

技術簡介

FlexiRemap 技術

最強軟體，可將快閃記憶體儲存設備的真實效能發揮得淋漓盡致

企業的資料存取效能有其瓶頸

企業的 IT 部門需具備高效能系統，才能處理當今存取巨量資料的流量，以及越來越多元的應用種類。在這種環境下，傳統的儲存系統自然會成為資料存取的瓶頸，因為這些傳統的儲存系統根本無法在瞬間內處理如此巨量的資料。單純增大儲存空間，卻不進行根本性的改變，會讓效能無法擴展，且 IT 預算很快就會超支。

比較：現代全快閃記憶體儲存陣列與傳統磁碟陣列

現代的全快閃記憶體儲存陣列經特別設計，可滿足使用者對擴展儲存空間效能的需求。全快閃記憶體儲存陣列 (All-flash array) 又名「固態裝置」 (Solid-state appliance，簡稱 SSA)，屬於企業儲存系統的一個新類別，以快閃記憶體構成，或是一般人熟知的固態硬碟 (SSD)。SSD 與傳統的硬碟磁碟機 (HDD) 不同；HDD 可執行的每秒隨機 I/O 作業數為數百個，SSD 可執行的每秒隨機 I/O 作業數則多達數萬個。此外，SSD 的耗電只有 HDD 的一半到三分之一，更加節能，資料延遲 (Latency) 亦可縮短至 HDD 的五十分之一。對於追求儲存效能的使用者來說，這些優點使得全快閃記憶體儲存陣列成為更理想的選擇。

不過，僅在傳統磁碟陣列中以 SSD 來取代 HDD，仍無法建立高效能的全快閃記憶體儲存陣列。

這種「解藥」無法提供最佳的儲存效能，而這種效能也不可能長久延續。這是因為傳統磁碟陣列中所使用的 RAID 演算法最初是專為 HDD 而設計，實際上 RAID 在全快閃記憶體儲存陣列中不僅會限制整體效能，往往還會縮短 SSD 的使用期限。

何謂 FlexiRemap 技術？

FlexiRemap 技術是由 AccelStor 開發的快閃記憶體導向軟體，在架構上經過重新設計，並且專為快閃記憶體的基本及內部運作方式而量身打造。FlexiRemap 與 RAID 演算法不同；FlexiRemap 會在適當的時候重新對應資料，然後再把資料傳到底層的快閃記憶體，避免產生不必要的額外負擔，同時延長 SSD 的使用年限。對隨機寫入的存取模式來說，這種重新對應作業特別重要，大多數的企業資料流量都由隨機寫入模式而構成，而且也是儲存裝置存取效能瓶頸的一大肇因。另外，由於 FlexiRemap 內建資料保護機制，因此可提供 RAID 演算法令人欣賞的優點，卻排除 RAID 為人詬病的缺點。

FlexiRemap 技術的運作方式

在說明 FlexiRemap 所用的方法之前，必須先瞭解快閃記憶體的運作方式。快閃記憶體是電子的非揮發性儲存媒體，可以用電子方式將其清除和重新編程。在 SSD 或全快閃記憶體儲存陣列中，通常會使用 NAND 快閃記憶體。NAND 快閃記憶體是由區塊 (Block) 所組成，每個區塊則含有一定數量的頁面 (Page)，每個頁面都可以執行資料寫入的操作。但與 HDD 不同的是，在 NAND 快閃記憶體的頁面上，新資料無法直接覆寫取代舊資料。現有的資料必須先被複製到新位置上，舊位置上的資料才會被清除，這稱為「空間回收」(Garbage collection)。完成之後，舊位置上才可以存放新的資料。讀取和編程的執行，都是以「頁面」為基礎，但清除卻只能針對「區塊」執行。這項差異遂產生儲存效能上的難題，特別是當新資料儲存在已存入資料的 SSD 中時。對儲存效能來說，隨機寫入操作更是一大障礙。

FlexiRemap 技術

FlexiRemap 技術能以獨特的方式克服儲存程序中的種種關卡。FlexiRemap 針對隨機寫入的速度採取一項有效率的策略，能加快資料處理的速度。FlexiRemap 會把所有的新資料先重新排列成連續的區塊位址，然後傳送至 SSD，隨後再由 SSD 接收和處理這些連續的資料。簡單來說，FlexiRemap 讓 SSD 能把序列資料直接寫入連續的頁面中，因此可減輕 SSD 的工作量。

在下列範例中，新資料不需要被儲存到原位置上，產生新的同位碼 (Parity) 也不需經過繁複的計算過程。所有的新資料都會被存放在新的位置上。因此，資料存取的效能將不會受到影響。標有 X (圖 1) 的區塊不會派上用場，並且會由 FlexiRemap 的空間回收功能負責完成清理；釋出的新空間即可供日後的新資料使用。這項機制大大排除了隨機寫入所產生的缺點。

Striping 0	P	X	2	3	4	5
Striping 1	6	P	7	8	X	10
Striping 2	X	12	P	13	14	15
Striping 3	16	17	X	P	19	20
Striping 4	21	22	23	24	P	X
Striping 5	1	9	11	18	25	P

SSD0 SSD1 SSD2 SSD3 SSD4 SSD5

圖 1：FlexiRemap 演算法會在全快閃記憶體儲存陣列上重新排列資料

比較：FlexiRemap 與 RAID

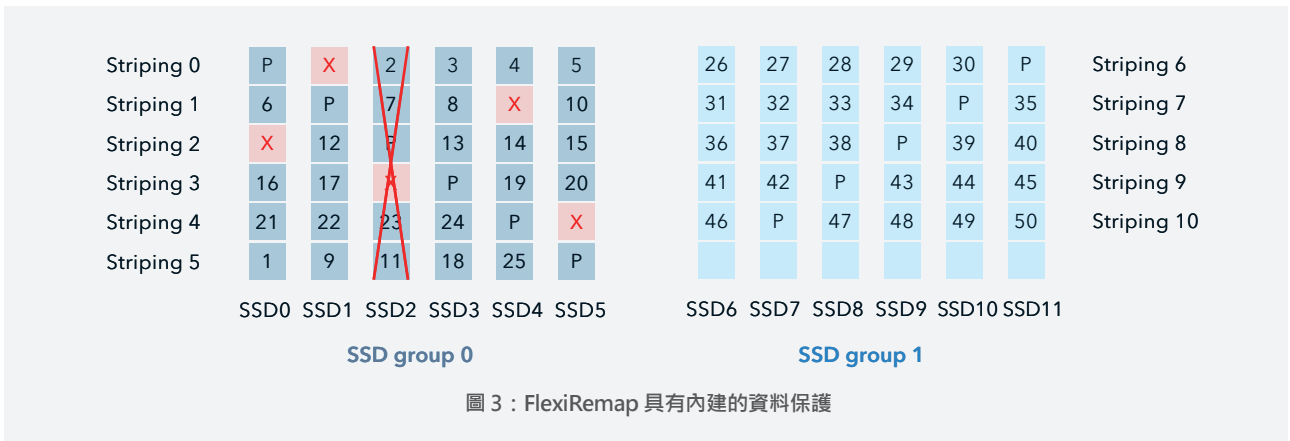
在儲存系統中，磁碟故障通常指同時失去讀取和寫入功能。因此 RAID 在多顆實體磁碟中進行分割，藉此回復故障磁碟中的資料。在多數的 RAID 演算法中，計算和儲存備援資料可用於重組一顆或多顆故障 HDD 中所遺失的資訊。不過，RAID 演算法並非全快閃記憶體儲存陣列的最佳解決方案。

在以下範例中，我們把六顆 SSD 組成一部 RAID 5 磁碟。當主機在位置 1、9、11、18 和 25 (圖 2) 上更新新資料時，目前的資料和同位碼就會與新資料被擷取和重新計算，以產生新的同位碼，隨後這些新同位碼就會被存入原來的位置上。額外讀取和寫入作業需要更長的處理時間，因此會減緩資料存取的速度。

Striping 0	P	1	2	3	4	5
Striping 1	6	P	7	8	9	10
Striping 2	11	12	P	13	14	15
Striping 3	16	17	18	P	19	20
Striping 4	21	22	23	24	P	25

SSD0 SSD1 SSD2 SSD3 SSD4 SSD5

圖 2：RAID 5 演算法會在全快閃記憶體儲存陣列上造成超額外負擔



使用 RAID 5 時，隨機寫入作業最繁雜，因為這其中通常涉及多次讀取和寫入操作，才能完成一項要求。為了提升系統效能，我們需要使用更有效率的方法。

FlexiRemap 技術具備哪些優點？

企業 IT 的挑戰始終在於資料保護、儲存效能，以及系統的耐久性。AccelStor 的 FlexiRemap 技術提供下列高價值功能，因此能滿足企業 IT 的需求：

▶ 資料保護

相較於傳統 RAID 以資料冗餘 (Data redundancy) 處理資料保護的問題，FlexiRemap 可為資料保護提供專用的演算法；它能把 SSD 磁碟分成兩組獨立的磁碟(圖 3)，資料會以單一磁碟故障的容錯平均分配給這兩組 SSD。

比方說，如果圖 3 中的 SSD 0 群組有一顆 SSD 故障，這個群組就會進入唯讀模式，以免後續寫入操作載入更多工作量。這能降低同一群組內第二顆磁碟發生故障的機率。也就是說，FlexiRemap 可避免對有故障 SSD 的群組執行 I/O 寫入作業，如在此例中的群組 0，藉此則可避免資料損毀。

▶ 高效能

當系統嘗試在快閃記憶體內更新資料時，RAID 演算法會先讀取數據條帶 (Striping) 內的現有資料和同位碼。完成計算後，新的資料和同位碼才會寫入至舊的分割位置。這項程序會為舊資料和同位碼產生兩項讀取作業，並且會為新資料和同位碼產生兩項寫入作業。RAID 無法有效率地處理大量的 I/O 要求，特別是在隨機寫入模式中，因此會劣化儲存效能。

FlexiRemap 可根據 NAND 快閃記憶體的物理性質，提供特有的資料重新對應演算法。FlexiRemap 不會把新資料寫入舊的位置，而是會把內送的寫入作業編排成連續的邏輯位址。4KB 的隨機寫入作業效能獲得了大幅改善；相較於使用 RAID 演算法時只能產生約 5K IOPS 到 78K IOPS，FlexiRemap 技術則可大幅提升整體效能至 1,020K IOPS (圖 4)。

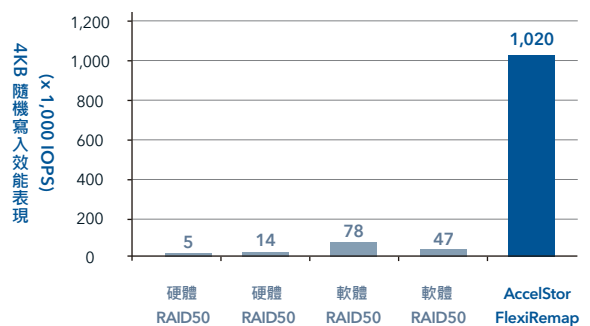


圖 4：比較 FlexiRemap 與 RAID 的隨機寫入作業效能

儘可能延長 SSD 的使用年限

SSD 與 HDD 的不同之處還包括 SSD 的抹寫次數 (P/E cycle) 是有限的。若對特定的 SSD 進行頻繁的編程作業，往往會對區塊造成不良影響，並且會縮短儲存裝置的使用壽命。FlexiRemap 的全域均衡損耗演算法 (Global wear leveling) 可讓資料平均分配於每個群組內的每顆 SSD 中。當儲存系統中重新寫入新資料時，FlexiRemap 會先檢查對應表的記錄，以確保每顆 SSD 都儲存約等量的資料。

對群組內的 SSD 平均分配資料的優點：

- 由於快閃記憶體的抹寫次數有限，因此這可延長 SSD 的耐久性。
- 改善效能。由於每顆 SSD 都儲存約等量的資料，因此較不會形成流量瓶頸。

RAID 不會解決快閃記憶體上重複編程和損耗的情形。RAID 的本質是在同樣的位置上產生重複的作業；頻繁的清除和寫入作業會導致部分 SSD 必須承受更沉重的工作。如下所述，測試結果顯示，在 RAID 系統中最常和最少使用的 SSD 之間，兩者收到的要求數約有 1.5 倍的差異。然而在 FlexiRemap 的管理之下，每顆 SSD 執行的要求數之間皆無太大差異 (圖 5)。

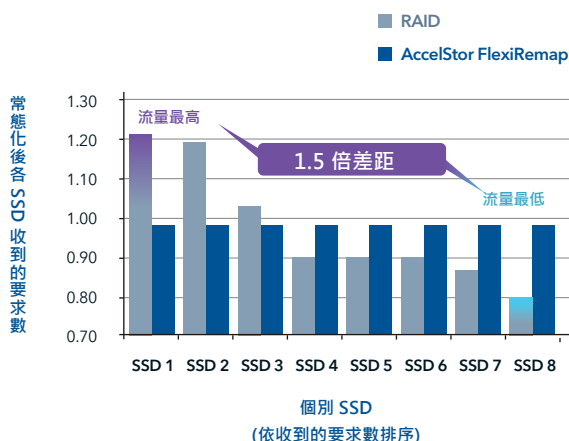


圖 5：比較 FlexiRemap 與 RAID - SSD 收到的要求數

結論

AccelStor FlexiRemap 是一個快閃記憶體導向的軟體。它能以獨特的演算法避免產生 I/O 作業的額外負擔，同時提供優異的儲存效能，更周全的資料保護功能，並可在最大程度內延長裝置使用年限。FlexiRemap 可讓 AccelStor 的產品確實滿足各種關鍵任務應用的最高效能要求。傳統的 HDD 磁碟陣列或 RAID 系統適合用於可管理的工作集合和連續寫入作業。不過，若是需進行統整的任務和高效能分析工作，使用 FlexiRemap 技術的全快閃記憶體儲存陣列能協助處理密集的隨機寫入作業和額外的 I/O 工作量，才是最佳的儲存解決方案。

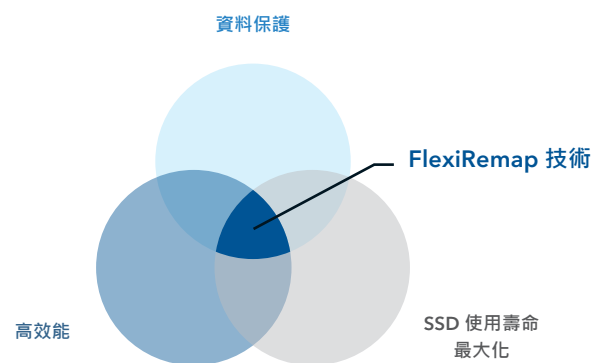


圖 6：FlexiRemap 集眾多優點於一身

本文中的測試結果是由 AccelStor 提供，在不同的測試環境下亦可能產生不同的測試結果。

AccelStor、FlexiRemap 及 NeoSapphire 是 AccelStor 在美國及/或其他國家的商標或註冊商標。

www.accelstor.com.tw
inquiry@accelstor.com

