



Oracle 資料保護

Enterprise applications

NetApp
May 09, 2024

目錄

Oracle 資料保護	1
使用 ONTAP 保護 Oracle 資料	1
Oracle 資料庫 RTO、RPO 和 SLA 規劃	1
ONTAP 提供 Oracle 資料庫可用度	3
Checksum 與 Oracle 資料庫完整性	4
備份與還原基礎知識	9

Oracle 資料保護

使用 ONTAP 保護 Oracle 資料

NetApp 知道資料庫中有最重要的關鍵任務資料。

企業無法在沒有資料存取權的情況下運作、有時資料會定義業務。此資料必須受到保護；然而、資料保護不只是確保可用的備份、還要快速可靠地執行備份、而且還要安全地儲存。

資料保護的另一面是資料恢復。當資料無法存取時、企業就會受到影響、而且可能無法運作、直到資料還原為止。此程序必須快速且可靠。最後、大多數資料庫都必須防範災難、這表示必須維護資料庫的複本。複本必須是最新的。複本也必須快速且簡單、才能使其成為完全運作的資料庫。



本文件取代先前發佈的技術報告 [_TR-4591](#)：Oracle 資料保護：備份、還原及複寫。

規劃

正確的企業資料保護架構、取決於資料保留、可恢復性、以及各種事件中斷的容忍度等業務需求。

例如、請考慮範圍內的應用程式、資料庫和重要資料集數量。為單一資料集建立備份策略、以確保符合典型 SLA 的要求相當簡單、因為沒有太多物件需要管理。隨著資料集數量增加、監控作業變得更複雜、系統管理員可能不得不花費更多時間來解決備份故障。當環境達到雲端與服務供應商的規模時、需要完全不同的方法。

資料集大小也會影響策略。例如、由於資料集太小、因此有許多選項可用於 100GB 資料庫的備份與還原。只要使用傳統工具從備份媒體複製資料、通常就能提供足夠的恢復 RTO。100TB 資料庫通常需要完全不同的策略、除非 RTO 允許多天中斷、否則傳統的複製備份與還原程序可能是可以接受的。

最後、除了備份與還原程序本身之外、還有其他因素。例如、是否有支援關鍵生產活動的資料庫、使恢復成為僅由熟練的 DBA 執行的罕見事件？或者、資料庫是否是大型開發環境的一部分、而在大型開發環境中、恢復是經常發生的事、並由一群通才的 IT 團隊負責管理？

Oracle 資料庫 RTO、RPO 和 SLA 規劃

ONTAP 可讓您輕鬆根據業務需求量身打造 Oracle 資料庫資料保護策略。

這些需求包括恢復速度、最大允許資料遺失量、以及備份保留需求等因素。資料保護計畫也必須考量資料保留與還原的各種法規要求。最後、必須考量不同的資料恢復情境、從使用者或應用程式錯誤所造成的典型和可預見的恢復、到包括站點完全遺失的災難恢復情境。

資料保護與還原原則的細微變更、可能會對儲存、備份與還原的整體架構造成重大影響。在開始設計工作之前、務必先定義並記錄標準、以免資料保護架構複雜化。不必要的功能或保護層級會導致不必要的成本和管理成本、而最初忽略的需求可能導致專案方向錯誤、或是需要最後一分鐘的設計變更。

恢復時間目標

恢復時間目標（RTO）定義了恢復服務所允許的最長時間。例如、人力資源資料庫的 RTO 可能為 24 小時、因為雖然在工作天內無法存取這些資料、但企業仍可繼續營運。相反地、支援銀行總分類帳的資料庫、其 RTO 會以分鐘甚至秒為單位來衡量。RTO 為零是不可能的、因為必須有辦法區分實際服務中斷和例行事件、例如遺失的網路封包。然而、典型的要求是接近零的 RTO。

恢復點目標

恢復點目標（RPO）定義了最大可容忍的資料遺失量。在許多情況下、RPO 完全取決於快照或 SnapMirror 更新的頻率。

在某些情況下、RPO 可能會變得更具侵略性、因此可選擇性地更頻繁地保護某些資料。在資料庫內容中、RPO 通常是在特定情況下可能遺失多少記錄資料的問題。在資料庫因產品錯誤或使用者錯誤而受損的典型還原案例中、RPO 應為零、表示不應有資料遺失。恢復程序包括還原較早的資料庫檔案複本、然後重新播放記錄檔、將資料庫狀態提升至所需的時間點。此作業所需的記錄檔應已在原始位置中。

在不尋常的情況下、記錄資料可能會遺失。例如、意外或惡意 `rm -rf *` 資料庫檔案可能會導致刪除所有資料。唯一的選擇是從備份還原、包括記錄檔、有些資料必然會遺失。在傳統備份環境中改善 RPO 的唯一選項是執行記錄資料的重複備份。然而、由於資料持續移動、而且難以將備份系統維持為持續運作的服務、因此這項功能也有其侷限性。進階儲存系統的優點之一、就是能夠保護資料、避免意外或惡意損壞檔案、進而在不移動資料的情況下提供更好的 RPO。

災難恢復

災難恢復包括在發生實體災難時恢復服務所需的 IT 架構、原則和程序。這可能包括洪水、火災或惡意或疏忽意圖的人。

災難恢復不只是一套恢復程序。這是識別各種風險、定義資料恢復和服務持續性需求、以及提供適當架構及相關程序的完整程序。

在建立資料保護需求時、必須區分典型的 RPO 和 RTO 需求、以及災難恢復所需的 RPO 和 RTO 需求。有些應用程式環境需要零 RPO 和近乎零的 RTO、才能因應資料遺失的情況、從相對正常的使用者錯誤到破壞資料中心的火災。然而、這些高層級的保護措施會產生成本和管理上的後果。

一般而言、非災難性資料恢復需求應嚴格、原因有兩個。首先、應用程式錯誤和使用者錯誤會導致資料受損、幾乎是不可避免的。其次、只要儲存系統未遭銷毀、就不難設計出可提供零 RPO 和低 RTO 的備份策略。沒有理由不解決容易補救的重大風險、這就是為何用於本機恢復的 RPO 和 RTO 目標應該積極主動的原因。

災難恢復 RTO 和 RPO 需求因災難可能性及相關資料遺失或業務中斷所造成的後果而異。RPO 和 RTO 需求應以實際業務需求為基礎、而非一般原則。他們必須考慮多種邏輯和實體災難案例。

邏輯災難

邏輯災難包括使用者、應用程式或作業系統錯誤所造成的資料毀損、以及軟體故障。邏輯災難也可能包括由外部人士利用病毒或蠕蟲或利用應用程式弱點發動的惡意攻擊。在這些情況下、實體基礎架構並未受損、但基礎資料已不再有效。

越來越常見的邏輯災難類型稱為勒索軟體、其中攻擊模式是用來加密資料。加密不會損壞資料、但會在付款給第三方之前無法使用。越來越多的企業被勒索軟體攻擊的目標鎖定。針對此威脅、NetApp 提供防竄改快照、即使儲存管理員也無法在設定的到期日之前變更受保護的資料。

實體災難

實體災難包括基礎架構元件故障、其超過備援功能、導致資料遺失或服務延長中斷。例如、RAID 保護可提供磁碟機備援、而使用 HBA 則可提供 FC 連接埠和 FC 纜線備援。這類元件的硬體故障是可預見的、不會影響可用度。

在企業環境中、通常可以使用備援元件來保護整個站台的基礎架構、直到唯一可預見的實體災難案例完全遺失站台為止。災難恢復規劃則取決於站台對站台的複寫。

同步與非同步資料保護

在理想的環境中、所有資料都會在地理上分散的站台之間同步複寫。這類複寫不一定可行、甚至可能、原因有幾個：

- 同步複寫不可避免地會增加寫入延遲、因為所有變更都必須複寫到兩個位置、應用程式 / 資料庫才能繼續處理。因此產生的效能影響有時是不可接受的、無法使用同步鏡射。
- 100% SSD 儲存設備的採用率增加、意味著更有可能注意到額外的寫入延遲、因為效能期望包括數十萬 IOPS 和低於毫秒的延遲。若要充分發揮 100% SSD 的效益、可能需要重新規劃災難恢復策略。
- 資料集會繼續以位元組為單位增加、因此必須確保足夠的頻寬來維持同步複寫、因此會產生挑戰。
- 資料集的複雜度也隨之增加、管理大規模同步複寫也會帶來挑戰。
- 雲端型策略通常需要較長的複寫距離和延遲、進一步排除同步鏡射的使用。

NetApp 提供的解決方案包括同步複寫、可滿足最嚴苛的資料恢復需求、並可提供更優異的效能與靈活度的非同步解決方案。此外、NetApp 技術也能與許多第三方複寫解決方案無縫整合、例如 Oracle DataGuard

保留時間

資料保護策略的最後一個層面是資料保留時間、這可能會大幅改變。

- 一般要求是在主要站台上進行 14 天夜間備份、以及儲存在次要站台上的 90 天備份。
- 許多客戶會建立獨立的季度歸檔、儲存在不同的媒體上。
- 持續更新的資料庫可能不需要歷史資料、而備份只需保留幾天。
- 法規要求可能需要在 365 天內恢復任何任意交易的點。

ONTAP 提供 Oracle 資料庫可用度

ONTAP 旨在提供最大的 Oracle 資料庫可用度。ONTAP 高可用度功能的完整說明不在本文件的範圍之內。然而、與資料保護一樣、在設計資料庫基礎架構時、對此功能的基本瞭解非常重要。

HA 配對

高可用度的基本單位是 HA 配對。每對都包含備援連結、可支援將資料複寫到 NVRAM。NVRAM 不是寫入快取。控制器內的 RAM 會做為寫入快取。NVRAM 的用途是暫時記錄資料、以防止發生非預期的系統故障。在這方面、它與資料庫重做記錄類似。

NVRAM 和資料庫重做記錄都可用來快速儲存資料、讓資料的變更能夠儘快提交。磁碟機（或資料檔案）上的持續資料更新直到稍後在 ONTAP 和大多數資料庫平台上稱為檢查點的程序期間才會進行。正常作業期間不會讀取 NVRAM 資料或資料庫重做記錄。

如果控制器突然故障、可能會有擱置中的變更、這些變更儲存在 NVRAM 中、但尚未寫入磁碟機。合作夥伴控制器會偵測故障、控制磁碟機、並套用儲存在 NVRAM 中的必要變更。

接管與恢復

接管與恢復是指在 HA 配對中的節點之間轉移儲存資源責任的程序。接管和恢復有兩個層面：

- 管理可存取磁碟機的網路連線能力
- 管理磁碟機本身

支援 CIFS 和 NFS 流量的網路介面、都是設定在主位置和容錯移轉位置。接管包括將網路介面移至與原始位置位於同一子網路的實體介面上的暫存主目錄。贈品包括將網路介面移回其原始位置。您可以視需要調整確切行為。

支援 SAN 區塊傳輸協定（例如 iSCSI 和 FC）的網路介面不會在接管和恢復期間重新定位。而是應使用包含完整 HA 配對的路徑來佈建 LUN、以產生主要路徑和次要路徑。



您也可以設定其他控制器的路徑、以支援在較大叢集中的節點之間重新放置資料、但這並不屬於 HA 程序的一部分。

接管與恢復的第二個層面是磁碟擁有權的轉移。確切的程序取決於多種因素、包括接管 / 恢復的原因、以及發出的命令列選項。目標是盡可能有效率地執行作業。雖然整體程序可能需要幾分鐘的時間、但磁碟機的實際擁有權從節點移轉至節點的時間通常只需幾秒鐘。

接管時間

主機 I/O 在接管和恢復作業期間會短暫暫停 I/O、但在正確設定的環境中不應發生應用程式中斷。I/O 延遲的實際轉換程序通常是以秒為單位來測量、但主機可能需要額外的時間來識別資料路徑的變更並重新提交 I/O 作業。

中斷的性質取決於傳輸協定：

- 支援 NFS 和 CIFS 流量的網路介面會在移轉至新實體位置後、向網路發出位址解析傳輸協定（ARP）要求。這會導致網路交換器更新其媒體存取控制（MAC）位址表、並繼續處理 I/O 在計畫性接管和恢復的情況下、中斷通常以秒為單位來衡量、在許多情況下都無法偵測到。有些網路可能會較慢、無法完全辨識網路路徑的變更、有些作業系統可能會在很短的時間內排入大量 I/O、因此必須重新嘗試。這可能會延長恢復 I/O 所需的時間
- 支援 SAN 通訊協定的網路介面不會轉換到新位置。主機作業系統必須變更使用中的路徑。主機觀察到 I/O 暫停的情形取決於多種因素。從儲存系統的角度來看、無法提供 I/O 的時間只有幾秒鐘。不過、不同的主機作業系統可能需要額外的時間、才能讓 I/O 逾時、再重試。較新的作業系統更能更快辨識路徑變更、但較舊的作業系統通常需要 30 秒才能辨識變更。

下表顯示儲存系統無法將資料提供給應用程式環境的預期接管時間。在任何應用程式環境中都不應發生任何錯誤、而是在 IO 處理過程中、接管應該會顯示為短暫的暫停。

	NFS	AFF	ASA
計畫性接管	15 秒	6-10 秒	2-3 秒
非計畫性接管	30 秒	6-10 秒	2-3 秒

Checksum 與 Oracle 資料庫完整性

ONTAP 及其支援的通訊協定包含多項保護 Oracle 資料庫完整性的功能、包括靜態資料和透過網路傳輸的資料。

ONTAP 內的邏輯資料保護包含三項關鍵需求：

- 資料必須受到保護、以免資料毀損。
- 資料必須受到保護、避免磁碟機故障。
- 資料變更必須受到保護、以免遺失。

以下各節將討論這三項需求。

網路毀損：Checksum

最基本的資料保護層級是 Checksum、這是儲存在資料旁的特殊錯誤偵測程式碼。在網路傳輸期間、會使用 Checksum 和（在某些情況下）多個 Checksum 來偵測資料毀損。

例如、FC 框架包含稱為循環備援檢查（CRC）的校驗和形式、以確保有效負載不會在傳輸過程中毀損。傳輸器會同時傳送資料和資料的 CRC。FC 訊框的接收器會重新計算接收資料的 CRC、以確保其符合傳輸的 CRC。如果新計算的 CRC 與附加至框架的 CRC 不相符、則資料會毀損、FC 框架會遭到捨棄或拒絕。iSCSI I/O 作業包括 TCP/IP 層和以太網路層的校驗和、此外、為了提供額外的保護、也可在 SCSI 層提供選用的 CRC 保護。TCP 層或 IP 層偵測到線路上的任何位元毀損、導致封包重新傳輸。與 FC 一樣、SCSI CRC 中的錯誤也會導致作業遭到捨棄或拒絕。

磁碟機毀損：Checksum

Checksum 也可用來驗證儲存在磁碟機上的資料完整性。寫入磁碟機的資料區塊會以 Checksum 功能儲存、產生與原始資料相關的不可預測數字。從磁碟機讀取資料時、會重新計算總和檢查碼、並與儲存的總和檢查碼進行比較。如果不相符、則資料已毀損、必須由 RAID 層還原。

資料毀損：寫入遺失

最難偵測的毀損類型之一是遺失或錯誤寫入。確認寫入後、必須將其寫入正確位置的媒體。就地資料毀損使用儲存在資料中的簡單檢查碼、相當容易偵測。但是、如果寫入資料只是遺失、則先前版本的資料可能仍然存在、而且總和檢查碼是正確的。如果寫入放置在錯誤的實體位置、則相關的 Checksum 將再次對儲存的資料有效、即使寫入已銷毀其他資料。

這項挑戰的解決方案如下：

- 寫入作業必須包含中繼資料、以指出寫入的預期位置。
- 寫入作業必須包含某種版本識別碼。

ONTAP 寫入區塊時、會包含區塊所屬的資料。如果後續讀取識別出某個區塊、但中繼資料指出該區塊在位置 456 找到時屬於位置 123、則表示該寫入已放錯位置。

更難偵測完全遺失的寫入。這項說明非常複雜、但基本上 ONTAP 是以寫入作業導致磁碟機上兩個不同位置的更新方式來儲存中繼資料。如果寫入遺失、後續的資料讀取和相關中繼資料會顯示兩個不同的版本識別。這表示磁碟機未完成寫入。

遺失或放錯位置的寫入毀損極少發生、但隨著磁碟機持續成長、資料集也逐漸擴充至 EB 規模、風險也會增加。任何支援資料庫工作負載的儲存系統都應包含遺失寫入偵測。

磁碟機故障：RAID、RAID DP 和 RAID-TEC

如果發現磁碟機上的資料區塊毀損、或整個磁碟機故障且完全無法使用、則必須重新建立資料。這是在 ONTAP 中使用同位元磁碟機來完成的。資料會在多個資料磁碟機之間進行等量分割、然後產生同位元檢查資料。這會與

原始資料分開儲存。

ONTAP 最初使用 RAID 4、每組資料磁碟機使用單一同位元檢查磁碟機。結果是群組中的任何一個磁碟機都可能發生故障、而不會導致資料遺失。如果同位元磁碟機故障、則沒有資料受損、也可以建構新的同位元磁碟機。如果單一資料磁碟機故障、其餘磁碟機可與同位元磁碟機搭配使用、以重新產生遺失的資料。

當磁碟機很小時、同時發生兩個磁碟機故障的統計機率可忽略不計。隨著磁碟機容量的增加、磁碟機故障後重建資料所需的時間也隨之增加。這增加了第二個磁碟機故障會導致資料遺失的時間。此外、重建程序也會在正常運作的磁碟機上建立許多額外的 I/O。隨著磁碟機老化、導致第二個磁碟機故障的額外負載風險也會增加。最後、即使持續使用 RAID 4、資料遺失的風險並未增加、資料遺失的後果也會更加嚴重。在 RAID 群組故障時遺失的資料越多、還原資料所需的時間就越長、業務中斷也就越長。

這些問題導致 NetApp 開發 NetApp RAID DP 技術、這是 RAID 6 的變體。此解決方案包含兩個同位元檢查磁碟機、表示 RAID 群組中的任何兩個磁碟機都可能發生故障、而不會造成資料遺失。磁碟機的大小持續成長、最終導致 NetApp 開發 NetApp RAID-TEC 技術、引進第三個同位元磁碟機。

某些歷史資料庫最佳實務做法建議使用 RAID-10、也稱為等量鏡射。因為有多個雙磁碟故障案例、因此資料保護功能比 RAID DP 更少、而在 RAID DP 中則沒有。

由於效能考量、有些歷史資料庫最佳實務做法表示 RAID-10 較 RAID-4/5/6 選項更為偏好。這些建議有時是指 RAID 罰款。雖然這些建議一般都是正確的、但不適用於在 ONTAP 中實作 RAID。效能考量與同位元重生有關。在傳統的 RAID 實作中、處理資料庫執行的例行隨機寫入作業需要多個磁碟讀取才能重新產生同位元資料並完成寫入。其懲罰定義為執行寫入作業所需的額外讀取 IOPS。

ONTAP 不會發生 RAID 損失、因為寫入會分段在記憶體中產生同位元檢查、然後以單一 RAID 等量磁碟寫入磁碟。完成寫入作業不需要讀取。

總而言之、相較於 RAID 10、RAID DP 和 RAID-TEC 可提供更多可用容量、更好的磁碟機故障防護、而且不會犧牲效能。

硬體故障保護：NVRAM

任何服務資料庫工作負載的儲存陣列、都必須儘快服務寫入作業。此外、寫入作業也必須受到保護、避免因電源故障等非預期事件而遺失。這表示任何寫入作業都必須安全地儲存在至少兩個位置。

AFF 和 FAS 系統仰賴 NVRAM 來滿足這些需求。寫入程序的運作方式如下：

1. 傳入寫入資料儲存在 RAM 中。
2. 必須對磁碟上的資料所做的變更、會同時記入本機節點和合作夥伴節點上的 NVRAM。NVRAM 不是寫入快取、而是類似資料庫重做記錄的日誌。在正常情況下、系統不會讀取。它僅用於恢復、例如在 I/O 處理期間發生電源故障後。
3. 然後寫入會被確認給主機。

此階段的寫入程序從應用程式的角度來看已完成、而且資料會受到保護、不會遺失、因為資料會儲存在兩個不同的位置。最後、變更會寫入磁碟、但此程序會從應用程式的觀點超出頻外、因為它會在寫入確認之後發生、因此不會影響延遲。此程序再次類似於資料庫記錄。對資料庫的變更會盡快記錄在重做記錄檔中、然後將變更確認為已認可。資料檔案的更新會在稍後進行、不會直接影響處理速度。

如果控制器發生故障、合作夥伴控制器會取得所需磁碟的所有權、並重新執行 NVRAM 中記錄的資料、以恢復發生故障時正在執行的任何 I/O 作業。

硬體故障保護：NVFAIL

如前所述、寫入必須先登入本機 NVRAM 及至少一個其他控制器上的 NVRAM、才會被確認。此方法可確保硬體故障或停電不會導致機內 I/O 遺失。如果本機 NVRAM 故障或連線至 HA 合作夥伴失敗、則無法再鏡射此傳輸中的資料。

如果本機 NVRAM 回報錯誤、節點會關機。此關機會導致容錯移轉至 HA 合作夥伴控制器。由於發生故障的控制器尚未確認寫入作業、因此不會遺失任何資料。

除非強制容錯移轉、否則 ONTAP 不允許在資料不同步時進行容錯移轉。以這種方式強制變更條件、即表示資料可能會留在原始控制器中、而且資料遺失是可以接受的。

如果強制容錯移轉、則資料庫特別容易受損、因為資料庫會在磁碟上保留大量的內部資料快取。如果發生強制容錯移轉、先前確認的變更將會有效捨棄。儲存陣列的內容會有效地及時向後跳轉、而且資料庫快取的狀態不再反映磁碟上資料的狀態。

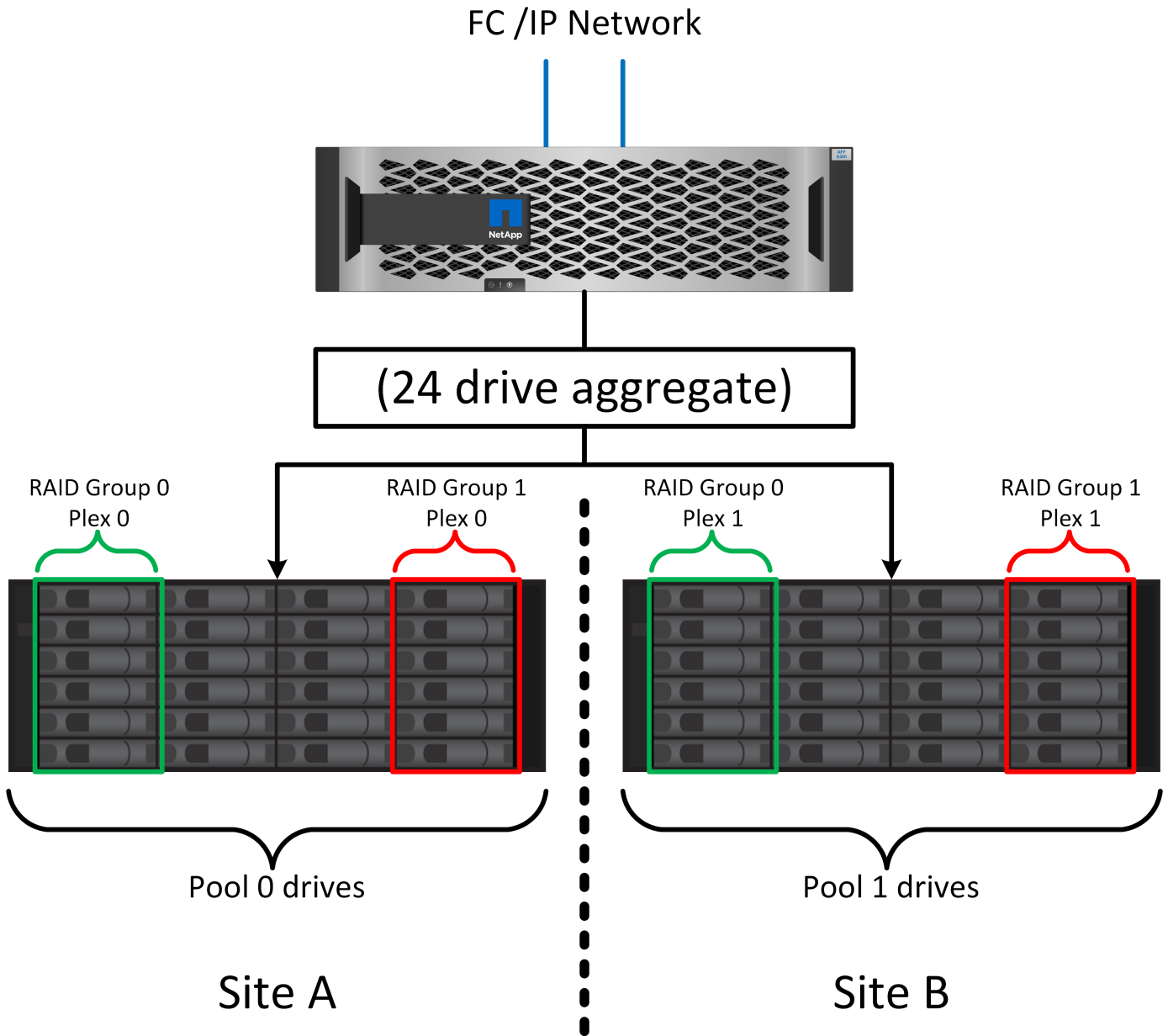
為了保護資料不受這種情況影響、ONTAP 允許設定磁碟區、以針對 NVRAM 故障提供特殊保護。觸發時、此保護機制會導致磁碟區進入稱為 NVFAIL 的狀態。此狀態會導致 I/O 錯誤、導致應用程式關機、使其不使用過時的資料。資料不應遺失、因為儲存陣列上應該存在任何已確認的寫入資料。

通常的後續步驟是讓系統管理員在手動將 LUN 和磁碟區重新上線之前、先完全關閉主機。雖然這些步驟可能涉及一些工作、但這種方法是確保資料完整性的最安全方法。並非所有資料都需要這項保護、因此 NVFAIL 行為可依每個磁碟區設定。

站台和機櫃故障保護：SyncMirror 和叢

SyncMirror 是一項鏡射技術、可增強但不取代 RAID DP 或 RAID-TEC。它會鏡射兩個不同 RAID 群組的內容。邏輯組態如下：

- 磁碟機會根據位置設定成兩個集區。一個集區由站台 A 上的所有磁碟機組成、第二個集區由站台 B 上的所有磁碟機組成
- 接著會根據鏡射的 RAID 群組集建立通用儲存池（稱為 Aggregate）。從每個站台擷取的磁碟機數量相等。例如、20 個磁碟機的 SyncMirror Aggregate 將由站台 A 的 10 個磁碟機和站台 B 的 10 個磁碟機組成
- 指定站台上的每組磁碟機都會自動設定為一個或多個完全備援的 RAID-DP 或 RAID-TEC 群組、而不受鏡像的使用影響。即使在站台遺失之後、也能提供持續的資料保護。



上圖說明 SyncMirror 組態範例。在控制器上建立了 24 個磁碟機的集合體、其中 12 個磁碟機來自於站台 A 上配置的機櫃、12 個磁碟機來自於站台 B 上配置的機櫃磁碟機分為兩個鏡射 RAID 群組。RAID 群組 0 包括站台 A 上的 6 磁碟機叢、鏡射到站台 B 上的 6 磁碟機叢同樣地、RAID Group 1 也包含站台 A 的 6 磁碟叢、鏡射到站台 B 的 6 磁碟叢

SyncMirror 通常用於提供 MetroCluster 系統的遠端鏡射、每個站台都有一份資料複本。有時候、它是用來在單一系統中提供額外的備援層級。特別是提供機架層級的備援。磁碟機櫃已包含雙電源供應器和控制器、整體上比金屬板稍多、但在某些情況下、可能需要額外的保護。例如、有一位 NetApp 客戶部署 SyncMirror、用於汽車測試期間使用的行動即時分析平台。系統由獨立的 UPS 系統提供獨立的電源供應器、分為兩個實體機架。

=checksum

對於習慣使用 Oracle RMAN 串流備份的 DBA 來說、檢查和主題特別重要、因為它們會移轉至快照型備份。RMAN 的一項功能是在備份作業期間執行完整性檢查。雖然這項功能有一些價值、但其主要優點是資料庫並未用於現代化的儲存陣列。當實體磁碟機用於 Oracle 資料庫時、幾乎可以確定磁碟機老化時最終會發生毀損、這是由真正儲存陣列中的陣列型校驗和所解決的問題。

使用真正的儲存陣列、資料完整性可在多個層級上使用校驗和加以保護。如果 IP 型網路中的資料毀損、傳輸控制傳輸協定 (TCP) 層會拒絕封包資料並要求重新傳輸。FC 傳輸協定包括校驗和、封裝的 SCSI 資料也一樣。在陣列上之後、ONTAP 就有 RAID 和 Checksum 保護。可能會發生毀損、但如同大多數企業陣列一樣、系統會偵測並修正毀損。一般而言、整個磁碟機都會故障、導致 RAID 重建、資料庫完整性也不會受到影響。ONTAP 偵測到 Checksum 錯誤的頻率較低、表示磁碟機上的資料已損壞。然後磁碟機故障、RAID 重建就會開始。資料完整性再次不受影響。

Oracle 資料檔案和重做記錄架構也設計成即使在極端情況下、也能提供最高程度的資料完整性。在最基本的層級、Oracle 區塊幾乎包含每個 I/O 的 Checksum 和基本邏輯檢查如果 Oracle 尚未當機或將資料表空間離線、則資料會保持不變。資料完整性檢查的程度可調整、Oracle 也可設定為確認寫入。因此、幾乎所有的當機和故障情況都可以恢復、而且在極罕見的不可恢復情況下、系統會立即偵測到毀損。

大多數使用 Oracle 資料庫的 NetApp 客戶在移轉至快照型備份後、都會停止使用 RMAN 和其他備份產品。仍有一些選項可讓 RMAN 使用 SnapCenter 執行區塊層級的還原。然而、日常使用的 RMAN、NetBackup 及其他產品只會偶爾用於建立每月或每季的歸檔複本。

有些客戶選擇執行 dbv 定期對現有資料庫執行完整性檢查。NetApp 不鼓勵這種做法、因為它會產生不必要的 I/O 負載。如上所述、如果資料庫先前沒有遇到問題、就有可能發生 dbv 偵測問題接近零、此公用程式會在網路和儲存系統上產生非常高的連續 I/O 負載。除非有理由相信存在毀損、例如暴露於已知的 Oracle 錯誤、否則沒有理由執行 dbv。

備份與還原基礎知識

Oracle 資料庫和快照型備份

NetApp Snapshot 技術是 ONTAP 上 Oracle 資料庫資料保護的基礎。

關鍵值如下：

- 簡易性。快照是特定時間點資料容器內容的唯讀複本。
- 效率。快照在建立時不需要任何空間。只有在資料變更時才會使用空間。
- 管理能力。由於快照是儲存作業系統的原生部分、因此以快照為基礎的備份策略很容易設定和管理。如果儲存系統已開機、就可以開始建立備份。
- 擴充性。最多可保留 1024 個檔案和 LUN 的單一容器備份。對於複雜的資料集、可透過一組一致的快照來保護多個資料容器。
- 無論磁碟區包含 1024 個快照或無、效能都不會受到影響。

雖然許多儲存廠商都提供快照技術、但 ONTAP 中的 Snapshot 技術是獨一無二的、可為企業應用程式和資料庫環境帶來顯著效益：

- Snapshot 複本是基礎 Write -Anywhere File Layout (WAFL) 的一部分。它們不是附加技術或外部技術。這簡化了管理、因為儲存系統是備份系統。
- 快照複本不會影響效能、但某些邊緣情況除外、例如當基礎儲存系統填滿的快照中儲存了大量資料時。
- 「一致性群組」一詞通常是指一組儲存物件、這些物件是以一致的資料集合來管理。特定 ONTAP 磁碟區的快照構成一致性群組備份。

ONTAP 快照的擴充能力也優於競爭技術。客戶可以儲存 5、50 或 500 個快照、而不會影響效能。磁碟區目前允許的最大快照數為 1024。如果需要額外的快照保留、則有選項可將快照串聯至其他磁碟區。

因此、保護託管在 ONTAP 上的資料集非常簡單且具有高度擴充性。備份不需要移動資料、因此備份策略可以根據業務需求量身打造、而非限制網路傳輸率、大量磁帶機或磁碟接移區域。

快照是否為備份？

將快照當作資料保護策略使用的常見問題之一、就是「真實」資料和快照資料位於同一個磁碟機上。遺失這些磁碟機將會導致主要資料和備份遺失。

這是一項有效的考量。本機快照是用於日常備份與還原需求、在這方面、快照是備份。在 NetApp 環境中、將近 99% 的還原案例都仰賴快照來滿足最嚴苛的 RTO 需求。

然而、本機快照不應是唯一的備份策略、因此 NetApp 提供 SnapMirror 和 SnapVault 複寫等技術、可快速有效地將快照複寫至一組不同的磁碟機。在架構正確的解決方案中、快照加上快照複寫功能可將磁帶的使用降至最低、甚至是每季歸檔、或完全消除。

快照型備份

使用 ONTAP Snapshot 複本保護資料有許多選項、快照是許多其他 ONTAP 功能的基礎、包括複寫、災難恢復和複製。快照技術的完整說明不在本文件的範圍內、但以下各節提供一般概觀。

建立資料集快照的主要方法有兩種：

- 損毀一致的備份
- 應用程式一致的備份

資料集的損毀一致備份是指在單一時間點擷取整個資料集結構。如果資料集儲存在單一 NetApp FlexVol Volume 中、則程序很簡單；您可以隨時建立 Snapshot。如果資料集橫跨多個磁碟區、則必須建立一致性群組（CG）快照。建立 CG 快照有多種選項、包括 NetApp SnapCenter 軟體、原生 ONTAP 一致性群組功能、以及使用者維護的指令碼。

當備份點還原足夠時、主要會使用損毀一致的備份。當需要更精細的恢復時、通常需要應用程式一致的備份。

「應用程式一致性」一詞通常是錯誤的。例如、將 Oracle 資料庫置於備份模式稱為應用程式一致的備份、但資料並未以任何方式保持一致或停止。資料會在整個備份過程中持續變更。相反地、大部分的 MySQL 和 Microsoft SQL Server 備份確實會在執行備份之前先將資料關閉。VMware 可能會或可能不會使某些檔案一致。

一致性群組

術語「一致性群組」是指儲存陣列將多個儲存資源視為單一映像來管理的能力。例如、資料庫可能包含 10 個 LUN。陣列必須能夠以一致的方式備份、還原及複寫這 10 個 LUN。如果 LUN 的映像備份時不一致、則無法還原。複寫這 10 個 LUN 需要所有複本彼此完全同步。

討論 ONTAP 時、「一致性群組」一詞並不常使用、因為一致性一向是 ONTAP 內 Volume 和 Aggregate 架構的基本功能。許多其他儲存陣列會將 LUN 或檔案系統視為個別單元進行管理。接著可選擇性地將它們設定為「一致性群組」、以保護資料、但這是組態中的額外步驟。

ONTAP 一向能夠擷取一致的本機和複寫資料映像。雖然 ONTAP 系統上的各種磁碟區通常並未正式描述為一致性群組、但這就是它們的名稱。該磁碟區的快照是一致性群組映像、該快照的還原是一致性群組還原、SnapMirror 和 SnapVault 都提供一致性群組複寫。

一致性群組快照

一致性群組快照（CG 快照）是基本 ONTAP Snapshot 技術的延伸。標準快照作業可在單一磁碟區內建立所有資料的一致映像、但有時必須在多個磁碟區甚至跨多個儲存系統建立一致的快照集。結果是一組快照、其使用方式與只有一個個別磁碟區的快照相同。它們可用於本機資料還原、複寫以進行災難恢復、或複製為單一一致的單元。

CG 快照的最大已知用途是資料庫環境、其大小約為 1PB、跨越 12 個控制器。在此系統上建立的 CG 快照已用於備份、恢復和複製。

大多數情況下、當資料集跨越磁碟區且必須保留寫入順序時、所選管理軟體會自動使用 CG 快照。在此情況下、無需瞭解 CG 快照的技術詳細資料。然而、在某些情況下、複雜的資料保護需求需要對資料保護和複寫程序進行詳細控制。自動化工作流程或使用自訂指令碼來呼叫 CG 快照 API 是其中的一些選項。若要瞭解最佳選項和 CG-snapshot 的角色、需要更詳細的技術說明。

建立一組 CG 快照是兩個步驟：

1. 在所有目標磁碟區上建立寫入屏障。
2. 在圍籬狀態下建立這些磁碟區的快照。

寫入隔離是連續建立的。這表示當隔離程序在多個磁碟區之間設定時、寫入 I/O 會依序凍結在第一個磁碟區上、因為它會繼續提交到稍後出現的磁碟區。這可能一開始就違反了保留寫入順序的要求、但這僅適用於非同步在主機上發出的 I/O、不需仰賴任何其他寫入。

例如、資料庫可能會發出許多非同步資料檔案更新、並允許作業系統重新排序 I/O、並根據其本身的排程器組態完成這些更新。這類 I/O 的順序無法保證、因為應用程式和作業系統已釋出保留寫入順序的要求。

以計數器為例、大部分的資料庫記錄活動都是同步的。在確認 I/O 之前、資料庫不會繼續進行記錄寫入、而且必須保留這些寫入的順序。如果記錄 I/O 到達圍籬式磁碟區、則不會予以確認、應用程式會在進一步寫入時加以封鎖。同樣地、檔案系統中繼資料 I/O 通常是同步的。例如、檔案刪除作業不得遺失。如果具有 xfs 檔案系統的作業系統刪除了檔案、而更新 xfs 檔案系統中繼資料的 I/O 則會移除位於圍籬磁碟區上的該檔案參照、則檔案系統活動就會暫停。這可確保 CG 快照作業期間檔案系統的完整性。

在目標磁碟區之間設定寫入屏障之後、就可以開始建立快照。由於磁碟區的狀態會從相關寫入點凍結、因此不需要同時精確建立快照。為了防範建立 CG 快照的應用程式中的瑕疵、初始寫入屏障包含可設定的逾時時間、ONTAP 會在定義的秒數後自動釋放隔離功能並繼續寫入處理。如果所有快照都是在逾時期間發生之前建立的、則產生的一組快照是有效的一致性群組。

相關寫入順序

從技術觀點來看、一致性群組的關鍵在於保留寫入順序、特別是根據寫入順序。例如、寫入 10 個 LUN 的資料庫會同時寫入所有 LUN。許多寫入都是以非同步方式發出、這表示完成的順序不重要、實際完成的順序會因作業系統和網路行為而異。

某些寫入作業必須存在於磁碟上、資料庫才能繼續進行其他寫入作業。這些關鍵寫入作業稱為「相關寫入」。後續寫入 I/O 則取決於磁碟上是否有這些寫入資料。這 10 個 LUN 的任何快照、恢復或複寫都必須確保相關寫入順序受到保證。檔案系統更新是寫入順序相關寫入的另一個範例。必須保留檔案系統變更的順序、否則整個檔案系統可能會毀損。

策略

以快照為基礎的備份主要有兩種方法：

- 損毀一致的備份
- 快照保護的熱備份

資料庫的損毀一致備份是指在單一時間點擷取整個資料庫結構、包括資料檔案、重做記錄和控制檔。如果資料庫儲存在單一 NetApp FlexVol Volume 中、則程序很簡單；您可以隨時建立 Snapshot。如果資料庫橫跨磁碟區、則必須建立一致性群組（CG）快照。建立 CG 快照有多種選項、包括 NetApp SnapCenter 軟體、原生 ONTAP 一致性群組功能、以及使用者維護的指令碼。

當備份點還原足夠時、主要會使用損毀一致的 Snapshot 備份。在某些情況下可以套用歸檔記錄檔、但如果需要更精細的時間點還原、則最好使用線上備份。

快照型線上備份的基本程序如下：

1. 將資料庫放入 backup 模式。
2. 建立所有託管資料檔案的磁碟區快照。
3. 結束 backup 模式。
4. 執行命令 `alter system archive log current` 強制記錄歸檔。
5. 為所有託管歸檔記錄的磁碟區建立快照。

此程序會產生一組快照、其中包含備份模式中的資料檔案、以及在備份模式中產生的重要歸檔記錄。這是恢復資料庫的兩項需求。控制檔等檔案也應受到保護、以方便使用、但唯一的絕對需求是保護資料檔案和歸檔記錄。

雖然不同的客戶可能有非常不同的策略、但幾乎所有這些策略最終都是以下列相同原則為基礎。

快照型還原

在設計 Oracle 資料庫的 Volume 配置時、第一個決定是是否使用 Volume NetApp SnapRestore（VBSR）技術。

Volume 型 SnapRestore 可讓磁碟區立即還原至較早的時間點。由於磁碟區上的所有資料都已還原、因此 VBSR 可能不適用於所有使用案例。例如、如果整個資料庫（包括資料檔案、重做記錄和歸檔記錄）儲存在單一磁碟區上、且此磁碟區使用 VBSR 還原、則資料會遺失、因為較新的歸檔記錄和重做資料會被捨棄。

還原不需要 VSR。許多資料庫都可以使用檔案型單一檔案 SnapRestore（SFSR）來還原、或只是將檔案從快照複製回作用中的檔案系統。

當資料庫非常大或必須盡快恢復時、最好使用 VBSR、而使用 VSR 需要隔離資料檔案。在 NFS 環境中、指定資料庫的資料檔案必須儲存在專用的磁碟區中、而這些磁碟區不會受到任何其他類型的檔案污染。在 SAN 環境中、資料檔案必須儲存在專用 FlexVol 磁碟區上的專用 LUN 中。如果使用 Volume Manager（包括 Oracle 自動儲存管理 [AS]）、則磁碟群組也必須專用於資料檔案。

以這種方式隔離資料檔案、可讓檔案還原至較早的狀態、而不會損壞其他檔案系統。

Snapshot保留

對於 SAN 環境中具有 Oracle 資料的每個 Volume `percent-snapshot-space` 應設為零、因為在 LUN 環境中保留快照空間並不實用。如果百分比保留設為 100、則具有 LUN 的磁碟區快照需要在磁碟區中有足夠的可用空間、但不包括快照保留空間、以吸收所有資料 100% 的營業額。如果將百分比保留設為較低的值、則需要相對較小的可用空間、但它一律會排除快照保留。這表示 LUN 環境中的快照保留空間會被浪費。

在 NFS 環境中、有兩個選項：

- 設定 `percent-snapshot-space` 根據預期的快照空間使用量。
- 設定 `percent-snapshot-space` 以歸零並統整管理作用中和快照空間使用量。

使用第一個選項、`percent-snapshot-space` 設為非零值、通常約 20%。然後、使用者就會隱藏此空間。不過、此值並不會限制使用率。如果具有 20% 保留的資料庫擁有 30% 的營業額、則快照空間可能會超出 20% 保留空間的範圍、並佔用無保留空間。

將保留設定為 20% 等值的主要優點是驗證某些空間永遠可供快照使用。例如、保留 20% 的 1TB 磁碟區只允許資料庫管理員 (DBA) 儲存 800GB 的資料。此組態保證至少有 200GB 的空間可供快照使用。

何時 `percent-snapshot-space` 設為零、則使用者可以使用磁碟區中的所有空間、以提供更好的可見度。DBA 必須瞭解、如果他 / 她看到 1TB 的磁碟區運用快照、則這 1TB 的空間會在使用中資料和 Snapshot 週轉之間共享。

終端使用者之間的選項 1 和選項 2 之間沒有明確的偏好設定。

ONTAP 和第三方快照

Oracle Doc ID 604683.1 說明第三方快照支援的需求、以及備份與還原作業的多種選項。

第三方廠商必須保證公司的快照符合下列要求：

- 快照必須與 Oracle 建議的還原與還原作業整合。
- 快照必須在快照點保持一致的資料庫損毀。
- 快照中的每個檔案都會保留寫入順序。

ONTAP 和 NetApp Oracle 管理產品符合這些要求。

使用 SnapRestore 快速恢復 Oracle 資料庫

NetApp SnapRestore 技術可從快照快速還原 ONTAP 中的資料。

當關鍵資料集無法使用時、關鍵業務營運就會中斷。磁帶可能會中斷、甚至是從磁碟型備份還原、在網路上傳輸速度可能會很慢。SnapRestore 可提供近乎即時的資料集還原功能、避免這些問題。即使是 PB 規模的資料庫、只要花幾分鐘的時間就能完全還原。

SnapRestore 有兩種形式：檔案 /LUN 型和磁碟區型。

- 無論是 2TB LUN 或 4KB 檔案、個別檔案或 LUN 都能在數秒內還原。
- 無論是 10GB 或 100TB 的資料、檔案或 LUN 的容器都能在數秒內還原。

「檔案或 LUN 的容器」通常指的是 FlexVol Volume。例如、您可以在單一磁碟區中擁有 10 個組成 LVM 磁碟群組的 LUN、或是一個磁碟區可以儲存 1000 位使用者的 NFS 主目錄。您可以將整個磁碟區還原為單一作業、而不需為每個個別檔案或 LUN 執行還原作業。此程序也適用於包含多個磁碟區的橫向擴充容器、例如 FlexGroup 或 ONTAP 一致性群組。

SnapRestore 之所以能如此快速有效地運作、是因為快照的本質、基本上是在特定時間點對磁碟區內容進行平行唯讀檢視。作用中區塊是可以變更的實際區塊、而快照則是建立快照時構成檔案和 LUN 之區塊狀態的唯讀檢

視。

ONTAP 僅允許唯讀存取快照資料、但可透過 SnapRestore 重新啟動資料。快照會重新啟用為資料的讀寫檢視、並將資料恢復至先前的狀態。SnapRestore 可以在磁碟區或檔案層級運作。這項技術基本上相同、但行為上有一些小差異。

Volume SnapRestore

Volume 型 SnapRestore 會將整個資料量傳回至較早的狀態。這項作業不需要資料移動、也就是說還原程序基本上是即時的、雖然 API 或 CLI 作業可能需要幾秒鐘的時間才能處理。還原 1GB 的資料並不比還原 1PB 的資料複雜或耗時。這項功能是許多企業客戶移轉至 ONTAP 儲存系統的主要原因。即使是最大的資料集、也能以秒為單位提供 RTO。

Volume 型 SnapRestore 的缺點之一、是因為磁碟區內的變更會隨時間累積。因此、每個快照和作用中檔案資料都取決於到該點之前的變更。將磁碟區還原為較早的狀態、表示捨棄所有後續對資料所做的變更。然而、不太明顯的是、這包括後續建立的快照。這並不總是理想的。

例如、資料保留 SLA 可能會指定每晚備份 30 天。將資料集還原至五天前以 Volume SnapRestore 建立的快照、將會捨棄前五天建立的所有快照、違反 SLA。

有許多選項可解決此限制：

1. 資料可從先前的快照複製、而非執行整個 Volume 的 SnapRestore。此方法最適合較小的資料集。
2. 您可以複製快照、而非還原快照。此方法的限制在於來源快照是複本的相依性。因此、除非也刪除複本、或將其分割成一個不同的 Volume、否則無法將其刪除。
3. 使用檔案型 SnapRestore。

File SnapRestore

檔案型 SnapRestore 是更精細的快照型還原程序。個別檔案或 LUN 的狀態會還原、而非還原整個磁碟區的狀態。不需要刪除快照、此作業也不需要對先前的快照建立任何相依性。檔案或 LUN 會立即在作用中磁碟區中可用。

SnapRestore 還原檔案或 LUN 時不需要移動資料。不過、需要進行一些內部中繼資料更新、以反映檔案或 LUN 中的基礎區塊現在同時存在於快照和作用中磁碟區中。不應影響效能、但此程序會封鎖快照的建立、直到快照完成為止。根據還原的檔案總大小、處理速度約為 5Gbps (18TB/小時)。

Oracle 資料庫線上備份

在備份模式中保護和恢復 Oracle 資料庫需要兩組資料。請注意、這不是唯一的 Oracle 備份選項、而是最常見的選項。

- 備份模式中資料檔案的快照
- 資料檔案處於備份模式時所建立的歸檔記錄

如果需要完整恢復 (包括所有已提交的交易)、則需要第三個項目：

- 一組目前的重做記錄

有許多方法可以推動線上備份的還原。許多客戶使用 ONTAP CLI 還原快照、然後使用 Oracle RMAN 或 sqlplus 來完成還原。這在大型正式作業環境中尤其常見、因為資料庫還原的可能性和頻率極低、而且任何還原程序都是

由熟練的 DBA 來處理。為了實現完整的自動化、NetApp SnapCenter 等解決方案包含 Oracle 外掛程式、其中包含命令列和圖形介面。

有些大型客戶已在主機上設定基本指令碼、以便在特定時間將資料庫置於備份模式、以準備排程的快照、藉此採取更簡單的方法。例如、排程命令 `alter database begin backup` 於 23 時 58 分、`alter database end backup` 於 00 : 02、然後於午夜直接在儲存系統上排程快照。如此一來、就能實現簡單易用、擴充性極高的備份策略、無需外部軟體或授權。

資料配置

最簡單的配置是將資料檔案隔離到一個或多個專用磁碟區。它們必須不受任何其他檔案類型的污染。這是為了確保資料檔案磁碟區可以透過 SnapRestore 作業快速還原、而不會破壞重要的重做記錄檔、控制檔或歸檔記錄。

SAN 對專用磁碟區內的資料檔案隔離有類似的需求。在 Microsoft Windows 等作業系統中、單一磁碟區可能包含多個資料檔案 LUN、每個 LUN 都有 NTFS 檔案系統。在其他作業系統中、通常會有邏輯 Volume Manager。例如、使用 Oracle ASM 時、最簡單的選項是將 ASM 磁碟群組的 LUN 限制在單一磁碟區、以便作為一個單元進行備份和還原。如果基於效能或容量管理的理由而需要額外的磁碟區、則在新磁碟區上建立額外的磁碟群組、將可簡化管理。

如果遵循這些準則、則可直接在儲存系統上排程快照、而無需執行一致性群組快照。原因是 Oracle 備份不需要同時備份資料檔案。線上備份程序旨在讓資料檔案在數小時內緩慢串流至磁帶、因此能夠繼續更新。

在使用分佈於不同磁碟區的 ASM 磁碟群組等情況下、會產生複雜性。在這種情況下、必須執行 CG 快照、以確保 ASM 中繼資料在所有組成磁碟區之間一致。

- 注意：* 驗證 ASM `spfile` 和 `passwd` 檔案不在主控資料檔案的磁碟群組中。這會影響選擇性還原資料檔案和僅還原資料檔案的能力。

本機恢復程序— NFS

此程序可以手動或透過 SnapCenter 等應用程式來驅動。基本程序如下：

1. 關閉資料庫。
2. 在所需還原點之前、立即將資料檔案磁碟區復原至快照。
3. 將歸檔記錄重播至所需的點。
4. 如果需要完整還原、請重新播放目前的重做記錄。

此程序假設所需的歸檔記錄檔仍存在於作用中的檔案系統中。如果沒有、則必須還原歸檔記錄、或將 RMAN/sqlplus 導向快照目錄中的資料。

此外、對於較小的資料庫、終端使用者可以直接從中復原資料檔案 `.snapshot` 目錄、無需自動化工具或儲存管理員協助執行 `snaprestore` 命令。

本機恢復程序— SAN

此程序可以手動或透過 SnapCenter 等應用程式來驅動。基本程序如下：

1. 關閉資料庫。
2. 將託管資料檔案的磁碟群組置於系統中。此程序會因所選的邏輯磁碟區管理程式而異。使用 ASM 時、此程序需要卸除磁碟群組。在 Linux 中、必須卸除檔案系統、且必須停用邏輯磁碟區和磁碟區群組。目標是停止要還原之目標 Volume 群組上的所有更新。

3. 在所需還原點之前、立即將資料檔案磁碟群組還原至快照。
4. 重新啟動新還原的磁碟群組。
5. 將歸檔記錄重播至所需的點。
6. 如果需要完整還原、請重新播放所有重做記錄。

此程序假設所需的歸檔記錄檔仍存在於作用中的檔案系統中。如果沒有、則必須將歸檔記錄 LUN 離線並執行還原、以還原歸檔記錄。這也是將歸檔記錄分割成專用磁碟區的範例。如果歸檔記錄與重做記錄共用一個磁碟區群組、則必須先將重做記錄複製到其他位置、才能還原整個 LUN 集。此步驟可防止這些最終記錄的交易遺失。

Oracle 資料庫儲存快照最佳化備份

當 Oracle 12c 發行時、快照型備份與還原變得更簡單、因為不需要將資料庫置於熱備份模式。結果是能夠直接在儲存系統上排程快照式備份、同時仍保留執行完整或時間點還原的能力。

雖然 DBA 較熟悉熱備份還原程序、但很長一段時間以來、它仍可使用資料庫處於熱備份模式時未建立的快照。恢復期間、Oracle 10g 和 11g 需要額外的步驟、才能使資料庫保持一致。使用 Oracle 12c、sqlplus 和 rman 包含額外的邏輯、可在非熱備份模式的資料檔案備份上重播歸檔記錄。

如前所述、復原快照型熱備份需要兩組資料：

- 在備份模式下建立的資料檔案快照
- 資料檔案處於熱備份模式時所產生的歸檔記錄

在還原期間、資料庫會從資料檔案讀取中繼資料、以選取所需的歸檔記錄進行還原。

儲存快照最佳化的還原需要稍微不同的資料集、才能達到相同的結果：

- 資料檔案的快照、加上一種識別快照建立時間的方法
- 從最新資料檔案檢查點的時間到快照的確切時間、都會歸檔記錄檔

在還原期間、資料庫會從資料檔案讀取中繼資料、以識別所需的最早歸檔記錄。可以執行完整或時間點恢復。執行時間點還原時、必須知道資料檔案快照的時間。指定的恢復點必須在快照建立時間之後。NetApp 建議您在快照時間中加入至少幾分鐘、以因應時鐘變化。

如需完整的詳細資料、請參閱 Oracle 各版本的 Oracle 12c 說明文件中有關「使用儲存 Snapshot 最佳化進行恢復」主題的 Oracle 文件。此外、請參閱 Oracle 文件 ID 文件 ID 604683.1、瞭解 Oracle 協力廠商快照支援。

資料配置

最簡單的配置是將資料檔案隔離為一個或多個專用磁碟區。它們必須不受任何其他檔案類型的污染。這是為了確保資料檔案磁碟區可以透過 SnapRestore 作業快速還原、而不會破壞重要的重做記錄檔、控制檔或歸檔記錄檔。

SAN 對專用磁碟區內的資料檔案隔離有類似的需求。在 Microsoft Windows 等作業系統中、單一磁碟區可能包含多個資料檔案 LUN、每個 LUN 都有 NTFS 檔案系統。在其他作業系統中、通常也會有邏輯 Volume Manager。例如、使用 Oracle ASM 時、最簡單的選項是將磁碟群組限制在單一磁碟區、以便作為一個單元進行備份和還原。如果基於效能或容量管理的理由而需要額外的磁碟區、則在新磁碟區上建立額外的磁碟群組、將可更輕鬆地進行管理。

如果遵循這些準則、則可直接在 ONTAP 上排程快照、而無需執行一致性群組快照。原因是快照最佳化備份不需要同時備份資料檔案。

在 ASM 磁碟群組等情況下、會發生複雜的情況、而 ASM 磁碟群組會分散在不同的磁碟區中。在這種情況下、必須執行 CG 快照、以確保 ASM 中繼資料在所有組成磁碟區之間一致。

[注意] 確認 ASM spfile 和 passwd 檔案不在主控資料檔案的磁碟群組中。這會影響選擇性還原資料檔案和僅還原資料檔案的能力。

本機恢復程序— NFS

此程序可以手動或透過 SnapCenter 等應用程式來驅動。基本程序如下：

1. 關閉資料庫。
2. 在所需還原點之前、立即將資料檔案磁碟區復原至快照。
3. 將歸檔記錄重播至所需的點。

此程序假設所需的歸檔記錄檔仍存在於作用中的檔案系統中。如果沒有、則必須還原歸檔記錄、或 rman 或 sqlplus 可導向至中的資料 .snapshot 目錄。

此外、對於較小的資料庫、終端使用者可以直接從中復原資料檔案 .snapshot 無需自動化工具或儲存管理員協助執行 SnapRestore 命令的目錄。

本機恢復程序— SAN

此程序可以手動或透過 SnapCenter 等應用程式來驅動。基本程序如下：

1. 關閉資料庫。
2. 將託管資料檔案的磁碟群組置於系統中。此程序會因所選的邏輯磁碟區管理程式而異。使用 ASM 時、此程序需要卸除磁碟群組。在 Linux 中、必須卸除檔案系統、並停用邏輯磁碟區和磁碟區群組。目標是停止要還原之目標 Volume 群組上的所有更新。
3. 在所需還原點之前、立即將資料檔案磁碟群組還原至快照。
4. 重新啟動新還原的磁碟群組。
5. 將歸檔記錄重播至所需的點。

此程序假設所需的歸檔記錄檔仍存在於作用中的檔案系統中。如果沒有、則必須將歸檔記錄 LUN 離線並執行還原、以還原歸檔記錄。這也是將歸檔記錄分割成專用磁碟區的範例。如果歸檔記錄與重做記錄共用磁碟區群組、則必須在還原整體 LUN 組之前、將重做記錄複製到其他位置、以免遺失最終記錄的交易。

完整恢復範例

假設資料檔案已毀損或毀損、且需要完整還原。執行程序如下：

```

[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                2924928 bytes
Variable Size             1040191104 bytes
Database Buffers         553648128 bytes
Redo Buffers              13848576 bytes
Database mounted.
SQL> recover automatic;
Media recovery complete.
SQL> alter database open;
Database altered.
SQL>

```

時間點恢復範例

整個恢復過程只需一個命令：recover automatic。

如果需要時間點恢復、則必須知道快照的時間戳記、並可識別如下：

```

Cluster01::> snapshot show -vserver vserver1 -volume NTAP_oradata -fields
create-time
vserver    volume          snapshot        create-time
-----
vserver1   NTAP_oradata   my-backup      Thu Mar 09 10:10:06 2017

```

快照建立時間列於 3 月 9 日和 10 : 10 : 06 。為了安全起見、快照時間會增加一分鐘：

```

[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                2924928 bytes
Variable Size             1040191104 bytes
Database Buffers         553648128 bytes
Redo Buffers              13848576 bytes
Database mounted.
SQL> recover database until time '09-MAR-2017 10:44:15' snapshot time '09-
MAR-2017 10:11:00';

```

恢復作業現在已啟動。它指定的快照時間為 10 : 11 : 00、記錄時間後一分鐘、以計算可能的時鐘差異、目標恢復時間為 10 : 44。接下來、sqlplus 會要求所需的歸檔記錄檔、以達到所需的 10 : 44 恢復時間。

```
ORA-00279: change 551760 generated at 03/09/2017 05:06:07 needed for
thread 1
ORA-00289: suggestion : /orlogs_nfs/arch/1_31_930813377.dbf
ORA-00280: change 551760 for thread 1 is in sequence #31
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 552566 generated at 03/09/2017 05:08:09 needed for
thread 1
ORA-00289: suggestion : /orlogs_nfs/arch/1_32_930813377.dbf
ORA-00280: change 552566 for thread 1 is in sequence #32
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 553045 generated at 03/09/2017 05:10:12 needed for
thread 1
ORA-00289: suggestion : /orlogs_nfs/arch/1_33_930813377.dbf
ORA-00280: change 553045 for thread 1 is in sequence #33
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 753229 generated at 03/09/2017 05:15:58 needed for
thread 1
ORA-00289: suggestion : /orlogs_nfs/arch/1_34_930813377.dbf
ORA-00280: change 753229 for thread 1 is in sequence #34
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
Log applied.
Media recovery complete.
SQL> alter database open resetlogs;
Database altered.
SQL>
```



使用來完成資料庫的快照還原 `recover automatic` 命令不需要特定授權、而是使用時間點還原 `snapshot time` 需要 Oracle 進階壓縮授權。

Oracle 資料庫管理與自動化工具

ONTAP 在 Oracle 資料庫環境中的主要價值來自核心 ONTAP 技術、例如即時 Snapshot 複本、簡單的 SnapMirror 複寫、以及快速建立 FlexClone Volume。

在某些情況下、直接在 ONTAP 上簡單設定這些核心功能即可滿足需求、但更複雜的需求則需要協調層。

SnapCenter

SnapCenter 是 NetApp 資料保護的旗艦產品。在極低的層級上、它與 SnapManager 產品在執行資料庫備份的方式上類似、但它是從頭開始打造、提供單一窗口來管理 NetApp 儲存系統上的資料保護。

SnapCenter 包括快照式備份與還原、SnapMirror 與 SnapVault 複寫等基本功能、以及大型企業大規模營運所需的其他功能。這些進階功能包括擴充的角色型存取控制 (RBAC) 功能、可與協力廠商協調化產品整合的

RESTful API、資料庫主機上 SnapCenter 外掛程式的不中斷中央管理、以及專為雲端規模環境設計的使用者介面。

休息

ONTAP 也包含豐富的 RESTful API 集。這可讓協力廠商建立資料保護及其他管理應用程式、並與 ONTAP 進行深度整合。此外、想要建立自己的自動化工作流程和公用程式的客戶也能輕鬆使用 RESTful API。

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。