vSphere Metro Storage 叢集搭配 ONTAP

Enterprise applications

NetApp
May 09, 2024

This PDF was generated from https://docs.netapp.com/zh-tw/ontap-apps-dbs/vmware/vmware_vmsc_overview.html on May 09, 2024. Always check docs.netapp.com for the latest.

目錄

vSphere Metro Storage 叢集搭配 ONTAP

vSphere Metro Storage 叢集搭配 ONTAP

VMware 領先業界的 vSphere Hypervisor 可部署為稱為 vSphere Metro Storage Cluster （ VMSC ）的延伸叢集。

NetApp ® MetroCluster ™和 SnapMirror 主動同步（以前稱爲 SnapMirror 業務連續性或 SMBC ）均支持 VMSC 解決方案，如果一個或多個故障域發生整體中斷，則可提供高級業務連續性。不同故障模式的恢復能力取決於您選擇的組態選項。

適用於 **vSphere** 環境的持續可用度解決方案

ONTAP 架構是靈活且可擴充的儲存平台、可為資料存放區提供 SAN （ FCP 、 iSCSI 和 NVMe of ）和 NAS （ NFS v3 和 v4.1 ）服務。NetApp AFF 、 ASA 和 FAS 儲存系統使用 ONTAP 作業系統來提供額外的通訊協定、以供 S3 和 SMB/CIFS 等來賓儲存設備存取。

NetApp MetroCluster 使用 NetApp 的 HA （控制器容錯移轉或 CFO ）功能來防範控制器故障。它還包括本機 SyncMirror 技術、災難時的叢集容錯移轉（隨需控制器容錯移轉或 CFOD ）、硬體備援、以及地理區隔、以達到高可用度。SyncMirror 會將資料寫入兩個叢中、以同步鏡射 MetroCluster 組態的兩個部份資料：本機叢（位於本機櫃上）主動提供資料、而遠端叢（位於遠端機櫃上）通常不會提供資料。所有 MetroCluster 元件（例如控制器、儲存設備、纜線、交換器（與 Fabric MetroCluster 搭配使用）和介面卡）均具備硬體備援功能。

NetApp SnapMirror 主動式同步可透過 FCP 和 iSCSI SAN 傳輸協定提供資料存放區精細保護、讓您只能選擇性地保護高優先順序的工作負載。它提供對本機和遠端站台的主動式存取、而 NetApp MetroCluster 則是主動式待命解決方案。目前、主動式同步是一種非對稱式解決方案、其中一端較另一端更偏好、提供更好的效能。這是使用 ALUA （非對稱邏輯單元存取）功能來達成的、此功能會自動通知 ESXi 主機偏好的控制器。不過、 NetApp 已宣佈啟用主動式同步功能、即將啟用完全對稱的存取。

若要跨兩個站台建立 VMware HA/DRS 叢集、 ESXi 主機會由 vCenter Server Appliance （ VCSA ）使用和管理。vSphere 管理、 VMotion ® 和虛擬機器網路是透過兩個站台之間的備援網路連線。管理 HA/DRS 叢集的 vCenter Server 可連線至兩個站台的 ESXi 主機、並應使用 vCenter HA 進行設定。

請參閱 "如何在 vSphere Client 中建立和設定叢集" 設定 vCenter HA 。

您也應該參閱 "VMware vSphere Metro儲存叢集建議實務做法"。

什麼是 **vSphere Metro Storage Cluster** ？

vSphere Metro Storage Cluster （ VMSC ）是經過認證的組態、可保護虛擬機器（ VM ）和容器免於故障。這是透過使用延伸儲存概念和 ESXi 主機叢集來達成的、這些主機分佈在不同的故障網域、例如機架、建築物、校園或甚至城市。NetApp MetroCluster 和 SnapMirror 主動同步儲存技術可分別為主機叢集提供 RPO = 0 或近乎 RPO = 0 的保護。VMSC 組態的設計是為了確保即使完整的實體或邏輯「站台」故障、資料仍可隨時使用。在成功通過 VMSC 認證程序之後、必須通過 VMSC 組態一部分的儲存裝置認證。所有支援的儲存裝置都可以在中找到 "VMware 儲存相容性指南"。

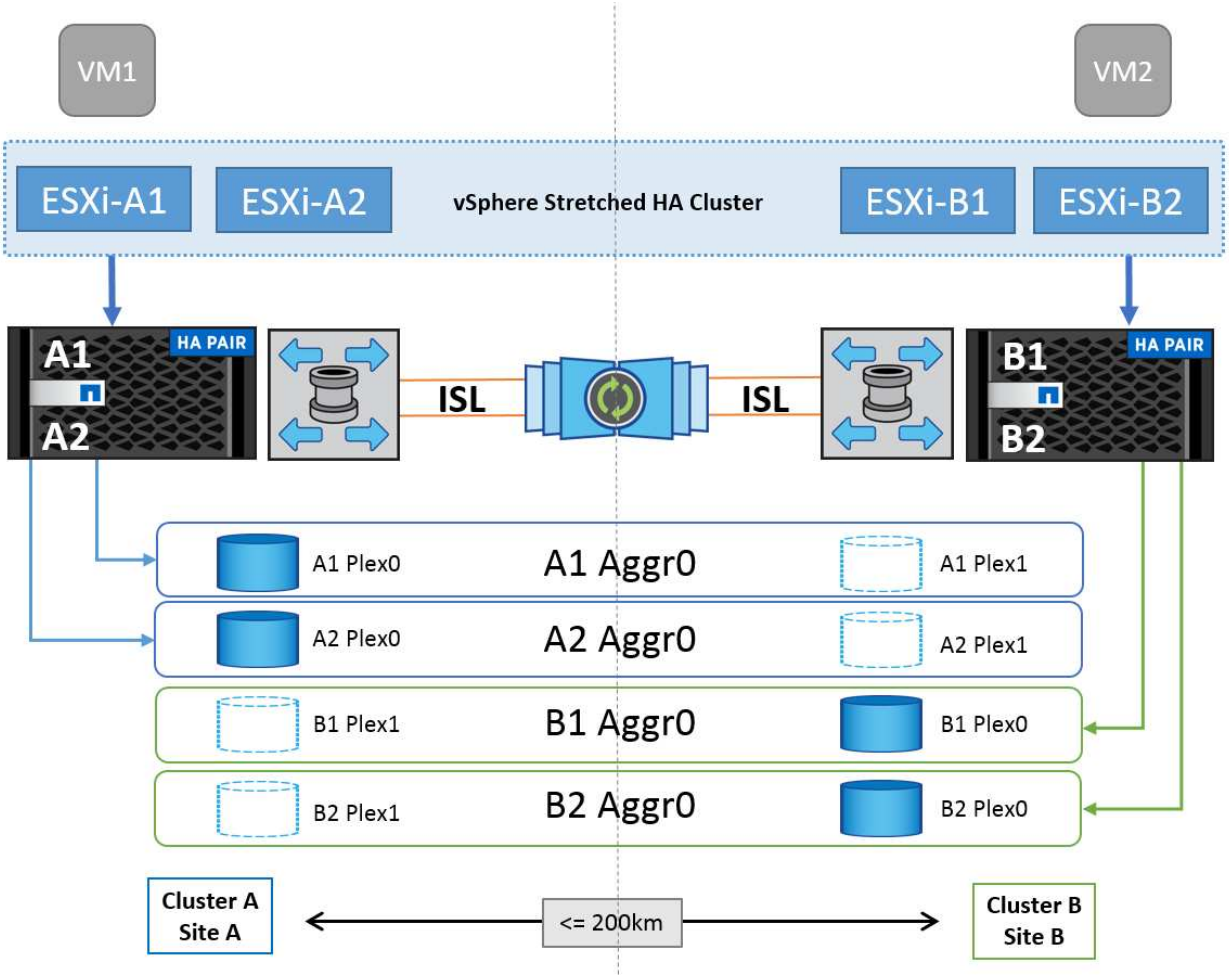
如果您想要更多有關 vSphere Metro Storage 叢集設計準則的資訊、請參閱下列文件：

- "VMware vSphere 支援 NetApp MetroCluster"
- "VMware vSphere 支援 NetApp SnapMirror 業務持續運作" （現在稱為 SnapMirror 主動同步）

視延遲考量因素而定、 NetApp MetroCluster 可部署在兩種不同的組態中、以搭配 vSphere 使用：

- Stretch MetroCluster
- Fabric MetroCluster

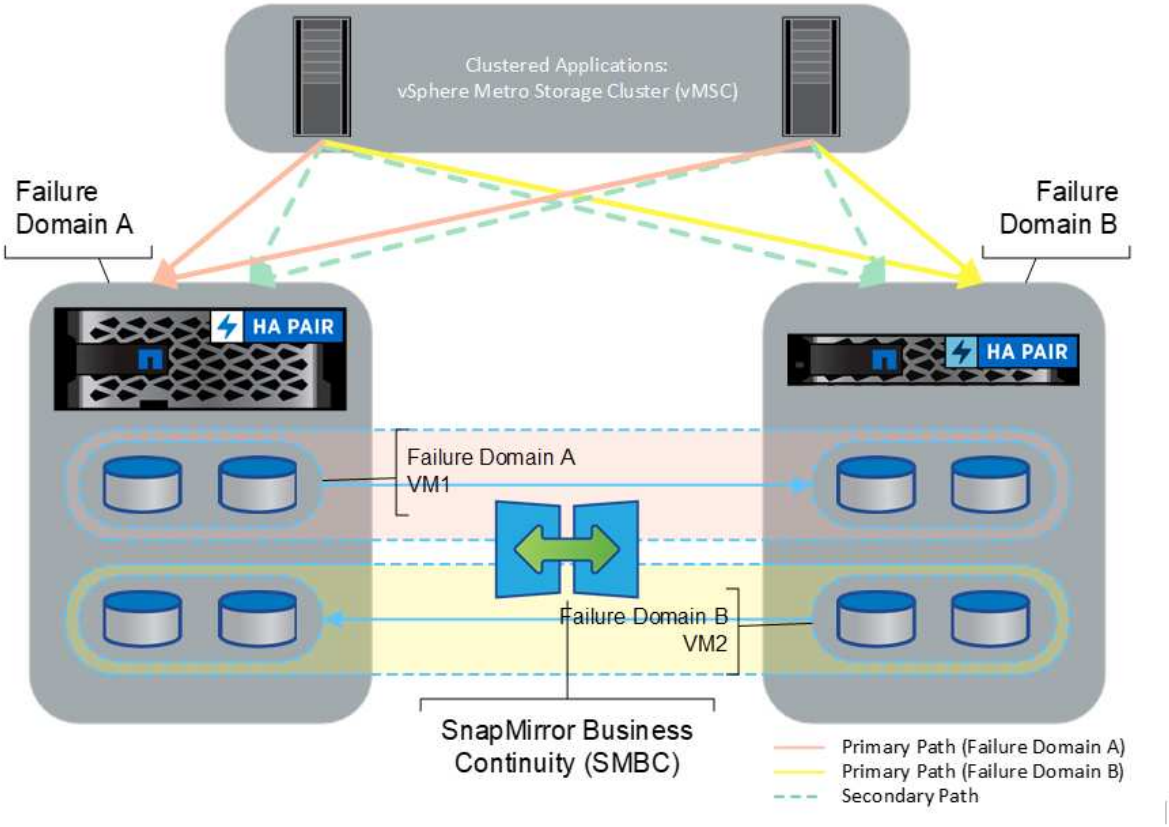
以下說明 Stretch MetroCluster 的高階拓撲圖。



請參閱 "本文檔MetroCluster" 取得 MetroCluster 的特定設計與部署資訊。

SnapMirror 主動式同步也可透過兩種不同方式部署。

- 非對稱
- 對稱（ ONTAP 9.14.1 中的私有預覽）



請參閱 "NetApp文件" 取得 SnapMirror 主動同步的特定設計與部署資訊。

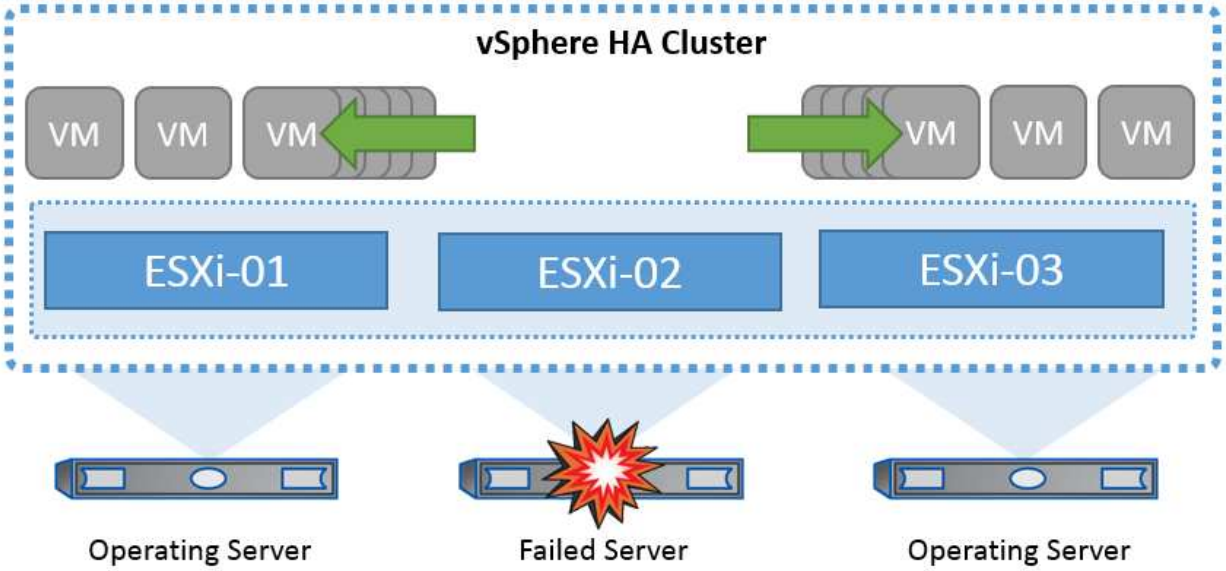
VMware vSphere 解決方案概觀

vCenter Server Appliance （ VCSA ）是強大的集中式管理系統、也是 vSphere 的單一窗口、可讓管理員有效操作 ESXi 叢集。它有助於執行重要功能、例如 VM 資源配置、 VMotion 作業、高可用度（ HA ）、分散式資源排程器（ DRS ）、 Tanzu Kubernetes Grid 等。它是 VMware 雲端環境中的重要元件、設計時應考慮到服務可用度。

vSphere 高可用度

VMware 的叢集技術可將 ESXi 伺服器分組到虛擬機器的共用資源集區中、並提供 vSphere High Availability （ HA ）。vSphere HA 可為在虛擬機器中執行的應用程式提供易於使用的高可用度。當叢集上啟用 HA 功能時、每部 ESXi 伺服器都會與其他主機保持通訊、以便在任何 ESXi 主機無回應或隔離時、 HA 叢集可在叢集中的未運作主機之間、協調在該 ESXi 主機上執行的虛擬機器的還原作業。萬一來賓作業系統發生故障、 vSphere HA 會在同一部實體伺服器上重新啟動受影響的虛擬機器。vSphere HA 可減少計畫性停機、避免非計畫性停機、並快速從停機中恢復。

vSphere HA 叢集可從故障伺服器還原 VM 。



請務必瞭解 VMware vSphere 不知道 NetApp MetroCluster 或 SnapMirror 主動同步、並視主機和 VM 群組關聯性組態而定、將 vSphere 叢集中的所有 ESXi 主機視為 HA 叢集作業的合格主機。

主機故障偵測

建立 HA 叢集之後、叢集中的所有主機都會參與選舉、其中一部主機即成為主主機。每個從屬設備都會對主主機執行網路活動訊號、而主設備則會在所有從屬主機上執行網路活動訊號。vSphere HA 叢集的主要主機負責偵測從屬主機的故障。

視偵測到的故障類型而定、在主機上執行的虛擬機器可能需要容錯移轉。

在 vSphere HA 叢集中、偵測到三種類型的主機故障：

- 故障：主機停止運作。
- 隔離：主機會變成網路隔離。
- 分割區 - 主機失去與主主機的網路連線。

主主機會監控叢集中的從屬主機。這種通訊是透過每秒交換網路訊號來完成。當主主機停止從從屬主機接收這些心跳時、它會在宣告主機故障之前先檢查主機的活動性。主要主機執行的活性檢查是判斷從屬主機是否與其中一個資料存放區交換活動訊號。此外、主主機會檢查主機是否回應傳送至其管理 IP 位址的 ICMP Ping 、以偵測其是否只是與主節點隔離、或完全與網路隔離。它會透過 ping 預設閘道來執行此作業。您可以手動指定一或多個隔離位址、以增強隔離驗證的可靠性。

最佳實務做法 _

NetApp 建議指定至少兩個額外的隔離位址、而且每個位址都是站台本機位址。這將提高隔離驗證的可靠性。

主機隔離回應

隔離回應是 vSphere HA 中的一項設定、可決定當 vSphere HA 叢集中的主機失去管理網路連線但仍繼續執行時、在虛擬機器上觸發的動作。此設定有三個選項：「已停用」、「關機並重新啟動 VM 」和「關機並重新啟動 VM 」。

「關機」比「關機」好、因為「關機」無法清除磁碟或認可交易的最新變更。如果虛擬機器在 300 秒內未關機、則會關閉電源。若要變更等待時間、請使用進階選項 das.isolationshutdowntimeout 。

在 HA 起始隔離回應之前、會先檢查 vSphere HA 主要代理程式是否擁有包含 VM 組態檔案的資料存放區。如果沒有、則主機不會觸發隔離回應、因為沒有主節點可重新啟動 VM 。主機會定期檢查資料存放區狀態、以判斷是否由擁有主角色的 vSphere HA 代理程式宣告。

最佳實務做法 _

NetApp 建議將「主機隔離回應」設定為「已停用」。

如果主機與 vSphere HA 主主機隔離或分割、而主主機無法透過心跳資料存放區或 ping 進行通訊、就可能發生分割腦部狀況。主機會宣告隔離的主機當機、並在叢集中的其他主機上重新啟動 VM 。現在存在分割腦狀況、因為有兩個執行中的虛擬機器執行個體、只有其中一個執行個體可以讀取或寫入虛擬磁碟。現在可以透過設定 VM 元件保護（ VMCP ）來避免發生大腦分裂的情況。

VM 元件保護（ VMCP ）

與 HA 相關的 vSphere 6 功能增強功能之一是 VMCP 。VMCP 針對區塊（ FC 、 iSCSI 、 FCoE ）和檔案儲存（ NFS ）、提供增強的保護、防止所有路徑中斷（ APD ）和永久裝置遺失（ PDL ）情況。

永久裝置遺失（ **PDL** ）

當儲存設備永久故障或被管理性移除、且不預期返回時、會發生 PDL 狀況。NetApp 儲存陣列會向 ESXi 發出 SCSI Sense 程式碼、聲明該裝置已永久遺失。在 vSphere HA 的「故障條件和 VM 回應」區段中、您可以設定在偵測到 PDL 條件後應回應的內容。

最佳實務做法 _

NetApp 建議將「使用 PDL 的資料存放區回應」設定為「 * 關閉並重新啟動 VM* 」。偵測到這種情況時、將會在 vSphere HA 叢集中的健全主機上立即重新啟動 VM 。

所有下行路徑（ **APD** ）

當主機無法存取儲存裝置、且沒有通往陣列的路徑可用時、便會發生 APD 狀況。ESXi 認為這是裝置的暫時性問題、因此預期裝置會再次出現。

偵測到 APD 狀況時、會啟動定時器。140 秒後、 APD 條件會正式宣告、且裝置會標示為 APD 逾時。140 秒過後、 HA 會開始計算 VM 容錯移轉 APD 延遲中指定的分鐘數。指定時間過後、 HA 會重新啟動受影響的虛擬機器。您可以設定 VMCP 在需要時以不同的方式回應（停用、問題事件、或關機和重新啟動 VM ）。

最佳實務做法 _

NetApp 建議將「使用 APD 的資料存放區回應」設定為「 * 關閉並重新啟動 VM （保守） * 」。

保守是指 HA 能夠重新啟動 VM 的可能性。如果設為保守、 HA 只會重新啟動受 APD 影響的 VM 、前提是它知道其他主機可以重新啟動。在積極的情況下、 HA 會嘗試重新啟動 VM 、即使它不知道其他主機的狀態。如果沒有可存取其所在資料存放區的主機、這可能導致 VM 無法重新啟動。

如果 APD 狀態已解決、且在逾時之前已還原對儲存設備的存取、則 HA 不會不必要地重新啟動虛擬機器、除非您明確將其設定為如此。如果即使環境已從 APD 條件恢復、仍需要回應、則 APD 逾時後的 APD 恢復回應應設定為重設虛擬機器。

最佳實務做法 _

NetApp 建議將 APD 逾時後的 APD 恢復回應設定為停用。

適用於 NetApp MetroCluster 的 VMware DRS 實作

VMware DRS 是一項功能、可將叢集中的主機資源集合在一起、主要用於在虛擬基礎架構中的叢集內進行負載平衡。VMware DRS 主要會計算 CPU 和記憶體資源、以便在叢集中執行負載平衡。由於 vSphere 不知道延伸叢集、因此在負載平衡時會考慮兩個站台中的所有主機。為了避免跨站台流量、 NetApp 建議您設定 DRS 關聯性規則、以管理虛擬機器的邏輯分隔。這可確保除非發生完整的站台故障、否則 HA 和 DRS 只會使用本機主機。

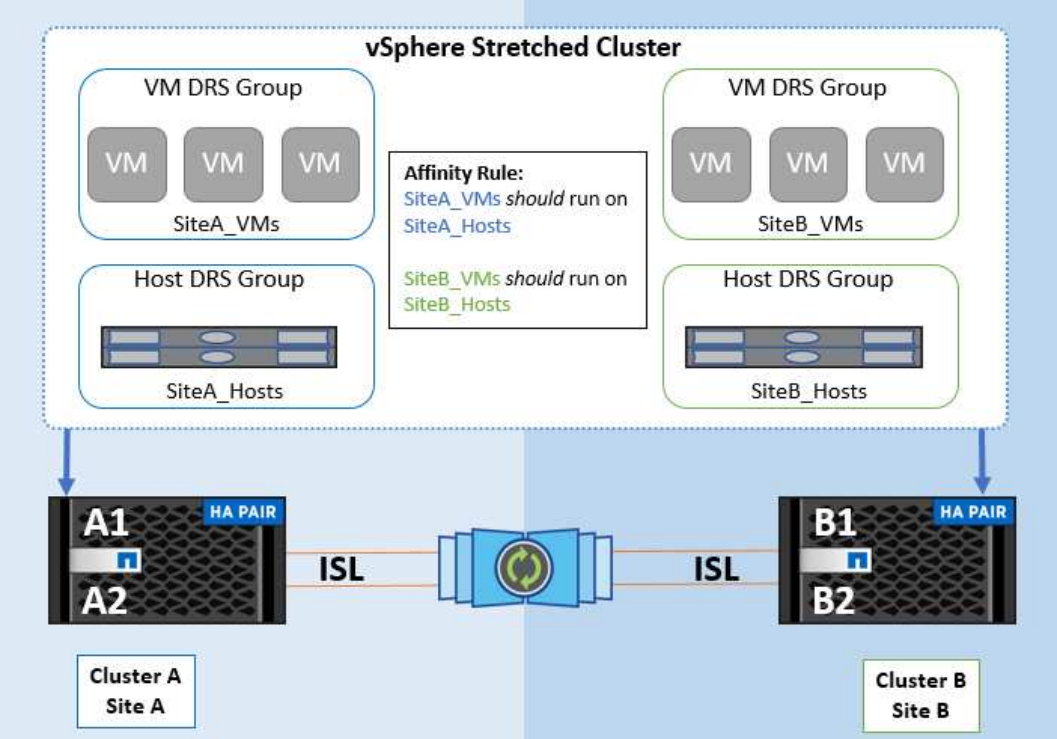
如果您為叢集建立 DRS 關聯性規則、您可以指定 vSphere 如何在虛擬機器容錯移轉期間套用該規則。

您可以指定 vSphere HA 容錯移轉行為的規則有兩種類型：

- VM 反關聯性規則會強制指定的虛擬機器在容錯移轉動作期間保持分離。
- VM 主機關聯性規則會在容錯移轉動作期間、將指定的虛擬機器放置在特定主機或已定義主機群組的成員上。

使用 VMware DRS 中的 VM 主機關聯性規則、可以在站台 A 和站台 B 之間有邏輯分隔、以便 VM 在主機上執行、而該主機與陣列是設定為指定資料存放區的主要讀取 / 寫入控制器。此外、 VM 主機關聯性規則可讓虛擬機器保持儲存設備的本機狀態、進而在站台之間發生網路故障時確定虛擬機器連線。

以下是 VM 主機群組和關聯規則的範例。



最佳實務做法 _

NetApp 建議實作「應該」規則、而非「必須」規則、因為在發生故障時、 vSphere HA 會違反這些規則。使用「必須」規則可能導致服務中斷。

服務的可用度應永遠高於效能。在完整資料中心故障的情況下、「必須」規則必須從 VM 主機關聯群組中選擇主機、而當資料中心無法使用時、虛擬機器將不會重新啟動。

使用 NetApp MetroCluster 實作 VMware Storage DRS

VMware Storage DRS 功能可將資料存放區集合至單一單元、並在超過儲存 I/O 控制臨界值時平衡虛擬機器磁碟。

依預設、啟用 Storage DRS 的 DRS 叢集會啟用儲存 I/O 控制。儲存 I/O 控制功能可讓管理員控制 I/O 壅塞期間分配給虛擬機器的儲存 I/O 數量、讓更重要的虛擬機器能夠優先選擇較不重要的虛擬機器來分配 I/O 資源。

Storage DRS 使用 Storage VMotion 將虛擬機器移轉至資料存放區叢集中的不同資料存放區。在 NetApp MetroCluster 環境中、必須在該站台的資料存放區內控制虛擬機器移轉。例如、在站台 A 的主機上執行的虛擬機器 A 、最好能在站台 A 的 SVM 資料存放區內移轉如果無法這麼做、虛擬機器將繼續運作、但效能降低、因為虛擬磁碟讀取 / 寫入將透過站台間連結來自站台 B 。

最佳實務做法 _

NetApp 建議針對儲存站台親和性建立資料存放區叢集、也就是說、站台 A 的站台親和性資料存放區不應與站台 B 具有站台親和性的資料存放區叢集混合使用

每當使用 Storage VMotion 新佈建或移轉虛擬機器時、 NetApp 建議手動更新這些虛擬機器的所有 VMware DRS 規則。這將確定主機和資料存放區在站台層級的虛擬機器關聯性、進而降低網路和儲存負荷。

VMSC 設計與實作準則

本文件概述 VMSC 搭配 ONTAP 儲存系統的設計與實作準則。

NetApp 儲存組態

NetApp MetroCluster 的設定指示（稱為 MCC 組態）可在以下網址取得： "資訊文件MetroCluster"。SnapMirror 主動同步的說明也可在以下網址取得： "SnapMirror營運不中斷總覽"。

設定 MetroCluster 之後、管理就像管理傳統的 ONTAP 環境一樣。您可以使用命令列介面（ CLI ）、系統管理員或 Ansible 等各種工具來設定儲存虛擬機器（ SVM ）。設定 SVM 後、在叢集上建立邏輯介面（生命體）、磁碟區和邏輯單元編號（ LUN ）、以用於正常作業。這些物件將會使用叢集對等網路自動複寫到其他叢集。

如果不使用 MetroCluster 、您可以使用 SnapMirror 主動式同步功能、在不同故障網域中的多個 ONTAP 叢集之間提供資料存放區精細保護和主動式存取。SnapMirror 主動式同步會使用一致性群組、確保一或多個資料存放區之間的寫入順序一致性、您可以根據應用程式和資料存放區需求、建立多個一致性群組。一致性群組對於需要在多個資料存放區之間進行資料同步的應用程式特別有用。SnapMirror 主動式同步也支援原始裝置對應（ RDM ）和來賓 iSCSI 啟動器的來賓連線儲存設備。如需更多關於一致性群組的資訊、請參閱 "一致性群組總覽"。

與 MetroCluster 相比、使用 SnapMirror 主動式同步管理 VMSC 組態有一些差異。首先、這是僅限 SAN 的組態、沒有 NFS 資料存放區可以使用 SnapMirror 主動式同步進行保護。其次、您必須將兩個 LUN 複本對應到 ESXi 主機、以便它們存取兩個故障網域中的複寫資料存放區。

VMware vSphere HA

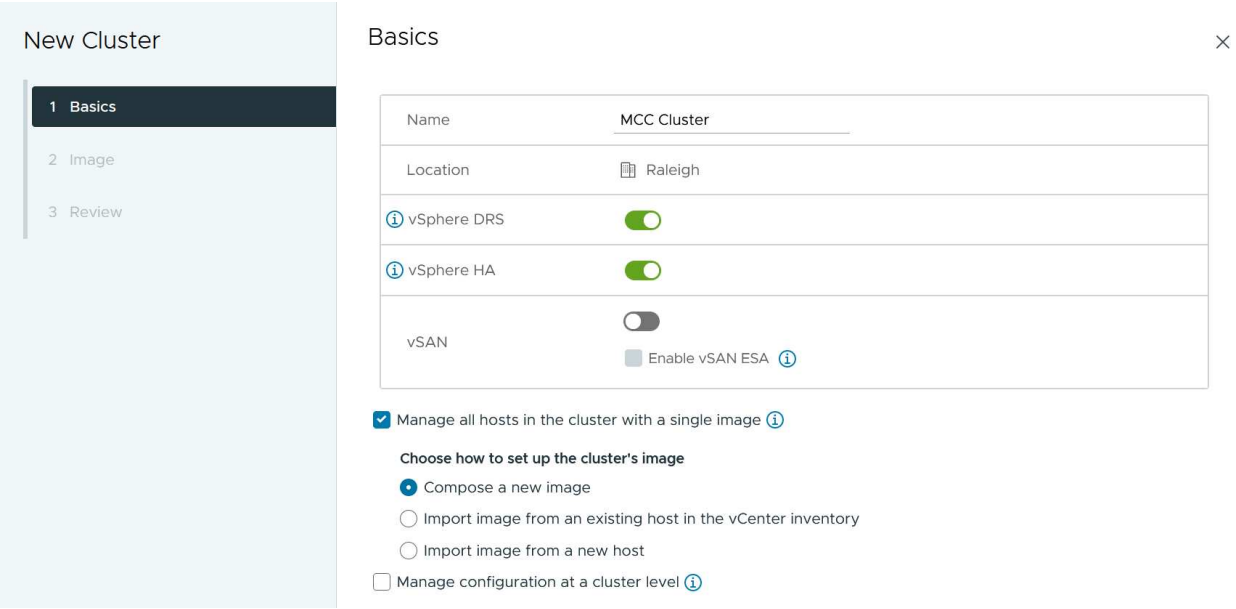
建立 **vSphere HA** 叢集

建立 vSphere HA 叢集是一個多步驟程序、完整記錄於 "如何在 docs.vmware.com 上的 vSphere Client 中建立和設定叢集"。簡言之、您必須先建立空叢集、然後使用 vCenter 新增主機、並指定叢集的 vSphere HA 和其他設定。

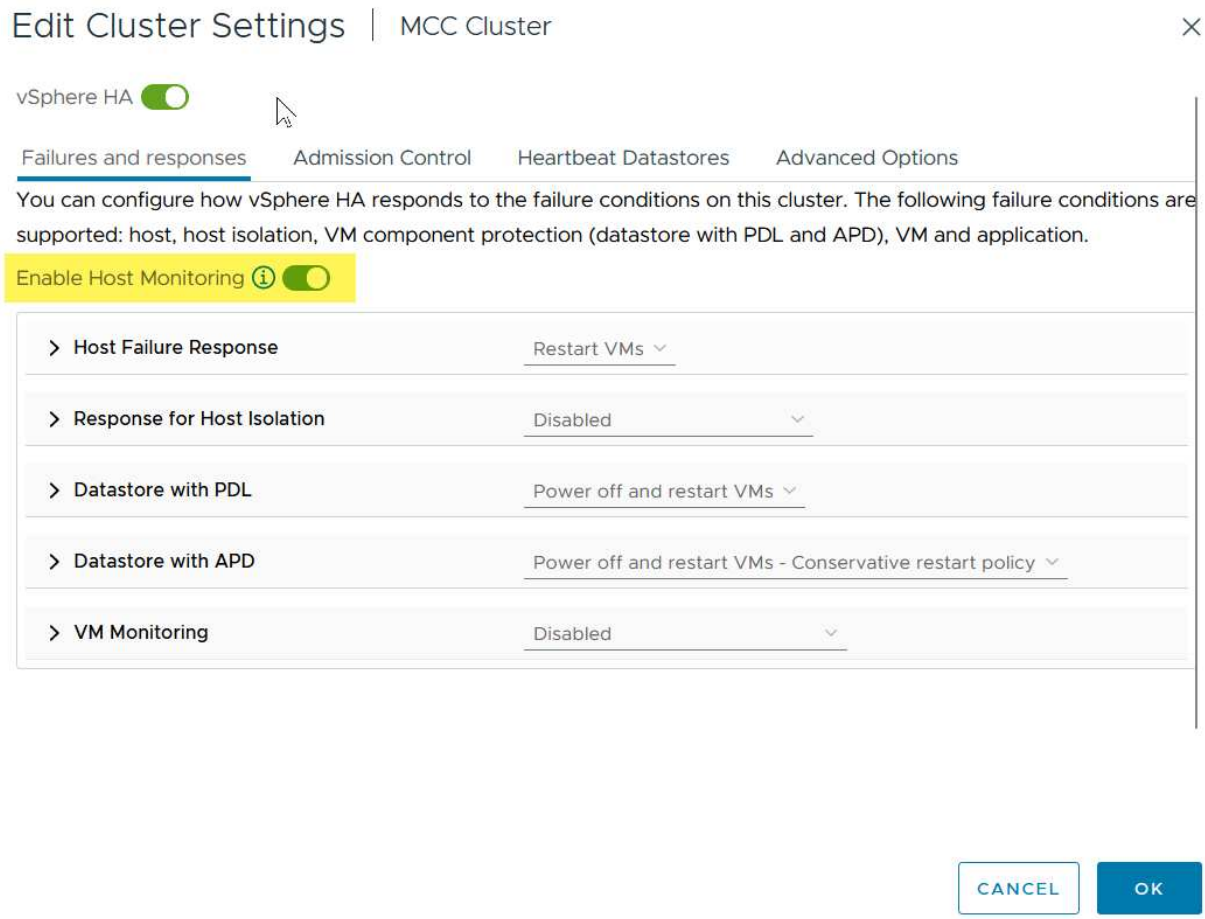
- 附註： * 本文件並無取代之處 "VMware vSphere Metro儲存叢集建議實務做法"

若要設定 HA 叢集、請完成下列步驟：

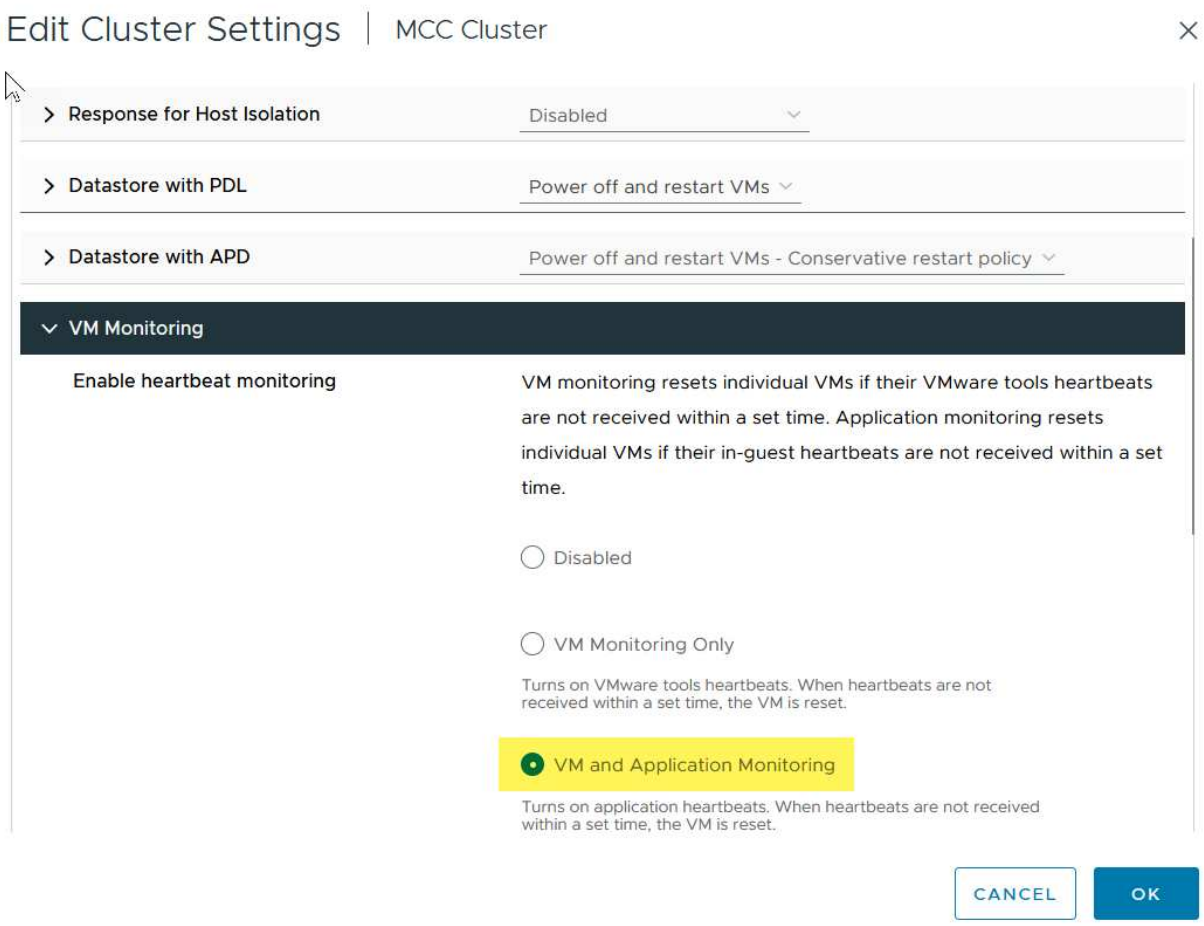
1. 連線至 vCenter UI 。
2. 在主機和叢集中、瀏覽至您要建立 HA 叢集的資料中心。
3. 以滑鼠右鍵按一下資料中心物件、然後選取新叢集。在基礎知識之下、確保您已啟用 vSphere DRS 和 vSphere HA 。完成精靈。



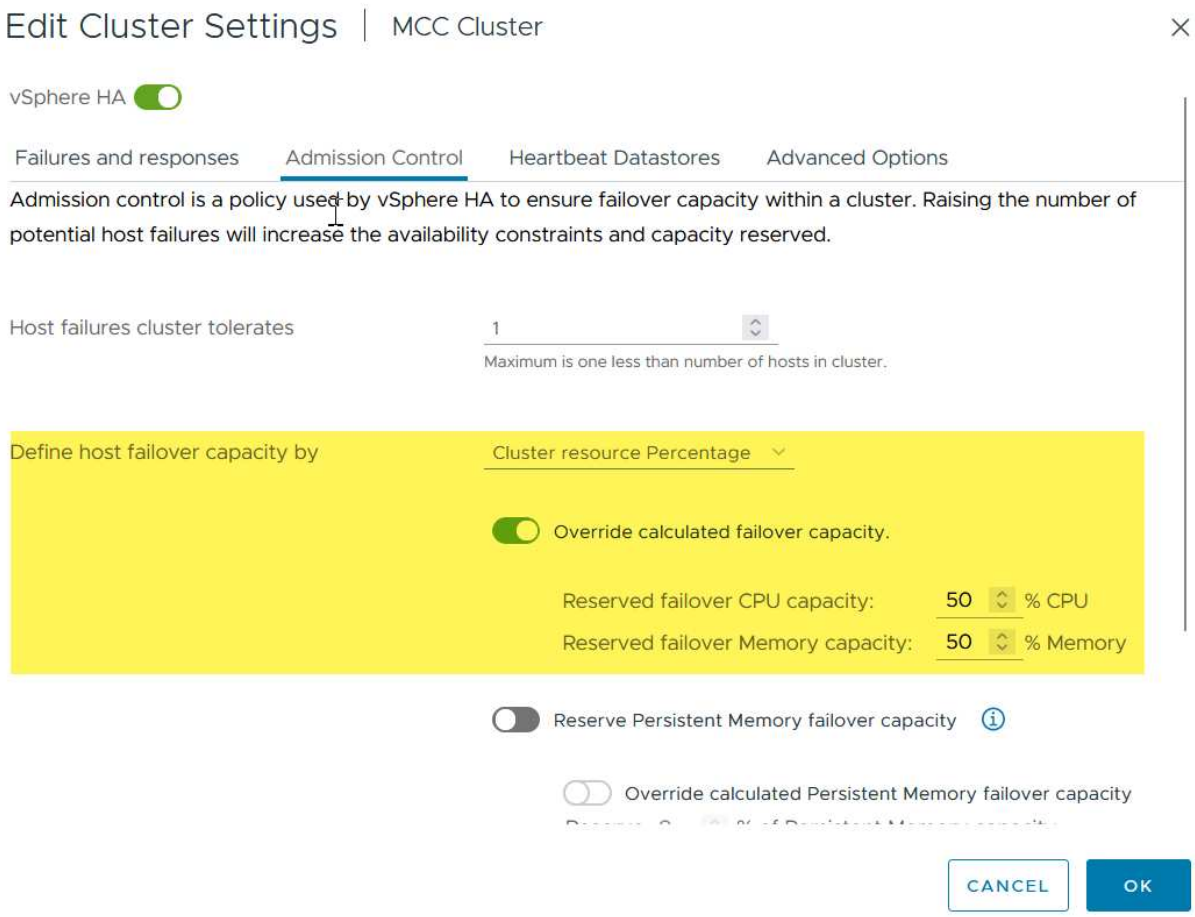
1. 選取叢集、然後移至「組態」標籤。選取 vSphere HA 、然後按一下編輯。
2. 在 [主機監控] 下，選取 [啟用主機監控] 選項。



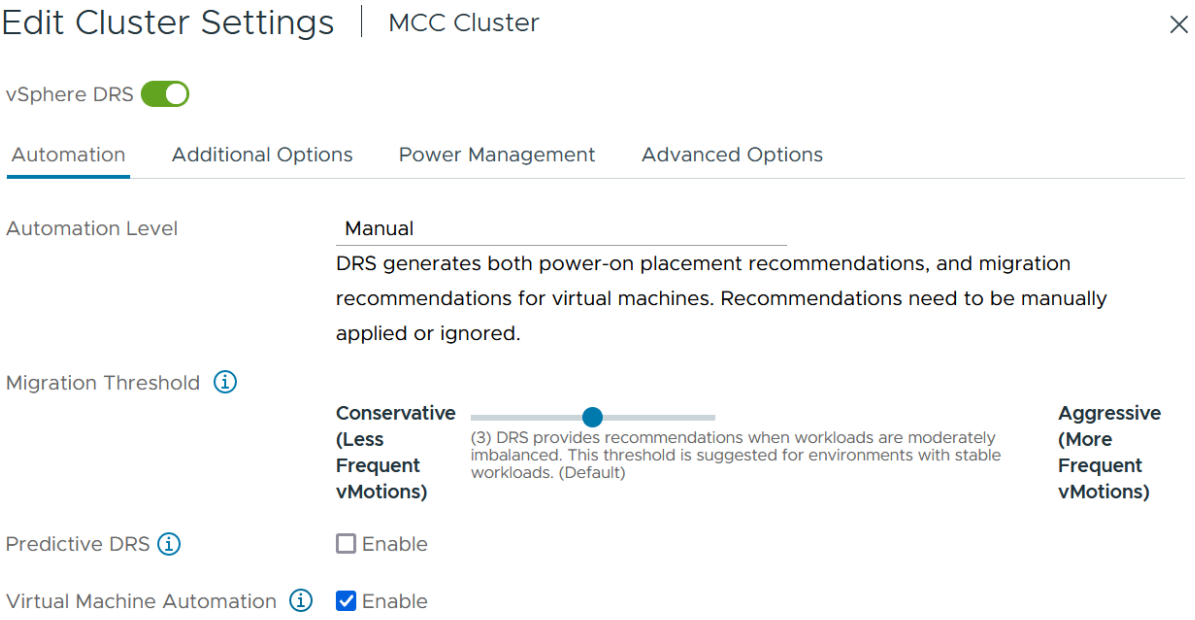
1. 在「故障與回應」標籤上、於「 VM 監控」下、選取「僅限 VM 監控」選項或「 VM 與應用程式監控」選項。



1. 在 [許可控制] 下，將 HA 接入控制選項設定為叢集資源保留；使用 50% 的 CPU/ MEM 。



1. 按一下「確定」。
2. 選取 DRS 、然後按一下編輯。
3. 除非應用程式要求、否則請將自動化層級設為手動。



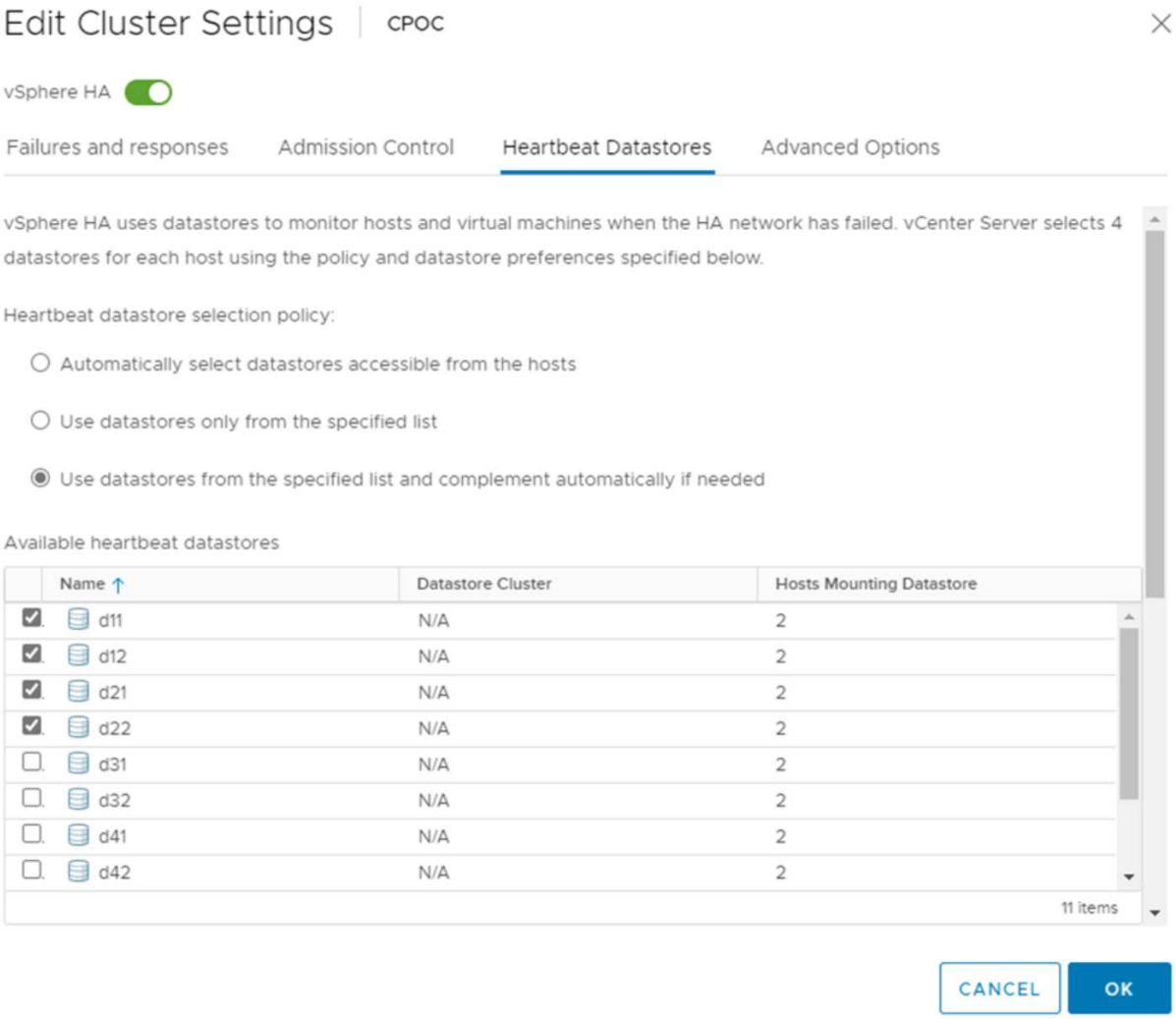
1. 啟用 VM 元件保護、請參閱 "docs.vmware.com"。
2. 建議使用 MCC 的 VMSC 使用下列其他 vSphere HA 設定：

故障	回應
主機故障	重新啟動 VM
主機隔離	已停用
永久裝置遺失（ PDL ）的資料存放區	關閉並重新啟動 VM
All Paths Down （ APD ）資料存放區	關閉並重新啟動 VM
客人不會心碎	重設 VM
VM 重新啟動原則	由虛擬機器的重要性決定
主機隔離的回應	關閉並重新啟動 VM
使用 PDL 的資料存放區回應	關閉並重新啟動 VM
對具有 APD 的資料存放區的回應	關閉並重新啟動 VM （保守）
APD 的 VM 容錯移轉延遲	3 分鐘
APD 逾時的 APD 恢復回應	已停用
VM 監控靈敏度	預設為高

設定資料存放區以進行心跳

當管理網路故障時、 vSphere HA 會使用資料存放區來監控主機和虛擬機器。您可以設定 vCenter 如何選取心跳資料存放區。若要設定資料存放區以進行心跳、請完成下列步驟：

1. 在資料存放區心跳區段中、從指定清單中選取使用資料存放區、並在需要時自動補充資料。
2. 選取您要 vCenter 從兩個站台使用的資料存放區、然後按下 OK 。

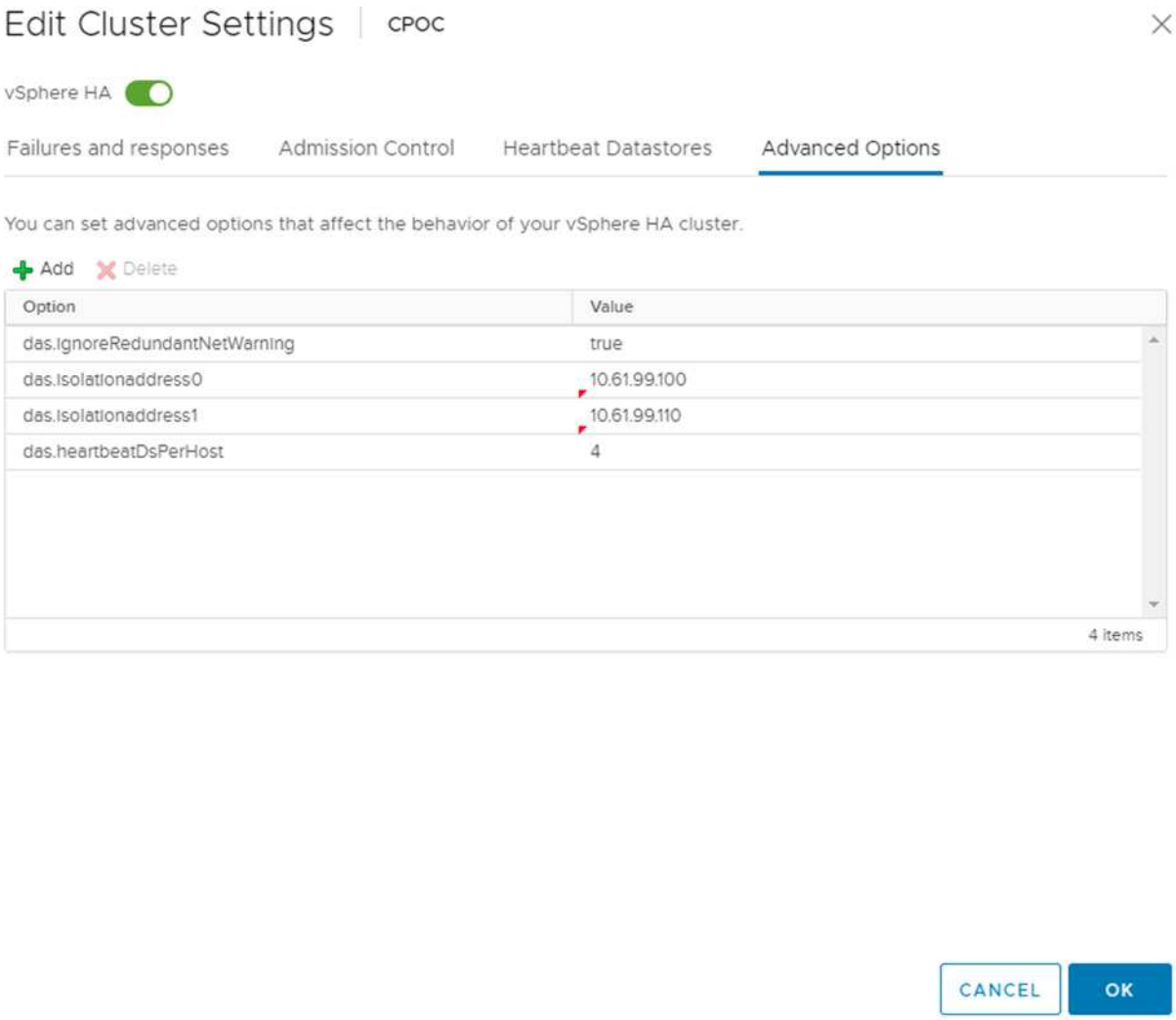


設定進階選項

- 主機故障偵測 *

當 HA 叢集內的主機無法連線至網路或叢集中的其他主機時、就會發生隔離事件。根據預設、 vSphere HA 會使用其管理網路的預設閘道做為預設隔離位址。不過、您可以為主機指定其他隔離位址來執行 ping 、以判斷是否應該觸發隔離回應。新增兩個可 ping 的隔離 IP 、每個站台一個。請勿使用閘道 IP 。使用的 vSphere HA 進階設定為 das.isolationaddress 。您可以將 ONTAP 或 Mediator IP 位址用於此用途。

請參閱 "core.vmware.com" 以取得更多資訊



新增稱為 das.心跳 DsPerHost 的進階設定、可能會增加心跳資料存放區的數量。使用四個心跳資料存放區（ HB DSS ）、每個站台兩個。使用「從清單中選取但輔助」選項。這是必要的、因為如果某個站台發生故障、您仍需要兩個 HB DSS 。但是、這些不需要透過 MCC 或 SnapMirror 主動同步來保護。

請參閱 "core.vmware.com" 以取得更多資訊

適用於 NetApp MetroCluster 的 VMware DRS 關聯性

在本節中、我們會為 MetroCluster 環境中的每個站台 \ 叢集、建立 VM 和主機的 DRS 群組。然後我們設定 VM\Host 規則、使 VM 主機與本機儲存資源的關聯性一致。例如、站台 A VM 屬於 VM 群組 sitea_vms 、站台 A 主機屬於主機群組 sitea_hosts 。接下來、在 VM\Host 規則中、我們指出 sitea_vms 應該在 sitea_hosts 中的主機上執行。

最佳實務做法 _

- NetApp 強烈建議在組 * 中的主機上運行規範 * ，而不是規範 * 必須在組 * 中的主機上運行。萬一站台 A 主機故障、站台 A 的 VM 需要透過 vSphere HA 在站台 B 的主機上重新啟動、但後者的規格不允許 HA 在站台 B 上重新啟動 VM 、因為這是硬規則。以前的規格是軟性規則、在 HA 發生時會違反、因此可提供可用度而非效能。
- 附註： * 您可以建立事件型警示、在虛擬機器違反 VM-Host 關聯性規則時觸發。在 vSphere Client 中、新增虛擬機器的警示、並選取「 VM 正在違反 VM-Host Affinity Rule 」作為事件觸發程序。如需建立及編輯警

示的詳細資訊、請參閱 "vSphere 監控與效能" 文件。

建立 **DRS** 主機群組

若要建立站台 A 和站台 B 專屬的 DRS 主機群組、請完成下列步驟：

1. 在 vSphere Web Client 中、以滑鼠右鍵按一下資源清冊中的叢集、然後選取「設定」。
2. 按一下 VM\Host Groups 。
3. 按一下「新增」
4. 輸入群組的名稱（例如、 sitea_hosts ）。
5. 從「類型」功能表中、選取「主機群組」。
6. 按一下「新增」、然後從站台 A 選取所需的主機、再按一下「確定」。
7. 重複這些步驟、為站台 B 新增另一個主機群組
8. 按一下「確定」。

建立 **DRS VM** 群組

若要建立站台 A 和站台 B 專屬的 DRS VM 群組、請完成下列步驟：

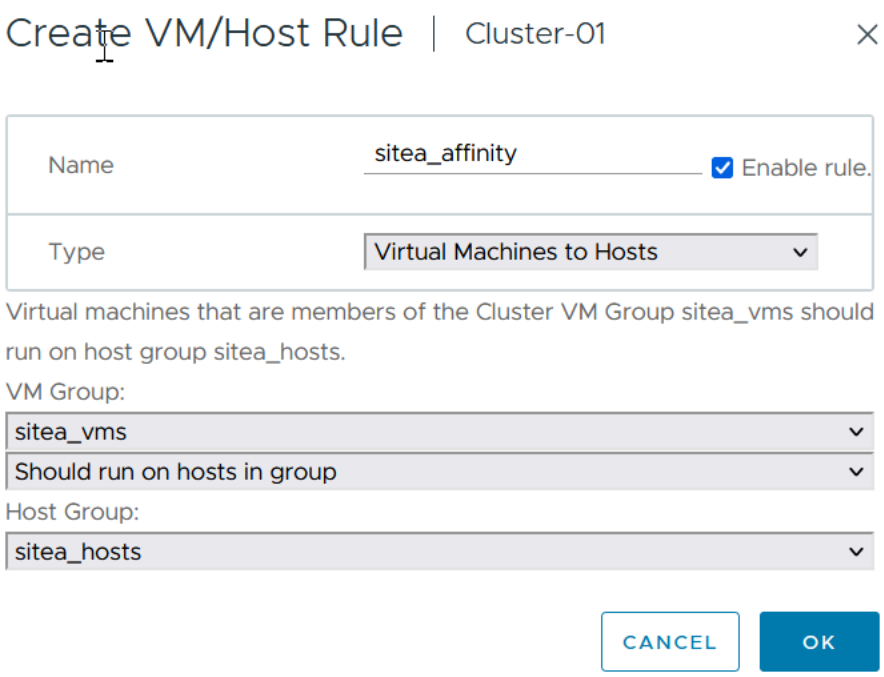
1. 在 vSphere Web Client 中、以滑鼠右鍵按一下資源清冊中的叢集、然後選取「設定」。
2. 按一下 VM\Host Groups 。
3. 按一下「新增」
4. 輸入群組的名稱（例如、 sitea_vms ）。
5. 從 Type （類型）功能表中、選取 VM Group （ VM 群組）。
6. 按一下「新增」、然後從站台 A 選取所需的 VM 、再按一下「確定」。
7. 重複這些步驟、為站台 B 新增另一個主機群組
8. 按一下「確定」。

建立 **VM Host** 規則

若要建立站台 A 和站台 B 特有的 DRS 關聯性規則、請完成下列步驟：

1. 在 vSphere Web Client 中、以滑鼠右鍵按一下資源清冊中的叢集、然後選取「設定」。
2. 按一下 VM\Host Rules 。
3. 按一下「新增」
4. 輸入規則的名稱（例如、 sitea_fit射 ）。
5. 確認已核取「啟用規則」選項。
6. 從 Type （類型）功能表中、選取 Virtual Machines to Hosts （虛擬機器至主機）。
7. 選取 VM 群組（例如、 sitea_vms ）。
8. 選取主機群組（例如、 sitea_hosts ）。

9. 重複這些步驟、為站台 B 新增另一個 VM\ 主機規則
10. 按一下「確定」。



VMware vSphere Storage DRS for NetApp MetroCluster

建立資料存放區叢集

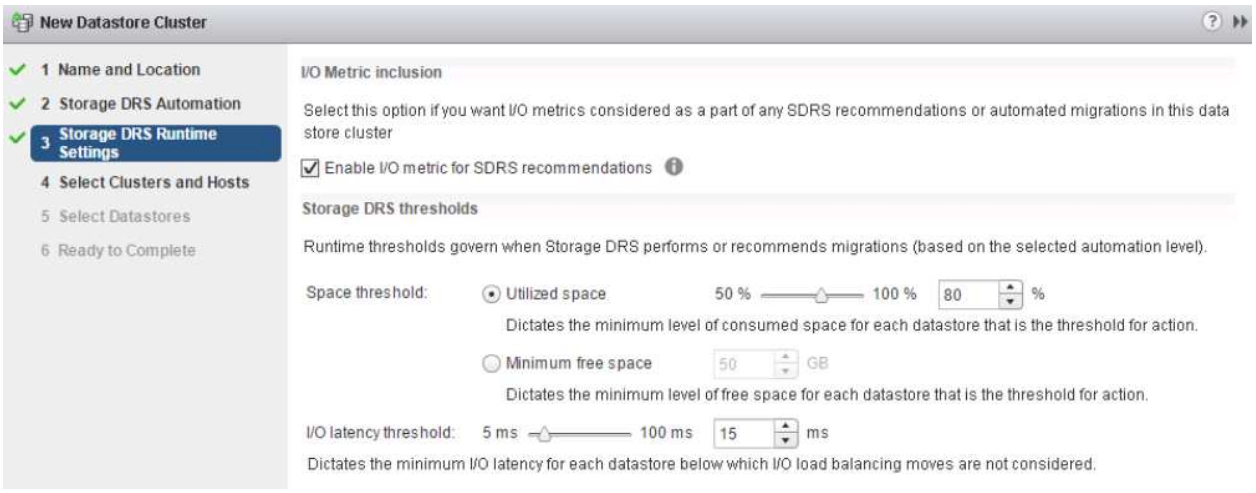
若要為每個站台設定資料存放區叢集、請完成下列步驟：

1. 使用 vSphere Web 用戶端、瀏覽至 HA 叢集位於 Storage 下的資料中心。
2. 以滑鼠右鍵按一下資料中心物件、然後選取儲存 > 新資料存放區叢集。
3. 選取「開啟 Storage DRS 」選項、然後按一下「下一步」。
4. 將所有選項設定為「無自動化（手動模式）」、然後按一下「下一步」。

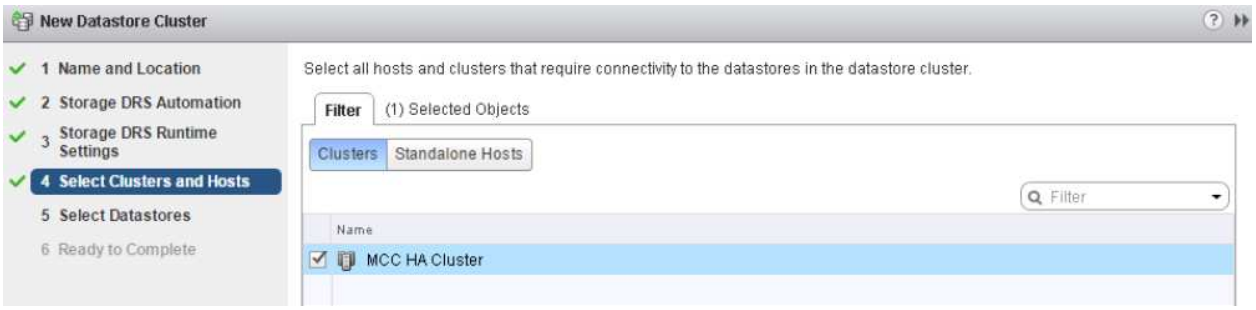
最佳實務做法 _

- NetApp 建議您將儲存 DRS 設定為手動模式、以便系統管理員決定並控制何時需要移轉。

1. 確認已核取「啟用 SDR 建議的 I/O 度量」核取方塊；度量設定可以保留預設值。



1. 選取 HA 叢集、然後按一下「下一步」。



1. 選取屬於站台 A 的資料存放區、然後按一下「下一步」。

1. 檢閱選項、然後按一下「完成」。
2. 重複這些步驟以建立站台 B 資料存放區叢集、並確認只選取站台 B 的資料存放區。

vCenter Server 可用度

您的 vCenter Server 應用裝置（ VCSA ）應使用 vCenter HA 加以保護。vCenter HA 可讓您在主動式被動式 HA 配對中部署兩個 VCSA 。每個故障網域各有一個。您可以在上閱讀更多有關 vCenter HA 的資訊 "docs.vmware.com"。

計畫性和非計畫性事件的恢復能力

NetApp MetroCluster 和 SnapMirror 主動同步是強大的工具、可增強 NetApp 硬體和 ONTAP ® 軟體的高可用度和不中斷營運。

這些工具可為整個儲存環境提供全站台保護、確保資料永遠可用。無論您是使用獨立式伺服器、高可用度伺服器叢集、 Docker 容器或虛擬化伺服器、 NetApp 技術都能在電力中斷、冷卻或網路連線中斷、儲存陣列關機或作業錯誤等情況下、無縫維持儲存可用度。

MetroCluster 和 SnapMirror 主動式同步提供三種基本方法、可在發生計畫性或非計畫性事件時維持資料連續性：

- 備援元件、可防止單一元件故障
- 本機 HA 接管、用於影響單一控制器的事件
- 完整的站台保護：將儲存設備和用戶端存取從來源叢集移至目的地叢集、以快速恢復服務

這表示在單一元件故障時、作業會順暢地繼續、並在更換故障元件時自動恢復至備援作業。

除了單節點叢集（例如 ONTAP Select 等軟體定義版本）之外、所有 ONTAP 叢集都具有稱為接管和恢復的內建 HA 功能。叢集中的每個控制器都會與另一個控制器配對、形成 HA 配對。這些配對可確保每個節點都在本機上連線至儲存設備。

接管是一種自動化程序、其中一個節點會接管另一個節點的儲存設備、以維護資料服務。GiveBack 是還原正常作業的反向程序。可以規劃接管、例如執行硬體維護或 ONTAP 升級、或是因節點緊急或硬體故障而非計畫性地進行。

在接管期間、 MetroCluster 組態中的網路附加儲存邏輯介面（ NAS 生命期）會自動容錯移轉。但是、儲存區域網路生命（ SAN 生命）不會容錯移轉；它們會繼續使用邏輯單元編號（ LUN ）的直接路徑。

如需 HA 接管與恢復的詳細資訊、請參閱 "HA配對管理總覽"。值得一提的是、這項功能並非 MetroCluster 或 SnapMirror 主動式同步的專屬功能。

當某個站台離線、或是作為整個站台維護的計畫活動時、就會使用 MetroCluster 進行站台切換。其餘站台則假設擁有離線叢集的儲存資源（磁碟和集合體）、而故障站台上的 SVM 則會在災難站台上線並重新啟動、保留其完整身分以供用戶端和主機存取。

有了 SnapMirror 主動式同步、由於兩個複本都是同時使用的、因此您現有的主機將繼續運作。NetApp Mediator 是確保站台容錯移轉正確進行所需的工具。

使用 **MCC** 的 **VMSC** 的失敗案例

以下各節概述 VMSC 和 NetApp MetroCluster 系統各種故障情況的預期結果。

單一儲存路徑故障

在這種情況下、如果 HBA 連接埠、網路連接埠、前端資料交換器連接埠或 FC 或乙太網路纜線等元件故障、 ESXi 主機會將該儲存裝置的特定路徑標記為已停用。如果在 HBA/ 網路 / 交換器連接埠上提供恢復功能、就能為儲存裝置設定多個路徑、 ESXi 理想情況下會執行路徑切換。在這段期間內、虛擬機器會持續執行而不會受到影響、因為提供多條路徑可通往儲存設備、因此可確保儲存設備的可用度。

- 附註： * 在此案例中、 MetroCluster 行為並無變更、所有資料存放區仍會保留在各自站台內。

最佳實務做法 _

在使用 NFS/iSCSI 磁碟區的環境中、 NetApp 建議在標準 vSwitch 中、至少為 NFS vmkernel 連接埠設定兩個網路上行鏈路、而在對應 NFS vmkernel 介面的連接埠群組中、則必須設定相同的上行鏈路。NIC 群組可在雙主

動式或雙主動式待命模式中進行設定。

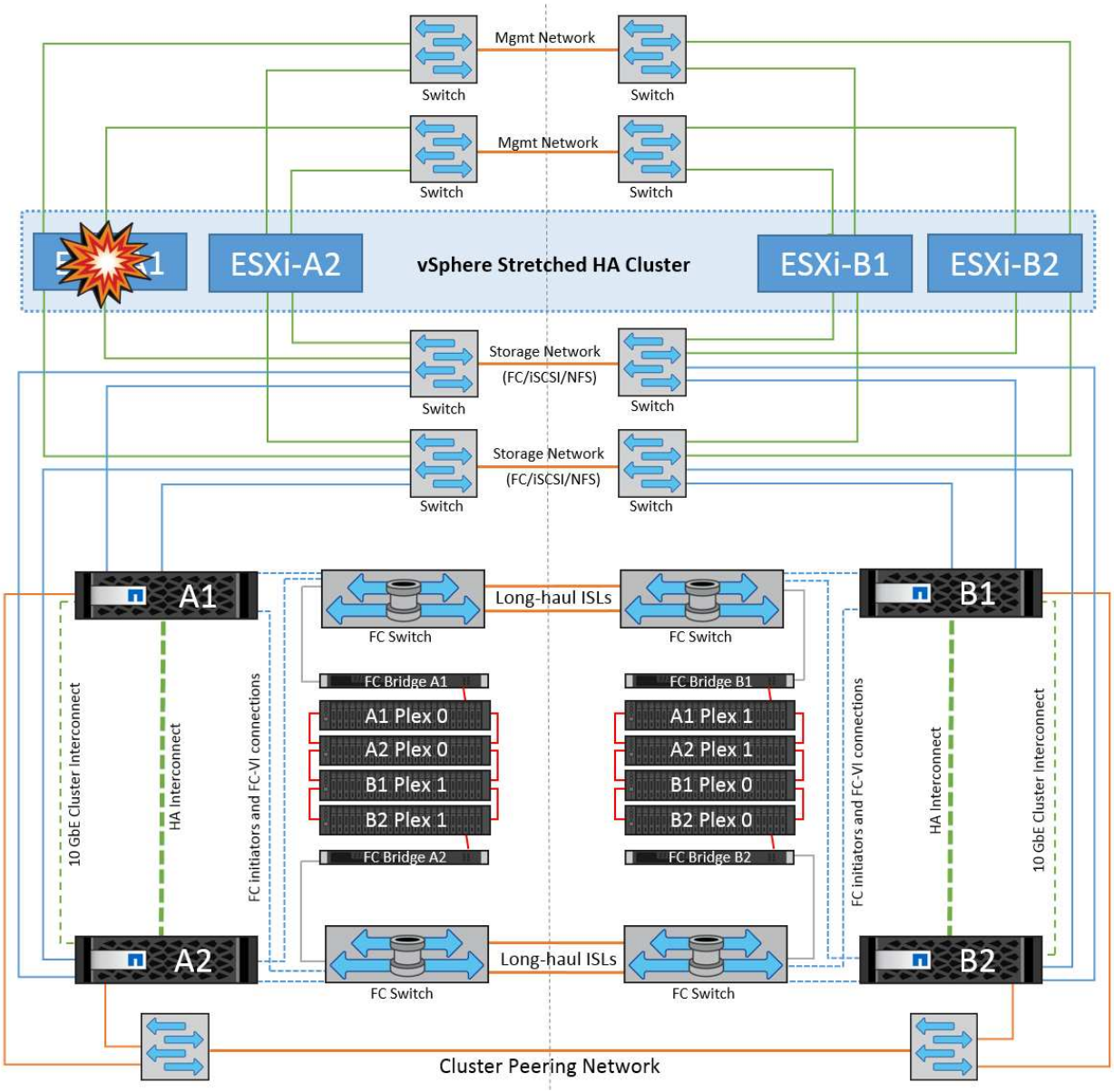
此外、對於 iSCSI LUN 、必須將 vmkernel 介面繫結至 iSCSI 網路介面卡、以設定多重路徑。如需詳細資訊、請參閱 vSphere 儲存文件。

最佳實務做法 _

在使用光纖通道 LUN 的環境中、 NetApp 建議至少有兩個 HBA 、以保證 HBA/ 連接埠層級的恢復能力。NetApp 也建議將單一啟動器分區至單一目標分區、以做為設定分區的最佳實務做法。

應使用虛擬儲存主控台（ VSC ）來設定多重路徑原則、因為它會為所有新的和現有的 NetApp 儲存裝置設定原則。

單一 ESXi 主機故障



在這種情況下、如果 ESXi 主機發生故障、 VMware HA 叢集中的主節點會偵測主機故障、因為主機不再接收到網路心跳。若要判斷主機是否真的停機或只是網路分割區、主節點會監控資料存放區的訊次、如果沒有、則會 ping 失敗主機的管理 IP 位址、以執行最終檢查。如果所有這些檢查都是負數、則主節點會將此主機宣告為故障

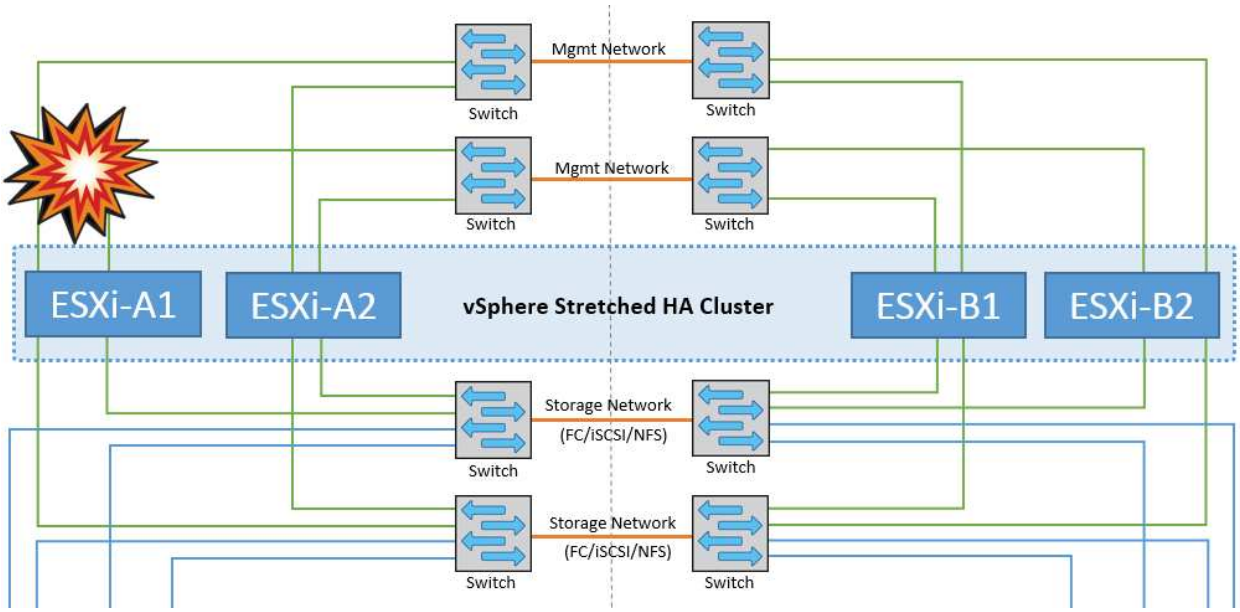
主機、而在該故障主機上執行的所有虛擬機器都會在叢集中的正常主機上重新開機。

如果已設定 DRS VM 和主機關聯性規則（ VM 群組 sitea_vms 中的 VM 應在主機群組 sitea_hosts 中執行主機）、則 HA 主機會先檢查站台 A 的可用資源如果站台 A 沒有可用的主機、則主主機會嘗試在站台 B 的主機上重新啟動 VM

如果本機站台有資源限制、則可能會在其他站台的 ESXi 主機上啟動虛擬機器。不過、如果將虛擬機器移轉回本機站台中任何仍在運作的 ESXi 主機、而違反任何規則、則定義的 DRS VM 和主機關聯性規則將會修正。如果 DRS 設定為手動、 NetApp 建議您啟動 DRS 、並套用建議來修正虛擬機器的放置位置。

在此案例中、 MetroCluster 行為並無任何變更、所有資料存放區仍會保持不變、不受其個別站台影響。

ESXi 主機隔離

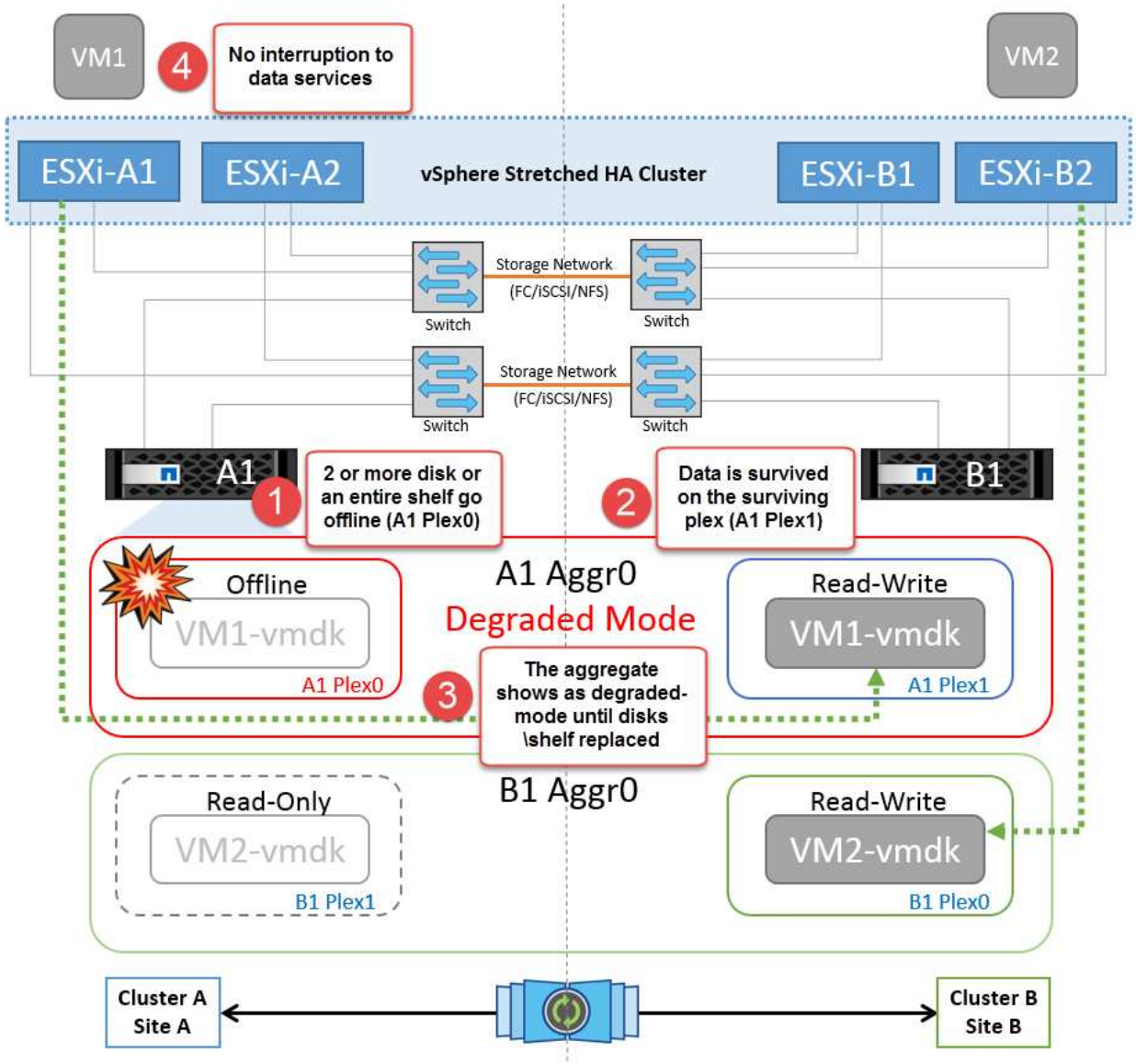


在此案例中、如果 ESXi 主機的管理網路中斷、 HA 叢集中的主節點將不會接收任何訊息、因此此主機會在網路中隔離。若要判斷它是否發生故障或只是隔離、主節點會開始監控資料存放區心跳。如果主機存在、則主機會由主節點宣告為隔離。根據設定的隔離回應、主機可能會選擇關閉、關閉虛擬機器、甚至讓虛擬機器保持開機。隔離回應的預設時間間隔為 30 秒。

在此案例中、 MetroCluster 行為並無任何變更、所有資料存放區仍會保持不變、不受其個別站台影響。

磁碟機櫃故障

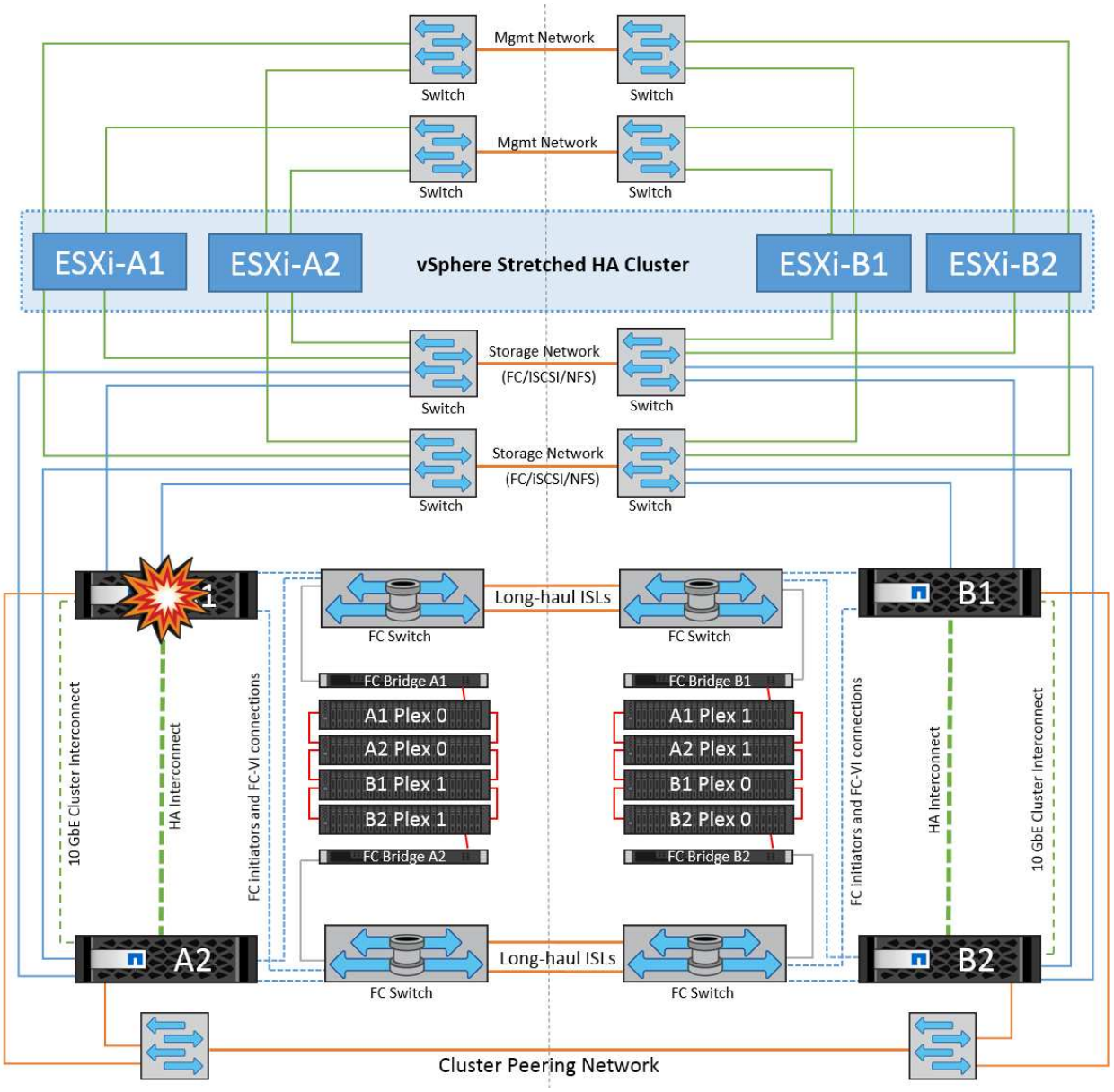
在此案例中、有兩個以上的磁碟或整個機櫃發生故障。資料是從仍在運作的複合環境提供、不會中斷資料服務。磁碟故障可能會影響本機或遠端叢。由於只有一個叢處於作用中狀態、因此集合體會顯示為降級模式。更換故障磁碟後、受影響的集合體將自動重新同步以重建資料。重新同步後、集合體將自動返回正常的鏡射模式。如果單一 RAID 群組中有兩個以上的磁碟發生故障、則必須從頭重建叢。



- 注意： * 在此期間、虛擬機器 I/O 作業不會受到影響、但效能會降低、因為資料是透過 ISL 連結從遠端磁碟機櫃存取。

單一儲存控制器故障

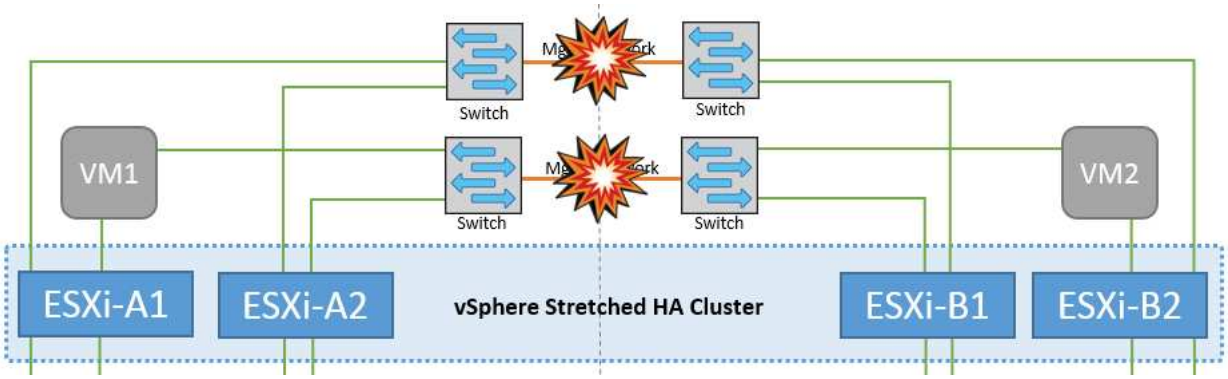
在這種情況下、兩個儲存控制器中的其中一個會在一個站台發生故障。由於每個站台都有 HA 配對、因此一個節點的故障會以透明方式自動觸發容錯移轉至另一個節點。例如、如果節點 A1 故障、其儲存設備和工作負載會自動傳輸至節點 A2 。虛擬機器將不會受到影響、因為所有的叢集都仍然可用。第二個站台節點（ B1 和 B2 ）不受影響。此外、 vSphere HA 將不會採取任何行動、因為叢集中的主節點仍會接收到網路心跳。



如果容錯移轉是循環災難的一部分（節點 A1 容錯移轉至 A2 ）、而且之後發生 A2 故障、或是站台 A 完全故障、則災難後的切換可能會發生在站台 B

交換器間連結故障

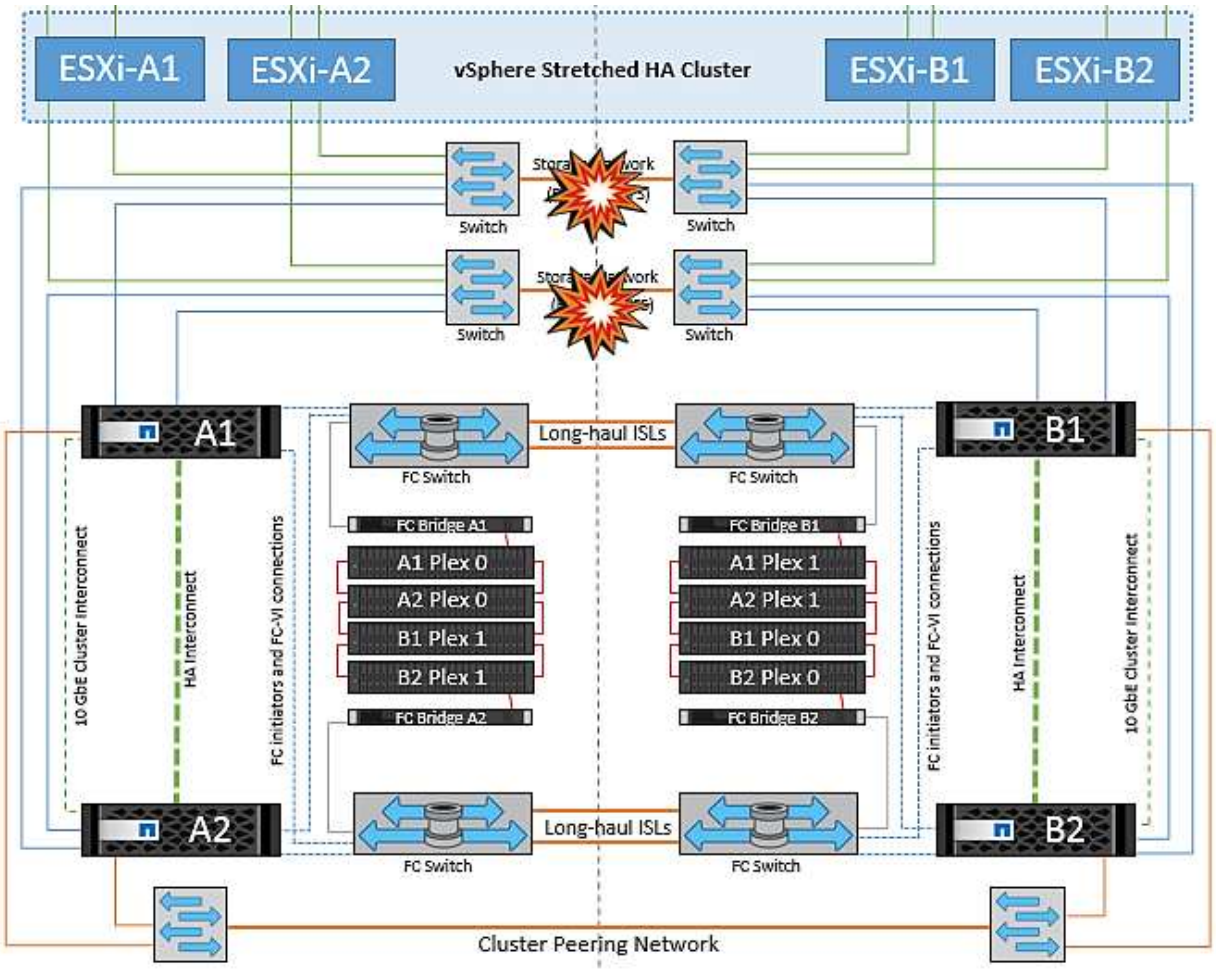
管理網路的交換器間連結故障



在此案例中、如果前端主機管理網路的 ISL 連結失敗、站台 A 的 ESXi 主機將無法與站台 B 的 ESXi 主機通訊這會導致網路分割區、因為特定站台的 ESXi 主機將無法將網路心跳傳送至 HA 叢集中的主節點。因此、由於分割區的緣故、將會有兩個網路區段、每個區段中都會有一個主節點、可保護 VM 免於特定站台內的主機故障。

- 附註： * 在此期間、虛擬機器仍在執行中、在此案例中、 MetroCluster 行為並無變更。所有的資料存放區都會繼續保持不變、不受其個別站台影響。

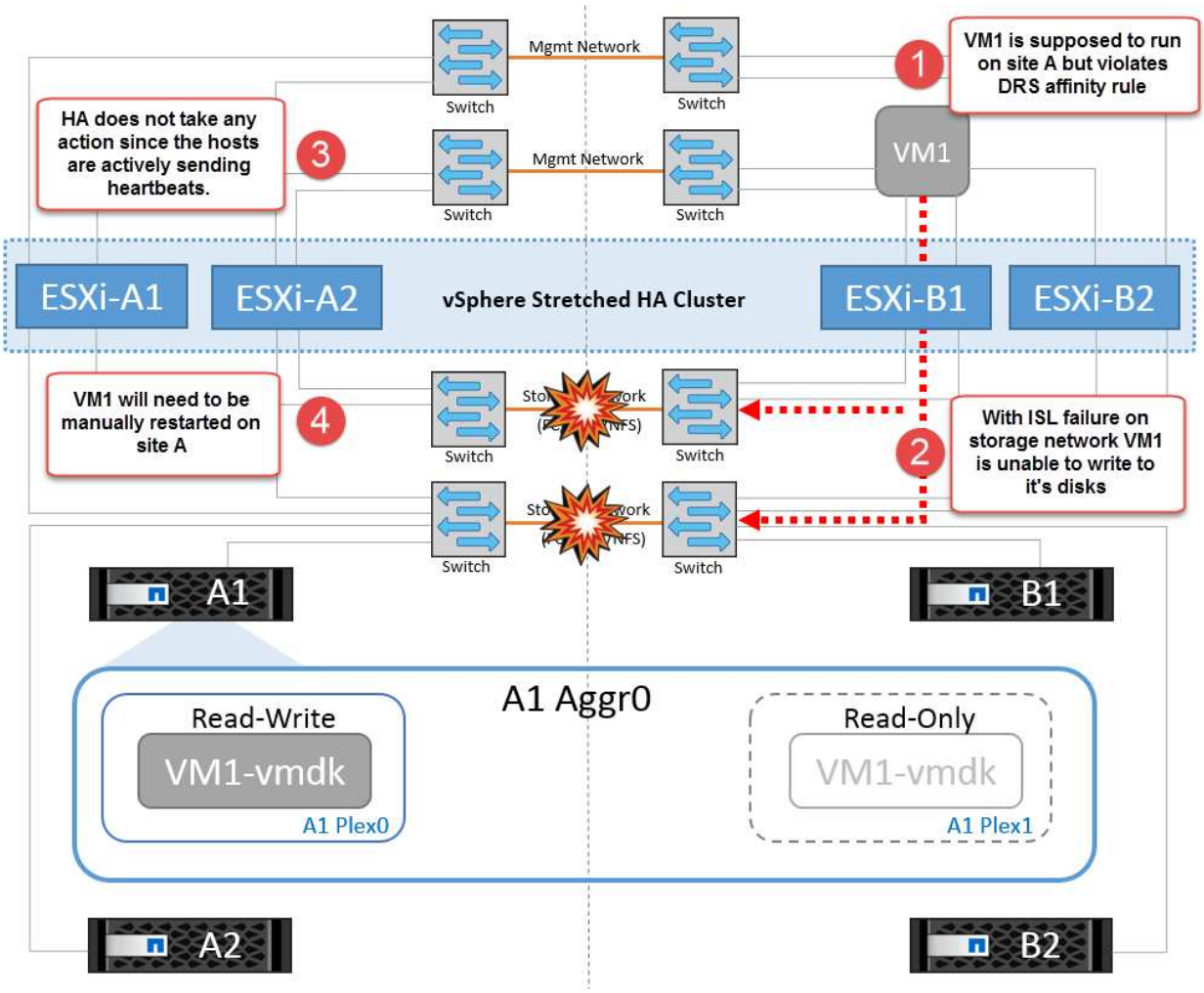
儲存網路的交換器間連結故障



在此案例中、如果後端儲存網路的 ISL 連結故障、站台 A 的主機將無法存取站台 B 的儲存磁碟區或叢集 B 的 LUN 、反之亦然。VMware DRS 規則的定義、是為了讓主機儲存站台的關聯性能讓虛擬機器在不影響站台的情況下執行。

在此期間、虛擬機器會繼續在各自的站台上執行、在此案例中、 MetroCluster 行為不會有任何變更。所有的資料存放區都會繼續保持不變、不受其個別站台影響。

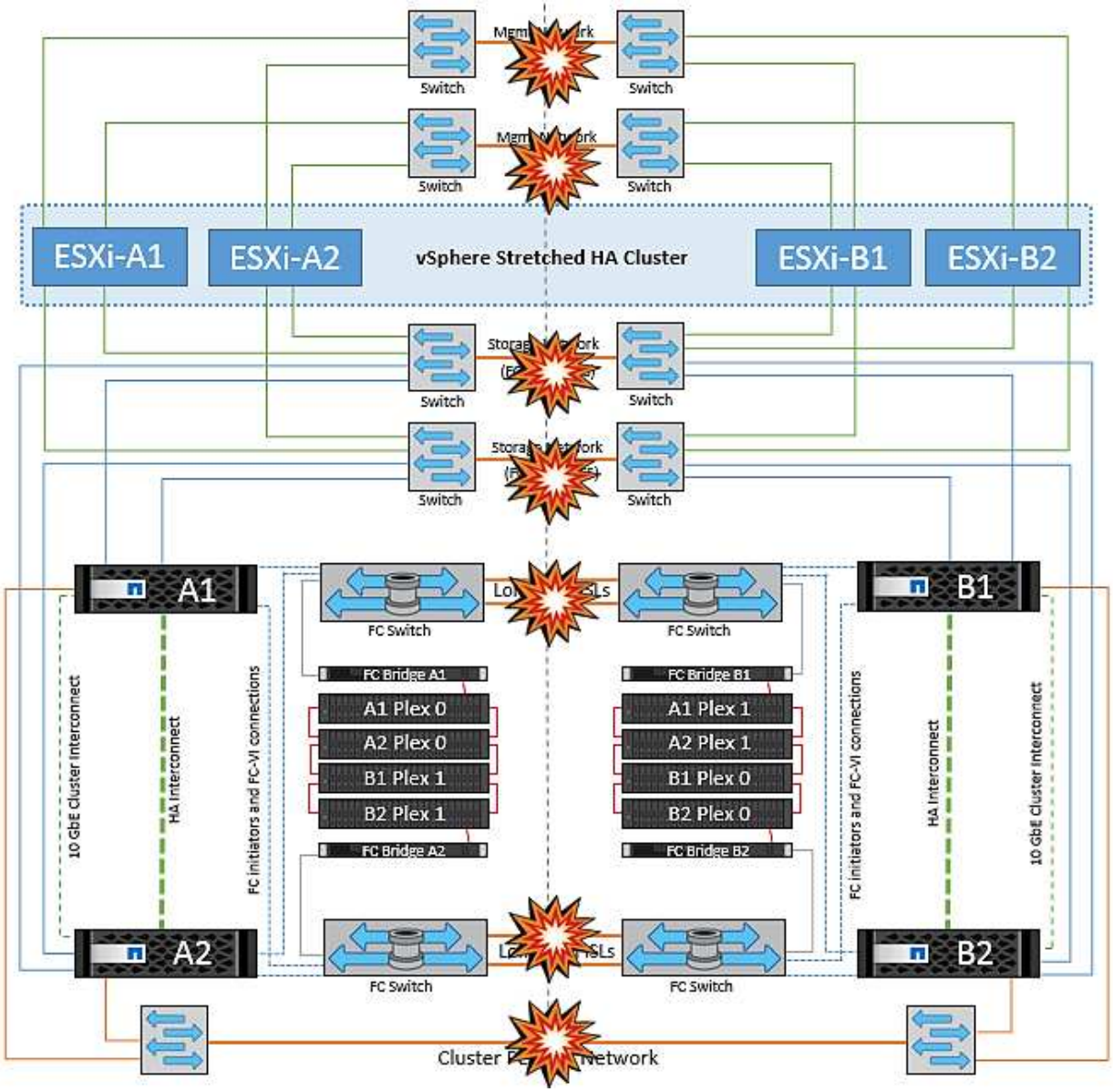
如果因為某種原因違反關聯規則（例如、 VM1 原本應從站台 A 執行、其磁碟位於本機叢集 A 節點上、而 VM1 則是在站台 B 的主機上執行）、則虛擬機器的磁碟將透過 ISL 連結遠端存取。由於 ISL 連結故障、在站台 B 執行的 VM1 將無法寫入其磁碟、因為通往儲存磁碟區的路徑已關閉、且該特定虛擬機器已關閉。在這些情況下、VMware HA 不會採取任何行動、因為主機正在主動傳送心跳。這些虛擬機器必須在各自的站台手動關閉並開啟電源。下圖說明違反 DRS 關聯性規則的虛擬機器。



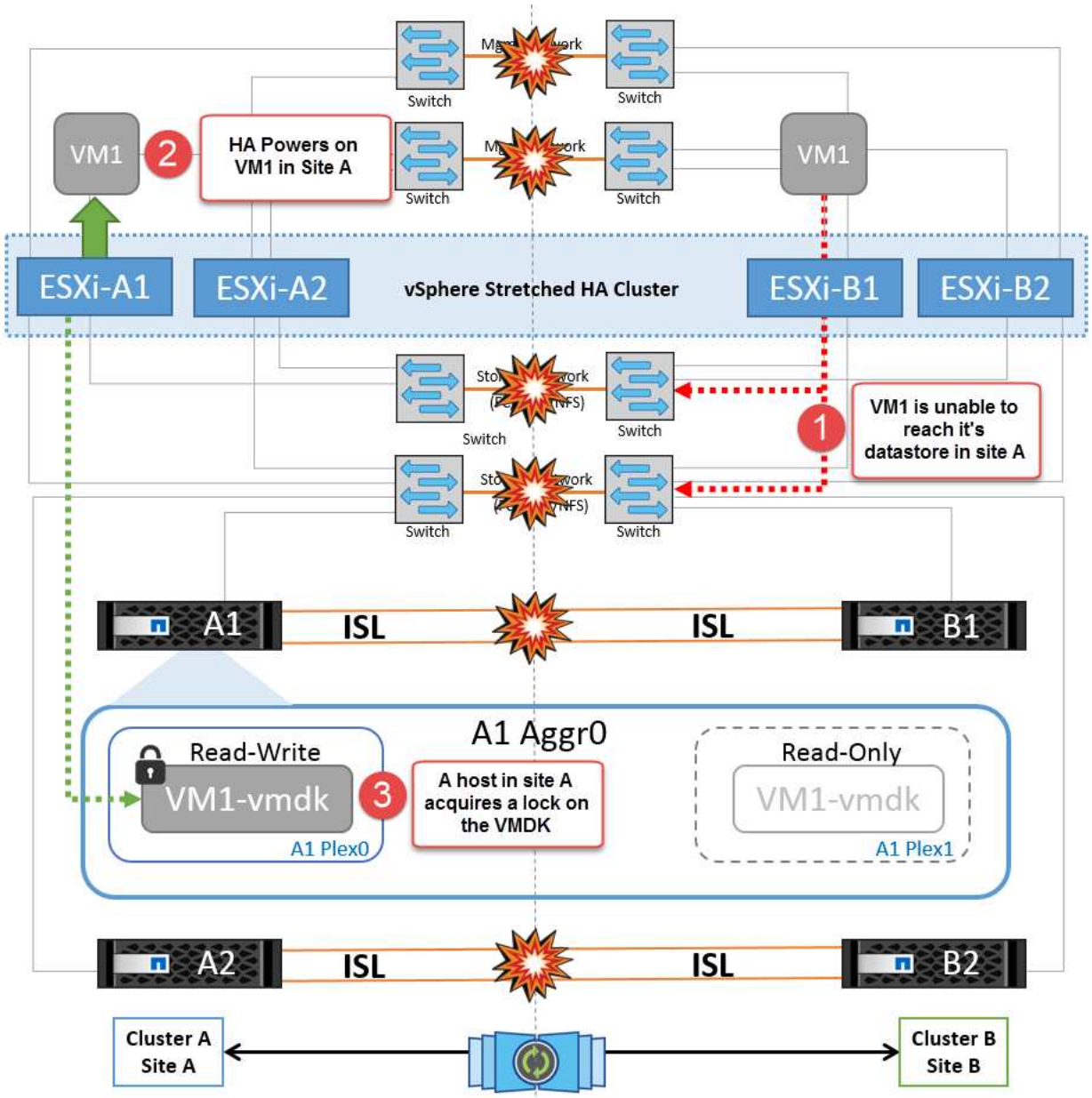
所有交換器間故障或完整資料中心分割區

在這種情況下、站台之間的所有 ISL 連結都會中斷、而且兩個站台彼此之間會隔離。如先前的案例所述、例如管理網路和儲存網路的 ISL 故障、虛擬機器在完全 ISL 故障時不會受到影響。

在站台之間分割 ESXi 主機之後、 vSphere HA 代理程式會檢查資料存放區心跳、而且在每個站台中、本機 ESXi 主機將能夠將資料存放區心跳更新至各自的讀寫磁碟區 /LUN 。站台 A 中的主機會假設站台 B 中的其他 ESXi 主機故障、因為沒有網路 / 資料存放區檢測信號。站台 A 的 vSphere HA 會嘗試重新啟動站台 B 的虛擬機器、最終會失敗、因為站台 B 的資料存放區因為儲存 ISL 故障而無法存取。站台 B 也會再次出現類似的情況



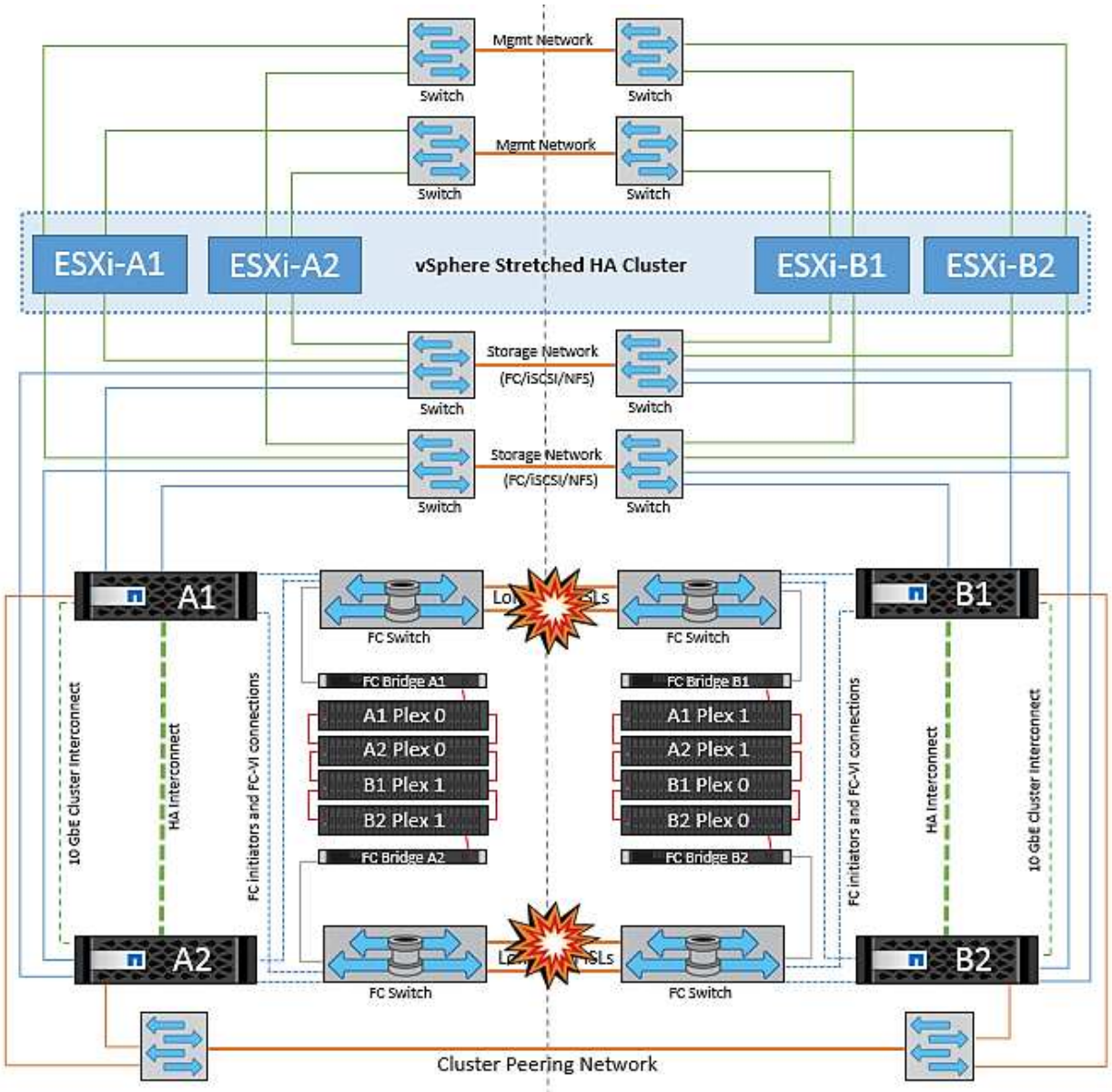
NetApp 建議判斷是否有任何虛擬機器違反 DRS 規則。從遠端站台執行的任何虛擬機器都會停機、因為它們將無法存取資料存放區、 vSphere HA 會在本機站台上重新啟動該虛擬機器。當 ISL 連結恢復上線後、在遠端站台上執行的虛擬機器將會停止運作、因為無法有兩個執行個體使用相同的 MAC 位址執行虛擬機器。



NetApp MetroCluster 中兩個 **Fabric** 上的交換器間連結故障

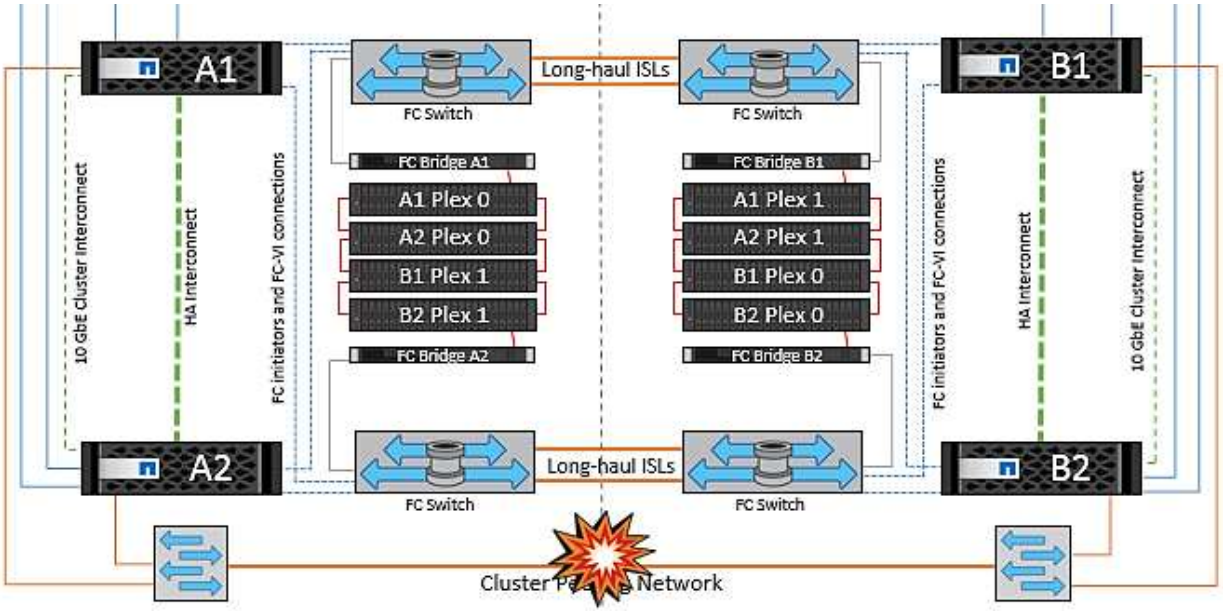
在一個或多個 ISL 故障的情況下、流量會繼續流經其餘的連結。如果兩個架構上的所有 ISL 都發生故障、使得儲存和 NVRAM 複寫站台之間沒有連結、則每個控制器都會繼續提供其本機資料。還原至少一個 ISL 時、所有的叢會自動重新同步。

在所有 ISL 停機之後所發生的任何寫入動作、都不會鏡射到另一個站台。當組態處於此狀態時、發生災難時的切入將會遺失尚未同步的資料。在這種情況下、需要手動介入才能在進行重新操作後恢復。如果很可能在較長的時間內沒有可用的 ISL 、系統管理員可以選擇關閉所有資料服務、以避免在發生災難時發生資料遺失的風險。在至少有一個 ISL 可供使用之前、應將執行此動作的可能性與需要進行重新操作的災難可能性進行權衡。或者、如果 ISL 在串聯案例中發生故障、系統管理員可能會在所有連結失敗之前、觸發已規劃的切換至其中一個站台。



已消除叢集連結故障

在對等叢集連結故障案例中、由於 Fabric ISL 仍處於作用中狀態、因此兩個站台的資料服務（讀取和寫入）都會繼續存在於兩個叢集。任何叢集組態變更（例如、新增 SVM 、在現有 SVM 中配置 Volume 或 LUN ）都無法傳播到其他站台。這些資料會保留在本機 CRS 中繼資料磁碟區中、並在還原對等叢集連結時自動傳播到其他叢集。如果必須強制切換才能還原對等叢集連結、則在切換程序中、仍在運作中的站台上、中繼資料磁碟區的遠端複寫複本會自動重新播放未完成的叢集組態變更。



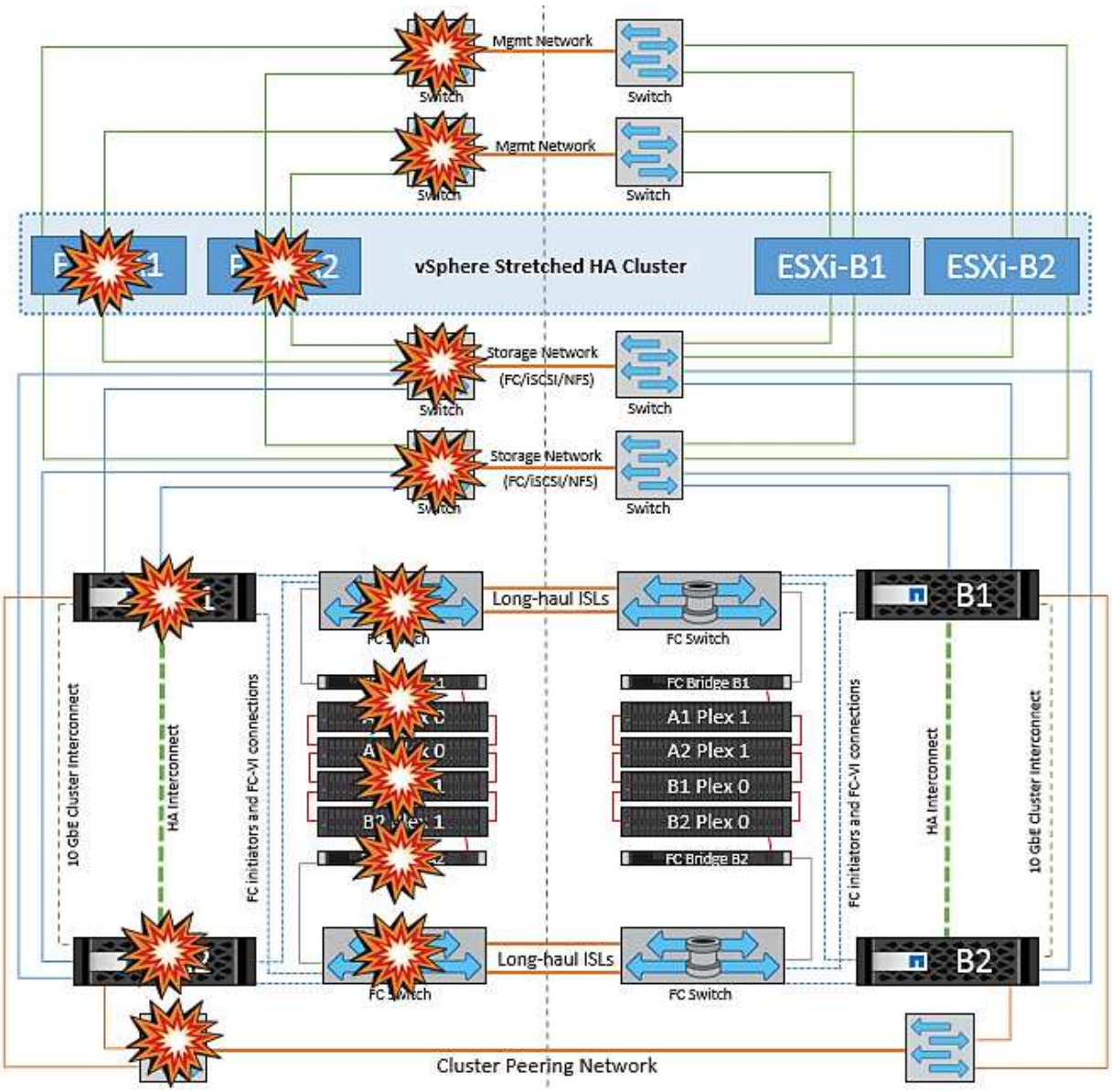
完成站台故障

在完整站台 A 故障案例中、站台 B 的 ESXi 主機因為故障而無法從站台 A 的 ESXi 主機取得網路心跳。站台 B 的 HA 主機會驗證資料存放區心跳不存在、宣告站台 A 的主機故障、並嘗試重新啟動站台 B 中的站台 A 虛擬機器在此期間、儲存管理員會執行一次轉換、以恢復仍在運作的站台上故障節點的服務、該站台將還原站台 B 上站台 A 的所有儲存服務站台 A 磁碟區或 LUN 在站台 B 上可用後、 HA 主代理程式會嘗試重新啟動站台 B 中的站台 A 虛擬機器

如果 vSphere HA 主要代理程式嘗試重新啟動虛擬機器（包括登錄及開機）失敗、則會在延遲後重試重新啟動。重新啟動之間的延遲時間最多可設定為 30 分鐘。vSphere HA 會嘗試重新啟動這些項目、最多嘗試次數（預設為六次）。

- 附註： * 在放置管理程式找到適當的儲存設備之前、 HA 主機不會開始重新啟動嘗試、因此在整個站台發生故障的情況下、這將是在執行切入之後。

如果站台 A 已切換、則可透過容錯移轉至正常運作的節點、無縫地處理其中一個仍在運作的站台 B 節點的後續故障。在這種情況下、四個節點的工作現在僅由一個節點執行。在這種情況下、恢復將包括執行恢復到本機節點的贈品。然後、當站台 A 還原時、會執行切換作業、以還原組態的穩定狀態作業。


Mgmt Network
Switch
Mgmt Network
Switch
vSphere Stretched HA Cluster
ESXi-B1
ESXi-B2
Storage Network
(FC/iSCSI/NFS)
Switch
Storage Network
(FC/iSCSI/NFS)
Switch
Long-haul ISLs
FC Switch
B1
FC Bridge B1
A1 Plex 1
A2 Plex 1
B1 Plex 0
B2 Plex 0
FC Bridge B2
Long-haul ISLs
FC Switch
B2
10 GbE Cluster Interconnect
HA Interconnect
FC initiators and FC-VI connections
FC initiators and FC-VI connections
HA Interconnect
10 GbE Cluster Interconnect
Cluster Peering Network

版權資訊

Copyright © 2024 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。