高可用性架構
ONTAP Select

NetApp
November 08, 2026

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

目錄

高可用性架構 . . . 1
- ONTAP Select 高可用性組態 . . . 1
  - 雙節點 HA 與多節點 HA . . . 2
  - 雙節點 HA 與雙節點延伸 HA（MetroCluster SDS） . . . 3
- ONTAP Select HA RSM 和鏡像 Aggregate . . . 3
  - 同步複寫 . . . 3
  - 鏡射 Aggregate . . . 3
  - 寫入路徑 . . . 4
- ONTAP Select HA 的資料保護增強功能 . . . 6
  - 磁碟活動訊號 . . . 6
  - HA 信箱張貼 . . . 6
  - HA 心跳 . . . 7
  - HA 容錯移轉與還原 . . . 7

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。

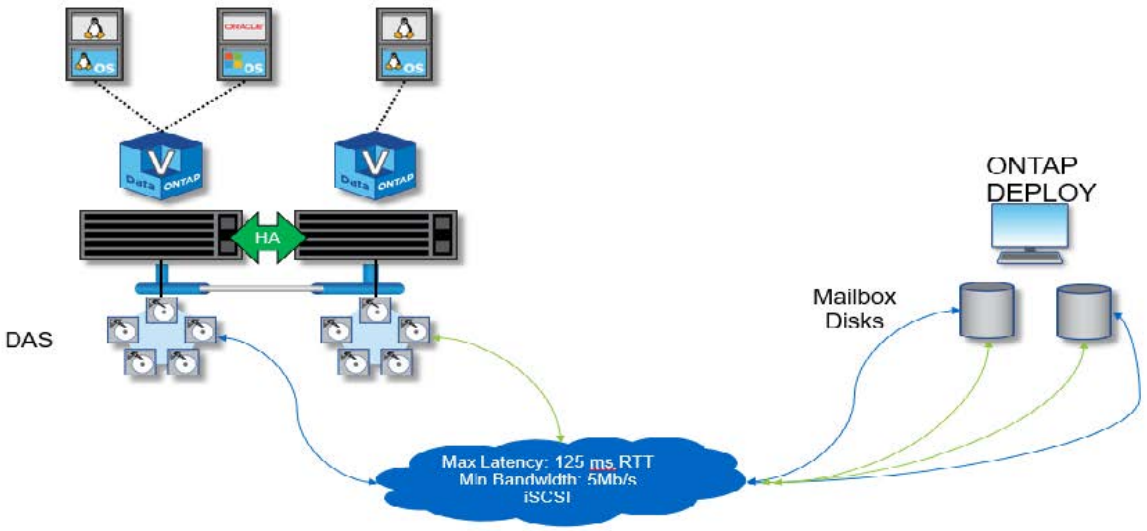
軟體定義儲存（SDS）市場很大一部分基於「無共享」儲存概念，透過軟體複寫在不同儲存孤島中儲存使用者資料的多個副本來提供資料恢復能力。ONTAP Select 在此基礎上，利用 ONTAP 提供的同步複寫功能（RAID SyncMirror）在叢集內儲存額外的使用者資料副本。這發生在 HA 配對的環境中。每個 HA 配對儲存兩個使用者資料副本：一個儲存在本地節點提供的儲存上，另一個儲存在 HA 合作夥伴提供的儲存上。在 ONTAP Select 叢集中，HA 和同步複寫緊密結合，二者的功能無法分離或獨立使用。因此，同步複寫功能僅在多節點版本中可用。

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。

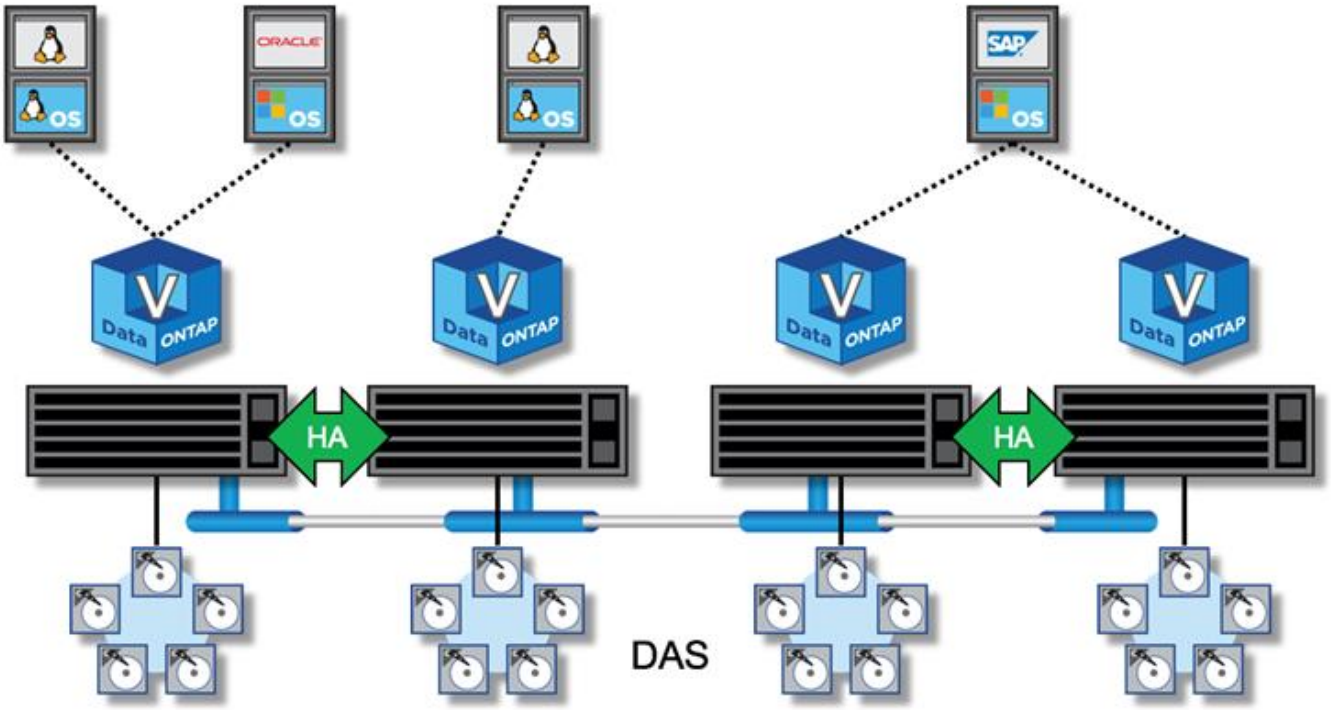
這兩種架構如下圖所示。

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。

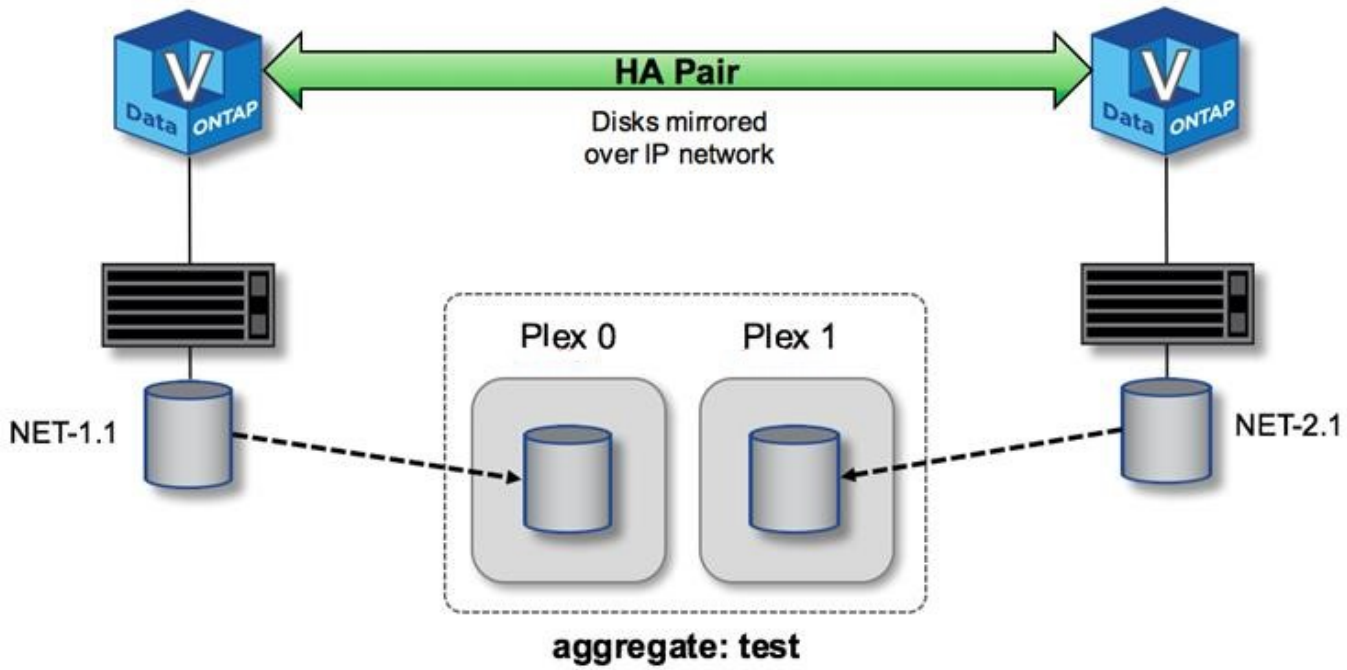
這種方法與標準 ONTAP 叢集的工作方式截然不同。這適用於 ONTAP Select 叢集中的所有根磁碟和資料磁碟。Aggregate 包含資料的本機副本和鏡像副本。因此，包含 N 個虛擬磁碟的 Aggregate 可提供 N/2 個磁碟的獨立儲存空間，因為第二個資料副本位於其自身的獨立磁碟上。

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：

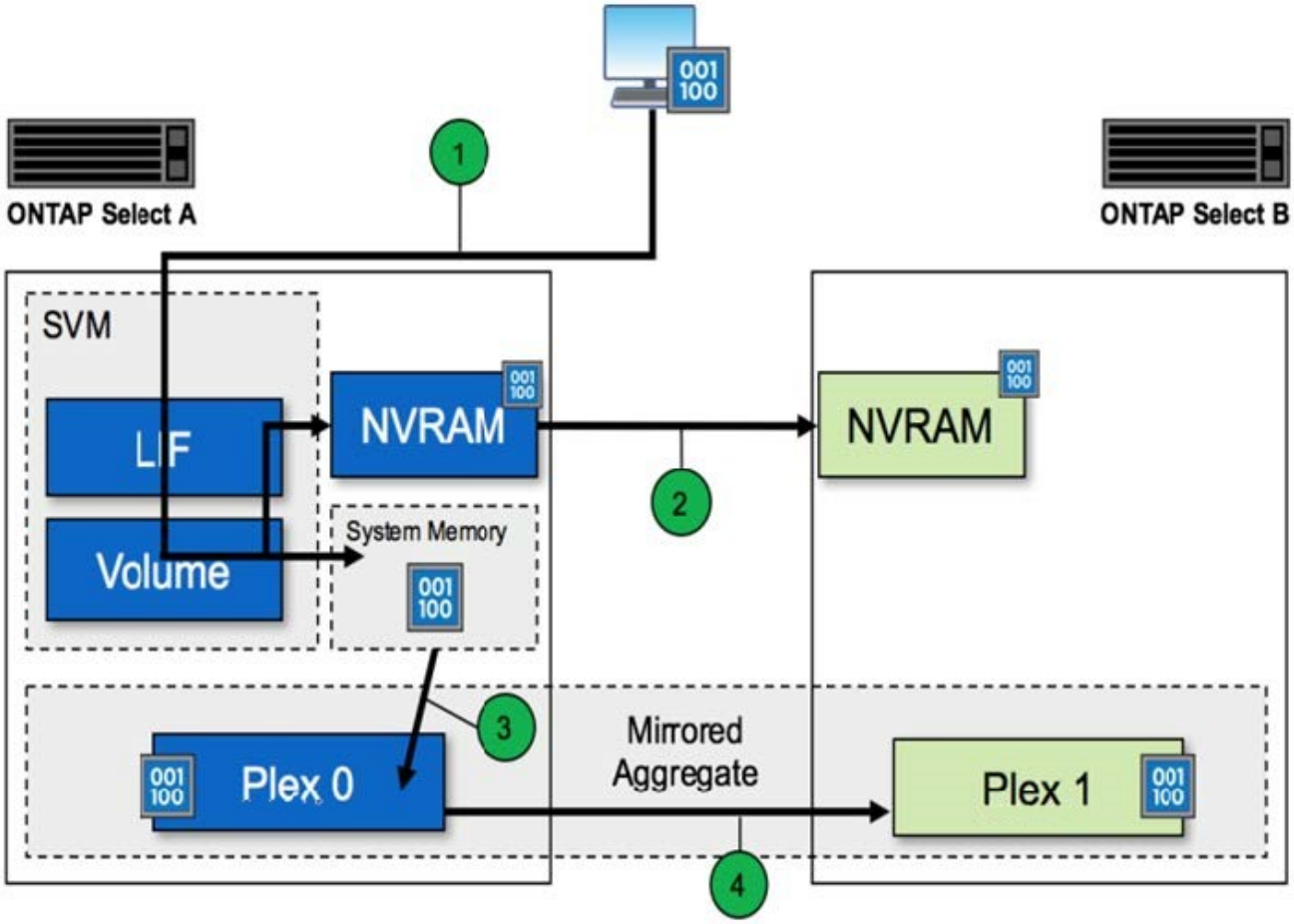
- 致謝
- 取消暫存

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。

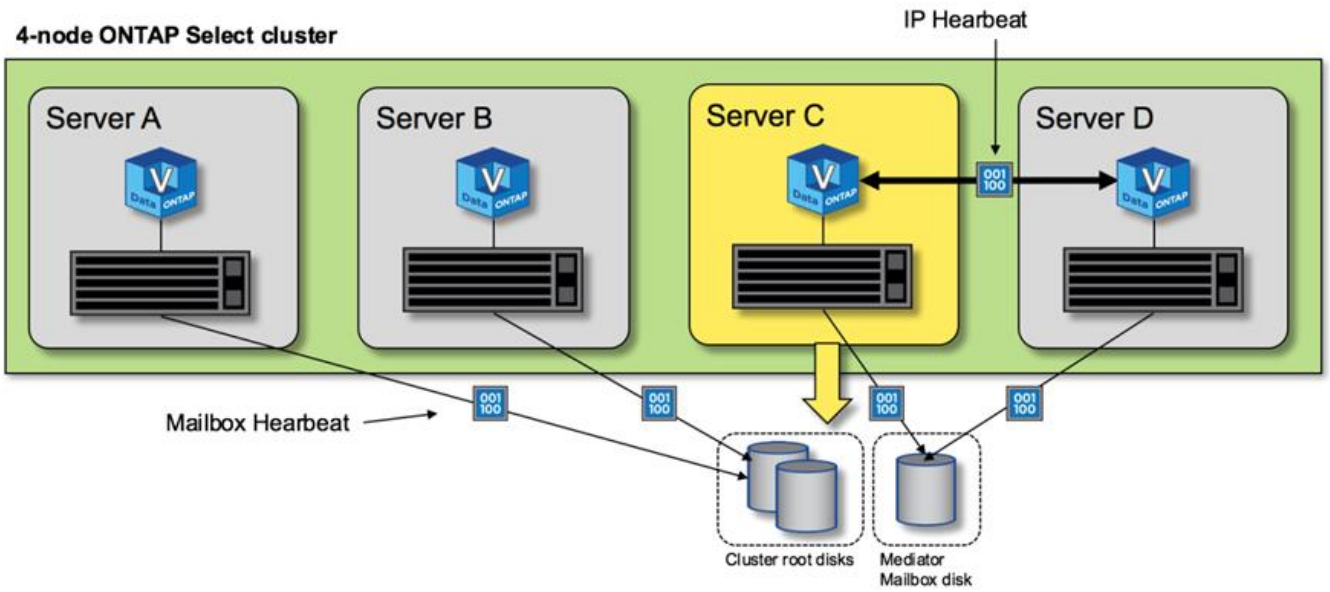
HA 心跳

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>
```

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。