有 `space guarantee = volume`的磁碟區在升級後看不到任何差異。
- 具有 `space guarantee = none`的磁碟區已啟用 Aggregate 背景重複資料刪除功能。
- 具有 `storage policy inline-only`的 Volume 的原則設定為自動。
- 除了具有 `space guarantee = none`的磁碟區之外，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。這些磁碟區已啟用 Aggregate 背景重複資料刪除。

網路

ONTAP Select 網路概念和特性

首先熟悉適用於 ONTAP Select 環境的一般網路概念。然後探索單節點和多節點叢集的具體特性和可用選項。

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。與實體網路相關的組態包括 Hypervisor 主機和更廣泛的交換網路環境。

主機 **NIC** 選項

每個 ONTAP Select Hypervisor 主機必須配置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：

- 叢集包含一台還是多台 ONTAP Select 主機
- 使用哪種 hypervisor 作業系統

- 虛擬交換器的設定方式
- 連結是否使用 LACP

實體交換器組態

您必須確保實體交換器的組態支援 ONTAP Select 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。主要考量包括下列項目：

- 您將如何保持內部網路和外部網路之間的隔離？
- 您是否會維持資料與管理網路之間的分離？
- 第二層 VLAN 將如何設定？

邏輯網路

ONTAP Select 使用兩個不同的邏輯網路，根據類型分隔流量。具體來說，流量可以在叢集內的主機之間流動，也可以流向儲存用戶端和叢集外的其他機器。由 hypervisor 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 節點透過隔離的「內部」網路進行通訊。此網路不會暴露在 ONTAP Select 叢集中的節點外部,也無法從外部存取。

內部網路僅存在於多節點叢集中。

內部網路具有以下特性：

- 用於處理 ONTAP 叢集內部流量，包括：
  - 叢集
  - 高可用性互連 （HA-IC）
  - RAID Synch Mirror（RSM）
- 基於 VLAN 的單層二層網路
- 靜態 IP 位址由 ONTAP Select 分配：
  - 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（含）進行調整

外部網路

外部網路處理 ONTAP Select 叢集節點與外部儲存用戶端以及其他機器之間的流量。外部網路是每個叢集部署的組成部分，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；選用 SVM）
- 叢集間（選用）

- 選擇性支援 VLAN：
  - 資料連接埠群組
  - 管理連接埠群組
- 根據管理員的組態選擇所指派的 IP 位址：
  - IPv4 或 IPv6
- 預設 MTU 大小為 1500 位元組（可調整）

所有規模的叢集都存在外部網路。

虛擬機器網路環境

Hypervisor 主機提供多種網路功能。

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 可以使用多個連接埠。這些連接埠的分配和使用取決於多種因素，包括叢集的大小。

虛擬交換器
: 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））會將虛擬機器暴露的連接埠與實體乙太網路 NIC 連接埠連接起來。您必須根據您的環境，為每個 ONTAP Select 主機設定 vSwitch。

ONTAP Select 單節點和多節點網路組態

ONTAP Select 支援單節點和多節點網路組態。

單節點網路組態

單節點 ONTAP Select 組態不需要 ONTAP 內部網路、因為沒有叢集、HA 或鏡射流量。

與 ONTAP Select 產品的多節點版本不同，每個 ONTAP Select VM 都包含三個虛擬網路介面卡，分別呈現給 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。

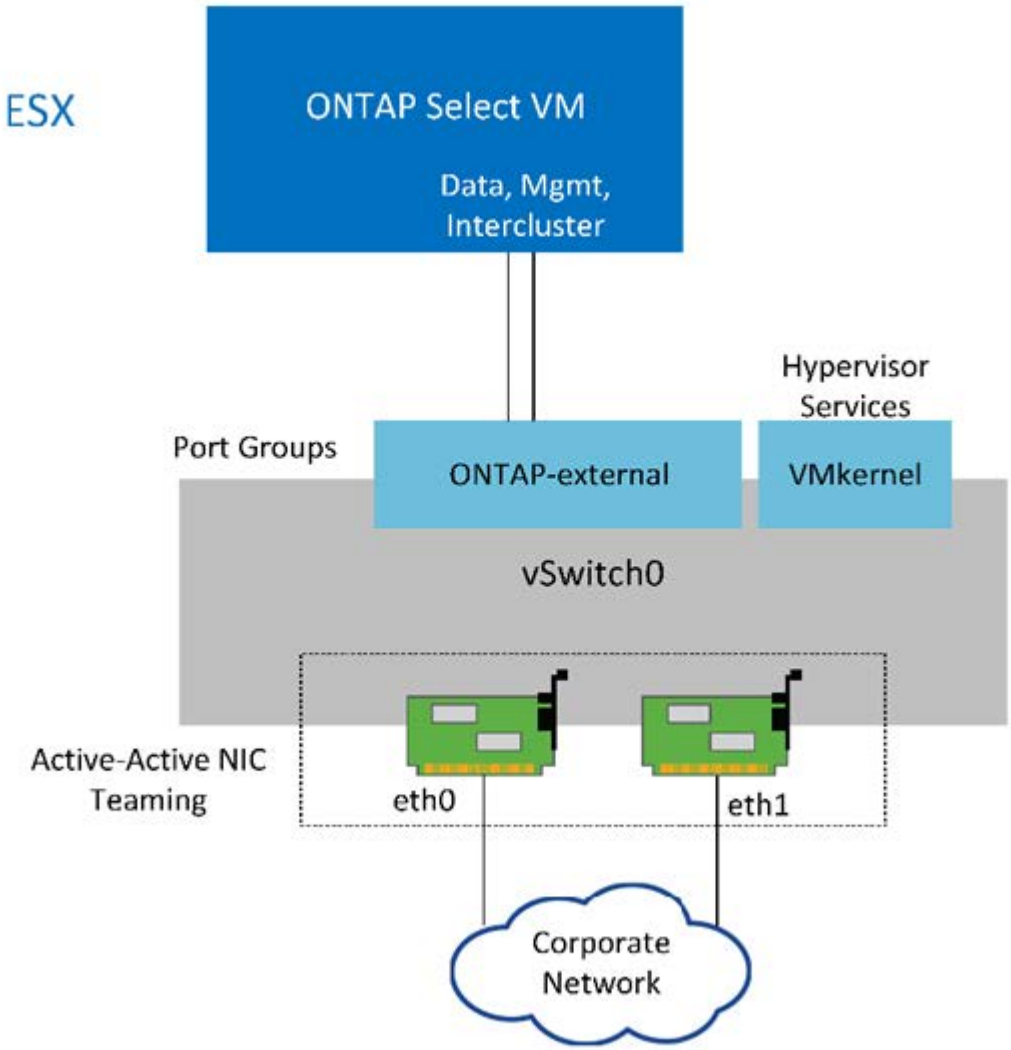
KVM

您可以將 ONTAP Select 部署為單節點叢集。Hypervisor 主機包含一個虛擬交換器，可存取外部網路。

ESXi

下圖展示了這些連接埠與底層實體適配器之間的關係。圖中描述了 ESXi 虛擬機器管理程式上的一個 ONTAP Select 叢集節點。

單節點 **ONTAP Select** 叢集的網路組態



即使單節點叢集只需要兩個介面卡，仍然需要 NIC 群組。

LIF 指派

如本文件多節點 LIF 指派區段所述，ONTAP 使用 IPspace 將叢集網路流量與資料和管理流量分開。此平台的單節點變體不包含叢集網路。因此，叢集 IPspace 中不存在任何連接埠。

在 ONTAP Select 叢集設定過程中，叢集和節點管理 LIF 會自動建立。您可以在部署後建立其餘的 LIF。

管理和資料 **LIF**（**e0a**、**e0b** 和 **e0c**）

ONTAP 連接埠 e0a、e0b 和 e0c 被委派為 LIF 的候選連接埠，用於承載以下類型的流量：

- 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

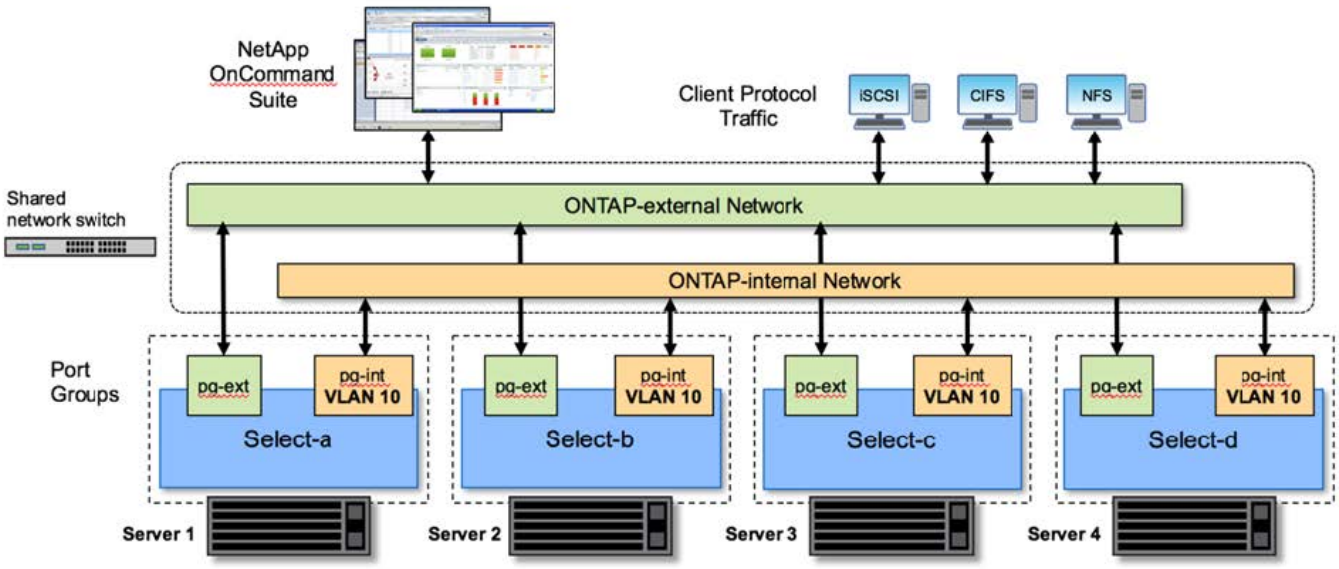
多節點 ONTAP Select 網路組態由兩個網路組成。

其中包括負責提供叢集和內部複寫服務的內部網路，以及負責提供資料存取和管理服務的外部網路。這兩個網路之間流量的端對端隔離對於建立適合叢集恢復能力的環境至關重要。

下圖展示了這些網路，圖中顯示的是一個運行在 VMware vSphere 平台上的四節點 ONTAP Select 叢集。六節點、八節點、十節點和十二節點叢集的網路佈局類似。

每個 ONTAP Select 執行個體都位於個別的實體伺服器上。內部和外部流量使用個別的網路連接埠群組進行隔離，這些連接埠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 多節點叢集網路組態總覽



每個 ONTAP Select VM 包含七個虛擬網路適配器，這些適配器以七個網路連接埠（e0a 到 e0g）的形式呈現給 ONTAP。雖然 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 新增虛擬網路介面卡。

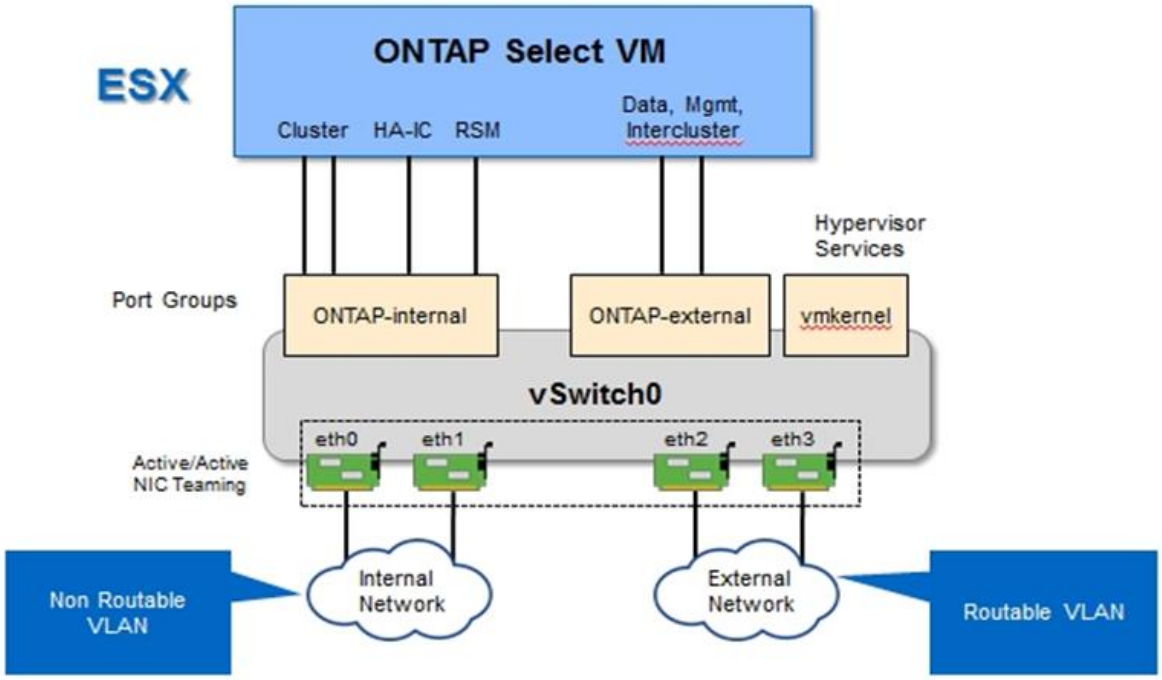
這些連接埠已預先設定為提供以下服務：

- e0a、e0b 和 e0g。管理和資料 LIF
- e0c、e0d。叢集網路 LIF
- e0e。RSM
- 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 執行多種不同的功能、但它們共同構成內部 Select 網路。在進行網路設計決策時、您應該將這些連接埠放置在單一第 2 層網路上。不需要將這些虛擬介面卡分散到不同的網路。

下圖展示了這些連接埠與底層實體適配器之間的關係，圖中描繪了 ESXi 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



將內部流量和外部流量隔離到不同的實體 NIC 上，可防止網路資源存取不足，從而避免系統延遲。此外，透過 NIC 綁定進行聚合，即使單一網路介面卡發生故障，ONTAP Select 叢集節點也能繼續存取網路。

外部網路連接埠群組和內部網路連接埠群組均以對稱方式包含全部四個 NIC 適配器。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

隨著 IPspaces 的引入，ONTAP 連接埠角色已被棄用。與 FAS 陣列類似，ONTAP Select 叢集包含預設 IPspace 和叢集 IPspace。透過將網路連接埠 e0a、e0b 和 e0g 放入預設 IPspace，並將連接埠 e0c 和 e0d 放入叢集 IPspace，這些連接埠實際上被隔離，無法承載不屬於該空間的 LIF。ONTAP Select 叢集中的其餘連接埠透過自動分配介面來提供內部服務。與 RSM 和 HA 互連介面不同，這些連接埠不會透過 ONTAP shell 公開。

並非所有 LIF 都可透過 ONTAP 指令 shell 存取。HA 互連和 RSM 介面對 ONTAP 隱藏，它們在內部用於提供各自的服務。

以下各節將詳細說明網路 port 和 LIF。

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

ONTAP 連接埠 e0a、e0b 和 e0g 被委派為 LIF 的候選連接埠，用於承載以下類型的流量：

- 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）

> 在 ONTAP Select 叢集設定過程中，叢集和節點管理 LIF 會自動建立。您可以在部署後建立其餘的 LIF。

叢集網路 LIF（e0c、e0d）

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

> 您無法為這些介面指派靜態 IP 位址，也不應該建立額外的叢集介面。

叢集網路流量必須透過低延遲、非路由的二層網路傳輸。由於叢集吞吐量和延遲要求，您應該將 ONTAP Select 叢集實體部署在彼此靠近的位置（例如，多包裝、單一資料中心）。不支援透過廣域網路或顯著的地理距離將高可用性節點分開來建構四節點、六節點、八節點、十節點或十二節點的跨叢集配置。支援使用中間節點的跨雙節點配置。

詳情請參閱相關章節"雙節點延伸 HA（MetroCluster SDS）最佳實務做法"。

> 為確保叢集網路流量的最大處理量，此網路連接埠配置為使用巨型訊框（7500 至 9000 MTU）。為確保叢集正常運作，請驗證所有向 ONTAP Select 叢集節點提供內部網路服務的上游虛擬和實體交換器是否已啟用巨型訊框。

RAID SyncMirror 流量 (e0e)

HA 夥伴節點之間的區塊同步複寫是透過位於網路連接埠 e0e 的內部網路介面進行。此功能會自動執行，使用 ONTAP 在叢集設定期間所設定的網路介面，不需要管理員進行任何組態設定。

> 連接埠 e0e 已被 ONTAP 用於內部複寫流量。因此，此連接埠及其託管的 LIF 在 ONTAP CLI 或 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，您無法指派備用 IP 位址。此網路連接埠需要使用巨型訊框（7500 至 9000 MTU）。

HA 互連（e0f）

NetApp FAS 陣列使用專用硬體在 ONTAP 叢集中的 HA 配對之間傳遞資訊。然而，軟體定義環境通常不具備此類設備（例如 InfiniBand 或 iWARP 裝置），因此需要替代解決方案。儘管考慮了多種可能性，但 ONTAP 對互連傳輸的要求決定了此功能必須在軟體中模擬。因此，在 ONTAP Select 叢集中，HA 互連功能（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 節點都配置了一個 HA 互連連接埠 e0f。此連接埠承載 HA 互連網路介面，負責兩個主要功能：

- 在 HA 配對之間鏡像 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號訊息

HA 互連流量透過此網路連接埠使用單一網路介面，透過在乙太網路封包中分層遠端直接記憶體存取（RDMA）訊框來實現。

與 RSM 連接埠（e0e）類似，使用者無法透過 ONTAP CLI 或 System Manager 存取此實體連接埠或託管網路介面。因此，您無法修改此介面的 IP 位址，也無法變更連接埠狀態。此網路連接埠需要使用巨型訊框（7500 至 9000 MTU）。

ONTAP Select 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

內部 ONTAP Select 網路僅存在於該產品的多節點版本中，負責為 ONTAP Select 叢集提供叢集通訊、HA 互連和同步複寫服務。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.*承載 HA 互連 LIF

此網路的處理量和延遲對於決定 ONTAP Select 叢集的效能和恢復能力至關重要。網路隔離是叢集安全性所必需的，並確保系統介面與其他網路流量分開。因此，此網路必須專供 ONTAP Select 叢集使用。

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

流經內部網路的網路封包必須位於專用的 VLAN 標籤第 2 層網路上。您可以完成下列其中一項工作來達成此目標：

- 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原生 VLAN，其中原生 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 ESXi 標準和分散式 vSwitches。不支援其他虛擬交換器或 ESXi 主機之間的直接連線。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透過稱為連接埠群組的虛擬二層網路物件進行隔離。正確 vSwitch 分配這些連接埠群組至關重要，尤其對於負責提供叢集、HA 互連和鏡像複寫服務的內部網路而言更是如此。這些網路連接埠的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、八節點、十節點和十二節點叢集要求內部 ONTAP Select 網路使用 10Gb 連線；不支援 1Gb NIC。但是，可以對外部網路進行權衡，因為限制流入 ONTAP Select 叢集的資料流量不會影響其可靠運作能力。

雙節點叢集可以使用四個 1Gb 連接埠用於內部流量，也可以使用單一 10Gb 連接埠取代四節點叢集所需的兩個 10Gb 連接埠。在伺服器環境不允許安裝四個 10Gb NIC 卡的情況下，可以使用兩個 10Gb NIC 卡用於內部網路，兩個 1Gb NIC 用於外部 ONTAP 網路。

內部網路驗證與疑難排解

可以使用網路連接檢查器功能來驗證多節點叢集的內部網路。此功能可以透過在 Deploy CLI 中執行 `network connectivity-check start` 命令來呼叫。

執行以下命令以檢視測試的輸出：

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

此工具僅適用於排查多節點 Select 叢集的內部網路故障。此工具不應用於排查單節點叢集（包括 vNAS 組態）、 ONTAP Deploy 與 ONTAP Select 連線問題或用戶端連線問題。

叢集建立精靈（ ONTAP Deploy 使用者介面的一部分）包含內部網路檢查器，作為建立多節點叢集期間的可選步驟。鑑於內部網路在多節點叢集中扮演的重要角色，將此步驟納入叢集建立工作流程可以提高叢集建立操作的成功率。

從 ONTAP Deploy 2.10 開始，內部網路使用的 MTU 大小可以設定為 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 等網路覆蓋層，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訊，因此在單節點和多節點組態中都存在。雖然此網路不像內部網路那樣具有嚴格定義的處理量需求，但管理員應小心避免在用戶端與 ONTAP VM 之間造成網路瓶頸，因為效能問題可能會被誤判為 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，ONTAP Select VM 本身也可以透過稱為 VGT 的程序對外部流量進行標記。如需更多詳細資訊，請參閱 "資料和管理流量分離" 章節。

下表列出了 ONTAP Select 內部網路和外部網路的主要差異。

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）	叢集間資料管理（SnapMirror 和 SnapVault）
網路隔離	必填	選用
訊框大小 (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 組隊

為確保內部和外部網路都具備提供高效能和容錯能力所需的頻寬和彈性特性，建議採用實體網路卡綁定。系統支援使用單條 10Gb 連結的雙節點叢集配置。然而，NetApp 建議的最佳實務做法是在 ONTAP Select 叢集的內部和外部網路上都使用 NIC 綁定。

MAC 位址產生

所有 ONTAP Select 網路連接埠的 MAC 位址均由內建部署公用程式自動產生。此公用程式使用特定於 NetApp 的平台專屬組織唯一識別碼（OUI），以確保與 FAS 系統不發生衝突。此位址的副本隨後儲存在 ONTAP Select 安裝虛擬機器（ONTAP Deploy）的內部資料庫中，以防止在後續節點部署期間意外重新分配。管理員在任何情

況下都不得修改已指派的網路連接埠 MAC 位址。

支援的 **ONTAP Select** 網路組態

選擇最佳硬體並設定網路，以最佳化效能與恢復能力。

伺服器廠商深知客戶需求各異，選擇至關重要。因此，在購買實體伺服器時，網路連線方面有多種選擇。大多數通用系統都配備多種 NIC 選項，提供單一連接埠和多連接埠選項，速度和吞吐量組合各異。這包括對 25Gb/s 和 40Gb/s NIC 適配器以及 VMware ESX 的支援。

由於 ONTAP Select VM 的效能直接取決於底層硬體的特性，因此選擇更高速度的 NIC 來提高 VM 的處理量，可提升叢集效能並改善整體使用者體驗。可以使用四個 10Gb NIC 或兩個更高速度（25/40 Gb/s）的 NIC 來實現高效能網路配置。此外，還支援其他多種組態。對於雙節點叢集，支援 4 個 1Gb 連接埠或 1 個 10Gb 連接埠。對於單節點叢集，支援 2 個 1Gb 連接埠。

網路最低和建議組態

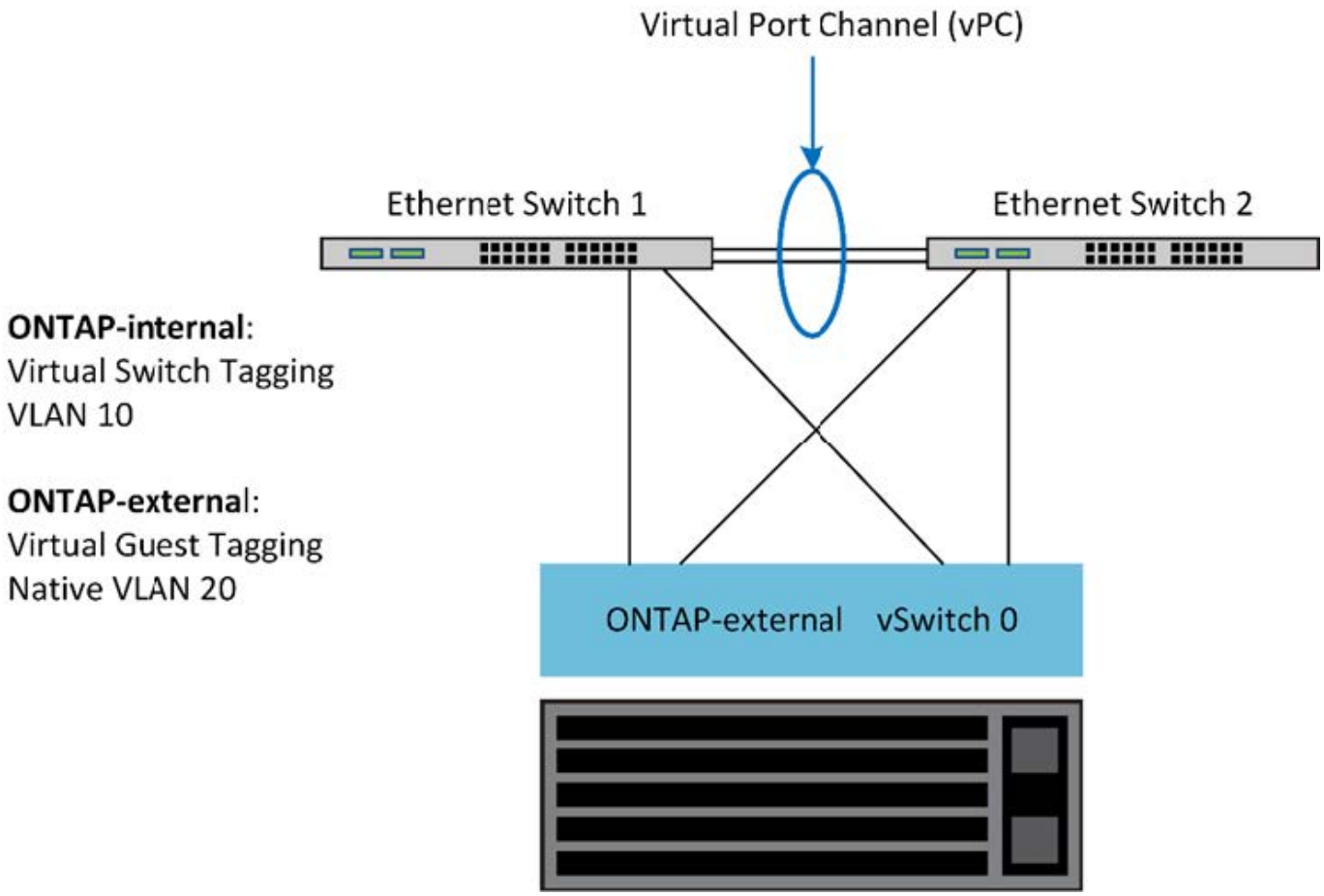
根據叢集大小，有數種支援的乙太網路組態。

叢集大小	最低需求	建議
單節點叢集	2 x 1GbE	2 x 10GbE
雙節點叢集或 MetroCluster SDS	4 x 1GbE 或 1 x 10GbE	2 x 10GbE
四節點、六節點、八節點、十節點或十二節點叢集	2 x 10GbE	4 x 10GbE 或 2 x 25/40GbE

由於每種拓撲可能需要在不同的 NIC 群組組態之間進行轉換，因此不支援在執行中的叢集上進行單一連結拓撲與多重連結拓撲之間的轉換。

使用多個實體交換器的網路組態

當有足夠的硬體可用時、NetApp 建議使用下圖所示的多交換器組態、因為這樣可以增加對實體交換器故障的保護。



ESXi 上的 ONTAP Select VMware vSphere vSwitch 組態

ONTAP Select vSwitch 雙網卡和四網卡組態的組態和負載平衡原則。

ONTAP Select 支援標準配置和分散式 vSwitch 配置。分散式 vSwitches 支援鏈路聚合結構（LACP）。鏈路聚合是一種常見的網路結構，用於聚合多個實體介面卡的頻寬。LACP 是一個廠商中立的標準。它為網路端點提供了一個開放協定，可將多個實體網路連接埠群組捆綁到單一邏輯通道中。ONTAP Select 可以與配置為鏈路聚合群組（LAG）的連接埠群組配合使用。但是，NetApp 建議將個別實體連接埠用作簡單的上行鏈路（主幹）連接埠，以避免使用 LAG 配置。在這種情況下，標準和分散式 vSwitches 的最佳實務做法是相同的。

本節介紹在雙網卡和四網卡配置中應使用的 vSwitch 設定和負載平衡策略。

當您為 ONTAP Select 設定連接埠群組時，請遵循下列最佳實務做法；連接埠群組層級的負載平衡原則為 Route Based on Originating Virtual Port ID。VMware 建議將連接到 ESXi 主機的交換器連接埠上的 STP 設定為 Portfast 模式。

所有 vSwitch 配置至少需要兩個實體網路介面卡，並將它們捆綁到一個網卡組（NIC Team）中。ONTAP Select 支援雙節點叢集使用單一 10Gb 連結。但是，NetApp 建議使用網卡聚合以確保硬體冗餘。

在 vSphere 伺服器上，NIC 團隊是用於將多個實體網路介面卡整合至單一邏輯通道的彙總架構，可讓網路負載在所有成員連接埠之間共享。請務必記住，建立 NIC 團隊時不需要實體交換器的支援。負載平衡和容錯移轉原則可直接套用至 NIC 團隊，而 NIC 團隊不會察覺上游交換器組態。在此情況下，原則僅套用至傳出流量。

ONTAP Select 不支援靜態連接埠通道。分散式 vSwitches 支援啟用 LACP 的通道，但使用

LACP LAG 可能會導致 LAG 成員之間的負載分佈不均。

對於單節點叢集，ONTAP Deploy 會將 ONTAP Select VM 設定為使用連接埠群組來連接外部網路，並使用相同的連接埠群組或（可選）不同的連接埠群組來處理叢集和節點管理流量。對於單節點叢集，您可以將所需數量的實體連接埠作為主動介面卡新增至外部連接埠群組。

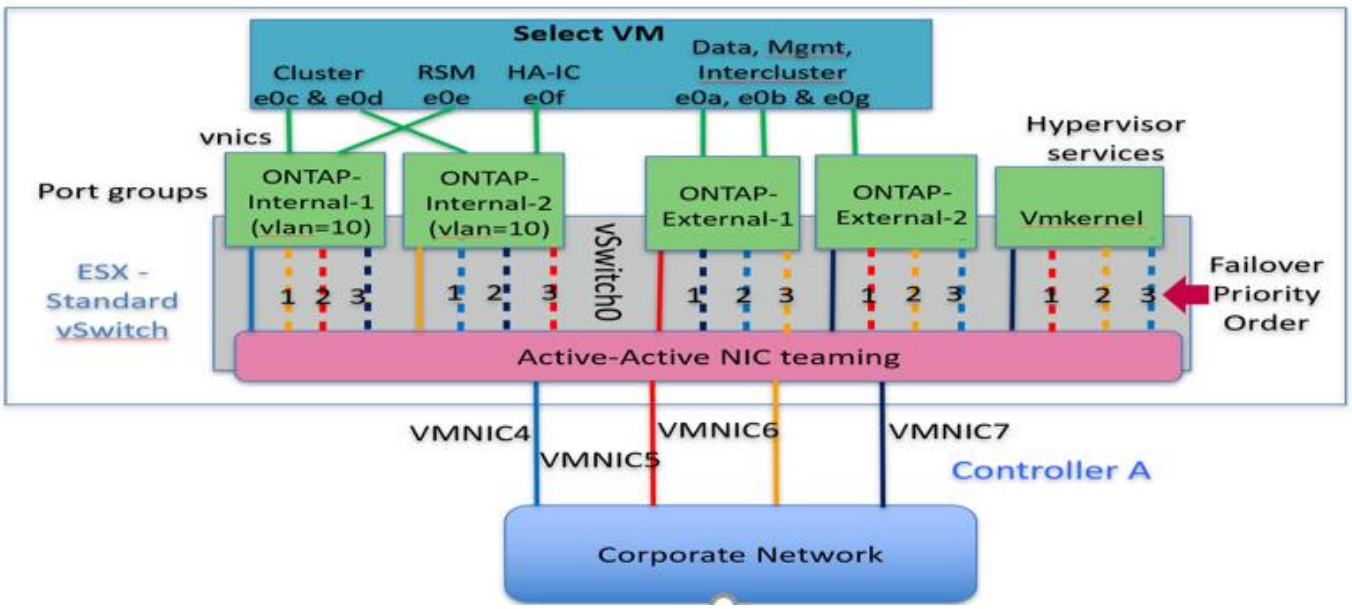
對於多節點叢集、ONTAP Deploy 會將每個 ONTAP Select VM 設定為內部網路使用一或兩個連接埠群組、外部網路則另外使用一或兩個連接埠群組。叢集和節點管理流量可以使用與外部流量相同的連接埠群組、或選擇性地使用個別的連接埠群組。叢集和節點管理流量無法與內部流量共用相同的連接埠群組。

ONTAP Select 最多支援四個 VMNIC。

標準或分散式 **vSwitch** 架構，每個節點四個實體連接埠

您可以在多節點叢集中為每個節點指派四個連接埠群組。每個連接埠群組都有一個作用中實體連接埠和三個待命實體連接埠，如下圖所示。

每個節點配備四個實體連接埠的 **vSwitch**



待命清單中連接埠的順序非常重要。下表提供四個連接埠群組中實體連接埠分佈的範例。

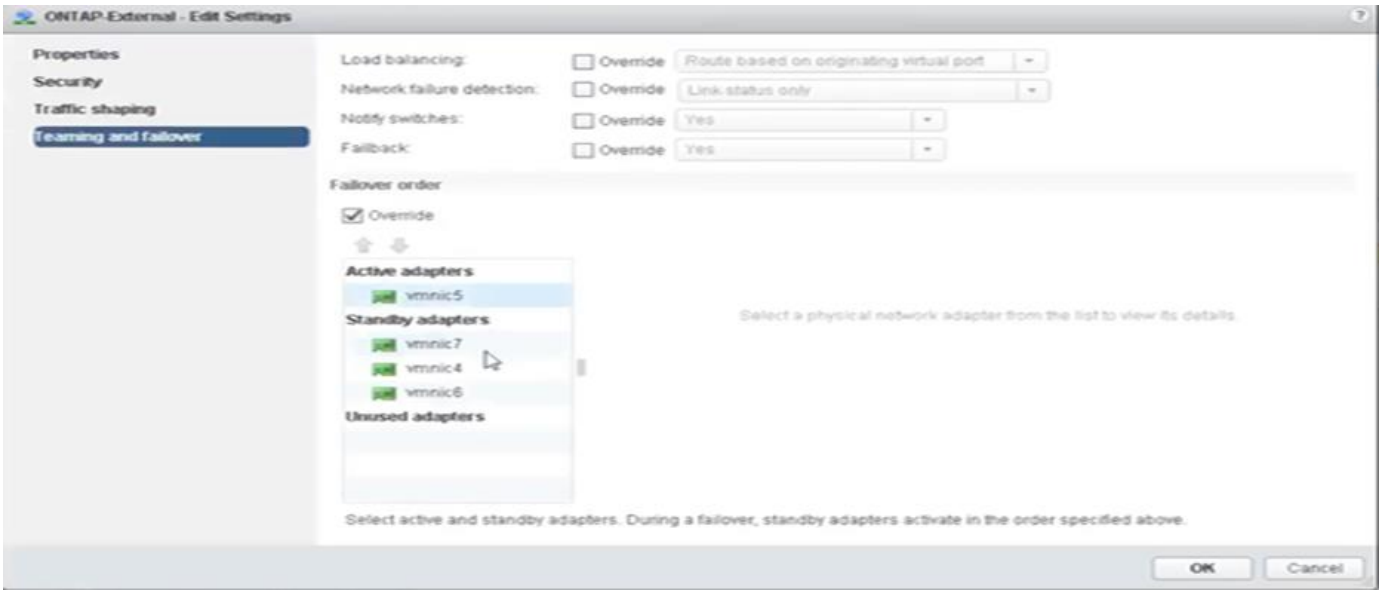
網路最低和建議組態

連接埠群組	外部 **1**	外部 **2**	內部 **1**	內部 **2**
作用中	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

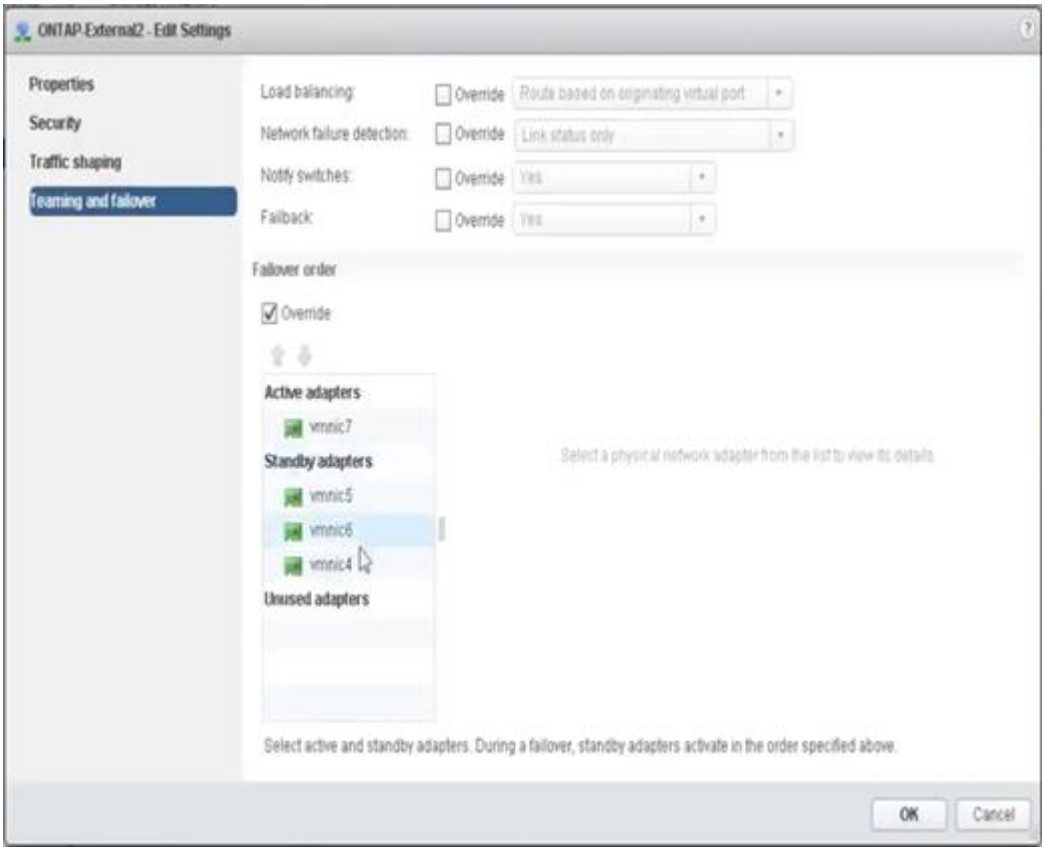
下圖展示了從 vCenter UI（ONTAP-External 和 ONTAP-External2）中外部網路連接埠群組的組態。請注意，活動配接器來自不同的網路卡。在此設定中、vmnic 4 和 vmnic 5 是同一實體 NIC 上的雙連接埠、而 vmnic 6 和

vminc 7 同樣是另一個 NIC 上的雙連接埠（此範例中未使用 vnmics 0 到 3）。待命介面卡的順序提供階層式容錯移轉、內部網路的連接埠位於最後。兩個外部連接埠群組的待命清單中內部連接埠的順序同樣會互換。

第一部分： **ONTAP Select** 外部連接埠群組組態



第二部分：**ONTAP Select** 外部連接埠群組組態

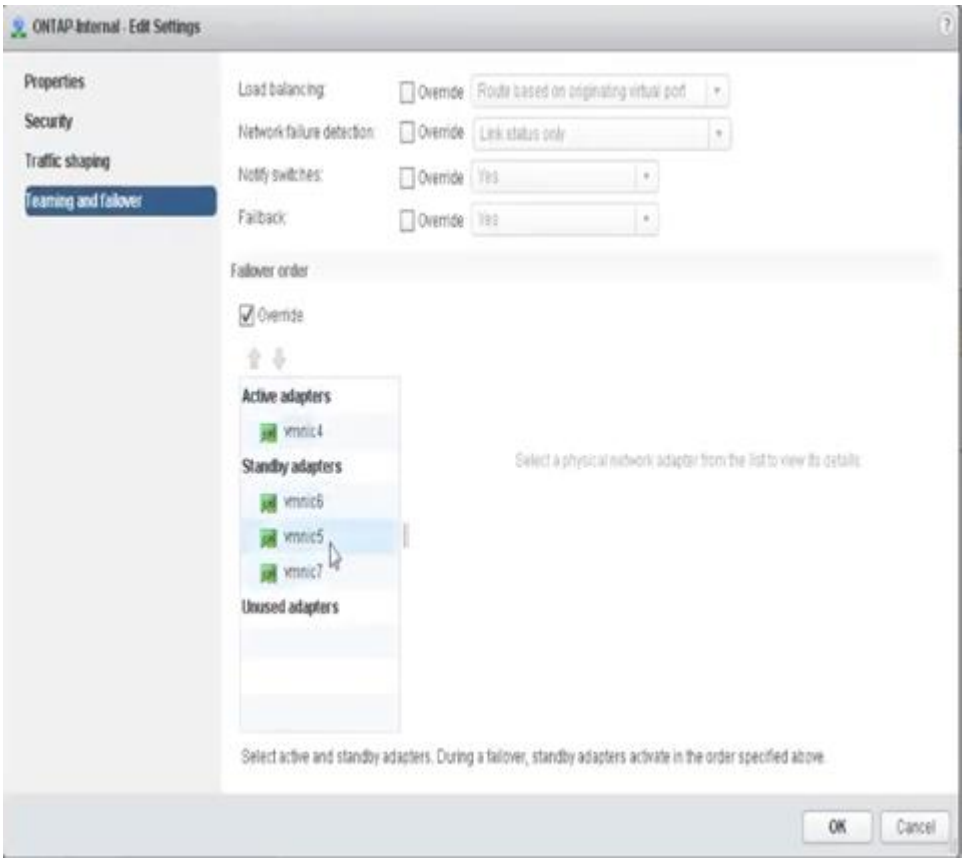


為了方便閱讀，指派如下：

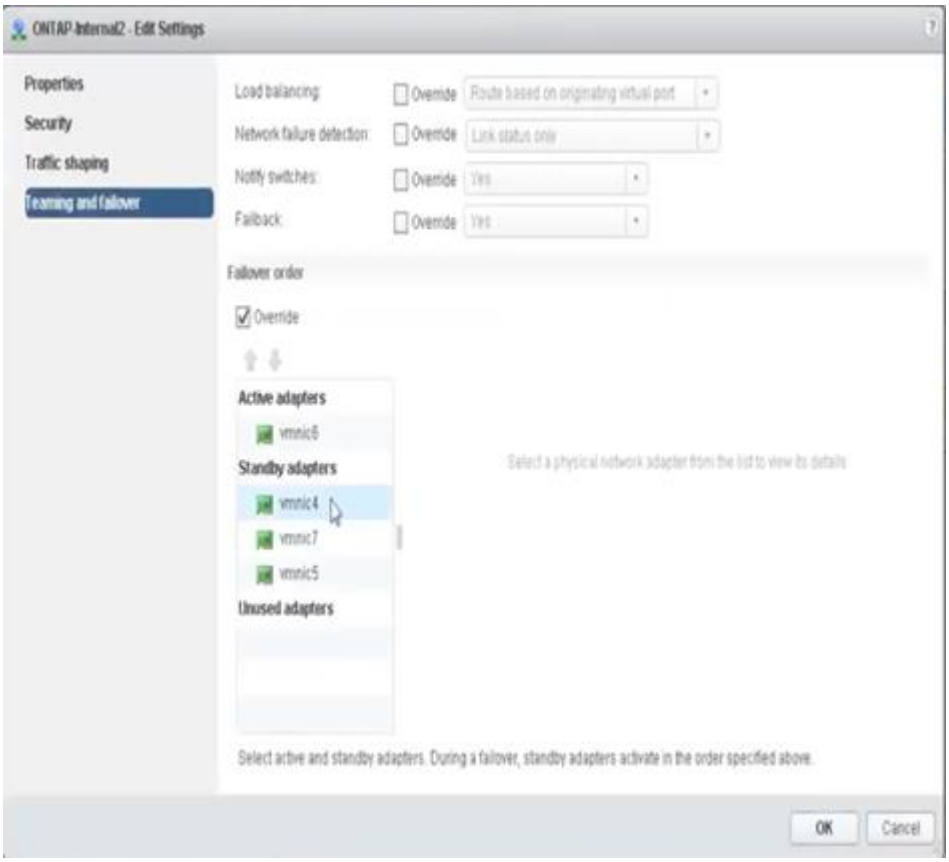
ONTAP-External	ONTAP-External2
作用中介面卡：vmnic5 待命介面卡：vmnic7、vmnic4、vmnic6	作用中介面卡：vmnic7 待命介面卡：vmnic5、vmnic6、vmnic4

下圖展示了內部網路連接埠群組（ONTAP-Internal 和 ONTAP-Internal2）的配置。請注意，活動配接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 上的雙連接埠，而 vmnic 6 和 vmnic 7 則是位於不同 ASIC 上的雙連接埠。備用配接器的順序提供了階層式容錯移轉，外部網路的連接埠位於最後。兩個內部連接埠群組的備用清單中，外部連接埠的順序也同樣互換。

第一部分：**ONTAP Select** 內部連接埠群組組態



第二部分：**ONTAP Select** 內部連接埠群組



為了方便閱讀，指派如下：

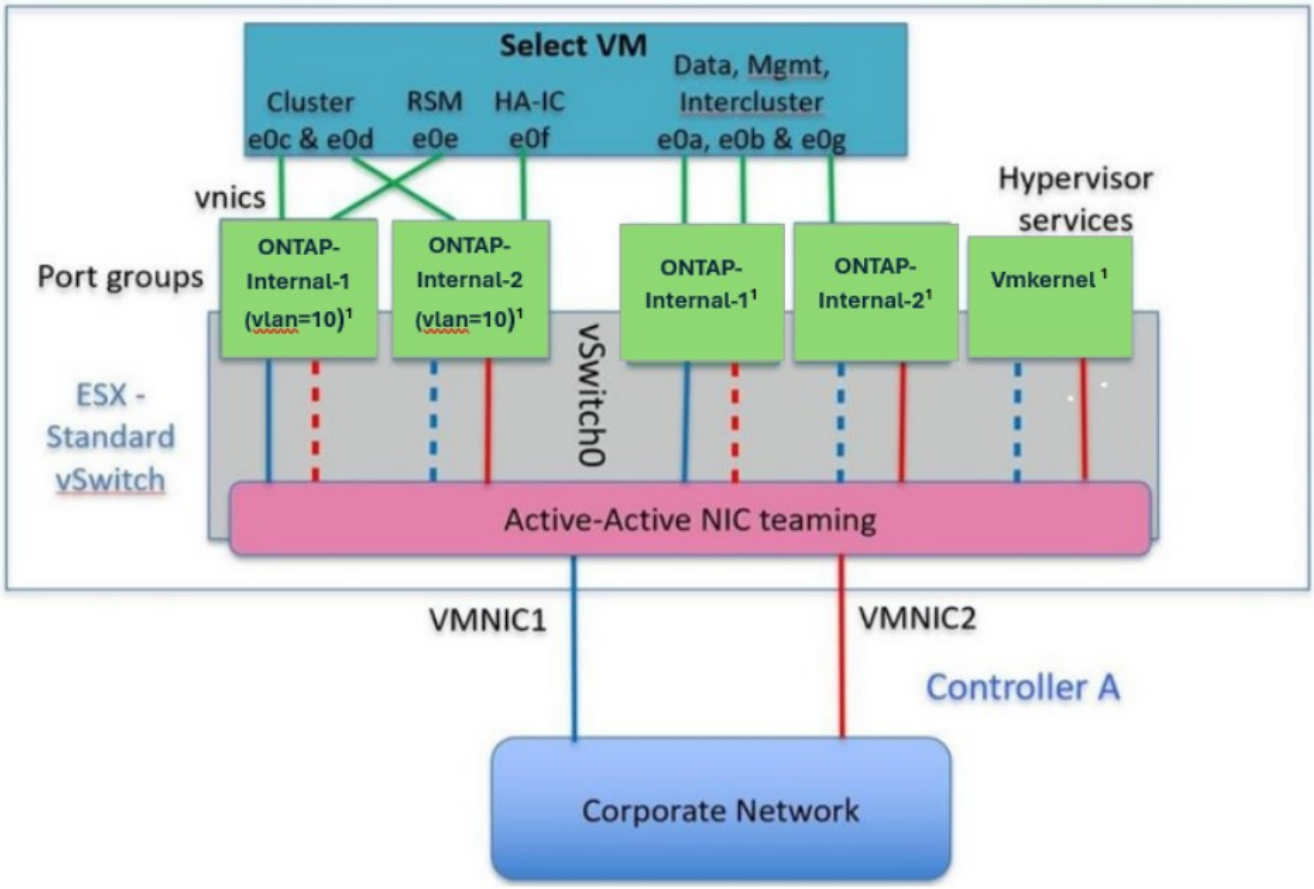
ONTAP-Internal	**ONTAP-Internal2**
作用中介面卡：vmnic4 待命介面卡：vmnic6、vmnic5、vmnic7	作用中介面卡：vmnic6 待命介面卡：vmnic4、vmnic7、vmnic5

標準或分散式 **vSwitch** 和每個節點兩個實體連接埠

當使用兩塊高速（25/40Gb）網路卡時，建議的連接埠組配置在概念上與使用四塊 10Gb 網路卡的配置非常相似。即使只使用兩塊實體網路卡，也應該使用四個連接埠組。連接埠組分配如下：

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

vSwitch 每個節點配備兩個高速（**25/40Gb**）實體連接埠

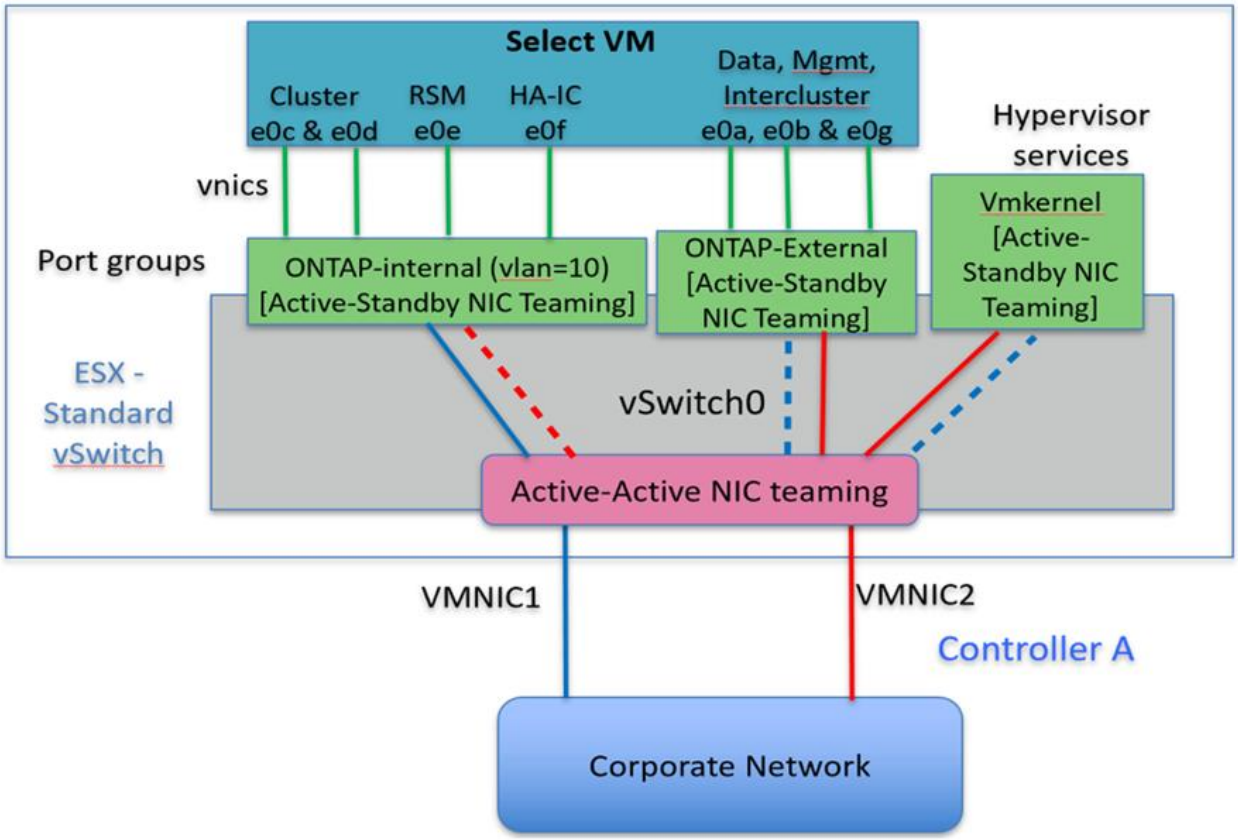


[1] The port groups attached to the virtual NICs are configured to use one NIC as active and the rest as standby.

當使用兩個實體連接埠（10Gb 或更低）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，二者配置方向相反。內部網路僅適用於多節點 ONTAP Select 叢集。對於單節點叢集，外部連接埠組中的兩個適配器均可配置為活動狀態。

以下範例展示了 vSwitch 的組態，以及負責處理多節點 ONTAP Select 叢集內部和外部通訊服務的兩個連接埠群組。由於內部網路 VMNIC 屬於此連接埠群組並設定為待命模式，因此在網路中斷時，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主動和待命 VMNIC，對於在網路中斷期間正確容錯移轉 ONTAP Select VM 至關重要。

每個節點配備兩個實體連接埠（**10Gb** 或更低）的 **vSwitch**

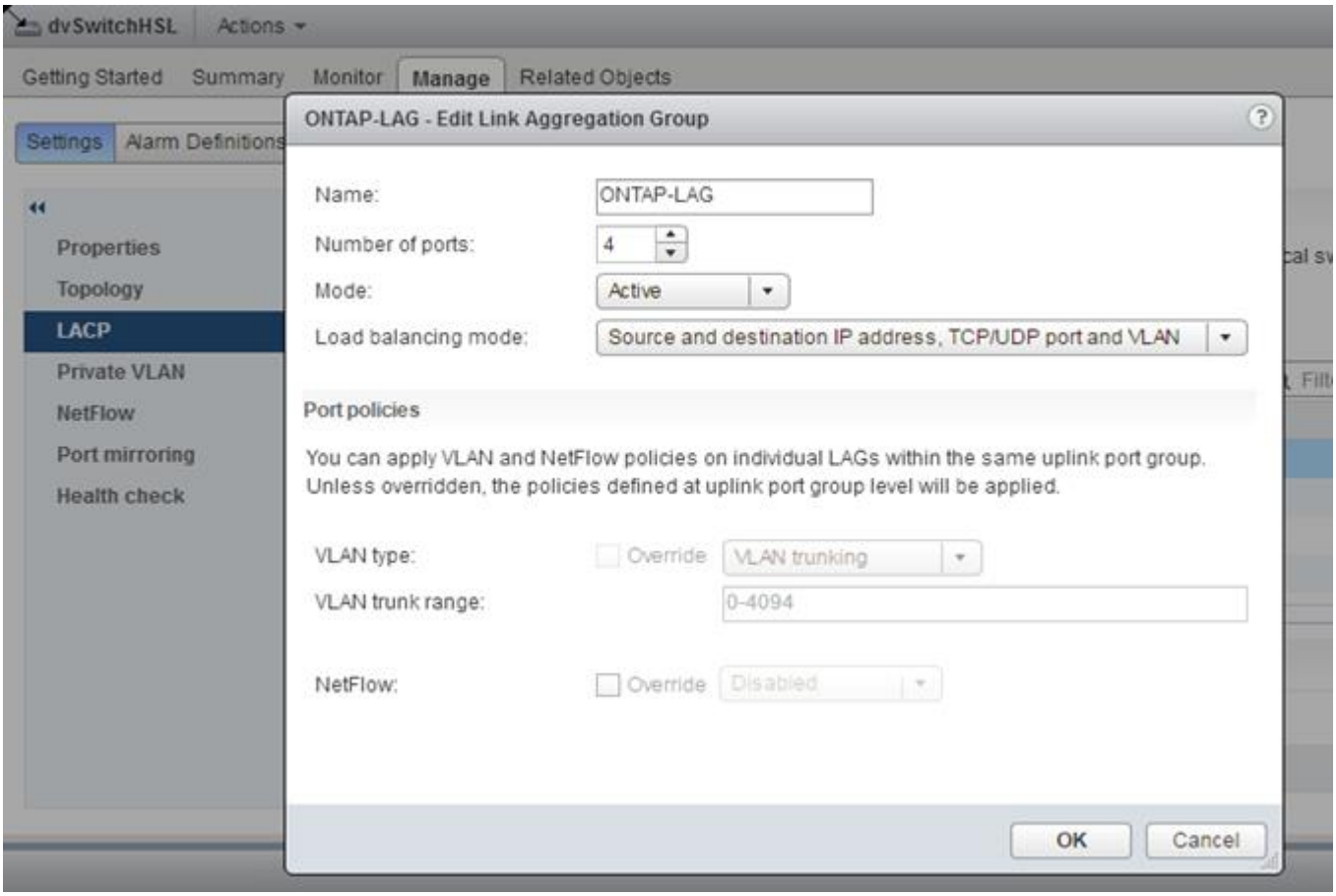


使用 **LACP** 的分散式 **vSwitch**

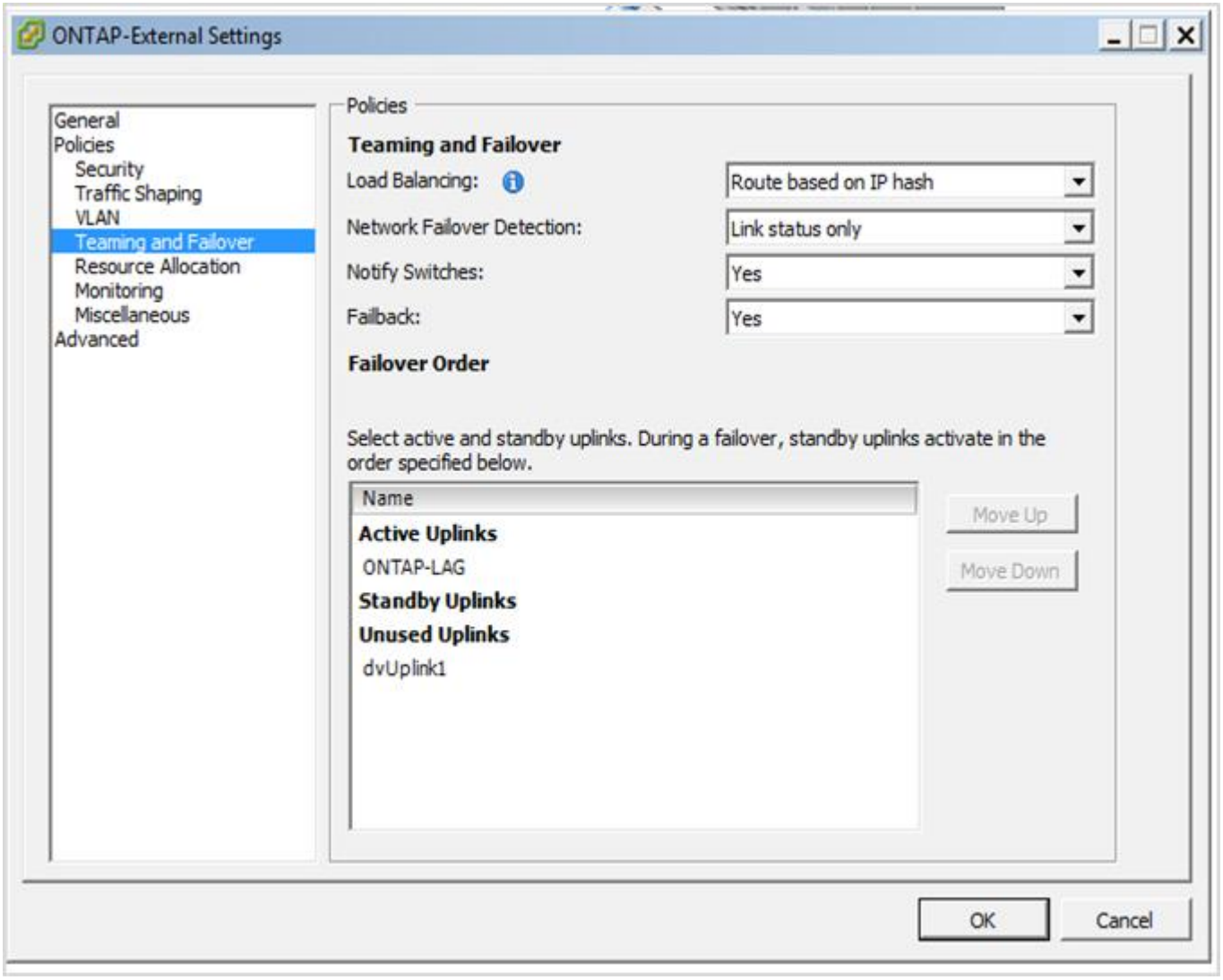
在組態中使用分散式 vSwitches 時，可以使用 LACP（儘管並非最佳實務做法）來簡化網路組態。唯一支援的 LACP 組態要求所有 VMNIC 都位於單一 LAG 中。上行實體交換器必須支援通道中所有連接埠的 MTU 大小在 7,500 到 9,000 之間。內部和外部 ONTAP Select 網路應在連接埠群組層級是隔離的。內部網路應使用不可路由（隔離的）VLAN。外部網路可以使用 VST、EST 或 VGT。

以下範例展示了使用 LACP 的分散式 vSwitch 設定。

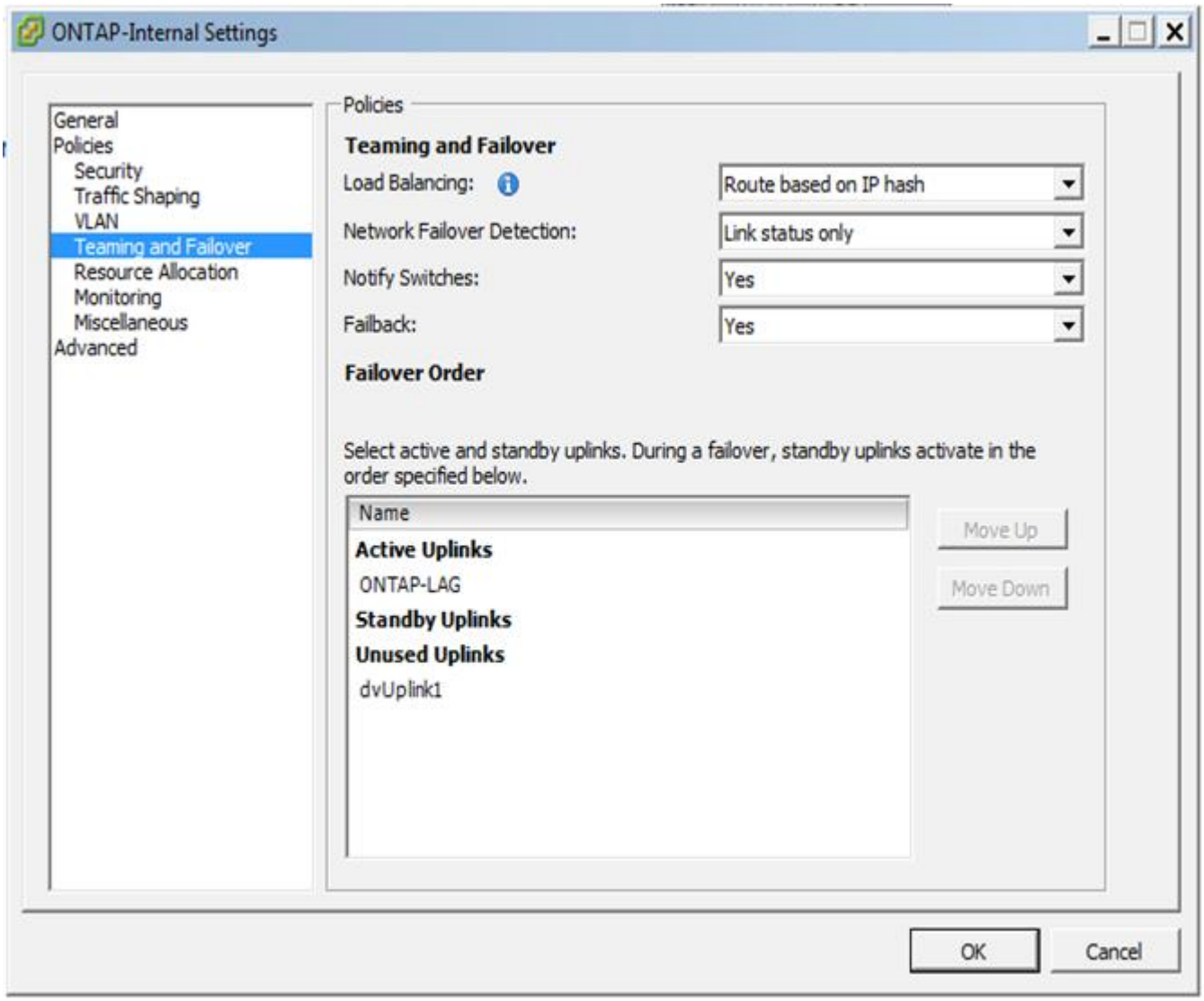
使用 **LACP** 時的 **LAG** 屬性



使用啟用 **LACP** 的分散式 **vSwitch** 進行外部連接埠群組組態



使用啟用 **LACP** 的分散式 **vSwitch** 進行內部連接埠群組組態



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 中啟用此設定之前，請確保已正確設定啟用 LACP 的連接埠通道。

ONTAP Select 實體交換器組態

基於單交換器和多交換器環境的上游實體交換器組態詳細資料。

在製定從虛擬交換器層到實體交換器的連線決策時，應仔細考慮。內部叢集流量與外部資料服務的隔離應延伸至上游實體網路層，並透過第 2 層 VLAN 提供的隔離來實現。

實體交換器連接埠應配置為主幹連接埠。ONTAP Select 外部流量可以透過兩種方式在多個第二層網路中進行分隔。一種方法是使用 ONTAP VLAN 標記虛擬連接埠與單一連接埠群組。另一種方法是在 VST 模式下為管理連接埠 e0a 指派單獨的連接埠群組。您還必須根據 ONTAP Select 版本以及單節點或多節點組態，將資料連接埠指派給 e0b 和 e0c/e0g。如果外部流量在多個第二層網路中進行分隔，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。

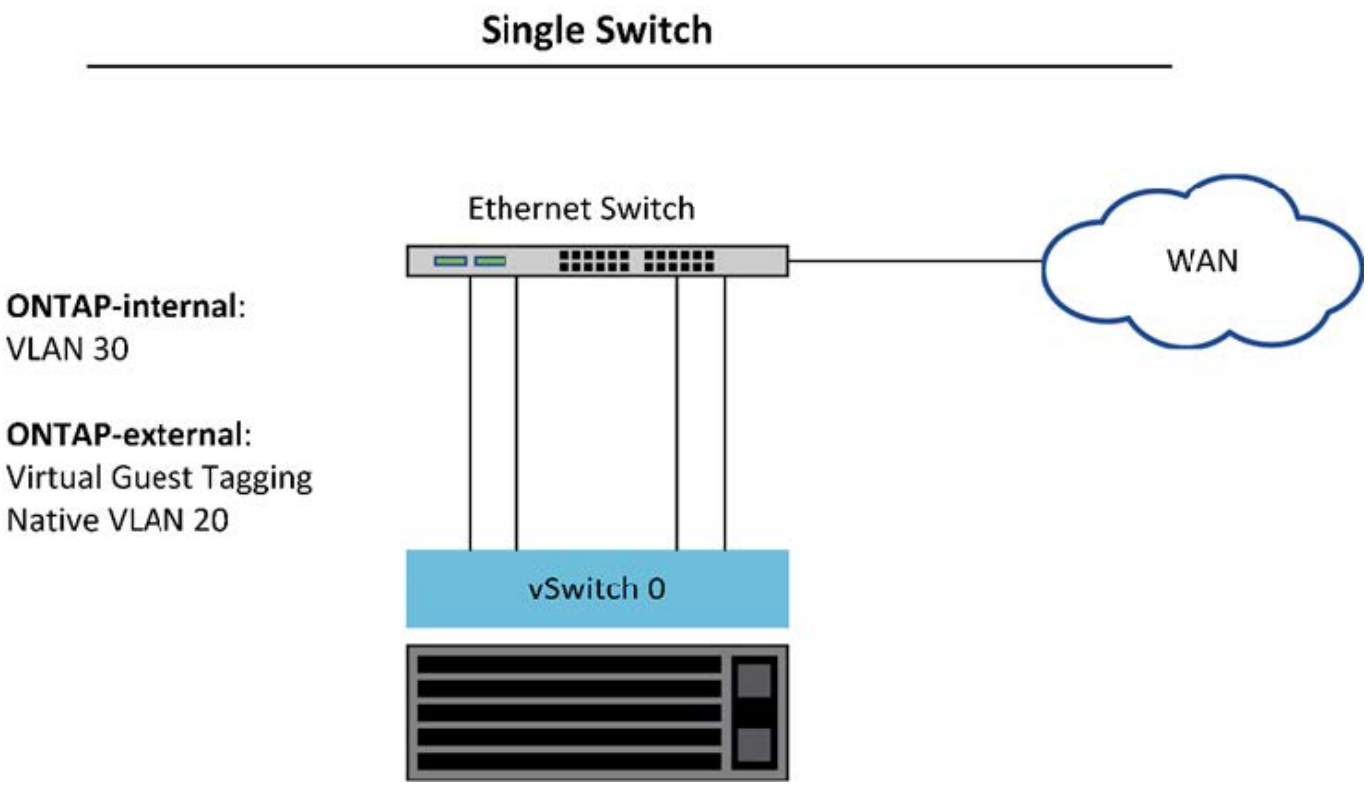
ONTAP Select 內部網路流量使用由連結本機 IP 位址定義的虛擬介面進行傳輸。由於這些 IP 位址不可路由，因此叢集節點之間的內部流量必須透過單一二層網路傳輸。ONTAP Select 叢集節點之間不支援路由躍點。

共享實體交換器

下圖展示了多節點 ONTAP Select 叢集中單一節點可能使用的交換器配置。在本例中，vSwitches 承載內部和外部網路連接埠群組的實體 NIC 均連接到同一上游交換器。交換器流量透過位於不同 VLAN 中的廣播網域進行隔離。

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

使用共用實體交換器進行網路組態

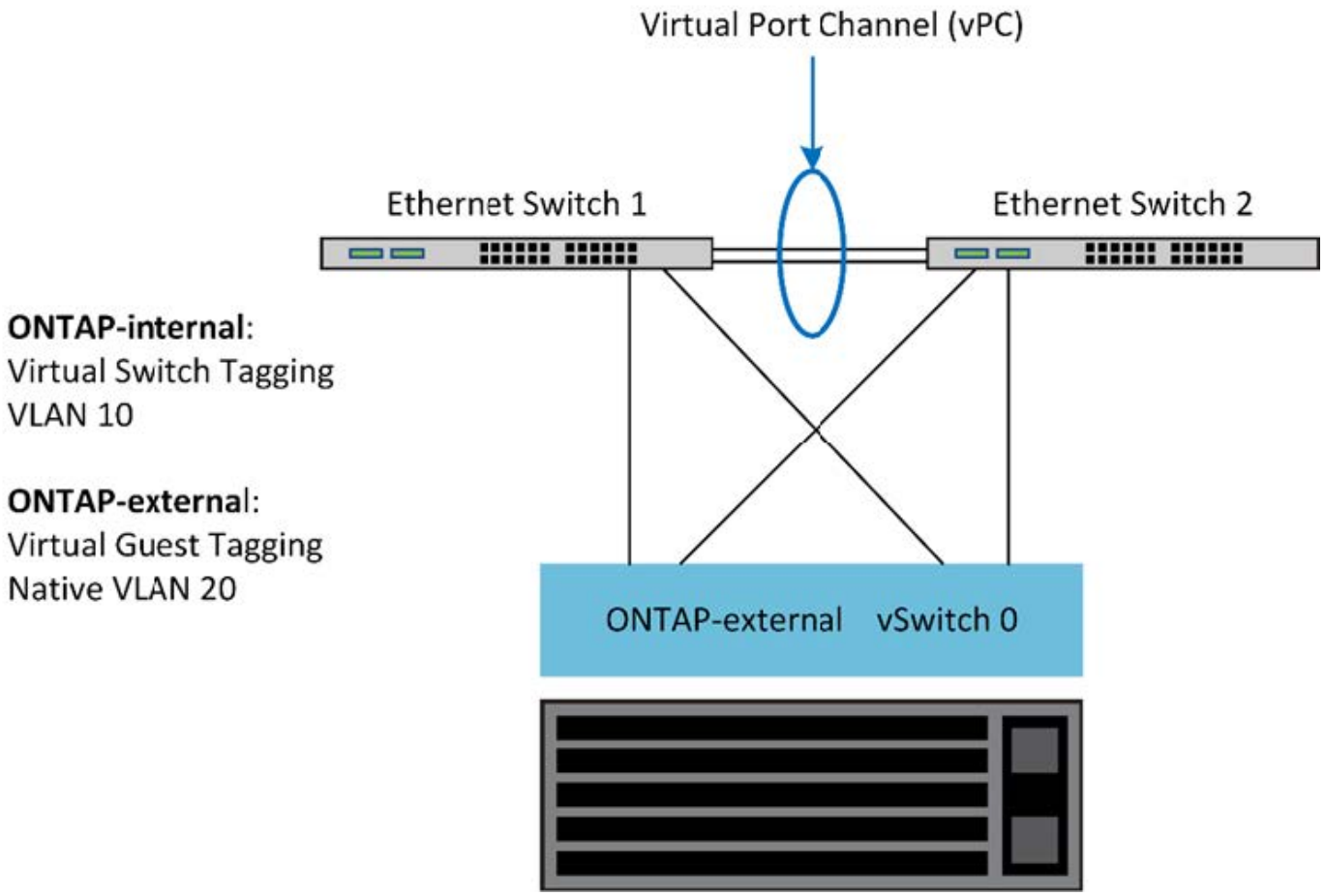


在此組態中，共享交換器會成為單點故障。如果可能、應使用多個交換器、以防止實體硬體故障導致叢集網路中斷。

多個實體交換器

當需要冗餘時，應使用多個實體網路交換器。下圖展示了多節點 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 NIC 分別連接到不同的實體交換器，從而避免單一硬體交換器故障對使用者造成影響。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組態



ONTAP Select 資料和管理流量分離

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）、管理和複寫（SnapMirror）流量。在 ONTAP 叢集中，每種流量類型都使用個別的邏輯介面，這些介面必須託管在虛擬網路連接埠上。在 ONTAP Select 的多節點組態中，這些連接埠被指定為 e0a 和 e0b/e0g。在單節點組態中，這些連接埠被指定為 e0a 和 e0b/e0c，其餘連接埠則保留給內部叢集服務使用。

NetApp 建議將資料流量與管理流量隔離到不同的第二層網路中。在 ONTAP Select 環境中，這是透過使用 VLAN 標籤來實現的。這可以透過將帶有 VLAN 標籤的埠群組指派給網路介面卡 1（埠 e0a）以進行管理流量來達成。然後，你可以將不同的埠群組指派給 e0b 和 e0c（單節點叢集）以及 e0b 和 e0g（多節點叢集）這些埠，以處理資料流量。

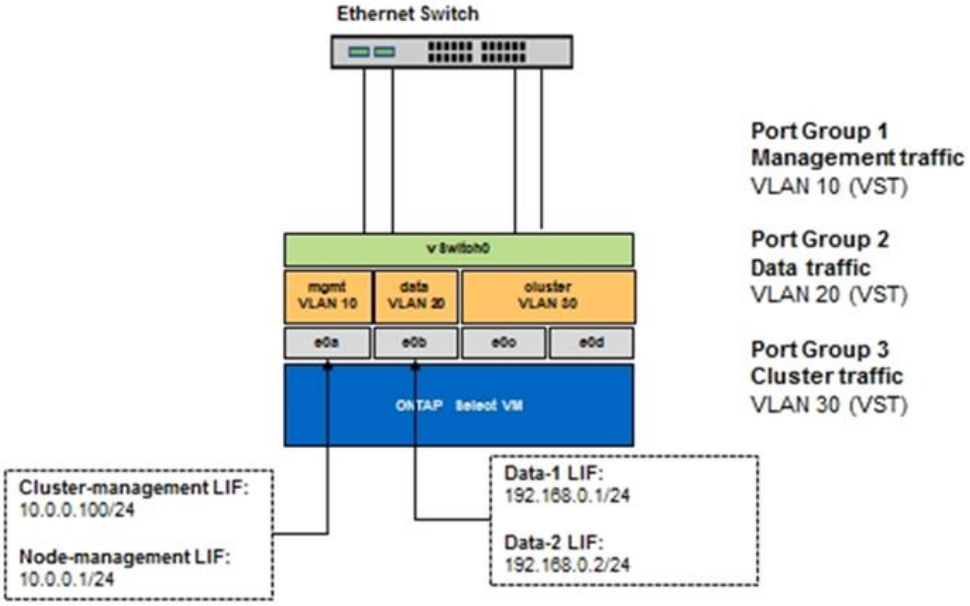
如果本文件前面所述的 VST 解決方案不足以滿足需求，則可能需要將資料 LIF 和管理 LIF 放在同一個虛擬連接埠上。為此，可以使用稱為 VGT 的流程，該流程由 VM 執行 VLAN 標記。

使用 ONTAP Deploy 公用程式時，無法透過 VGT 實現資料網路和管理網路的分離。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，需要注意以下事項。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軟體的連線。因此，對應到節點管理 LIF（連接埠 e0a）的連接埠群組僅支援 EST 和 VST 標記。此外，如果管理流量和資料流量都使用相同連接埠群組，則整個雙節點叢集僅支援 EST/VST。

VST 和 VGT 兩種組態選項均受支援。下圖展示了第一種方案 VST，其中流量透過分配的連接埠群組在 vSwitch 層級上進行標記。在此組態中，叢集和節點管理 LIF 被指派到 ONTAP 連接埠 e0a，並透過指派的連接埠群組標記 VLAN ID 10。資料 LIF 被分配到連接埠 e0b 以及 e0c 或 e0g，並使用第二個連接埠群組賦予 VLAN ID 20。叢集連接埠使用第三個連接埠群組，並位於 VLAN ID 30 上。

使用 **VST** 進行資料和管理分離

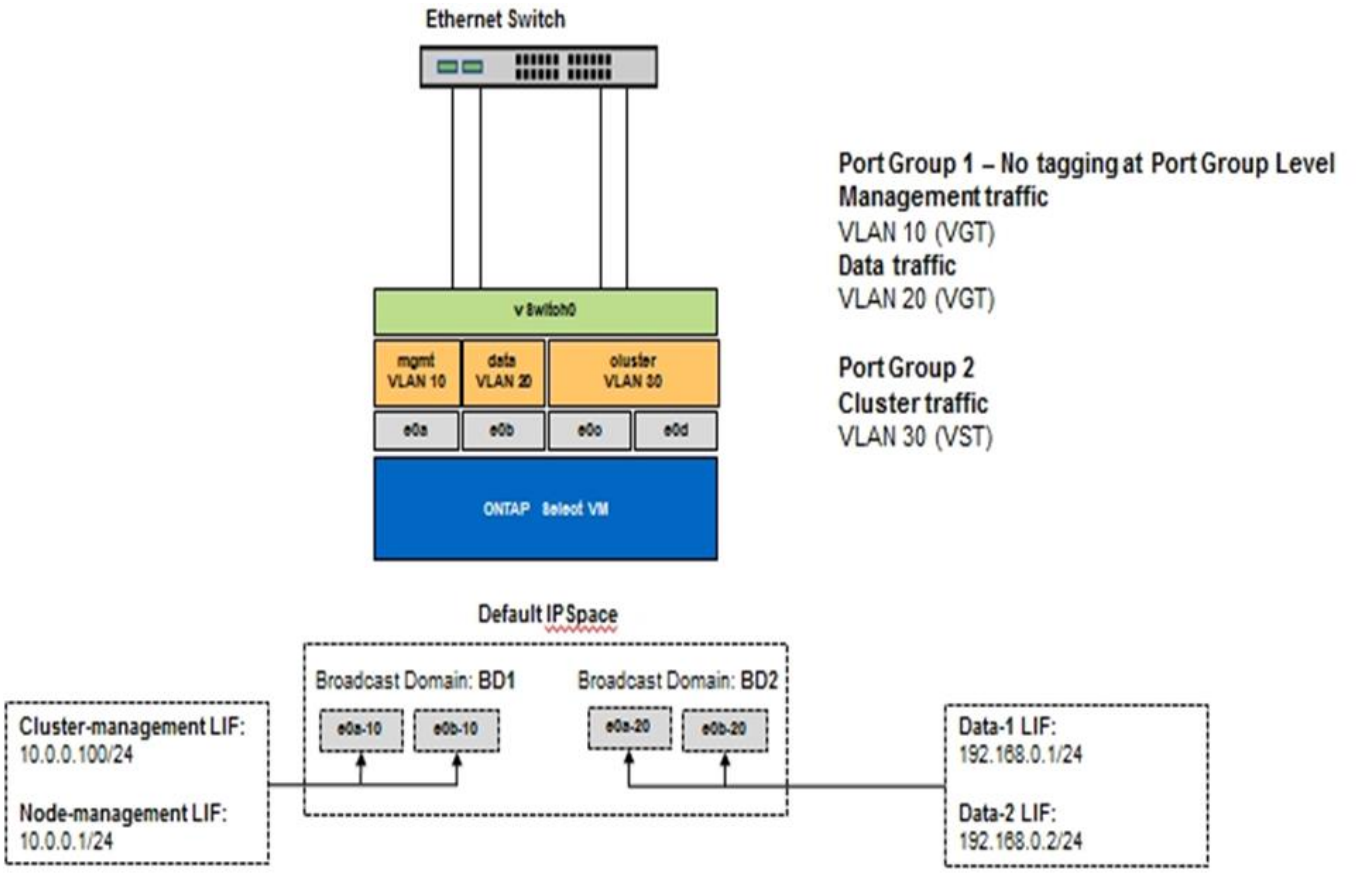


下圖展示了第二種方案，即 VGT，其中 ONTAP VM 使用位於不同廣播網域中的 VLAN 連接埠對流量進行標記。在本例中，虛擬連接埠 e0a-10/e0b-10/(e0c 或 e0g)-10 和 e0a-20/e0b-20 分別位於 VM 連接埠 e0a 和 e0b 之上。這種組態允許直接在 ONTAP 內部執行網路標記，而無需在 vSwitch 層進行標記。管理和資料 LIF 位於這些虛擬連接埠上，允許在單一 VM 連接埠內進行進一步的第 2 層細分。叢集 VLAN（VLAN ID 30）仍然在連接埠群組上進行標記。

附註：

- 當使用多個 IPspace 時，這種組態方式特別理想。如果需要進一步的邏輯隔離和多租戶，可將 VLAN 連接埠分組到個別的自訂 IPspace 中。
- 若要支援 VGT，ESXi/ESX 主機網路介面卡必須連接至實體交換器上的 trunk port。連接至虛擬交換器的 port group 必須將其 VLAN ID 設為 4095，才能在 port group 上啟用 trunk 功能。

使用 **VGT** 進行資料和管理分離



高可用性架構

ONTAP Select 高可用性組態

探索高可用性選項，為您的環境選擇最佳的 HA 組態。

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在通用硬體上執行的軟體型解決方案，但對於恢復能力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 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所導致的資料遺失。

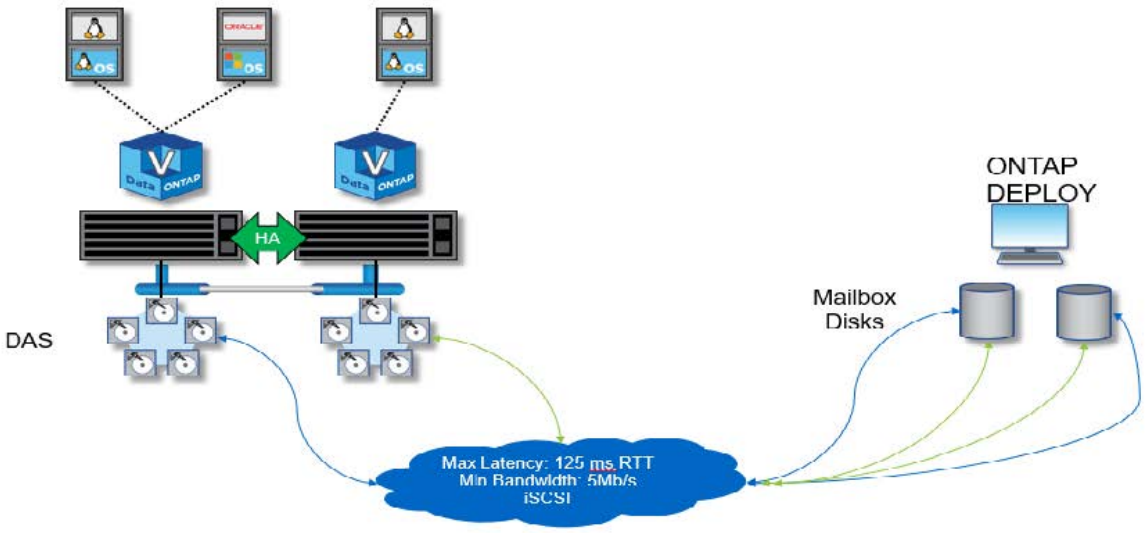
軟體定義儲存（SDS）市場很大一部分基於「無共享」儲存概念，透過軟體複寫在不同儲存孤島中儲存使用者資料的多個副本來提供資料恢復能力。ONTAP Select 在此基礎上，利用 ONTAP 提供的同步複寫功能（RAID SyncMirror）在叢集內儲存額外的使用者資料副本。這發生在 HA 配對的環境中。每個 HA 配對儲存兩個使用者資料副本：一個儲存在本地節點提供的儲存上，另一個儲存在 HA 合作夥伴提供的儲存上。在 ONTAP Select 叢集中，HA 和同步複寫緊密結合，二者的功能無法分離或獨立使用。因此，同步複寫功能僅在多節點版本中可用。

> ⓘ 在 ONTAP Select 叢集中，同步複寫功能是 HA 實作的一部分，而不是非同步 SnapMirror 或 SnapVault 複寫引擎的替代品。同步複寫不能脫離 HA 單獨使用。

ONTAP Select HA 部署模型有兩種：多節點叢集（四節點、六節點、八節點、十節點或十二節點）和雙節點叢集。雙節點 ONTAP Select 叢集的主要特點是使用外部中介服務來解決腦裂問題。ONTAP Deploy VM 作為其配置的所有雙節點 HA 配對的預設中介。

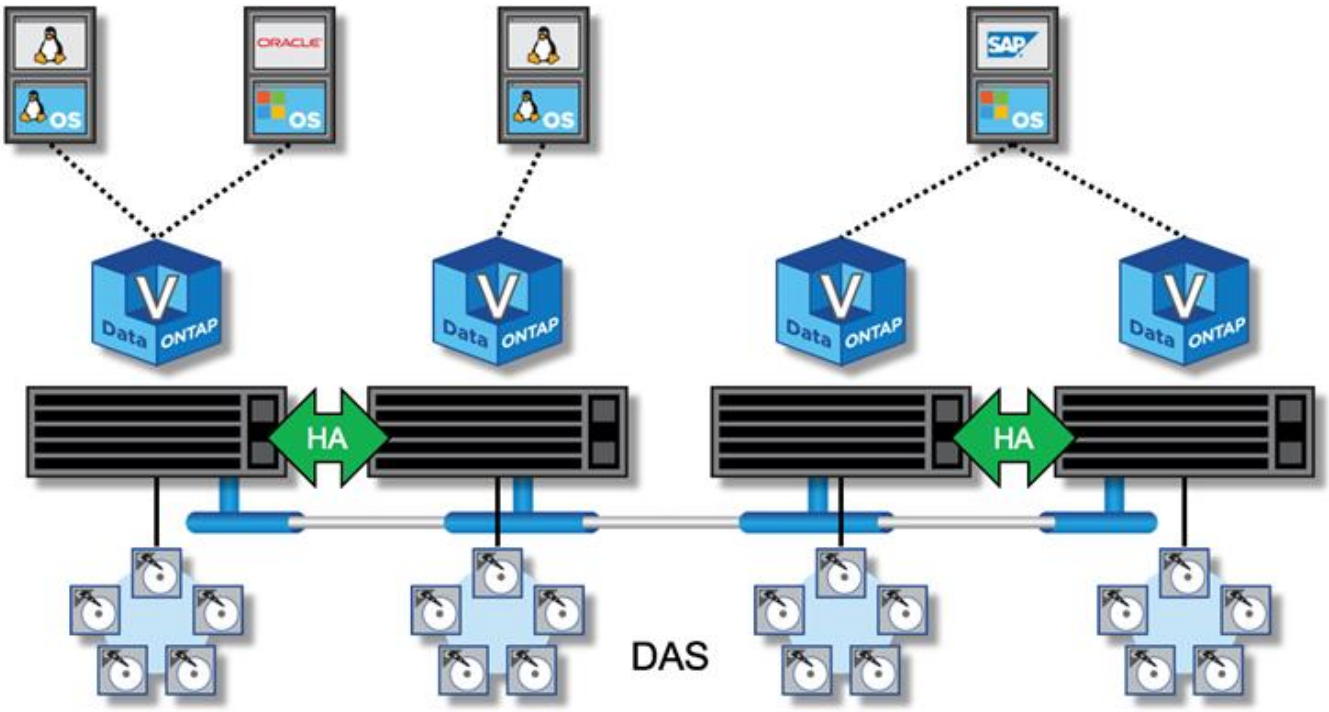
這兩種架構如下圖所示。

採用遠端中介器與本地附加儲存的雙節點 **ONTAP Select** 叢集



雙節點 ONTAP Select 叢集由一個 HA 配對和一個中介節點組成。在 HA 配對內，每個叢集節點上的資料集合體是同步鏡射的，因此，在發生容錯移轉時，不會遺失資料。

使用本機附加儲存設備的四節點 **ONTAP Select** 叢集



- 四節點 ONTAP Select 叢集由兩個 HA 配對組成。六節點、八節點、十節點和十二節點叢集分別由三個、四個、五個和六個 HA 配對組成。在每個 HA 配對內，每個叢集節點上的資料集合體都會進行同步鏡射，因此在發生容錯移轉時，不會遺失任何資料。
- 使用 DAS 儲存時，一台實體伺服器上只能存在一個 ONTAP Select 執行個體。ONTAP Select 需要對系統本機 RAID 控制器擁有非共用存取權，並且旨在管理本機連接的磁碟，如果沒有與儲存裝置的實體連接，則無

法實現此功能。

雙節點 **HA** 與多節點 **HA**

與 FAS 陣列不同，HA 配對中的 ONTAP Select 節點僅透過 IP 網路進行通訊。這意味著 IP 網路是單點故障（SPOF），因此防止網路分割和裂腦情況成為設計的重要層面。多節點叢集可以承受單節點故障，因為叢集仲裁可由三個或更多倖存節點建立。雙節點叢集依賴 ONTAP Deploy VM 託管的中介服務來達成相同的結果。

ONTAP Select 節點與 ONTAP Deploy 中介服務之間的心跳網路流量極少且具有彈性,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中。

> ⓘ 當 ONTAP Deploy 虛擬機器作為雙節點叢集的中介節點時，它就成為該叢集不可或缺的一部分。如果中介節點服務無法使用，雙節點叢集仍會繼續提供資料服務，但 ONTAP Select 叢集的儲存故障轉移功能將會停用。因此，ONTAP Deploy 中介節點服務必須與 HA 配對中的每個 ONTAP Select 節點保持持續通訊。為了確保叢集仲裁正常運作，需要至少 5Mbps 的頻寬和 125ms 的最大往返時間（RTT）延遲。

如果作為中介節點的 ONTAP Deploy VM 暫時或可能永久無法使用，則可以使用次要 ONTAP Deploy VM 來還原雙節點叢集仲裁。這會導致新的 ONTAP Deploy VM 無法管理 ONTAP Select 節點的組態，但它可以成功參與叢集仲裁演算法。ONTAP Select 節點和 ONTAP Deploy VM 之間的通訊是透過 IPv4 上的 iSCSI 協定進行。ONTAP Select 節點管理 IP 位址是啟動器、ONTAP Deploy VM IP 位址是目標。因此，在建立雙節點叢集時，無法支援節點管理 IP 位址的 IPv6 位址。ONTAP Deploy 託管的信箱磁碟會在建立雙節點叢集時自動建立並遮罩至適當的 ONTAP Select 節點管理 IP 位址。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 ONTAP Deploy 執行個體是該叢集的預設中介節點。

如果需要變更原始中介器位置，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 建議您在每次實例化雙節點叢集後備份 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 叢集擴展到更大的距離，甚至可以將每個節點放置在不同的資料中心。雙節點叢集和雙節點擴展叢集（也稱為 MetroCluster SDS）之間的唯一區別在於節點之間的網路連線距離。

雙節點叢集是指兩個節點位於同一資料中心內，且距離不超過 300 公尺的叢集。通常情況下，兩個節點都透過上行鏈路連接到同一網路交換器或一組交換器間連結（ISL）網路交換器。

雙節點 MetroCluster SDS 是指節點之間物理距離超過 300 公尺（位於不同的房間、不同的建築物或不同的資料中心）的叢集。此外，每個節點的上行鏈路都連接到獨立的網路交換器。MetroCluster SDS 不需要專用硬體。但是，環境必須滿足延遲要求（往返時間最大為 5 毫秒，抖動最大為 5 毫秒，總計 10 毫秒）。

MetroCluster SDS 是一項進階功能，需要 Premium 授權或 Premium XL 授權。Premium 授權支援建立小型和中型 VM，以及 HDD 和 SSD 媒體。Premium XL 授權還支援建立 NVMe 磁碟機。

> ⓘ MetroCluster SDS 同時支援本地附加儲存 (DAS) 和共用儲存 (vNAS)。請注意，由於 ONTAP Select 虛擬機與共用儲存之間存在網路，vNAS 配置通常具有較高的固有延遲。MetroCluster SDS 配置必須確保節點之間的延遲（包括共用儲存延遲）不超過 10 毫秒。換句話說，僅測量 Select 虛擬機器之間的延遲是不夠的，因為對於這些配置而言，共用儲存延遲不可忽略。

ONTAP Select HA RSM 和鏡像 **Aggregate**

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

同步複寫

ONTAP HA 模型是基於 HA 合作夥伴的概念建構。ONTAP Select 將此架構擴展到非共享的通用伺服器領域，它利用 ONTAP 中存在的 RAID SyncMirror (RSM) 功能在叢集節點之間複製資料區塊，從而在 HA 對中提供兩份使用者資料副本。

一個包含中間節點的雙節點叢集可以跨越兩個資料中心。如需更多資訊，請參閱 "雙節點延伸 HA（MetroCluster SDS）最佳實務做法" 章節。

鏡射 **Aggregate**

ONTAP Select 叢集由 2 到 12 個節點組成。每個高可用性（HA）配對包含兩份使用者資料副本，透過 IP 網路在節點間同步鏡像。這種鏡像對使用者而言是透明的，它是資料 Aggregate 的屬性，在資料 Aggregate 建立過程中自動設定。

ONTAP Select 叢集中的所有 Aggregate 都必須進行鏡像，以確保在節點容錯移轉時資料可用，並避免硬體故障導致的單點故障（SPOF）。ONTAP Select 叢集中的 Aggregate 由 HA 配對中每個節點提供的虛擬磁碟構建，並使用下列磁碟：

- 本機磁碟集（由目前 ONTAP Select 節點提供）
- 一組鏡像磁碟（由目前節點的 HA 合作夥伴貢獻）

> ⓘ 用於建構鏡像 Aggregate 的本機磁碟和鏡像磁碟必須大小相同。這些 Aggregate 分別稱為 plex 0 和 plex 1（分別表示本機鏡像配對和遠端鏡像配對）。實際的 plex 編號在您的安裝中可能有所不同。

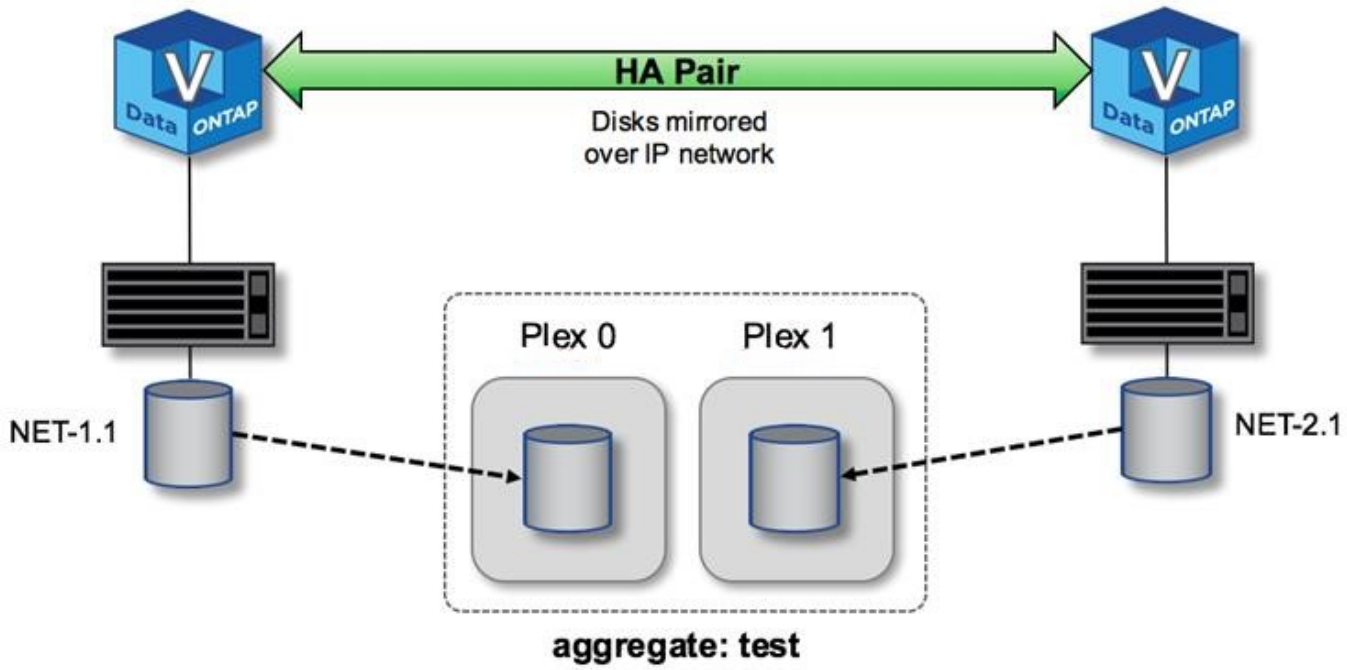
這種方法與標準 ONTAP 叢集的工作方式截然不同。這適用於 ONTAP Select 叢集中的所有根磁碟和資料磁碟。Aggregate 包含資料的本機副本和鏡像副本。因此，包含 N 個虛擬磁碟的 Aggregate 可提供 N/2 個磁碟的獨立儲存空間，因為第二個資料副本位於其自身的獨立磁碟上。

下圖顯示四節點 ONTAP Select 叢集中的 HA 配對。此叢集中有一個 Aggregate（test），使用來自兩個 HA 合作夥伴的儲存設備。此資料 Aggregate 由兩組虛擬磁碟組成：本機集由擁有叢集節點的 ONTAP Select（Plex 0）提供，遠端集由容錯移轉合作夥伴（Plex 1）提供。

Plex 0 是存放所有本機磁碟的儲存桶。Plex 1 是存放鏡像磁碟的儲存桶，這些磁碟負責儲存使用者資料的第二個複寫副本。擁有 Aggregate 的節點向 Plex 0 提供磁碟，而該節點的 HA 合作夥伴節點則向 Plex 1 提供磁碟。

在下圖中、有一個包含兩個磁碟的鏡射 Aggregate。此 Aggregate 的內容會在兩個叢集節點之間進行鏡射、本機磁碟 NET-1.1 放置在 Plex 0 儲存區中、遠端磁碟 NET-2.1 放置在 Plex 1 儲存區中。在此範例中、Aggregate test 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。

*ONTAP Select 鏡像
Aggregate*



部署 ONTAP Select 叢集時，系統上的所有虛擬磁碟都會自動指派到正確的 plex，使用者無需進行任何額外的磁碟分配操作。這可以防止磁碟被意外分配到錯誤的 plex，並提供最佳的鏡像磁碟配置。

寫入路徑

叢集節點之間資料區塊的同步鏡像、以及系統故障時不遺失資料的要求、對於傳入寫入作業在 ONTAP Select 叢集中傳播時所採取的路徑有重大影響。此程序包含兩個階段：

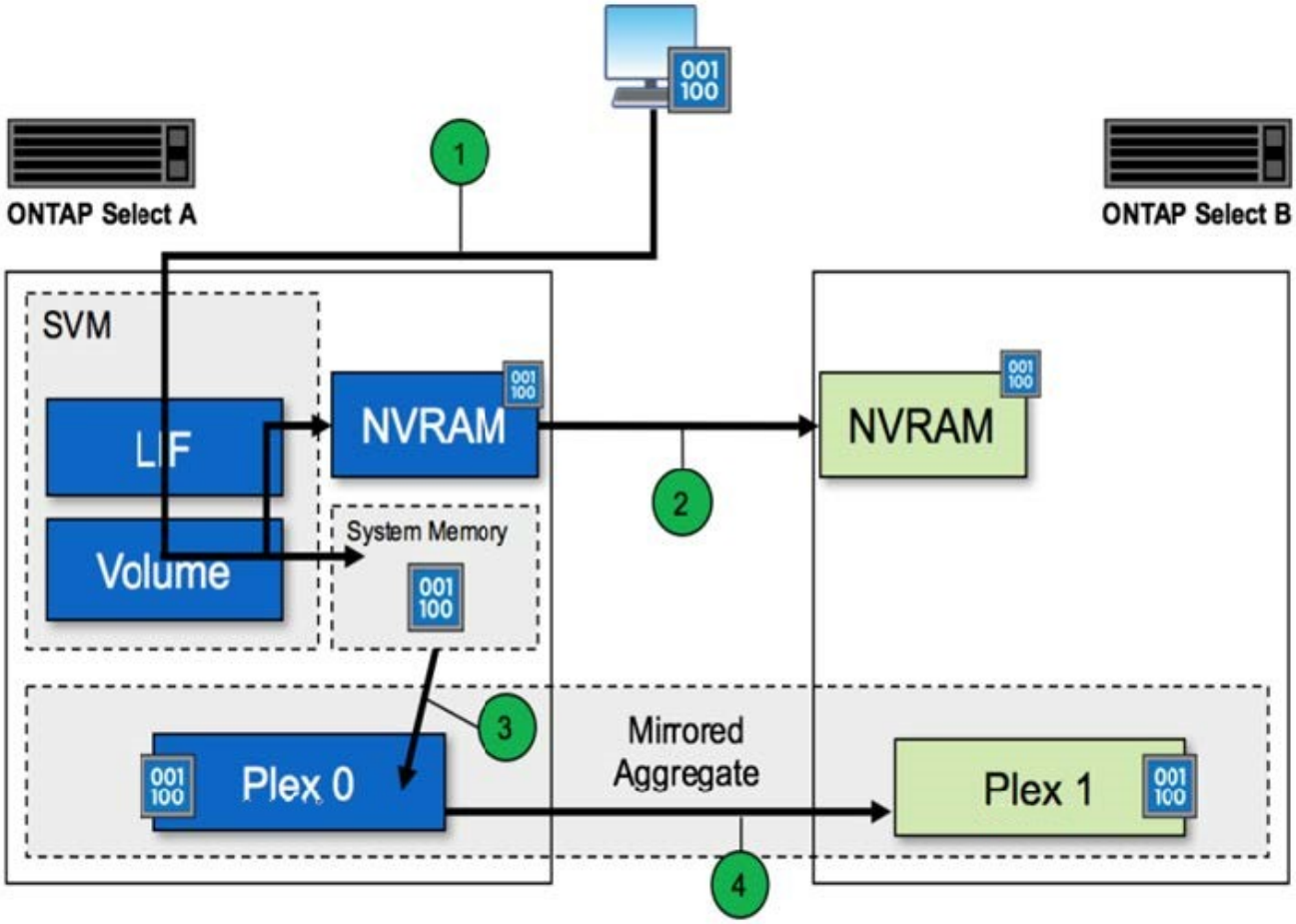
- 致謝
- 取消暫存

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，之後才會確認並傳回給客戶端。在 HA 配置中，由於這些 NVRAM 寫入作業在確認之前會立即鏡像到目標磁碟區擁有者的 HA 夥伴節點，因此會增加一個步驟。此程序可確保在原始節點發生硬體故障時，HA 夥伴節點上的檔案系統一致性。

寫入作業提交到 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「移出」。此程序僅在擁有目標磁碟區的叢集節點上執行一次，不會在 HA 合作夥伴節點上執行。

下圖顯示傳入 ONTAP Select 節點之寫入要求的寫入路徑。

ONTAP Select 寫入路徑工作流程



傳入寫入確認包括以下步驟：

- 寫入操作透過 ONTAP Select 節點 A 擁有的邏輯介面進入系統。
- 寫入作業會提交到節點 A 的 NVRAM,並鏡像到 HA 合作夥伴節點 B。
- 當 I/O 請求同時出現在兩個 HA 節點上時，該請求就會被確認並傳回給用戶端。

ONTAP Select 從 NVRAM 移轉至資料 Aggregate（ONTAP CP）包括下列步驟：

- 寫入操作從虛擬 NVRAM 轉移到虛擬資料集合體。
- 鏡像引擎會同步將區塊複寫到兩個 plex。

ONTAP Select HA 的資料保護增強功能

高可用性（HA）磁碟活動訊號、HA 信箱、HA 活動訊號、HA 容錯移轉和恢復可共同增強資料保護。

磁碟活動訊號

儘管 ONTAP Select HA 架構利用了傳統 FAS 陣列的許多程式碼路徑，但仍存在一些例外。其中一個例外是磁碟心跳機制的實作。磁碟心跳機制是一種非網路通訊方法，叢集節點使用它來防止網路隔離導致腦裂行為。腦裂是叢集分區的結果，通常由網路故障引起，在這種情況下，叢集中的一方會誤認為另一方已當機，並試圖接管叢集資源。

企業級 HA 實作必須能夠妥善處理此類場景。ONTAP 透過一種客製化的、基於磁碟的心跳機制來實現這一點。這是 HA 信箱的工作，它是實體儲存上的一個位置，叢集節點使用該位置來傳遞心跳訊息。這有助於叢集確定連線狀態，從而在發生容錯移轉時定義仲裁。

對於採用共享儲存高可用性架構的 FAS 陣列，ONTAP 透過以下方式解決腦裂問題：

- SCSI 持續保留
- 持續性 HA 中繼資料
- HA 狀態透過 HA 互連傳送

然而，在 ONTAP Select 叢集的無共用架構中，節點只能存取自身的本機儲存，而無法存取高可用性（HA）夥伴的本機儲存。因此，當網路分區將 HA 對的兩端隔離時，先前用於確定叢集仲裁和故障轉移行為的方法將不再適用。

儘管現有的腦裂檢測和規避方法無法使用，但仍需要一種調解方法，該方法必須符合無共享環境的限制。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機制。由於共用儲存不可用，調解是透過 NAS 存取郵箱磁碟來實現的。這些磁碟分佈在整個叢集中，包括雙節點叢集中的調解節點，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高可用性夥伴之外的其他節點的郵箱磁碟，則該節點很可能處於正常運作狀態。

ONTAP Select 之所以需要四個獨立的節點或一個用於雙節點叢集的中介器，是因為其採用了郵箱架構和基於磁碟的活動訊號機制來解決叢集仲裁和腦裂問題。

HA 信箱張貼

HA 郵箱架構使用訊息張貼模型。叢集節點會定期向叢集中所有其他郵箱磁碟（包括中介器）張貼訊息，表示該節點已啟動並正在執行。在健全叢集的任何時間點，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張貼的訊息。

每個 Select 叢集節點都連接著一個專門用於共用信箱存取的虛擬磁碟。該磁碟被稱為中介信箱磁碟，因為它的主要功能是在節點故障或網路分割時充當叢集的中介機制。此信箱磁碟包含每個叢集節點的分割區，並透過 iSCSI 網路由其他 Select 叢集節點掛載。這些節點會定期將健全狀況發佈到信箱磁碟的適當分割區。透過使用分佈在整個叢集中的網路可存取信箱磁碟，您可以透過可連線性矩陣推斷節點健全狀況。例如，叢集節點 A 和 B 可以發佈到叢集節點 D 的信箱，但無法發佈到節點 C 的信箱。此外，叢集節點 D 也無法發佈到節點 C 的信箱，因此節點 C 很可能已當機或網路隔離，應予以接管。

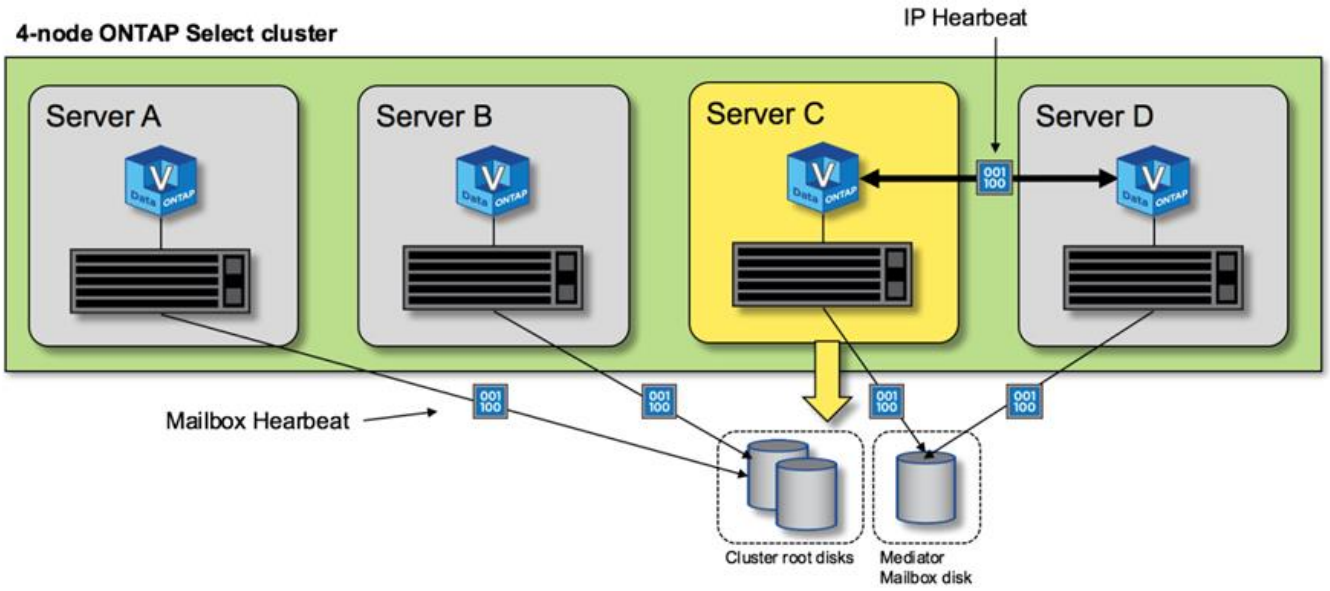
HA 心跳

與 NetApp FAS 平台類似，ONTAP Select 會定期透過 HA 互連發送 HA 心跳訊息。在 ONTAP Select 叢集內部，此程序透過 HA 夥伴之間的 TCP/IP 網路連線完成。此外，基於磁碟的心跳訊息也會傳送到所有 HA 郵箱磁碟，包括中介郵箱磁碟。這些訊息每隔幾秒鐘發送一次，並定期被讀取。這種頻繁的發送和接收使得 ONTAP Select 叢集能夠在大約 15 秒內偵測到 HA 故障事件，與 FAS 平台的偵測視窗相同。當不再讀取心跳訊息時，將觸發故障轉移事件。

下圖顯示從單一 ONTAP Select 叢集節點（節點 C）的角度，透過 HA 互連和中介磁碟傳送和接收活動訊號訊息的程序。

網路活動訊號會透過 HA 互連傳送至 HA 合作夥伴節點 D，而磁碟活動訊號則會使用所有叢集節點 A、B、C 和 D 的信箱磁碟。

四節點叢集中的 **HA** 心跳：穩定狀態



HA 容錯移轉與還原

在故障轉移操作期間，倖存節點會使用其高可用性 (HA) 夥伴節點資料的本機副本，承擔對等節點的資料服務職責。客戶端 I/O 可以不間斷地繼續運行，但必須先將資料變更複製回對等節點，才能進行復原操作。請注意，ONTAP Select 不支援強制恢復，因為這會導致倖存節點上儲存的變更遺失。

當重新開機的節點重新加入叢集時，會自動觸發同步回復作業。同步回復所需的時間取決於多種因素。這些因素包括必須複寫的變更數量、節點之間的網路延遲、以及每個節點上磁碟子系統的速度。同步回復所需的時間可能會超過 10 分鐘的自動回復時間範圍。在這種情況下，需要在同步回復之後手動回復。可以使用下列命令來監控同步回復的進度：

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

了解 **ONTAP Select** 的效能

ONTAP Select 叢集的效能會因底層硬體和組態的特性而有顯著差異。具體的硬體組態是影響特定 ONTAP Select 執行個體效能的最主要因素。以下是一些影響特定 ONTAP Select 執行個體效能的因素：

- 核心頻率。一般而言，頻率越高越好。
- 單路與多路。ONTAP Select 不使用多路功能，但支援多路配置的 Hypervisor 負荷會導致總效能出現一定偏差。
- **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量。
- **Hypervisor** 版本和修補程式等級。

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。