



AFF / FAS 系統上的儲存組態

Enterprise applications

NetApp
February 11, 2026

目錄

AFF / FAS 系統上的儲存組態	1
總覽	1
資料儲存設計	1
集合體	1
磁碟區	1
LUN	2
資料庫檔案和檔案群組	2
儲存效率	6
壓縮	6
資料壓縮	7
重複資料刪除	8
效率與精簡配置	8
效率最佳實務做法	9
資料庫壓縮	9
空間回收	9
資料保護	10
SnapCenter	10
使用 T-SQL 快照保護資料庫	10
SQL Server 可用度群組搭配 SnapCenter	11
災難恢復	12
災難恢復	12
SnapMirror	13
MetroCluster	13
SnapMirror 主動同步	18

AFF / FAS 系統上的儲存組態

總覽

ONTAP 儲存解決方案與 Microsoft SQL Server 的結合、可實現企業級資料庫儲存設計、滿足現今最嚴苛的應用程式需求。

在 ONTAP 解決方案上最佳化 SQL Server 需要瞭解 SQL Server I/O 模式和特性。SQL Server 資料庫設計完善的儲存配置必須支援 SQL Server 的效能需求、同時提供整個基礎架構的最大管理能力。良好的儲存配置也能讓初始部署成功、並隨著業務成長、環境隨時間而順利成長。

資料儲存設計

對於不使用 SnapCenter 支援功能執行備份的 SQL Server 資料庫、Microsoft 建議將資料和記錄檔放在不同的磁碟機上。對於同時更新和要求資料的應用程式、記錄檔會密集寫入、而且資料檔（視應用程式而定）會密集讀寫。對於資料擷取、不需要記錄檔。因此、您可以從放在自己磁碟機上的資料檔案來滿足資料要求。

當您建立新資料庫時、Microsoft 建議您為資料和記錄指定個別的磁碟機。若要在資料庫建立之後移動檔案、資料庫必須離線。如需更多 Microsoft 建議、請參閱 ["將資料和記錄檔放在不同的磁碟機上"](#)。

集合體

Aggregate 是 NetApp 儲存組態的最低層級儲存容器。網際網路上存在一些舊版文件、建議將 IO 分隔到不同的基礎磁碟機集。ONTAP 不建議這麼做。NetApp 已使用資料檔案和交易記錄檔分離的共用和專用集合體、執行各種 I/O 工作負載特性分析測試。測試結果顯示、一個大型集合體含有更多 RAID 群組和磁碟機、可最佳化和改善儲存效能、而且管理員更容易管理、原因有兩個：

- 一個大型集合體可讓所有檔案都能使用所有磁碟機的 I/O 功能。
- 一個大型 Aggregate 可讓您以最有效率的方式使用磁碟空間。

若要獲得高可用度（HA）、請將 SQL Server Always On Availability Group 次要同步複本放在 Aggregate 中的獨立儲存虛擬機器（SVM）上。為了進行災難恢復、請將非同步複本放在 DR 站台中屬於獨立儲存叢集的集合上、並使用 NetApp SnapMirror 技術複寫內容。NetApp 建議在集合體中至少有 10% 的可用空間、以獲得最佳的儲存效能。

磁碟區

磁碟區會建立並位於集合體內。此術語有時會造成混淆、因為 ONTAP 磁碟區不是 LUN。ONTAP Volume 是資料的管理容器。磁碟區可能包含檔案、LUN、甚至 S3 物件。磁碟區不會佔用空間、只會用於管理內含的資料。

Volume 設計考量

在您建立資料庫 Volume 設計之前、請務必瞭解 SQL Server I/O 模式和特性的差異、視工作負載及備份與還原需求而定。請參閱下列 NetApp 建議的彈性磁碟區：

- 避免在主機之間共用磁碟區。例如、雖然可以在單一磁碟區中建立 2 個 LUN、並將每個 LUN 共用至不同的主機、但這應該避免、因為這樣可能會使管理複雜化。在同一主機上執行多個 SQL Server 執行個體的情況下、除非您接近節點上的磁碟區限制、否則請避免共用磁碟區、而每個主機的每個執行個體都有個別的磁碟

區，以便於進行資料管理。

- 請使用 NTFS 掛載點而非磁碟機代號、以超越 Windows 中 26 個磁碟機代號的限制。使用 Volume 掛載點時、一般建議將 Volume 標籤命名為與掛載點相同的名稱。
- 適當時、請設定 Volume 自動調整大小原則、以協助避免空間不足的情況。
- 如果您在 SMB 共用上安裝 SQL Server、請確定已在 SMB 磁碟區上啟用 Unicode 以建立資料夾。
- 將磁碟區中的快照保留值設為零、以便從作業角度進行監控。
- 停用快照排程和保留原則。而是使用 SnapCenter 來協調 SQL Server 資料磁碟區的 Snapshot 複本。
- 將 SQL Server 系統資料庫放在專用磁碟區上。
- Tempdb 是 SQL Server 用來做為暫用工作區的系統資料庫、特別是用於 I/O 密集型 DBCC CHECKDB 作業。因此、請將此資料庫放在具有獨立磁碟集的專用磁碟區上。在磁碟區數是一項挑戰的大型環境中、您可以將 Tempdb 整合為較少的磁碟區、並在經過仔細規劃之後、將其儲存在與其他系統資料庫相同的磁碟區中。對 tempdb 的資料保護不是高優先順序、因為每次重新啟動 SQL Server 時都會重新建立此資料庫。
- 將使用者資料檔案放(.mdf、在不同的磁碟區上、因為它們是隨機讀取 / 寫入工作負載。建立交易記錄備份的頻率通常高於資料庫備份。因此(.ldf、請將交易記錄檔放在不同的磁碟區或 VMDK 上、以便為每個資料檔案建立獨立的備份排程。這種分隔方式也能將記錄檔的連續寫入 I/O 與資料檔案的隨機讀寫 I/O 隔離、大幅提升 SQL Server 效能。

LUN

- 請確定使用者資料庫檔案和用於儲存記錄備份的記錄目錄位於不同的磁碟區、以防止保留原則在與 SnapVault 技術搭配使用時覆寫快照。
- 請勿在同一個 LUN 上混合使用資料庫和非資料庫檔案、例如與全文搜尋相關的檔案。
- 將資料庫次要檔案（作為檔案群組的一部分）放在不同的磁碟區、可改善 SQL Server 資料庫的效能。只有當資料庫的檔案未與任何其他檔案共用其 LUN .mdf 時、此分隔才有效 .mdf。
- 如果您使用 DiskManager 或其他工具建立 LUN、請確保在格式化 LUN 時、將分割區的分配單元大小設為 64K。
- 請參閱 ["適用於現代 SAN 的 ONTAP 最佳實務做法下的 Microsoft Windows 和原生 MPIO"](#) 將 Windows 上的多重路徑支援套用至 MPIO 內容中的 iSCSI 裝置。

資料庫檔案和檔案群組

在 ONTAP 上正確放置 SQL Server 資料庫檔案是初始部署階段的關鍵。如此可確保最佳效能、空間管理、備份與還原時間、並可根據您的業務需求進行設定。

理論上、SQL Server（64 位元）支援每個執行個體 32,767 個資料庫、以及 524272TB 的資料庫大小、雖然一般安裝通常有多個資料庫。不過 SQL Server 可以處理的資料庫數量取決於負載和硬體。看到 SQL Server 執行個體託管數十個、數百個甚至數千個小型資料庫、並不罕見。

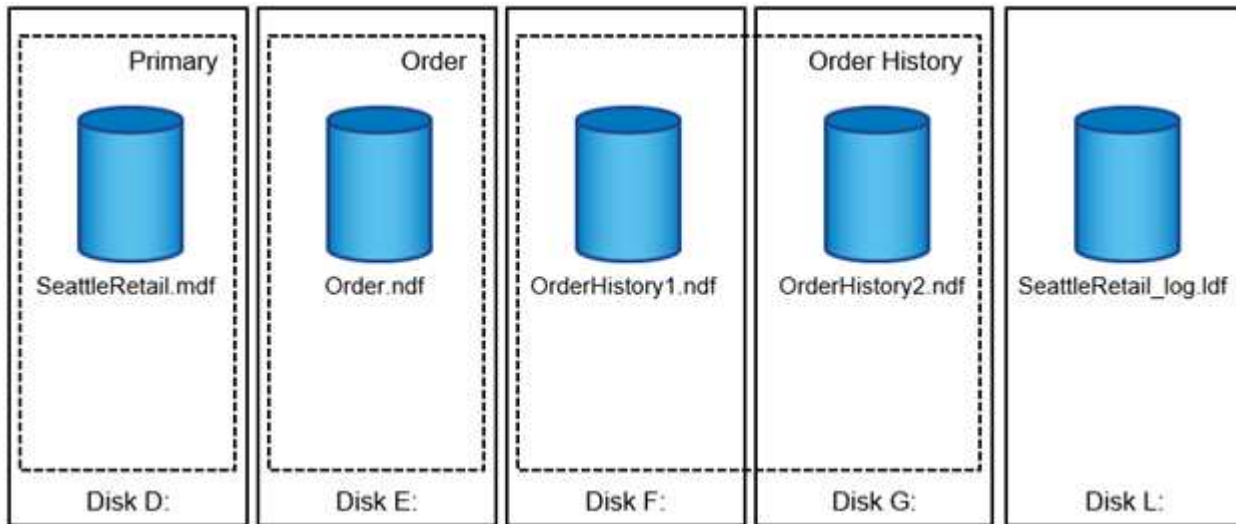
資料庫檔案與檔案群組

每個資料庫都包含一或多個資料檔案、以及一或多個交易記錄檔。交易記錄會儲存資料庫交易的相關資訊、以及每個工作階段所做的所有資料修改。每次修改資料時、SQL Server 都會在交易記錄中儲存足夠的資訊、以復原（復原）或重做（重新執行）動作。SQL Server 交易記錄是 SQL Server 在資料完整性和健全性方面聲譽的重要一環。交易記錄對於 SQL Server 的原子性、一致性、隔離和耐用性（ACID）功能至關重要。一旦資料頁發生任何變更、SQL Server 就會立即寫入交易記錄檔。每個 Data 操縱語言（DML）陳述式（例如、SELECT

、INSERT、UPDATE 或 DELETE) 都是完整的交易、而且交易記錄會確保整個以 Set 為基礎的作業都能進行、確保交易的完整性。

每個資料庫都有一個主要資料檔案、預設會有 .mdf 副檔名。此外、每個資料庫都可以有次要資料庫檔案。根據預設、這些檔案的副檔名為 .NDF。

所有資料庫檔案都會分組為檔案群組。檔案群組是邏輯單元、可簡化資料庫管理。它們允許在邏輯物件放置和實體資料庫檔案之間進行分隔。當您建立資料庫物件表格時、您可以在檔案群組中指定它們應該放置的位置、而無需擔心基礎資料檔案組態。



將多個資料檔案放入檔案群組的功能可讓您將負載分散到不同的儲存裝置、有助於改善系統的 I/O 效能。與此相反的是、由於 SQL Server 會循序寫入交易記錄檔、因此無法從多個檔案中獲益。

檔案群組中的邏輯物件放置與實體資料庫檔案之間的分隔、可讓您微調資料庫檔案配置、充分發揮儲存子系統的效益。支援給定工作負載的資料檔案數量可視需要而有所不同、以支援 I/O 需求和預期容量、而不會影響應用程式。資料庫配置的這些變化對應用程式開發人員來說是透明的、他們將資料庫物件放置在檔案群組中、而非資料庫檔案中。



* NetApp 建議 * 避免將主要檔案群組用於系統物件以外的任何項目。為使用者物件建立個別的檔案群組或一組檔案群組、可簡化資料庫管理和災難恢復、尤其是大型資料庫。

資料庫執行個體檔案初始化

您可以在建立資料庫或將新檔案新增至現有資料庫時指定初始檔案大小和自動成長參數。SQL Server 在選擇要將資料寫入哪個資料檔案時、會使用比例填滿演算法。它會將大量資料按比例寫入檔案中的可用空間。檔案中的可用空間越大、其處理的寫入次數就越多。



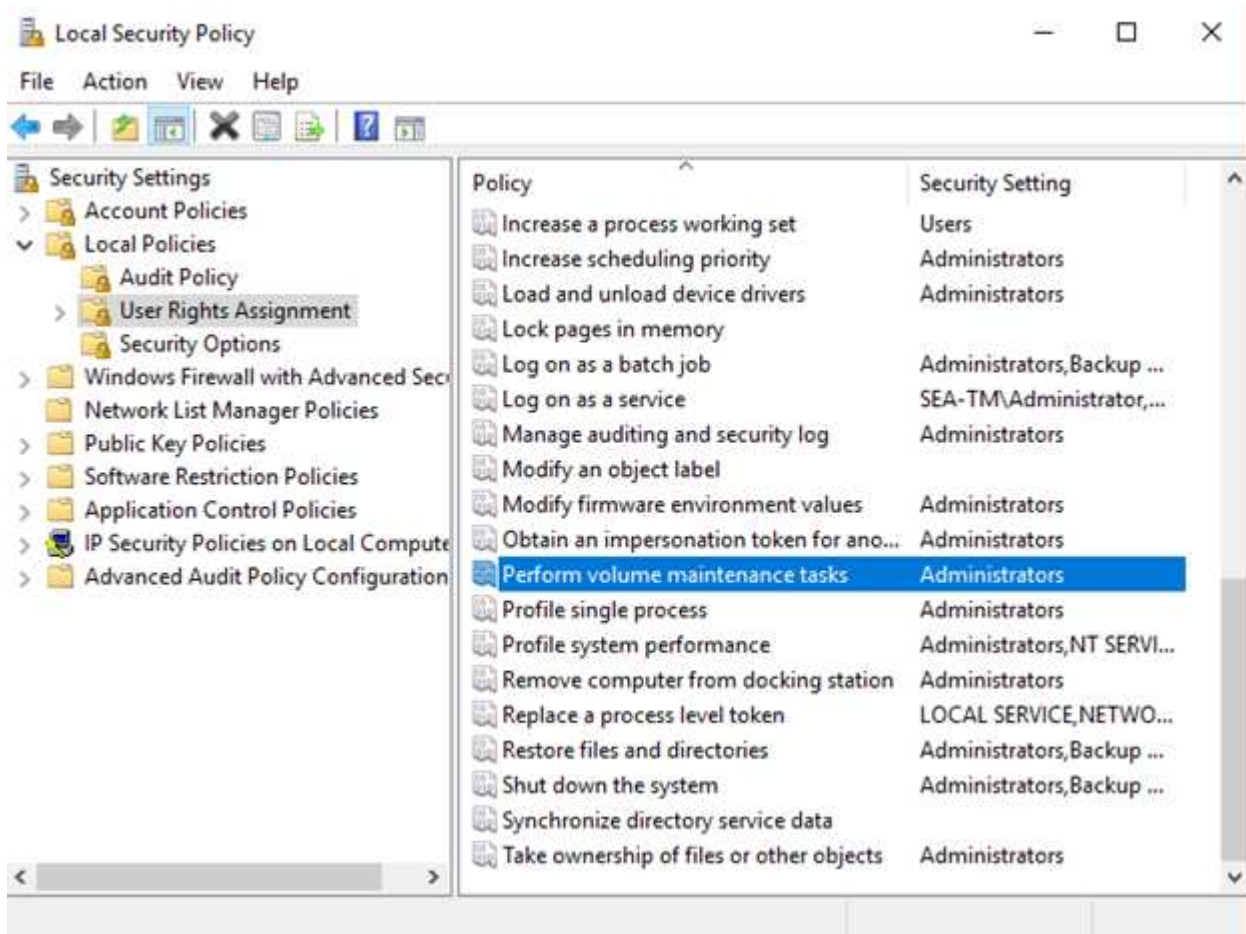
* NetApp 建議 * 單一檔案群組中的所有檔案都具有相同的初始大小和自動成長參數、其成長大小以 MB 為單位定義、而非百分比。這有助於比例填滿演算法在資料檔案之間平均平衡寫入活動。

每次 SQL Server 增加檔案時、都會以零填滿新分配的空間。該程序會封鎖所有需要寫入對應檔案的工作階段、或在交易記錄增加時產生交易記錄。

SQL Server 一律會將交易記錄檔歸零、而且該行為無法變更。不過、您可以啟用或停用即時檔案初始化來控制資料檔案是否正在歸零。啟用即時檔案初始化有助於加速資料檔案成長、並縮短建立或還原資料庫所需的時間。

與即時檔案初始化有關的安全風險較小。啟用此選項時、資料檔案的未分配部分可能會包含先前刪除的作業系統檔案資訊。資料庫管理員可以檢查這類資料。

您可以將 SA_SA_SAM_VOLUM_NAME 權限（也稱為「執行 Volume 維護工作」）新增至 SQL Server 啟動帳戶、以啟用即時檔案初始化。您可以在本機安全性原則管理應用程式（secpol.msc）下執行此動作、如下圖所示。開啟「執行 Volume 維護工作」權限的內容、並將 SQL Server 啟動帳戶新增至該處的使用者清單。



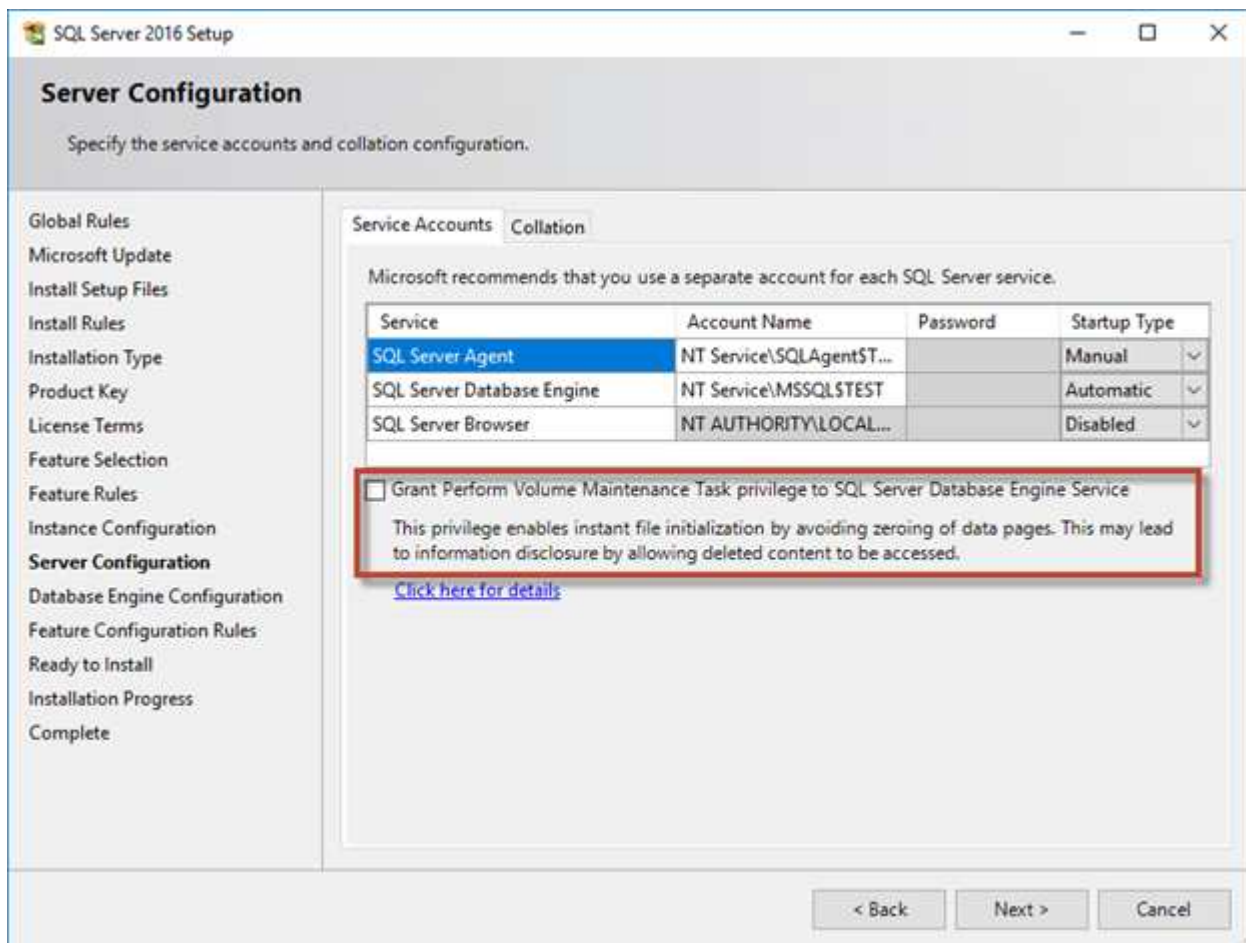
若要檢查權限是否已啟用、您可以使用下列範例中的程式碼。此程式碼會設定兩個追蹤旗標、強制 SQL Server 將其他資訊寫入錯誤記錄檔、建立小型資料庫、以及讀取記錄內容。

```
DBCC TRACEON(3004,3605,-1)
GO
CREATE DATABASE DelMe
GO
EXECUTE sp_readerrorlog
GO
DROP DATABASE DelMe
GO
DBCC TRACEOFF(3004,3605,-1)
GO
```

如果未啟用即時檔案初始化、SQL Server 錯誤記錄會顯示 SQL Server 除了將 ldf 記錄檔歸零之外、還會將 MDF 資料檔案歸零、如下例所示。當啟用即時檔案初始化時、它只會顯示記錄檔的零位。

	LogDate	ProcessInfo	Text
365	2017-02-09 08:10:07.660	spid53	Ckpt dbid 3 flush delta counts.
366	2017-02-09 08:10:07.660	spid53	Ckpt dbid 3 logging active xact info.
367	2017-02-09 08:10:07.750	spid53	Ckpt dbid 3 phase 1 ended (8)
368	2017-02-09 08:10:07.750	spid53	About to log Checkpoint end.
369	2017-02-09 08:10:07.880	spid53	Ckpt dbid 3 complete
370	2017-02-09 08:10:08.130	spid53	Starting up database 'DelMe'.
371	2017-02-09 08:10:08.150	spid53	FixupLog (ail(progress) zeroing C:\Program Files\Microsoft SQL Server\
372	2017-02-09 08:10:08.160	spid53	Zeroing C:\Program Files\Microsoft SQL Server\MSSQL
373	2017-02-09 08:10:08.170	spid53	Zeroing completed on C:\Program Files\Microsoft SQL
374	2017-02-09 08:10:08.710	spid53	Ckpt dbid 6 started
375	2017-02-09 08:10:08.710	spid53	About to log Checkpoint begin.

執行 Volume 維護工作在 SQL Server 2016 中已簡化、稍後會在安裝過程中提供選項。此圖顯示選項、可授予 SQL Server 資料庫引擎服務執行 Volume 維護工作的權限。



控制資料庫檔案大小的另一個重要資料庫選項是自動壓縮。啟用此選項時、SQL Server 會定期縮減資料庫檔案、減少檔案大小、並釋出空間給作業系統。這項作業需要大量資源、而且很少有用、因為當新資料進入系統時、資料庫檔案會在一段時間後再次增加。資料庫上不應啟用自動壓縮。

記錄目錄

記錄目錄是在 SQL Server 中指定、用於在主機層級儲存交易記錄備份資料。如果您使用 SnapCenter 來備份記錄檔、則 SnapCenter 使用的每個 SQL Server 主機都必須設定一個主機記錄目錄、才能執行記錄備份。由於包

含資料庫儲存庫、因此與備份、還原或複製作業相關的中繼資料會儲存在中央資料庫儲存庫中SnapCenter。

主機記錄目錄的大小計算方式如下：

主機記錄目錄大小 = ((最大 DB LDF 大小 x 每日記錄變更率 %)) x (快照保留) ÷ (1 - LUN 額外負荷空間 %)

主機記錄目錄大小調整公式假設 LUN 負荷空間為 10%

將記錄目錄放在專用磁碟區或 LUN 上。主機記錄目錄中的資料量取決於備份的大小和保留備份的天數。SnapCenter 每個 SQL Server 主機只允許一個主機記錄目錄。您可以在 SnapCenter → 主機 → 組態外掛程式中設定主機記錄目錄。



- NetApp 建議 * 下列主機記錄目錄：
- 請確定主機記錄目錄未被任何其他可能毀損備份快照資料的資料類型共用。
- 請勿將使用者資料庫或系統資料庫放置在裝載點的 LUN 上。
- 在 SnapCenter 複製交易記錄檔的專用磁碟區上建立主機記錄目錄。
- 使用 SnapCenter 精靈將資料庫移轉至 NetApp 儲存設備、以便將資料庫儲存在有效位置、進而成功執行 SnapCenter 備份與還原作業。請記住、移轉程序會中斷運作、並可能導致資料庫在移轉進行中時離線。
- SQL Server 的容錯移轉叢集執行個體 (FCI) 必須符合下列條件：
 - 如果您使用容錯移轉叢集執行個體、則主機記錄目錄 LUN 必須是與要備份 SnapCenter 的 SQL Server 執行個體位於同一個叢集群組中的叢集磁碟資源。
 - 如果您使用容錯移轉叢集執行個體、則使用者資料庫必須放置在共用 LUN 上、這些 LUN 是指派給與 SQL Server 執行個體相關聯的叢集群組的實體磁碟叢集資源。

儲存效率

ONTAP 儲存效率經過最佳化、可儲存及管理 SQL Server 資料、但所需的儲存空間最少、而且不會影響效能。

空間效率功能 (例如壓縮、壓縮和重複資料刪除) 的設計、是為了增加符合特定實體儲存量的邏輯資料量。結果是降低成本和管理成本。

在高層級、壓縮是一種數學程序、可偵測及編碼資料模式、以減少空間需求。相反地、重複資料刪除功能會偵測實際重複的資料區塊、並移除額外的複本。資料實作可讓多個邏輯區塊在媒體上共用相同的實體區塊。



請參閱以下關於精簡配置的章節、以瞭解儲存效率與部分保留之間互動的說明。

壓縮

在提供 All Flash 儲存系統之前、以陣列為基礎的壓縮價值有限、因為大多數 I/O 密集的工作負載都需要大量磁碟來提供可接受的效能。儲存系統的容量總是比所需的容量大得多、這是大量磁碟機的副作用。固態儲存設備的興起、改變了這種情況。不再需要純粹為了獲得良好效能而大幅過度配置磁碟機。儲存系統中的磁碟機空間可與實際容量需求相符。

固態硬碟機 (SSD) 的 IOPS 容量增加、幾乎總是比旋轉硬碟機節省成本、但壓縮技術可以增加固態媒體的有效容量、進而進一步節省成本。

壓縮資料的方法有好幾種。許多資料庫都包含自己的壓縮功能、但在客戶環境中很少會發現這種情況。其原因通常是對壓縮資料的 * 變更 * 效能會受到影響、而對於某些應用程式而言、資料庫層級壓縮的授權成本較高。最後、對資料庫作業的整體效能影響。對於執行資料壓縮與解壓縮的 CPU、而非實際的資料庫工作、支付高昂的每 CPU 授權成本是不合理的。更好的選擇是將壓縮工作卸載到儲存系統。

自適應壓縮

即使在以微秒為單位測量延遲的 All Flash 環境中、主動式壓縮也已針對企業工作負載進行徹底測試、且未對效能產生任何影響。有些客戶甚至報告使用壓縮技術時效能會提高、因為資料會保持在快取中的壓縮、有效增加控制器中可用的快取數量。

ONTAP 以 4KB 單位管理實體區塊。自適應壓縮使用 8KB 的預設壓縮區塊大小、也就是以 8KB 為單位壓縮資料。這與關係式資料庫最常使用的 8KB 區塊大小相符。隨著將更多資料壓縮成單一單元、壓縮演算法就會變得更有效率。32 KB 壓縮區塊大小比 8 KB 壓縮區塊單元更具空間效率。這表示使用預設 8KB 區塊大小的調適式壓縮確實會導致效率稍微降低、但使用較小的壓縮區塊大小也有很大的好處。資料庫工作負載包括大量的覆寫活動。若要覆寫 32 KB 壓縮資料區塊的 8KB 資料、必須讀回整個 32 KB 的邏輯資料、將其解壓縮、更新所需的 8 KB 區域、重新壓縮、然後將整個 32 KB 寫入磁碟機。這對儲存系統來說是非常昂貴的作業、也是因為某些競爭儲存陣列以較大的壓縮區塊大小為基礎、也會對資料庫工作負載造成重大效能損失的原因。



調適式壓縮所使用的區塊大小最多可增加至 32KB。這可能會改善儲存效率、而且當大量的這類資料儲存在陣列上時、應該考慮用於靜態檔案、例如交易記錄檔和備份檔案。在某些情況下、使用 16KB 或 32KB 區塊大小的作用中資料庫、也可能因為增加適應式壓縮的區塊大小而受惠。請洽詢 NetApp 或合作夥伴代表、瞭解這是否適合您的工作負載。



在串流備份目的地上、不應將大於 8KB 的壓縮區塊大小與重複資料刪除一起使用。原因是備份資料的細微變更會影響 32KB 壓縮時間。如果視窗移動、則產生的壓縮資料會在整個檔案中有所不同。重複資料刪除是在壓縮之後進行、這表示重複資料刪除引擎會以不同的方式檢視每個壓縮備份。如果需要重複資料刪除串流備份、則只應使用 8KB 區塊調適性壓縮。調適性壓縮較為理想、因為它的區塊大小較小、不會中斷重複資料刪除的效率。由於類似的原因、主機端壓縮也會影響重複資料刪除的效率。

壓縮對齊

資料庫環境中的調適性壓縮需要考量壓縮區塊對齊。這樣做只是對隨機覆寫非常特定區塊的資料的考量。這種方法的概念與整體檔案系統對齊方式類似、檔案系統的開始必須與 4K 裝置邊界對齊、檔案系統的區塊大小必須是 4K 的倍數。

例如、只有在檔案與檔案系統本身的 8KB 邊界對齊時、才會壓縮寫入 8KB 檔案。這表示它必須落在檔案的前 8KB、檔案的第二 8KB 等。確保正確對齊的最簡單方法是使用正確的 LUN 類型、建立的任何分割區都應該與 8K 的倍數裝置開始偏移、並使用資料庫區塊大小的倍數檔案系統區塊大小。

備份或交易記錄等資料會循序寫入跨越多個區塊的作業、所有這些區塊都會被壓縮。因此、不需要考慮對齊。唯一令人擔憂的 I/O 模式是隨機覆寫檔案。

資料壓縮

資料壓縮技術可改善壓縮效率。如前所述、僅有調適式壓縮功能、最多可節省 2 : 1、因為它僅限於在 4KB WAFL 區塊中儲存 8KB I/O。較大區塊大小的壓縮方法可提供更好的效率。不過、這些資料不適合受到小型區塊覆寫的資料。解壓縮 32KB 的資料單元、更新 8KB 部分、重新壓縮及回寫磁碟機、都會產生額外的負荷。

資料壓縮的運作方式是允許將多個邏輯區塊儲存在實體區塊內。例如、含有高度壓縮資料（例如文字或部分完整區塊）的資料庫、可能會從 8KB 壓縮至 1KB。如果沒有壓縮、1KB 的資料仍會佔用整個 4KB 區塊。即時資料

壓縮功能可將 1KB 的壓縮資料與其他壓縮資料一起儲存在 1KB 的實體空間中。這不是一項壓縮技術、只是在磁碟機上分配空間的一種更有效率的方法、因此不應產生任何可偵測的效能影響。

節省的程度各不相同。已壓縮或加密的資料通常無法進一步壓縮、因此資料集無法從資料壓縮中獲益。相反地、新初始化的資料檔案僅包含區塊中繼資料和零、最多可壓縮至 80 : 1。

對溫度敏感的儲存效率

ONTAP 9.8 及更新版本均提供溫度敏感儲存效率（TSSE）。它仰賴區塊存取熱圖來識別不常存取的區塊、並以更高效率加以壓縮。

重複資料刪除

重複資料刪除是從資料集移除重複的區塊大小。例如、如果 10 個不同的檔案中存在相同的 4KB 區塊、重複資料刪除會將所有 10 個檔案中的 4KB 區塊重新導向至相同的 4KB 實體區塊。結果是該資料的效率提升 10 : 1。

VMware 來賓開機 LUN 等資料通常會極好地刪除重複資料、因為這些資料包含相同作業系統檔案的多個複本。效率達到 100 : 1 以上。

部分資料不包含重複資料。例如、Oracle 區塊包含資料庫的全域唯一標頭、以及近乎唯一的標尾。因此、Oracle 資料庫的重複資料刪除功能很少能節省 1% 以上的成本。使用 MS SQL 資料庫進行重複資料刪除的效果稍微好一些、但區塊層級的獨特中繼資料仍是一項限制。

在某些情況下、使用 16KB 和大型區塊大小的資料庫可節省高達 15% 的空間。每個區塊的初始 4KB 包含全域唯一的標頭、最後 4KB 區塊則包含近乎獨特的標尾。內部區塊是重複資料刪除的候選項目、但實際上、這幾乎完全歸功於重複資料刪除零位資料。

許多競爭陣列都宣稱能夠根據資料庫複製多次的假設來刪除重複的資料庫。在這方面、也可以使用 NetApp 重複資料刪除技術、但 ONTAP 提供更好的選擇：NetApp FlexClone 技術。最終結果相同；資料庫的多個複本會建立共用大部分基礎實體區塊。使用 FlexClone 比花時間複製資料庫檔案然後刪除複製檔案更有效率。實際上、它是不重複數據刪除，而不是重複數據刪除，因為從一開始就不會創建重複數據。

效率與精簡配置

效率功能是精簡配置的形式。例如、佔用 100GB 磁碟區的 100GB LUN 可能會壓縮至 50GB。由於磁碟區仍為 100GB、因此尚未實現實際節省。必須先縮小磁碟區的大小、才能將儲存的空間用於系統的其他位置。如果稍後變更為 100GB LUN、則資料的壓縮性會降低、LUN 的大小會增加、而且磁碟區可能會填滿。

強烈建議採用精簡配置、因為它可以簡化管理、同時大幅改善可用容量、並節省相關成本。原因很簡單：資料庫環境通常包含大量的空空間、大量的磁碟區和 LUN、以及可壓縮的資料。如果磁碟區和 LUN 的儲存空間有一天 100% 滿、而且包含 100% 不可壓縮的資料、則大量資源配置會導致保留空間。這種情況不太可能發生。精簡配置可回收空間並在其他地方使用、並可讓容量管理以儲存系統本身為基礎、而非許多較小的磁碟區和 LUN。

有些客戶偏好針對特定工作負載使用完整資源配置、或是根據既定的營運和採購實務做法。



如果磁碟區是完整配置的，則必須注意完全停用該磁碟區的所有效率功能，包括使用命令解壓縮及移除重複資料刪除 `sis undo`。輸出中不應出現 `Volume volume efficiency show`。如果有、則磁碟區仍會部分設定為使用效率功能。因此、覆寫保證會以不同的方式運作、這會增加組態超視導致磁碟區意外用盡空間的機會、進而導致資料庫 I/O 錯誤。

效率最佳實務做法

- NetApp 建議 * 下列事項：

AFF 預設值

在 All Flash AFF 系統上執行的 ONTAP 上建立的磁碟區會自動精簡佈建、並啟用所有的內嵌效率功能。雖然資料庫通常無法從重複資料刪除中獲益、而且可能包含不可壓縮的資料、但預設設定仍適用於幾乎所有的工作負載。ONTAP 旨在有效處理所有類型的資料和 I/O 模式、無論是否能節省成本。只有在充分瞭解理由且有偏離的好處時、才應變更預設值。

一般建議

- 如果磁碟區和（或） LUN 並未精簡配置，您必須停用所有效率設定，因為使用這些功能並不會節省成本，而將複雜資源配置與啟用空間效率的組合，可能會導致非預期的行為，包括空間不足的錯誤。
- 如果資料不需要覆寫、例如備份或資料庫交易記錄檔、您可以在冷卻週期較短的情況下啟用 TSSE、以達到更高的效率。
- 某些檔案可能包含大量不可壓縮的資料、例如、當檔案的應用程式層級已啟用壓縮時、就會進行加密。如果上述任何情況屬實、請考慮停用壓縮、以便在包含可壓縮資料的其他磁碟區上執行更有效率的作業。
- 請勿將 32KB 壓縮和重複資料刪除同時用於資料庫備份。請參閱一節 [\[自適應壓縮\]](#) 以取得詳細資料。

資料庫壓縮

SQL Server 本身也具備可壓縮及有效管理資料的功能。SQL Server 目前支援兩種類型的資料壓縮：資料列壓縮和頁面壓縮。

資料列壓縮會變更資料儲存格式。例如、它會將整數和小數位數變更為可變長度格式、而非原生固定長度格式。它也會消除空白、將固定長度字元字串變更為可變長度格式。頁面壓縮會實作列壓縮及其他兩種壓縮策略（前置壓縮和字典壓縮）。您可以在中找到有關頁面壓縮的詳細資料 "[頁面壓縮實作](#)"。

SQL Server 2008 及更新版本的 Enterprise、Developer 及 Evaluation 版本目前支援資料壓縮。雖然壓縮可以由資料庫本身執行、但在 SQL Server 環境中很少會發生這種情況。

以下是管理 SQL Server 資料檔案空間的建議

- 在 SQL Server 環境中使用自動精簡配置、以提高空間使用率、並在使用空間保證功能時、降低整體儲存需求。
 - 對於大多數常見的部署組態、請使用自動擴充、因為儲存管理員只需要監控集合體中的空間使用量。
- 請勿在包含 SQL Server 資料檔案的 FAS 上的任何磁碟區上啟用重複資料刪除功能，除非已知該磁碟區包含相同資料的多個複本，例如將資料庫從備份還原至單一磁碟區。

空間回收

空間回收可定期啟動、以恢復 LUN 中未使用的空間。有了 SnapCenter、您可以使用下列 PowerShell 命令來啟動空間回收。

```
Invoke-SdHostVolumeSpaceReclaim -Path drive_path
```

如果您需要執行空間回收、則此程序應在低活動期間執行、因為它最初會在主機上使用週期。

資料保護

資料庫備份策略應以確定的業務需求為基礎、而非理論能力。透過結合 ONTAP 的 Snapshot 技術並運用 Microsoft SQL Server API、無論使用者資料庫的大小為何、您都能快速進行應用程式一致的備份。如需更進階或橫向擴充的資料管理需求、NetApp 提供 SnapCenter。

SnapCenter

SnapCenter 是適用於企業應用程式的 NetApp 資料保護軟體。使用適用於 SQL Server 的 SnapCenter 外掛程式、以及由 SnapCenter Plug-in for Microsoft Windows 管理的作業系統作業、即可快速輕鬆地保護 SQL Server 資料庫。

SQL Server 執行個體可以是獨立式安裝、容錯移轉叢集執行個體、也可以永遠位於可用性群組。結果是、從單一窗口即可保護、複製及還原主要或次要複本的資料庫。SnapCenter 可以同時管理內部部署、雲端和混合式組態的 SQL Server 資料庫。資料庫複本也可以在幾分鐘內在原始主機或替代主機上建立、以供開發或報告之用。

SQL Server 也需要作業系統與儲存設備之間的協調、以確保建立時快照中有正確的資料。在大多數情況下、唯一安全的方法是使用 SnapCenter 或 T-SQL。在沒有這種額外協調的情況下建立的快照、可能無法可靠地恢復。

如需適用於 SnapCenter 的 SQL Server 外掛程式的詳細資訊、請參閱 ["TR-4714：使用 NetApp SnapCenter 的 SQL Server 最佳實務指南"](#)。

使用 T-SQL 快照保護資料庫

在 SQL Server 2022 中、Microsoft 推出了 T-SQL 快照、可提供建立指令碼和自動化備份作業的路徑。您可以為資料庫準備快照、而非執行完整大小的複本。資料庫準備好備份之後、您就可以利用 ONTAP REST API 來建立快照。

以下是備份工作流範例：

1. 使用 ALTER 命令凍結資料庫。如此一來、資料庫就能在基礎儲存設備上準備一致的快照。凍結之後、您可以使用備份命令來解凍資料庫並記錄快照。
2. 使用新的備份群組和備份伺服器命令、同時在儲存磁碟區上執行多個資料庫的快照。
3. 執行完整備份或僅複製完整備份。這些備份也會記錄在 msdb 中。
4. 使用快照完整備份後以正常串流方式進行的記錄備份、執行時間點還原。如果需要、也支援串流差異備份。

若要深入瞭解、請參閱 ["瞭解 T-SQL 快照的 Microsoft 文件"](#)。



* NetApp 建議 * 使用 SnapCenter 來建立 Snapshot 複本。上述 T-SQL 方法也能正常運作，但 SnapCenter 提供完整的備份，還原及複製程序自動化功能。它也會執行探索、以確保建立正確的快照。不需要預先設定。

SQL Server 可用度群組搭配 SnapCenter

SnapCenter 支援備份使用 Windows 容錯移轉叢集設定的 SQL Server 可用性群組資料庫。

Microsoft SQL Server 的 SnapCenter 外掛程式必須安裝在 Windows 伺服器容錯移轉叢集的所有節點上。請參閱["文件"](#)先決條件和設定 SnapCenter 外掛程式的步驟。

SnapCenter 會探索 Windows 主機中的所有資料庫，執行個體和可用度群組，並在 SnapCenter 資源頁面上列舉資源。

保護 **Always On Availability** 群組中的資料庫

可用度群組中的資料庫可透過多種方式受到保護。

- 資料庫層級備份：選取資料庫資源頁面的可用度資料庫，新增包含完整 / 記錄備份的原則，排程備份。無論資料庫角色是主要複本或次要複本，SnapCenter 都會進行備份。您也可以將資料庫新增至資源群組來設定保護。
- 執行個體層級備份：選取執行個體，並根據選取的原則保護執行個體上執行的所有資料庫。所有資料庫，包括執行為主要或次要複本的可用度資料庫，都會使用 SnapCenter 進行備份。也可以將執行個體新增至資源群組來設定保護。
- 可用度群組層級備份：在設定原則時，SnapCenter 有進階選項可用於可用性群組層級備份。原則中的可用性群組設定可讓使用者選取備份的複本喜好設定。您可以選取主要複本，次要複本或全部複本。預設選項是根據 SQL Server 可用性群組組態中的備份複本集。

只有在使用可用性群組層級備份來保護可用性群組資料庫，且不適用於資料庫或執行個體層級備份時，SnapCenter 原則中的可用性群組設定才會套用。



* NetApp 建議 * 使用可用度層級備份，在 NetApp ONTAP 儲存設備上執行的所有複本之間進行備份。

在 **SnapCenter** 中設定記錄備份

如果可用性群組是在獨立式 SQL Server 安裝程式上設定，則必須在 Windows 伺服器容錯移轉叢集的每個節點上掛載專用磁碟。應使用專用磁碟來設定記錄目錄，以儲存交易記錄備份。

如果在 SQL Server 容錯移轉叢集上設定可用性群組，則應在 SQL Server 容錯移轉叢集執行個體上建立叢集磁碟，以將其移轉至主機記錄目錄。

使用 **SnapCenter** 還原可用度群組設定中的資料庫

- SnapCenter 提供重新植入選項，可從次複本可用的最新快照中自動恢復資料庫。重新連上的作業會自動還原，並將資料庫備份加入可用性群組。
- 在可用性群組中還原複本資料庫的另一種方法是中斷可用性群組並執行完整的完整和記錄還原。使用 SnapCenter 以 NORECOVERY 模式還原資料庫，然後使用 SQL Server Management Studio 或 T-SQL 將資料庫加入可用性群組。
- 若要只恢復資料子集，可以使用 SnapCenter 的複製功能來建立資料庫的複本。資料庫複本是使用 SnapCenter 在幾分鐘內建立，然後使用 SQL Server 原生工具將資料匯出至主要複本。

如需設定資料庫儲存配置以符合 RTO 和 RPO 需求的最佳實務做法["TR-4714 使用 NetApp SnapCenter 的"](#)



SnapCenter 不支援分散式可用性群組和內含可用性群組。

災難恢復

災難恢復

企業資料庫和應用程式基礎架構通常需要複寫、才能在最短的停機時間內、避免自然災難或非預期的業務中斷。

SQL Server 全年無休可用性群組複寫功能是絕佳的選項、NetApp 提供多種選項、可將資料保護與全年無休整合。不過、在某些情況下、您可能會想要考慮使用 ONTAP 複寫技術。有三個基本選項。

SnapMirror

SnapMirror 技術提供快速靈活的企業解決方案、可在 LAN 和 WAN 上複寫資料。SnapMirror 技術只會在建立初始鏡射後、將變更的資料區塊傳輸到目的地、大幅降低網路頻寬需求。可以在同步或非同步模式中設定。

NetApp MetroCluster 和 SnapMirror 作用中同步

對於許多客戶而言，災難恢復不僅需要擁有遠端資料複本，還需要能夠快速使用該資料。NetApp 提供兩種技術來滿足這種需求：MetroCluster 和 SnapMirror 主動同步

MetroCluster 指的是硬體組態中的 ONTAP、其中包括低階同步鏡射儲存設備和許多其他功能。MetroCluster 等整合式解決方案可簡化現今複雜的橫向擴充資料庫、應用程式及虛擬化基礎架構。它以一個簡單的中央儲存陣列取代多種外部資料保護產品和策略。它也能在單一叢集式儲存系統中提供整合式備份、還原、災難恢復和高可用度（HA）。

SnapMirror 主動同步是以 SnapMirror Synchronous 為基礎。使用 MetroCluster、每個 ONTAP 控制器都負責將其磁碟機資料複寫到遠端位置。有了 SnapMirror 主動式同步、您基本上擁有兩個不同的 ONTAP 系統、可維護 LUN 資料的獨立複本、但可以合作呈現該 LUN 的單一執行個體。從主機的角度來看、這是單一 LUN 實體。

SM 與 MCC 比較

SM-AS 和 MetroCluster 在整體功能上相似，但在實作 RPO = 0 複寫的方式及其管理方式上有重要差異。SnapMirror 非同步和同步也可作為 DR 計畫的一部分使用，但它們並非設計為 HA 重新安裝技術。

- MetroCluster 組態更像是一個整合式叢集，節點分散在各個站台上。SM-AS 的運作方式類似於兩個原本是相互不同的叢集，這些叢集正在合作為選取的 RPO = 0 同步複寫 LUN 提供服務。
- MetroCluster 組態中的資料只能在任何指定時間從特定站台存取。另一個資料複本位於另一個站台，但資料是被動的。如果沒有儲存系統容錯移轉，就無法存取。
- MetroCluster 和 SM-AS 執行鏡像會在不同層級執行。MetroCluster 鏡射是在 RAID 層執行。低階資料會使用 SyncMirror 以鏡射格式儲存。在 LUN，磁碟區和傳輸協定層，使用鏡像幾乎是不可見的。
- 相反地，SM-AS 鏡射則發生在傳輸協定層。這兩個叢集都是整體上的不相關叢集。資料的兩個複本同步後，兩個叢集只需鏡射寫入。當某個叢集發生寫入時，它會複寫到另一個叢集。只有當兩個站台的寫入作業完成時，才會將寫入內容確認給主機。除了這種傳輸協定分割行為之外，這兩個叢集都是正常的 ONTAP 叢集。
- MetroCluster 的主要角色是大規模複寫。您可以使用 RPO=0 和接近零的 RTO 來複寫整個陣列。這簡化了

容錯移轉程序，因為只有一件事需要容錯移轉，而且在容量和 IOPS 方面的擴充能力極佳。

- SMAS 的一個關鍵使用案例是精細複寫。有時候您不想將所有資料複寫為單一單元，或者您需要選擇性地容錯移轉特定工作負載。
- 另一個用於 SM-AS 的重要使用案例是用於雙主動式作業，您想要在兩個不同位置的兩個不同叢集上提供完全可用的資料複本，而且效能特性相同，如果需要，也不需要在各個站台之間擴充 SAN。您可以讓應用程式同時在兩個站台上執行，以降低容錯移轉作業期間的整體 RTO。

SnapMirror

以下是 SnapMirror for SQL Server 的建議：

- 如果使用 SMB、目的地 SVM 必須是來源 SVM 所屬 Active Directory 網域的成員、如此一來、在從災難恢復期間、NAS 檔案中儲存的存取控制清單（ACL）就不會中斷。
- 不需要使用與來源 Volume 名稱相同的目的地 Volume 名稱、但可讓將目的地 Volume 掛載至目的地的程序更容易管理。如果使用 SMB、您必須在來源命名空間的路徑和目錄結構中、使目的地 NAS 命名空間相同。
- 為了一致性的目的、請勿從控制器排程 SnapMirror 更新。而是在完整備份或記錄備份完成後、從 SnapCenter 啟用 SnapMirror 更新以更新 SnapMirror。
- 將包含 SQL Server 資料的磁碟區分散到叢集中的不同節點、以允許所有叢集節點共用 SnapMirror 複寫活動。此套裝作業系統可最佳化節點資源的使用。
- 在需要快速資料恢復的情況下使用同步複寫、而非同步解決方案則可在 RPO 中靈活運用。

如需 SnapMirror 的詳細資訊、請參閱 ["TR-4015：ONTAP 9 的 SnapMirror 組態與最佳實務做法指南"](#)。

MetroCluster

架構

Microsoft SQL Server 部署搭配 MetroCluster 環境需要對 MetroCluster 系統的實體設計進行一些說明。

MetroCluster 會同步鏡射不同位置或故障網域中兩個 ONTAP 叢集之間的資料和組態。MetroCluster 會自動管理兩個目標、為應用程式提供持續可用的儲存空間：

- 透過同步鏡射寫入叢集的資料、將還原點目標（RPO）歸零。
- 鏡射組態並自動化第二個站台的資料存取、讓還原時間目標（RTO）接近零。

MetroCluster 可在兩個站台的兩個不同叢集之間自動鏡射資料和組態、提供簡易性。在一個叢集內配置儲存設備時、會自動鏡射到第二個站台的第二個叢集。NetApp SyncMirror® 以零 RPO 提供所有資料的完整複本。這表示來自單一站台的工作負載可隨時切換至另一站台、並持續提供資料而不會遺失資料。MetroCluster 可管理切換程序、以便在第二個站台存取 NAS 和 SAN 資源配置的資料。MetroCluster 的驗證解決方案設計包含規模調整和組態、可在通訊協定逾時期間或更快（通常少於 120 秒）內執行轉換。這會導致接近零的 RPO、應用程式可以繼續存取資料而不會發生故障。MetroCluster 有多種由後端儲存架構定義的變化。

MetroCluster 可在 3 種不同組態中使用

- HA 可與 IP 連線配對
- HA 可與 FC 連線配對

- 單一控制器、具備 FC 連線能力



術語「連線」是指用於跨站台複寫的叢集連線。它並不指主機協定。無論叢集間通訊所使用的連線類型為何、MetroCluster 組態中的所有主機端通訊協定都會如常支援。

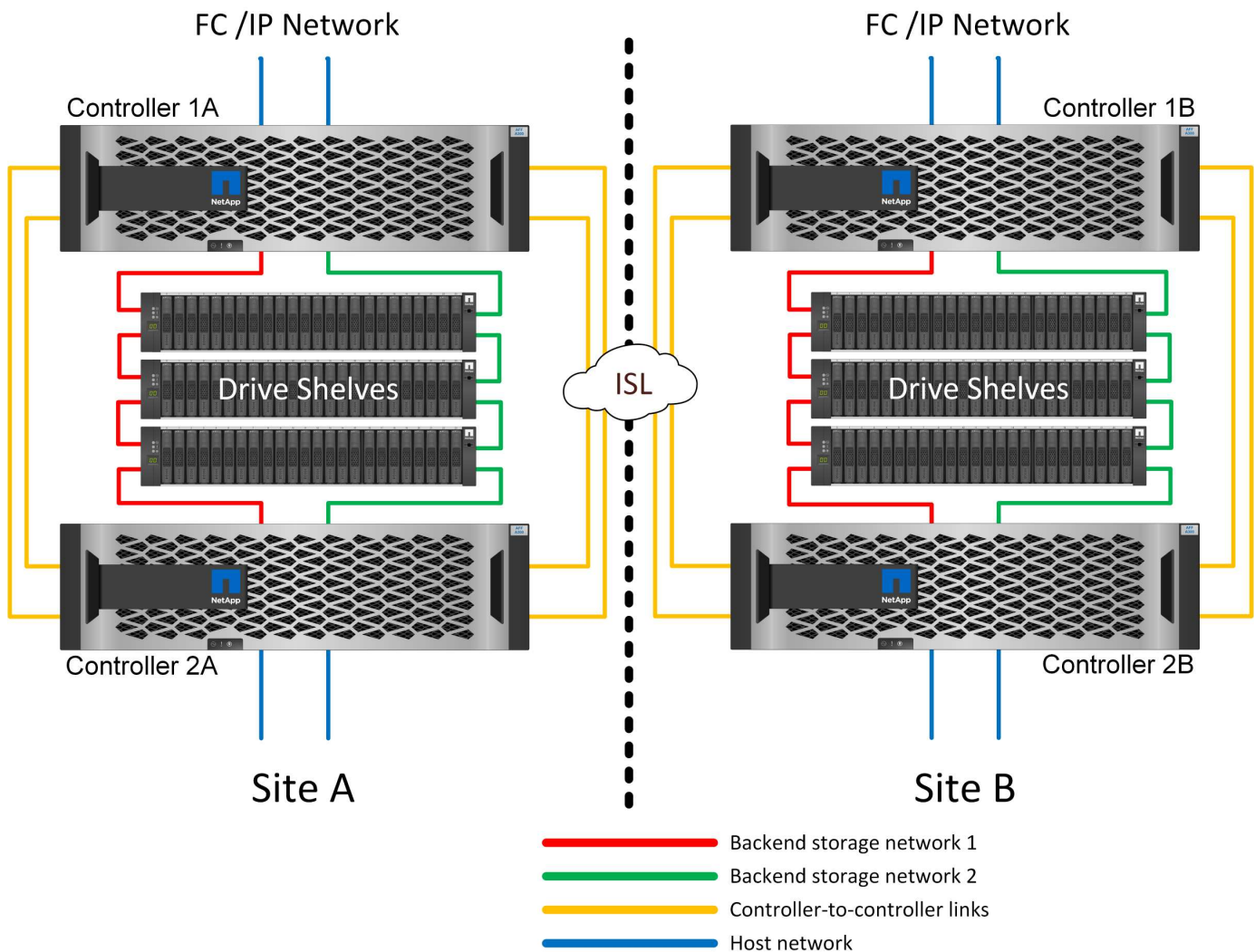
知識產權 MetroCluster

HA 配對 MetroCluster IP 組態每個站台使用兩或四個節點。此組態選項可增加與雙節點選項相關的複雜度和成本、但它提供重要的優點：站台內備援。簡單的控制器故障不需要透過 WAN 存取資料。透過替代本機控制器、資料存取仍保持在本機狀態。

大多數客戶都選擇 IP 連線、因為基礎架構需求較為簡單。過去、高速跨站台連線通常較容易使用深色光纖和 FC 交換器進行配置、但如今、高速、低延遲的 IP 電路更容易使用。

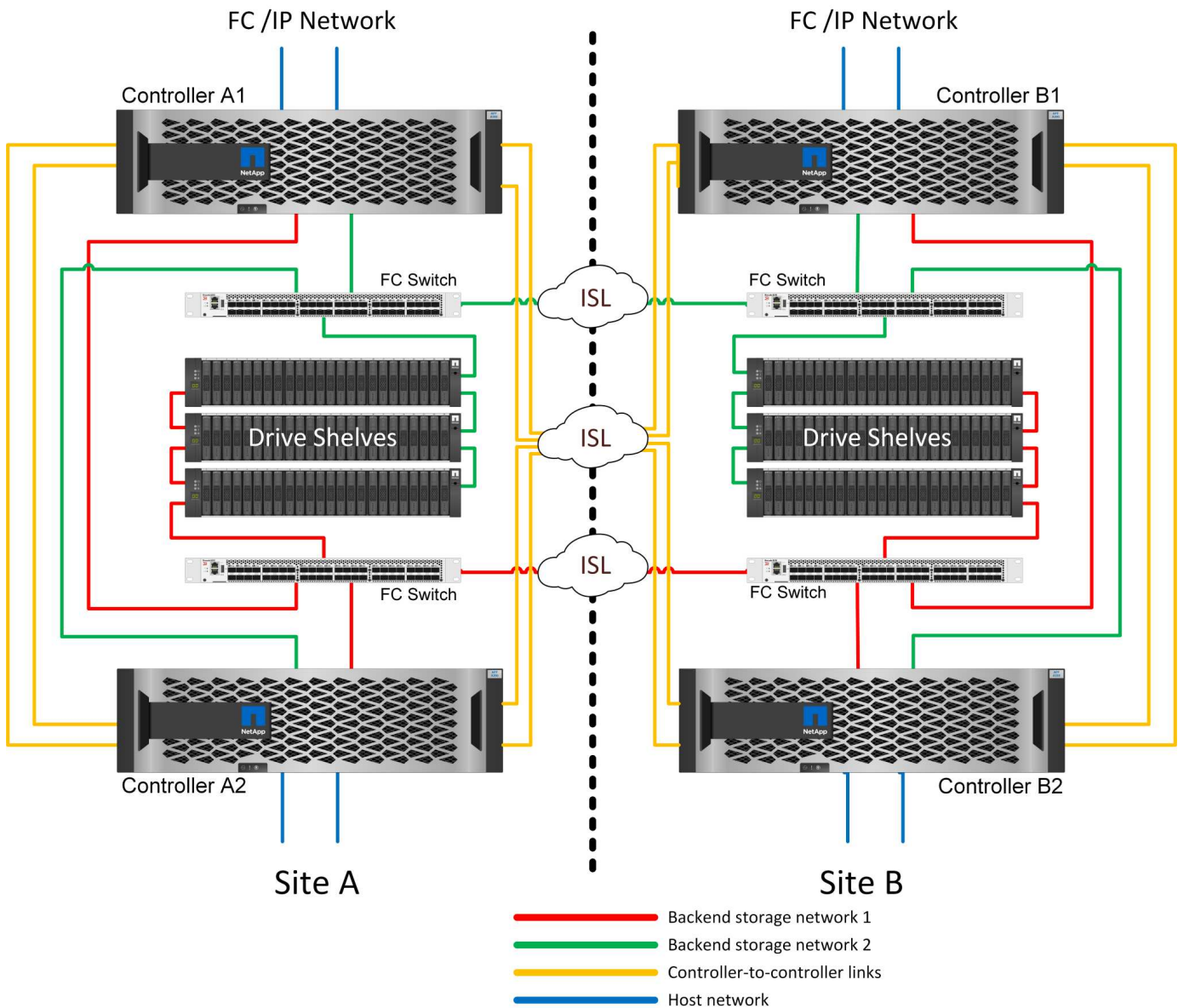
由於唯一的跨站台連線適用於控制器、因此架構也更簡單。在 FC SAN 附加 MetroCluster 中、控制器會直接寫入另一個站台上的磁碟機、因此需要額外的 SAN 連線、交換器和橋接器。相反地、IP 組態中的控制器會透過控制器寫入相對的磁碟機。

如需其他資訊、請參閱 ONTAP 正式文件和 "[SIP 解決方案架構與設計 MetroCluster](#)"。



HA 配對 FC SAN 附加 MetroCluster

HA 配對 MetroCluster FC 組態每個站台使用兩個或四個節點。此組態選項可增加與雙節點選項相關的複雜度和成本、但它提供重要的優點：站台內備援。簡單的控制器故障不需要透過 WAN 存取資料。透過替代本機控制器、資料存取仍保持在本機狀態。



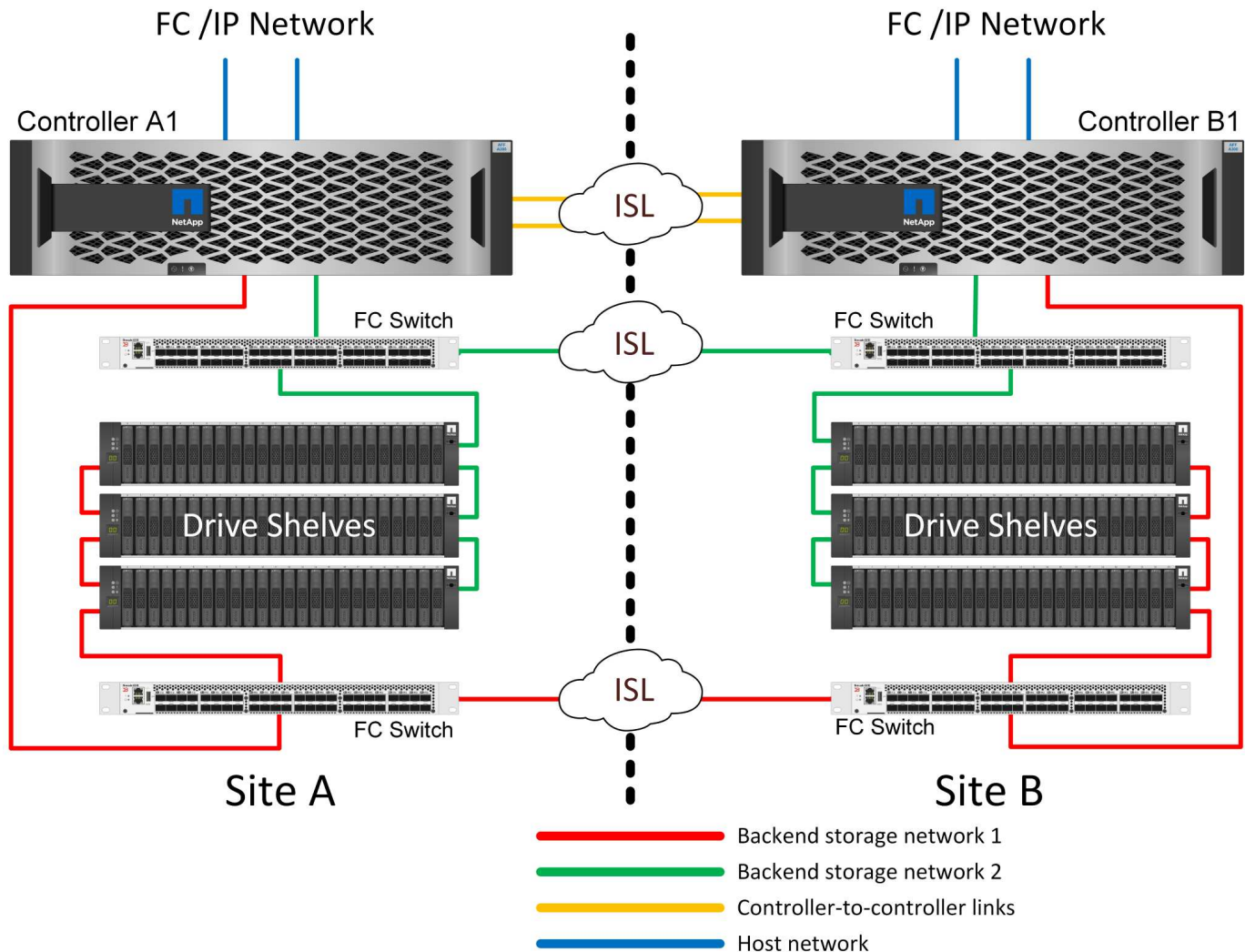
有些多站台基礎架構並非設計用於主動式作業、而是更多用於主要站台和災難恢復站台。在這種情況下、HA 配對 MetroCluster 選項通常較為理想、原因如下：

- 雖然雙節點 MetroCluster 叢集是 HA 系統、但控制器意外故障或規劃的維護作業需要資料服務必須在相反的站台上線。如果站台之間的網路連線能力不支援所需的頻寬、效能就會受到影響。唯一的選項是將各種主機作業系統和相關服務容錯移轉至替代站台。HA 配對 MetroCluster 叢集可消除此問題、因為遺失控制器會導致同一個站台內的簡單容錯移轉。
- 有些網路拓撲並非設計用於跨站台存取、而是使用不同的子網路或隔離的 FC SAN。在這種情況下、雙節點 MetroCluster 叢集不再作為 HA 系統運作、因為替代控制器無法將資料提供給位於相反站台的伺服器。HA 配對 MetroCluster 選項是提供完整備援的必要條件。
- 如果將雙站台基礎架構視為單一的高可用度基礎架構、則雙節點 MetroCluster 組態很適合。不過、如果系統

在站台故障後必須長時間運作、則最好使用 HA 配對、因為它會繼續在單一站台內提供 HA。

雙節點 FC SAN 附加 MetroCluster

雙節點 MetroCluster 組態每個站台僅使用一個節點。此設計比 HA 配對選項簡單、因為要設定和維護的元件較少。此外、它也降低了佈線和 FC 交換方面的基礎架構需求。最後、它能降低成本。



這項設計的明顯影響是、控制器在單一站台上故障、表示資料可從另一個站台取得。這種限制不一定是個問題。許多企業都有多站台資料中心作業、並有延伸、高速、低延遲的網路、基本上是一個基礎架構。在這些情況下、MetroCluster 的雙節點版本是慣用的組態。多家服務供應商目前以 PB 規模使用雙節點系統。

MetroCluster 恢復功能

MetroCluster 解決方案沒有單點故障：

- 每個控制器都有兩條通往本機站台磁碟櫃的路徑。
- 每個控制器都有兩條通往遠端站台磁碟機櫃的路徑。
- 每個控制器都有兩條通往另一個站台上控制器的路徑。
- 在 HA 配對組態中、每個控制器都有兩條路徑通往本機合作夥伴。

總而言之、您可以移除組態中的任何一個元件、而不會影響 MetroCluster 提供資料的能力。這兩個選項之間恢復能力的唯一差異是 HA 配對版本在站台故障後仍是整個 HA 儲存系統。

SyncMirror

使用 MetroCluster 保護 SQL Server 是以 SyncMirror 為基礎、提供最高效能的橫向擴充同步鏡射技術。

使用 SyncMirror 保護資料

在最簡單的層級上、同步複寫表示必須先對鏡射儲存設備的兩側進行任何變更、然後才會被確認。例如、如果資料庫正在寫入記錄檔、或是正在修補 VMware 來賓作業系統、則寫入作業絕不能遺失。作為一種協議級別，在兩個站點上的非易失性介質被認可之前，存儲系統不得確認寫入內容。只有這樣、在不遺失資料的風險下繼續作業是安全的。

使用同步複寫技術是設計和管理同步複寫解決方案的第一步。最重要的考量是瞭解在各種計畫性和非計畫性失敗案例中可能發生的情況。並非所有同步複寫解決方案都提供相同的功能。如果您需要提供零恢復點目標（RPO）的解決方案、亦即零資料遺失、則必須考慮所有故障情況。特別是、當站台之間的連線中斷而無法進行複寫時、預期會產生什麼結果？

SyncMirror 資料可用度

MetroCluster 複寫是以 NetApp SyncMirror 技術為基礎、其設計旨在有效率地切換至同步模式及從同步模式切換到同步模式。這項功能符合要求同步複寫、但也需要高可用度資料服務的客戶需求。例如、如果中斷與遠端站台的連線、通常最好讓儲存系統繼續以非複寫狀態運作。

許多同步複寫解決方案只能以同步模式運作。這種類型的全或全無複寫有時稱為 Domino 模式。這類儲存系統會停止提供資料、而不允許資料的本機和遠端複本進行非同步處理。如果複寫被強制中斷、重新同步可能會非常耗時、而且可能會讓客戶在重新建立鏡像期間暴露在完全資料遺失的風險中。

SyncMirror 不僅可以在無法連線到遠端站台時、無縫切換至同步模式、也可以在連線恢復時、快速重新同步至 RPO = 0 狀態。遠端站台的資料過時複本也可在重新同步期間保留為可用狀態、以確保資料的本機和遠端複本隨時都存在。

在需要 Domino 模式的情況下、NetApp 提供 SnapMirror 同步（SM-S）。應用程式層級選項也存在、例如 Oracle DataGuard 或 SQL Server Always On Availability Groups。作業系統層級的磁碟鏡射可以是一個選項。如需其他資訊和選項、請洽詢您的 NetApp 或合作夥伴客戶團隊。

SQL Server 搭配 MetroCluster

使用零 RPO 來保護 SQL Server 資料庫的其中一個選項是 MetroCluster。MetroCluster 是一項簡單、高效能的 RPO = 0 複寫技術、可讓您輕鬆複寫整個基礎架構到各個站台。

SQL Server 可在單一 MetroCluster 系統上擴充至數千個資料庫。可能有 SQL Server 獨立執行個體或容錯移轉叢集執行個體、MetroCluster 系統不一定會新增或變更管理資料庫的任何最佳實務做法。

MetroCluster 的完整說明不在本文件的範圍內、但原則很簡單。MetroCluster 可提供 RPO=0 複寫解決方案、並可快速容錯移轉。您在這個基礎上建置的內容取決於您的需求。

例如、突然站台遺失之後的基本快速災難恢復程序可能會使用下列基本步驟：

- 強制 MetroCluster 進行
- 執行 FC/iSCSI LUN 的探索（僅限 SAN）

- 掛載檔案系統
- 啟動 SQL Services

此方法的主要需求是在遠端站台上執行作業系統。它必須預先設定 SQL Server 安裝程式、並應以同等版本進行更新。SQL Server 系統資料庫也可以鏡射到遠端站台、並在宣告災難時掛載。

如果在切換之前、主控虛擬化資料庫的磁碟區、檔案系統和資料存放區並未在災難恢復站台上使用、則不需要在相關的磁碟區上設定 `dr-force-nvfail`。

SnapMirror 主動同步

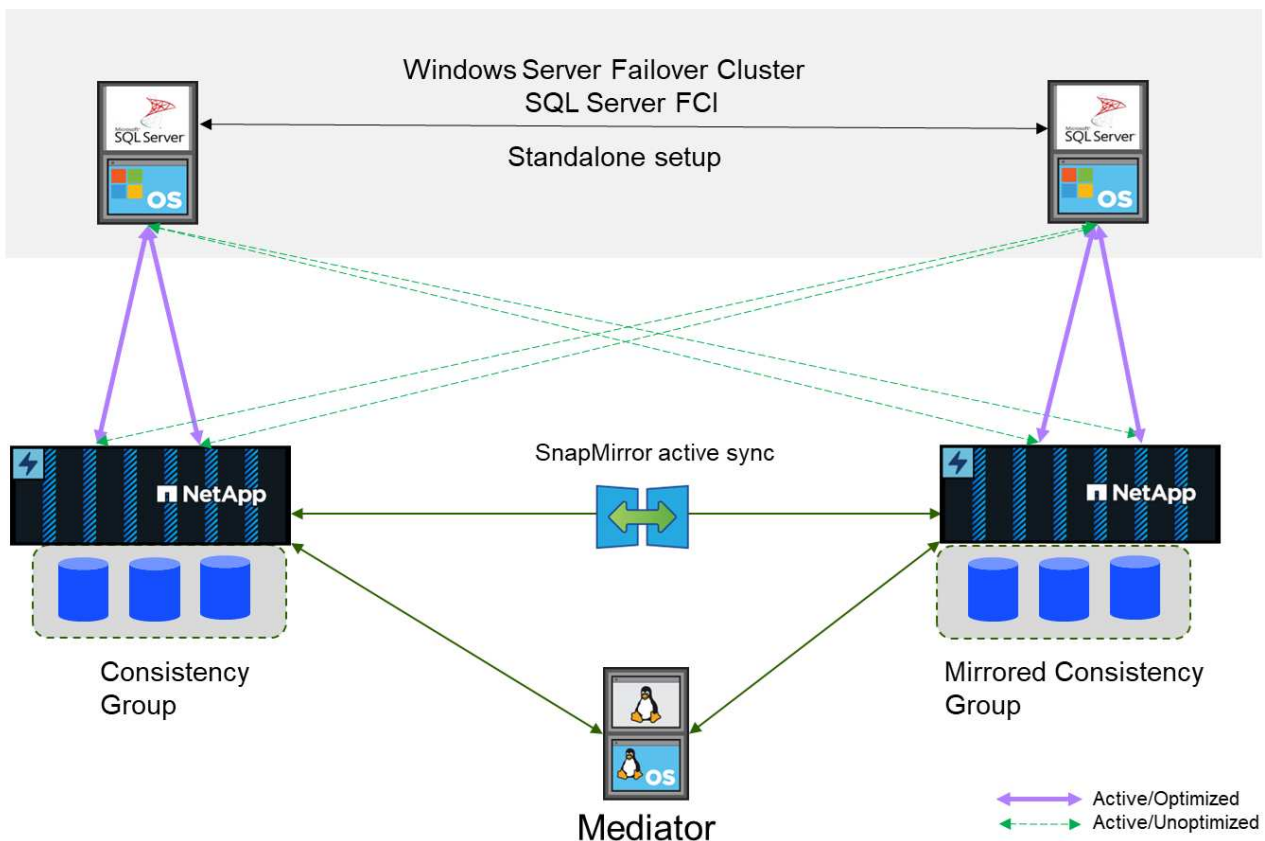
總覽

SnapMirror 主動式同步功能可讓個別 SQL Server 資料庫和應用程式在儲存設備和網路中斷期間繼續作業、並可透過透明的儲存容錯移轉、無需任何手動介入。

從 ONTAP 9.15.1 開始、SnapMirror 主動式同步除了支援現有的非對稱組態之外、還支援對稱式主動 / 主動式架構。對稱式主動 / 主動式功能可提供同步雙向複寫、以利業務持續運作和災難恢復。它可協助您保護關鍵 SAN 工作負載的資料存取、同時跨多個故障網域同時讀寫資料、確保不中斷營運、並將災難或系統故障期間的停機時間降至最低。

SQL Server 主機使用光纖通道（FC）或 iSCSI LUN 存取儲存設備。在每個託管複寫資料複本的叢集之間進行複寫。由於此功能是儲存層級複寫、因此在獨立主機或容錯移轉叢集執行個體上執行的 SQL Server 執行個體可以執行任何叢集的讀取 / 寫入作業。有關規劃和組態步驟、請參閱["SnapMirror Active Sync 的 ONTAP 文件"](#)。

SnapMirror 主動式與對稱式主動 / 主動式同步的架構



- 同步複寫 **

在正常作業中、每個複本都是 RPO=0 同步複本、但有一個例外。如果資料無法複寫、ONTAP 將釋出複寫資料的需求、並在另一個站台上的 LUN 離線時、繼續在一個站台上提供 IO 服務。

- 儲存硬體 **

與其他儲存災難恢復解決方案不同、SnapMirror 主動式同步提供非對稱式平台靈活度。每個站台的硬體不一定相同。此功能可讓您調整支援 SnapMirror 主動同步所用硬體的大小。如果遠端儲存系統需要支援完整的正式作業工作負載、則它可以與主要站台相同、但如果災難導致 I/O 減少、遠端站台上較小的系統可能會更具成本效益。

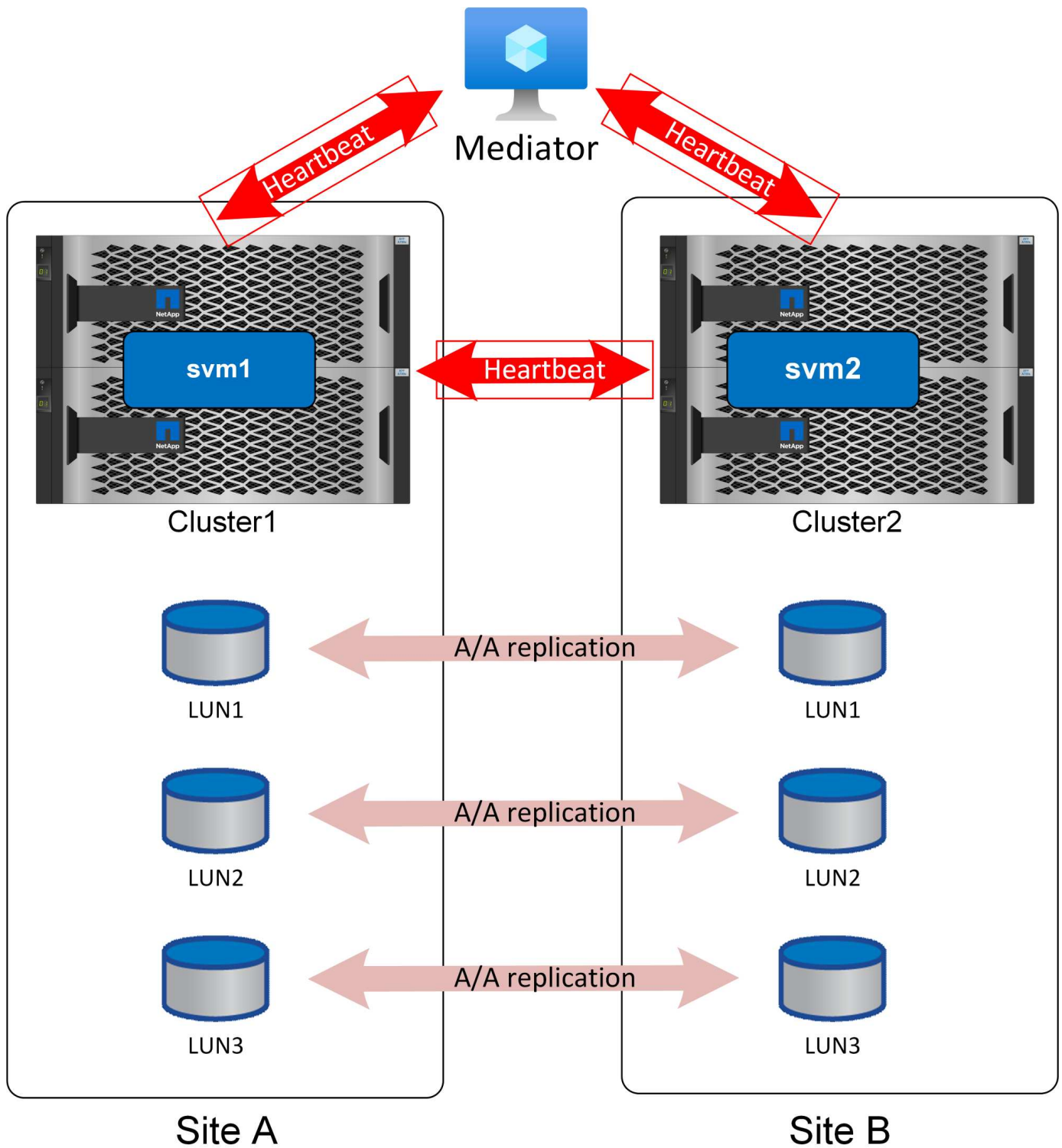
- ONTAP 調解器 **

ONTAP Mediator 是從 NetApp 支援下載的軟體應用程式、通常部署在小型虛擬機器上。ONTAP Mediator 不是一種斷路器。它是參與 SnapMirror 主動同步複寫之兩個叢集的替代通訊通道。自動化作業由 ONTAP 根據合作夥伴透過直接連線和協調員所收到的回應來驅動。

中間器ONTAP

安全地自動化容錯移轉需要 Mediator。理想情況下、它會放置在獨立的第三站台、但如果與參與複寫的叢集之一共存、則仍能滿足大多數需求。

調解員實際上並不是決勝者，儘管它實際上發揮了這樣的作用。中介器有助於確定叢集節點的狀態，並在站點發生故障時協助自動切換流程。中介在任何情況下都不會傳輸資料。



自動化容錯移轉的第一項挑戰是大腦分離問題、如果兩個站台彼此之間的連線中斷、就會發生這個問題。應該發生什麼事？您不想讓兩個不同的網站自行指定為資料的保存複本、但如何讓單一網站分辨相對網站的實際損失與無法與相對網站通訊的差異？

這是調解者輸入圖片的地方。如果放置在第三個站台上、而且每個站台都有與該站台的個別網路連線、則每個站台都有額外的路徑來驗證對方的健全狀況。請再次查看上圖、並思考下列案例。

- 如果調解器故障或無法從一個或兩個站台連線、會發生什麼情況？

- 這兩個叢集仍可透過複寫服務所使用的相同連結彼此通訊。
- 資料仍以 RPO = 0 保護提供
- 如果站台 A 故障會發生什麼情況？
 - 站台 B 會看到兩個通訊通道都中斷。
 - 站台 B 將接管資料服務、但不使用 RPO=0 鏡射
- 如果站台 B 故障會發生什麼情況？
 - 站台 A 會看到兩個通訊通道都中斷。
 - 站台 A 會接管資料服務、但不會使用 RPO=0 鏡射

還有一個案例需要考量：資料複寫連結遺失。如果站台之間的複寫連結遺失、RPO=0 鏡射顯然是不可能的。那麼應該發生什麼事？

這是由偏好的站台狀態所控制。在 SM 合夥關係中、其中一個站台是次要站台。這對正常作業沒有影響、所有資料存取都是對稱的、但如果複寫中斷、則必須中斷連結才能恢復作業。結果是首選站台將在不進行鏡射的情況下繼續作業、而次要站台將停止 IO 處理、直到複寫通訊恢復為止。

偏好的網站

SnapMirror 主動式同步處理行為是對稱的、但有一個重要的例外是偏好的站台組態。

SnapMirror 作用中同步將一個站台視為「來源」、另一個則視為「目的地」。這表示單向複寫關係、但這不適用於 IO 行為。複寫是雙向的、對稱的、而且在鏡像的兩側、IO 回應時間相同。

`\source\` 指定是控制偏好的站台。如果複寫連結遺失、來源複本上的 LUN 路徑將繼續提供資料、而目的地複本上的 LUN 路徑將無法使用、直到 SnapMirror 重新建立複寫並重新進入同步狀態為止。然後路徑將恢復服務資料。

來源 / 目的地組態可透過 SystemManager 檢視：

Relationships		
Local destinations Local sources		
Search Download Show/hide Filter		
Source	Destination	Policy type
▼ jfs_as1:/cg/jfsAA	jfs_as2:/cg/jfsAA	Synchronous

或在 CLI：

```
Cluster2::> snapmirror show -destination-path jfs_as2:/cg/jfsAA

Source Path: jfs_as1:/cg/jfsAA
Destination Path: jfs_as2:/cg/jfsAA
Relationship Type: XDP
Relationship Group Type: consistencygroup
SnapMirror Schedule: -
SnapMirror Policy Type: automated-failover-duplex
SnapMirror Policy: AutomatedFailOverDuplex
Tries Limit: -
Throttle (KB/sec): -
Mirror State: Snapmirrored
Relationship Status: InSync
```

關鍵在於來源為叢集 1 上的 SVM。如上所述、「來源」和「目的地」兩詞並未說明複寫資料的流程。這兩個站台都可以處理寫入作業、並將其複寫到另一個站台。實際上、兩個叢集都是來源和目的地。將一個叢集指定為來源的效果、只是控制在複寫連結遺失時、哪個叢集仍保留為讀寫儲存系統。

網路拓撲

統一存取

統一存取網路意味著主機可以存取兩個站台（或同一個站台內的故障網域）上的路徑。

SM — as 的一項重要功能是能夠設定儲存系統、以瞭解主機的位置。將 LUN 對應至指定主機時、您可以指出 LUN 是否接近指定的儲存系統。

特殊警示點設定

特殊警示是指每個叢集的組態、表示特定主機 WWN 或 iSCSI 啟動器 ID 屬於本機主機。這是設定 LUN 存取的第二個選用步驟。

第一步是一般的 igroup 組態。每個 LUN 都必須對應至包含需要存取該 LUN 之主機的 WWN/iSCSI ID 的 igroup。這會控制哪些主機擁有對 LUN 的 *access*。

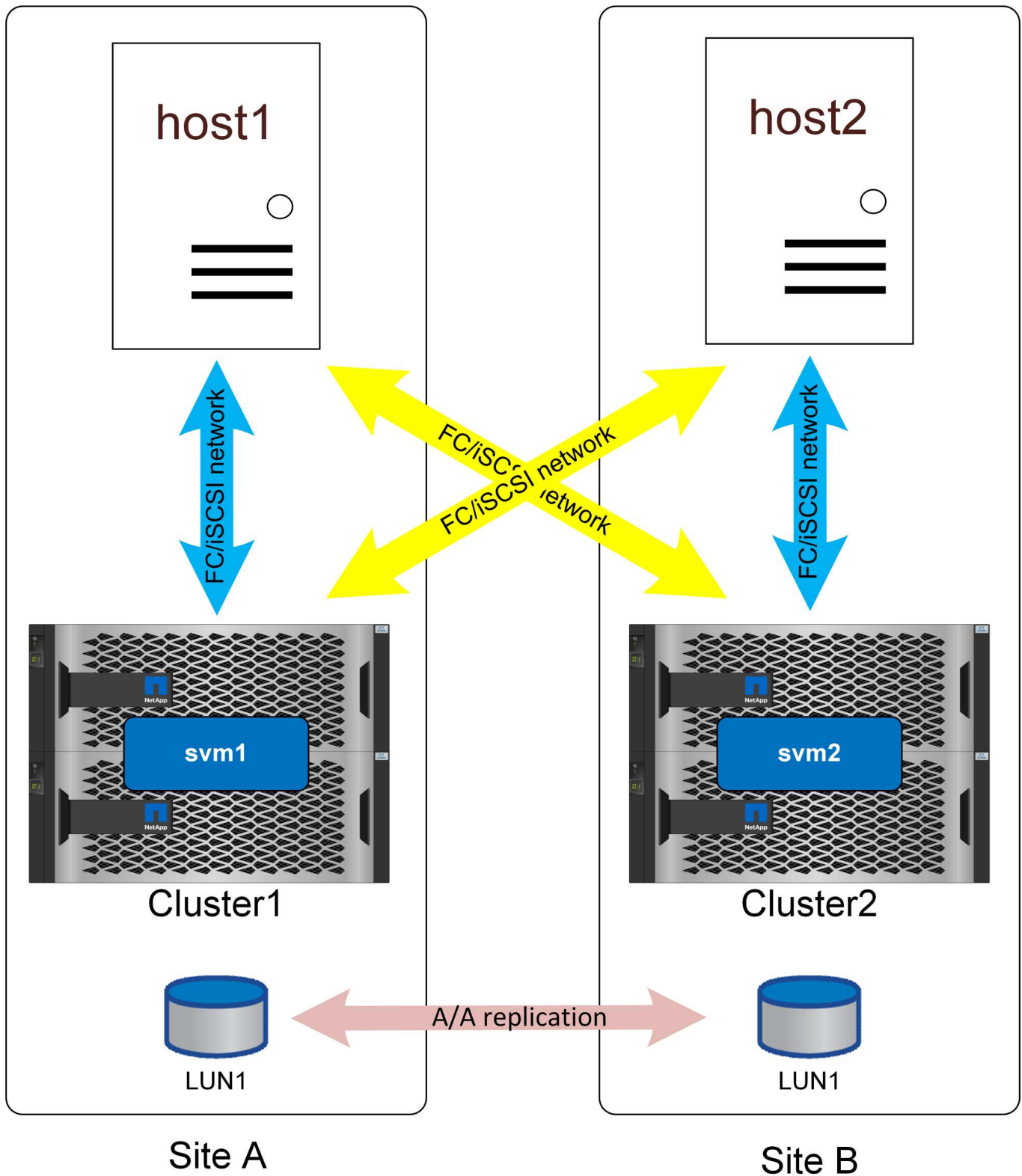
第二個選用步驟是設定主機鄰近度。這無法控制存取、而是控制 *priority*。

例如、站台 A 的主機可能設定為存取受 SnapMirror 主動式同步保護的 LUN、而且由於 SAN 延伸至站台、因此該 LUN 可使用站台 A 上的儲存設備或站台 B 上的儲存設備來存取路徑

如果沒有特殊警示點設定、則該主機會同時使用兩個儲存系統、因為這兩個儲存系統都會通告主動 / 最佳化的路徑。如果站台之間的 SAN 延遲和 / 或頻寬受到限制、這可能無法進行設計、您可能希望確保在正常作業期間、每個主機都優先使用本機儲存系統的路徑。這是透過將主機 WWN/iSCSI ID 新增至本機叢集做為近端主機來設定的。這可以在 CLI 或 SystemManager 上完成。

AFF

在 AFF 系統中、設定主機鄰近時、路徑會如下所示。



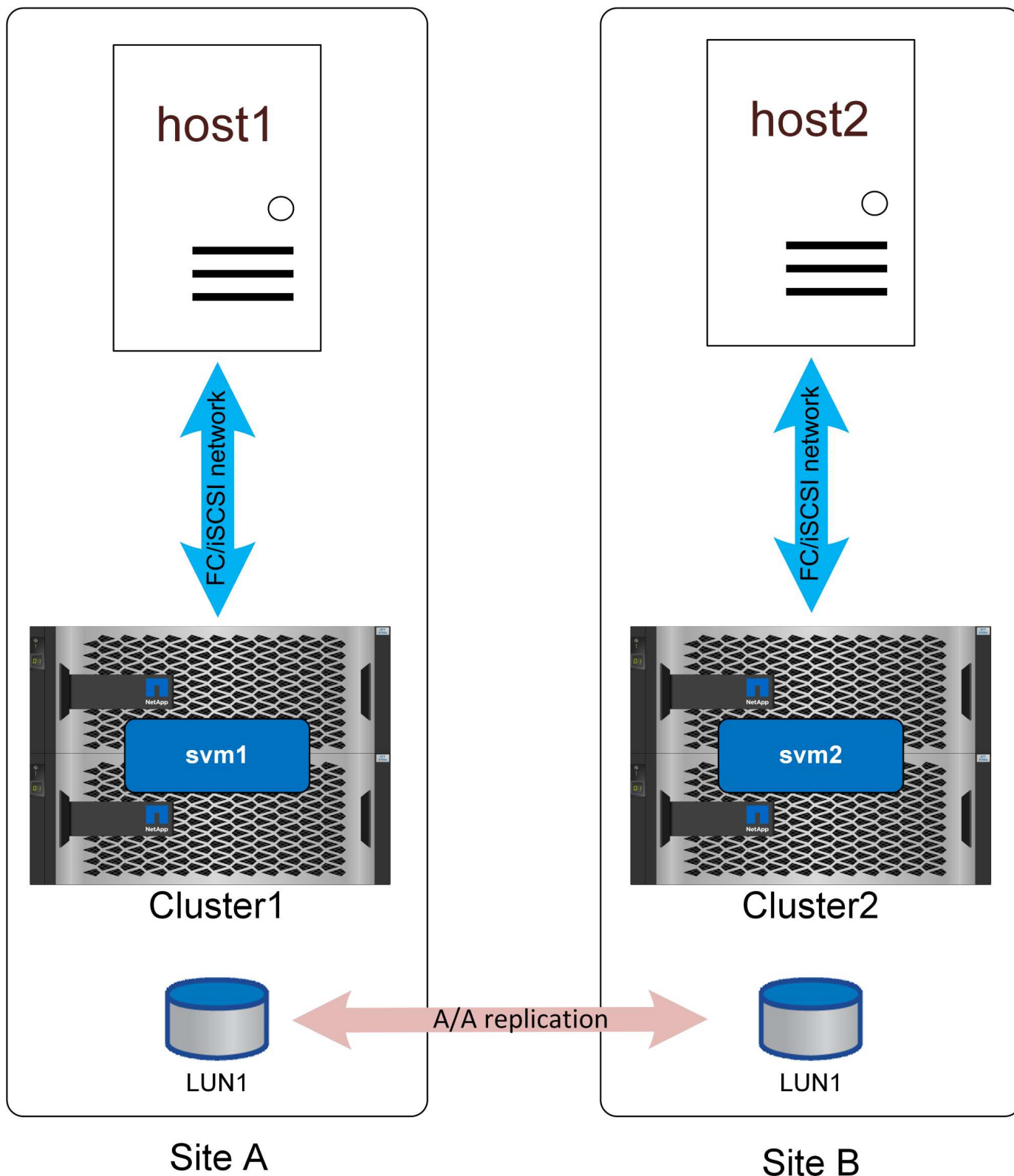
在正常作業中、所有 IO 都是本機 IO 。從本機儲存陣列提供讀取和寫入服務。寫入 IO 當然也需要由本機控制器複寫到遠端系統、然後才會被確認、但所有讀取 IO 都會在本機上提供服務、而且不會因在站台之間瀏覽 SAN 連結而產生額外延遲。

只有當所有主動 / 最佳化路徑都遺失時、才會使用非最佳化路徑。例如、如果站台 A 上的整個陣列失去電力、站台 A 的主機仍能存取站台 B 上陣列的路徑、因此仍可繼續運作、雖然延遲會較高。

由於簡單起見、本機叢集有多個備援路徑未顯示在這些圖表中。ONTAP 儲存系統本身就是 HA 、因此控制器故障不應導致站台故障。只會導致受影響網站上使用本機路徑的變更。

不一致的存取

非統一存取網路表示每部主機只能存取本機儲存系統上的連接埠。SAN 不會延伸至站台（或同一站台內的故障網域）。



Active/Optimized Path

這種方法的主要優點是 SAN 簡易性、您無需透過網路擴充 SAN。有些客戶在站台之間沒有足夠的低延遲連線、或缺乏基礎架構、無法透過站台間網路來通道 FC SAN 流量。

不一致存取的缺點是、某些失敗情況（包括遺失複寫連結）會導致部分主機失去儲存設備的存取權。以單一執行個體執行的應用程式、例如原本只在任何指定掛載的單一主機上執行的非叢集資料庫、如果本機儲存連線中斷、就會失敗。資料仍會受到保護、但資料庫伺服器將無法再存取。需要在遠端站台上重新啟動、最好是透過自動化程序來重新啟動。例如、VMware HA 可偵測一部伺服器上的所有路徑停機情況、並在另一部可用路徑的伺服器上重新啟動 VM。

相反地、叢集式應用程式（例如 Oracle RAC）可提供在兩個不同站台同時可用的服務。失去站台並不代表整個應用程式服務都會遺失。執行個體仍可在仍正常運作的站台上執行。

在許多情況下、透過站台對站台連結存取儲存設備的應用程式額外延遲成本是不可接受的。這表示統一網路的可用度提升到最低、因為站台上的儲存設備遺失、可能導致仍需要關閉該故障站台上的服務。



由於簡單起見、本機叢集有多個備援路徑未顯示在這些圖表中。ONTAP 儲存系統本身就是 HA、因此控制器故障不應導致站台故障。只會導致受影響網站上使用本機路徑的變更。

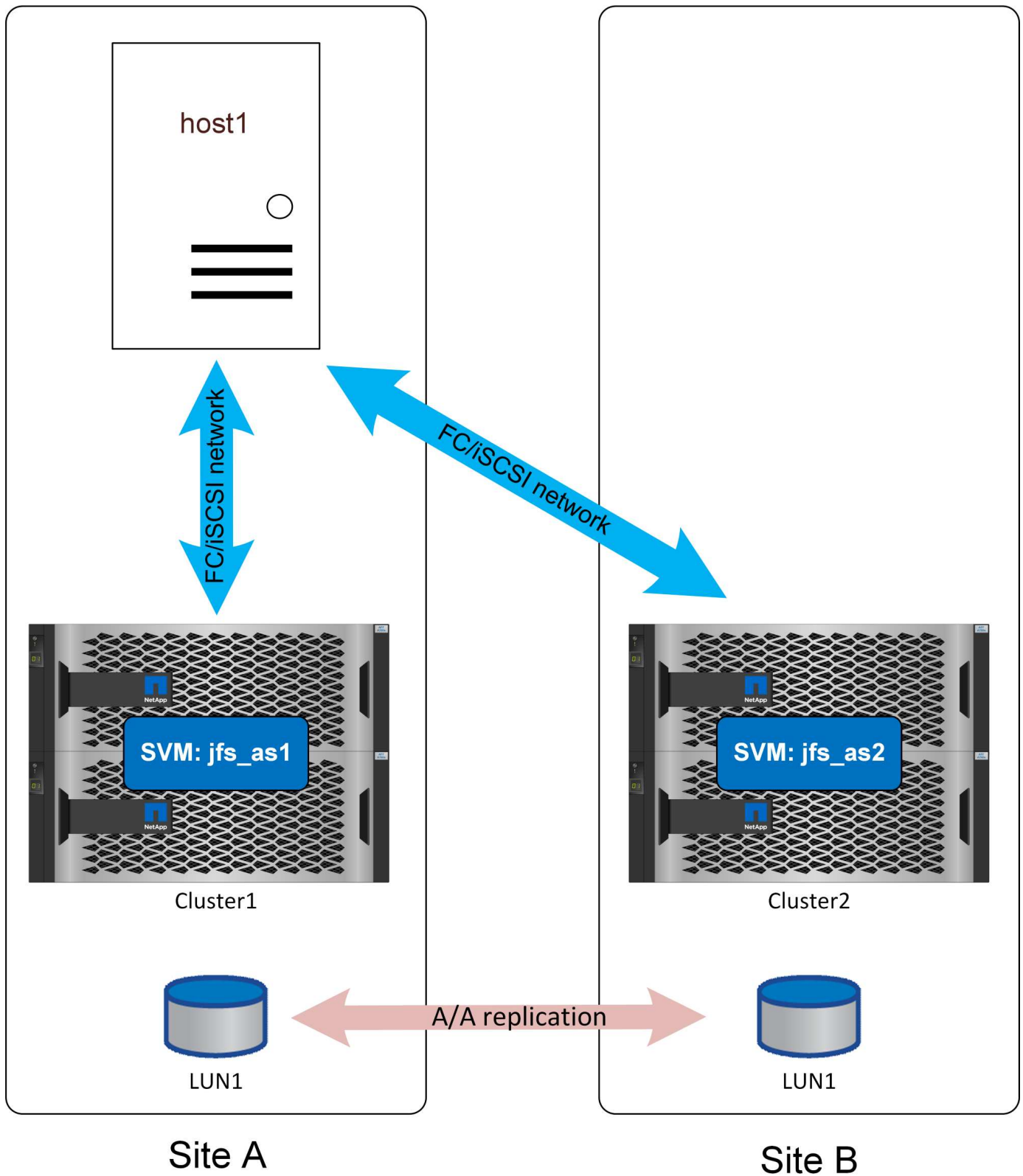
總覽

SQL Server 可設定為以多種方式搭配 SnapMirror 作用中同步處理。正確的答案取決於可用的網路連線能力、RPO 需求和可用度需求。

SQL Server 獨立執行個體

檔案配置和伺服器組態的最佳實務做法與文件中建議的做法相同"[ONTAP 上的 SQL Server](#)"。

有了獨立安裝程式、SQL Server 只能在一個站台上執行。可能"[統一](#)"會使用存取權。



在統一存取的情況下、任一站台的儲存設備故障都不會中斷資料庫作業。包含資料庫伺服器的站台發生完整的站台故障、當然會導致停機。

有些客戶可以使用預先設定的 SQL Server 安裝程式來設定在遠端站台上執行的作業系統、並以與正式作業執行個體相同的建置版本進行更新。容錯移轉需要在替代站台上啟動該 SQL Server 獨立執行個體、探索 LUN、以及啟動資料庫。完整程序可透過 Windows PowerShell Cmdlet 自動化、因為儲存端不需要任何作業。

"不一致"也可以使用存取、但如果資料庫伺服器所在的儲存系統故障、則會導致資料庫中斷、因為資料庫沒有可用的儲存路徑。在某些情況下、這點仍可接受。SnapMirror 主動式同步仍會提供 RPO = 0 資料保護、萬一站台發生故障、仍在運作中的複本就會開始運作、並可使用與上述統一存取相同的程序來恢復作業。

使用虛擬化主機可更輕鬆地設定簡單且自動化的容錯移轉程序。例如、如果 SQL Server 資料檔案與開機 VMDK 同步複寫至次要儲存設備、則在發生災難時、可在其他站台啟動完整的環境。系統管理員可以手動啟動仍在運作的站台上的主機、或是透過 VMware HA 等服務來自動化程序。

SQL Server 容錯移轉叢集執行個體

SQL Server 容錯移轉執行個體也可以裝載在實體伺服器或虛擬伺服器上執行的 Windows 容錯移轉叢集上、做為客體作業系統。這種多主機架構可提供 SQL Server 執行個體和儲存恢復能力。這類部署在尋求健全容錯移轉程序、同時維持強化效能的高需求環境中非常有用。在容錯移轉叢集設定中、當主機或主要儲存設備受到影響時、SQL Services 會容錯移轉至次要主機、同時次要儲存設備也可用於 IO。無需自動化指令碼或系統管理員介入。

故障案例

規劃完整的 SnapMirror 主動式同步應用程式架構時、需要瞭解 SM-AS 在各種計畫性和非計畫性容錯移轉案例中的回應方式。

針對下列範例、假設站台 A 已設定為慣用站台。

喪失複寫連線能力

如果 SM-AS 複寫中斷、寫入 IO 就無法完成、因為叢集無法將變更複寫到相反的站台。

站台 A (慣用站台)

偏好的站台上的複寫連結失敗、在寫入 IO 處理中會有大約 15 秒的暫停、因為 ONTAP 會在判斷複寫連結確實無法連線之前、重試複寫的寫入作業。15 秒後、站台 A 系統會恢復讀寫 IO 處理。SAN 路徑不會變更、LUN 也會保持連線。

站台 B

由於站台 B 不是 SnapMirror 作用中同步偏好的站台、因此其 LUN 路徑將在大約 15 秒後變成無法使用。

儲存系統故障

儲存系統故障的結果與遺失複寫連結的結果幾乎完全相同。當仍在運作的站台發生 IO 暫停約 15 秒。一旦超過 15 秒、IO 就會像往常一樣繼續在該站台上進行。

調解員遺失

中介服務無法直接控制儲存作業。它可作為叢集之間的替代控制路徑。它主要用於自動化容錯移轉、而不會有發生分裂的風險。在正常作業中、每個叢集都會將變更複寫到其合作夥伴、因此每個叢集都可以驗證合作夥伴叢集是否在線上並提供資料。如果複寫連結失敗、複寫就會停止。

安全自動容錯移轉需要協調員、因為否則儲存叢集就無法判斷雙向通訊是否因為網路中斷或實際儲存設備故障而中斷。

中介程序為每個叢集提供替代路徑、以驗證其合作夥伴的健全狀況。案例如下：

- 如果叢集可以直接聯絡其合作夥伴、複寫服務就可以運作。無需採取任何行動。
- 如果偏好的站台無法直接聯絡其合作夥伴或透過中介人聯絡、則會假設該合作夥伴實際上無法使用、或是被隔離、並已將其 LUN 路徑離線。接著、偏好的站台會繼續釋放 RPO=0 狀態、並繼續處理讀取和寫入 IO。
- 如果非偏好的站台無法直接聯絡其合作夥伴、但可以透過協調器聯絡、則會使其路徑離線、並等待複寫連線的恢復。
- 如果非偏好的站台無法直接或透過營運協調員聯絡其合作夥伴、則會假設該合作夥伴實際上無法使用、或是被隔離、並已將其 LUN 路徑離線。然後、非偏好的站台會繼續釋放 RPO = 0 狀態、並繼續處理讀取和寫入 IO。它將扮演複寫來源的角色、並將成為新的慣用站台。

如果調解器完全無法使用：

- 複寫服務因任何原因而失敗、包括非慣用站台或儲存系統故障、將導致偏好的站台釋放 RPO = 0 狀態、並恢復讀寫 IO 處理。非慣用站台將使其路徑離線。
- 偏好的站台故障將導致中斷、因為非偏好的站台將無法驗證相對站台是否確實離線、因此非偏好的站台無法安全恢復服務。

還原服務

解決故障（例如還原站台對站台連線或啟動故障系統）後、SnapMirror 作用中同步端點會自動偵測是否存在錯誤的複寫關係、並將其恢復至 RPO=0 狀態。重新建立同步複寫後、故障路徑將再次上線。

在許多情況下、叢集式應用程式會自動偵測失敗路徑的傳回、這些應用程式也會重新上線。在其他情況下、可能需要主機層級的 SAN 掃描、或是需要手動將應用程式恢復上線。這取決於應用程式及其設定方式、一般而言、這類工作可以輕鬆自動化。ONTAP 本身具有自我修復功能、不應需要任何使用者介入、即可恢復 RPO = 0 儲存作業。

手動容錯移轉

變更偏好的站台需要簡單的操作。IO 會暫停一秒或兩秒、作為叢集之間複寫行為切換的權限、但 IO 不會受到影響。

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。