

SMR(疊瓦式磁紀錄)硬碟 於監控儲存系統之應用優勢



隨著監控市場需求持續擴增，監控儲存系統也更趨向採用高容量、高效能的硬碟 (HDD)。這些硬碟被用於不同格式的影像儲存系統中。

為滿足不斷增長的市場需求，Toshiba 強化了 S300 系列監控型硬碟，採用疊瓦式磁紀錄 (SMR) 技術，並且提供 4TB 和 6TB 儲存容量的選項。

S300 監控型硬碟具有以下主要功能：

- 在 3.5 吋標準外型規格中，提供更高的儲存密度。
- 具備高效性能，能同時與多達 64 個攝影機連線，並且無需使用 MC (Media Cache)，能將影像資料傳送並即時記錄到硬碟的 SMR 紀錄區。
- 提供額外的傳統磁紀錄 (CMR) 區域供使用者存取，以直接記錄較小的系統資料。

1. 前言

隨著監控影像畫面的解析度提高，加上人工智慧 (AI) 影像處理技術逐漸成熟，裝設監控攝影機系統變得越來越普遍。在歐洲，監控系統可用於確保民眾及其資產的安全，並用於處理為數龐大的商業和智慧城市分析的功能。

監控攝影機產生的資料量正在急劇增加，這是因為影像解析度提高所導致。同時，市場對於能夠即時記錄多個視訊串流資料的需求也不斷提升。在這兩項因素的推動下，監控系統中的數位錄影機 (DVR) 和網路錄影機 (NVR) 所採用的硬碟，都需要具有較高的容量來進行儲存資料。

為滿足監控市場不斷提升的資料儲存需求，Toshiba 推出採用 SMR 紀錄技術的 3.5 吋監控型硬碟。這些經過應用最佳化的硬碟，可以高效的同步記錄來自多個監控鏡頭的視訊資料，並且在過程中無需利用快取 (media cache)。主要是透過將碟片的部分區域保留為 CMR 記錄區，較小的資料量且能夠直接寫入該區域來達成效能。

以下報告說明實現 SMR 監控型硬碟所面臨的挑戰，以及用於克服這些挑戰的技術。

2. 將SMR 硬碟整合到監控攝影機系統所面臨的挑戰

2.1 韌體規格

要能夠與現有檔案系統，以及監控攝影機系統中的 DVRs(數位錄影機) 和 NVRs(網路錄影機) 配合使用，硬碟需要能與現有命令集相容。此外，透過 Drive-Managed (DM) 韌體，達成以高儲存密度的 SMR 硬碟取代現有的 CMR 硬碟。經由 DM 韌體，來自主機系統的資料被暫時存放於 MC (Media Cache) 空間中。當硬碟的可用 MC 空間不足時，MC 中的資料會被寫入一個 band，或循序寫入的重疊的磁軌區域 (如下圖 1 所示)。如果這種情況是由於硬碟長時間連續運作而導致，資料從 MC 傳輸至 band 的情況將會頻繁發生。這將導致隨機寫入的效能下降⁽¹⁾，是必須解決的關鍵挑戰之一。

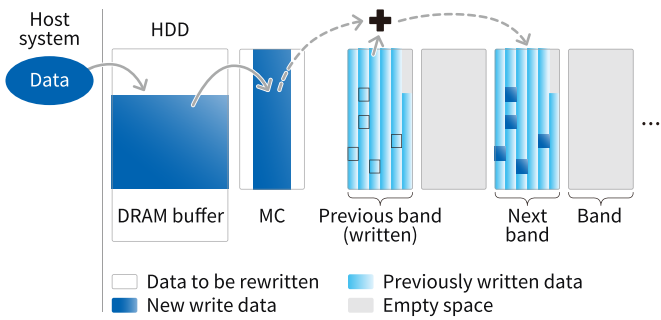


圖 1. 將資料重新覆寫至SMR硬碟的過程 - MC中的資料將會隨機寫入現有磁碟空間中。

2.2 監控攝影機系統具體的儲存使用狀況

監控攝影機系統在單顆硬碟或獨立磁碟冗餘陣列 (RAID；又稱「磁碟陣列」) 硬碟配置中，可同步記錄來自多個攝影機的視訊資料串流，通常按照下述方式執行。磁碟空間作為 Cluster 的組合來進行管理 (Cluster 組合是指數百 MB 至 GB 的邏輯區塊位址 (LBA) 空間的單位)。每個 Cluster 將分配給一部攝影機。來自特定攝影機的視訊資料串流將從 Cluster 的開頭進行循序寫入。當儲存空間接近滿載，系統會將該現有 Cluster 附近的新 Cluster 分配給該攝影機。重複此流程，將能記錄數小時的視訊內容 (如下圖 2 所示)。針對可擴增攝影機連線數量的監控系統儲存解決方案，其市場需求不斷成長。

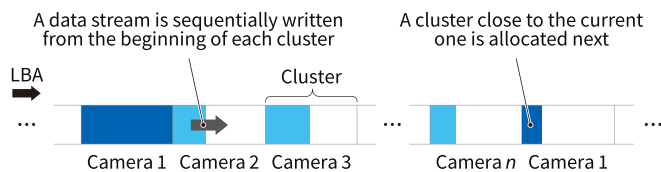


圖 2. 以監控攝影機系統中的資料儲存排列為例 - 來自某特定攝影機的視訊資料串流將從Cluster的開頭進行寫入。當現有Cluster空間使用完畢，將會分配給鄰近的Cluster來接續寫入。

3. S300 系列概述

表 1 顯示 Toshiba S300 3.5 吋 SMR 監控型硬碟 4TB 和 6TB 的主要規格。為了開發這些型號，需要對磁碟的記錄層和讀寫頭的效能進行增強，以便運用 SMR 提高磁軌密度的同時，能提升線性記錄密度。這些 S300 型號將現有硬碟指令與 DM 韌體結合。此韌體專為實現高串流效能而打造 (可支援多達 64 個監控攝影機的連線)。後續章節將介紹開發 S300 系列所採用的技術。

S300 監控型硬碟	6 TB	4 TB
	HDWT860UZSVA	HDWT840UZSVA
外型規格	3.5 吋	
介面	SATA 6 Gbit/s	
紀錄技術	SMR	
支援的攝影機數量	高達64台	
轉速	5,400 rpm	
最大資料傳輸速度	184 MB/s 持續 (typical值)	
緩衝區大小	256 MB	
工作負載	180 TB/年	
MTTF (小時)	1,000,000	
耗電量 (運作閒置)	2.81 W (Typical值)	2.33 W (Typical值)
外型規格/尺寸 (公釐)	147 (長) × 101.85 (寬) × 26.1 (高)	
重量	680 g	650 g

表 1. S300 4TB / 6TB 之主要規格。

4. 視訊資料記錄方法

4.1 使用備用band和bypass寫入的 SMR

S300 系列採用 SMR 技術並以 band 為單位記錄資料，並在 band 之間保留未使用的區域，以防止寫入特定 band 時破壞相鄰 band 上的資料。多個 band 的集合稱為 band group。雖然 band group 中的所有 band 具有相同的容量，但 band 的儲存容量取決於所處位置是外圍軌道還是內軌道。每個 band group 都有一個備用 band，做為沒有儲存資料的工作空間（如下圖 3 所示）。

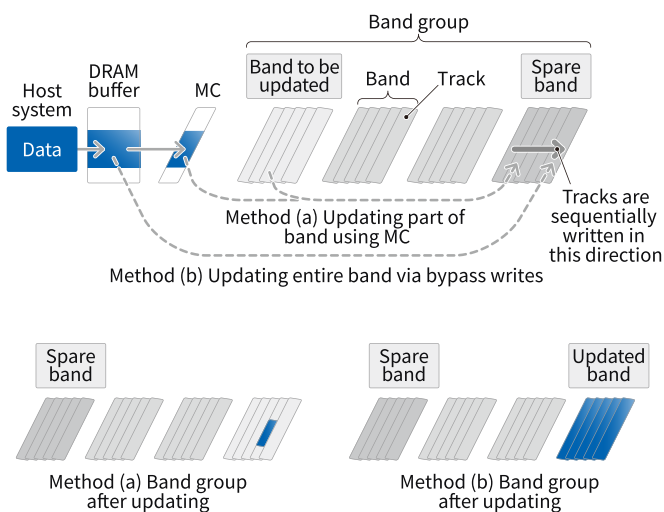


圖 3. 使用備用band進行資料更新操作 - 將新資料寫入備用band。當韌體會將更新的band標記為備用band，並將前一個備用band標記為有效band時，即可以安全地進行資料更新。

來自主機系統的資料首先會傳到 DRAM 緩衝區。接著，經過緩衝的資料透過兩種不同方式傳輸到備用 band。採用哪種方式將取決於資料大小。兩種選項分別是：

(a) 透過更新部分 band - 此處 DRAM 緩衝區中的資料首先寫入 MC (Media Cache)。這些資料可以與要更新的 band 中之現有資料合併，然後移動到備用 band。之後，韌體會將原始 band 標記為備用 band，並將前一個備用 band 標記為有效 band。

(b) 更新一個 band 中的所有資料 - 當大量連續資料串流來自主機系統時，DRAM 緩衝區中的資料將直接寫入備用 band 中。當到達備用 band 的尾端時，韌體會標記表示該 band 應更新為備用 band，並將前一個備用 band 標記為有效 band，然後重複此流程。

此方法 (b) 不使用 MC 而直接將資料傳輸到 band，稱為 bypass 寫入。儘管 bypass 寫入只能用於更新 band 中的所有資料，但只需極少量的操作即可完成。由於監控攝影機系統中的多數視訊資料都是連續的，因此提高寫入效能的關鍵在於盡可能的使用 bypass 寫入來記錄更多資料。

4.2 使用空間bands的bypass寫入

如第 2.2 節所述，監控攝影機系統同步記錄來自附近 band 的多個攝影機的視訊資料流。為有效處理此類資料，必須同時對附近的 bands (即同一 band group 中的多個 bands) 進行 bypass 寫入。4.1 節說明使用備用 band 進行 bypass 寫入的範例。

此方法的缺點是對於給定的 band group，只能執行一個 bypass 寫入串流。因此，Toshiba S300 系列的設計不是透過備用 band，而是允許使用空間 band (即儲存過時資料且可以覆蓋的 band)，以便同時進行對給定 band 群組的多個 bypass 寫入 (如下圖 4 所示)⁽²⁾。

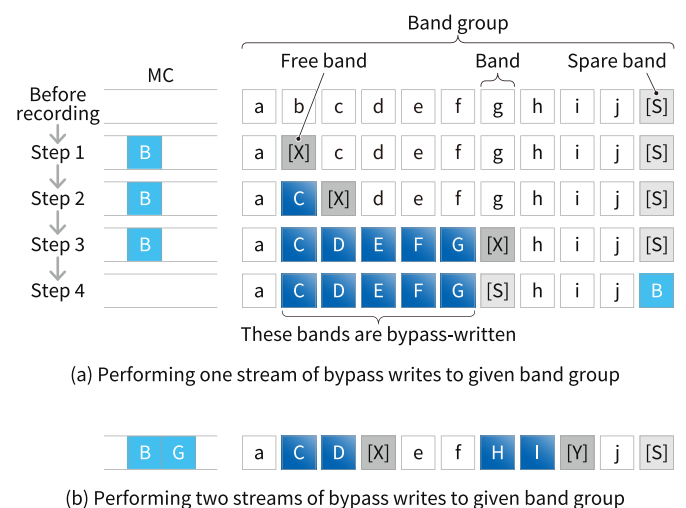


圖 4. 使用空間band對band group進行bypass寫入 - 將資料寫入 MC 時會產生空間band。使用空間band執行bypass寫入。

首先，讓我們檢視如何對 band group 執行 bypass 寫入串流 (如上圖 4 (a) 所示)。字母 a 到 j 表示每個 band 中儲存的現有資料，而 [S] 和 [X] 分別表示備用 band 和空閒 band。

以下段落說明用 B 到 G 順序更新 b 到 g 的連續資料的程序。

- 步驟 1：B 寫入 MC。此時，b 已過時，無需保留。因此，儲存 b 的 band 被標記為空閒 band [X]，以便用於記錄其他資料。
- 步驟 2：C 寫入 [X]。這是以 bypass 寫入執行的，同時保持現有資料 c 不變。因此，即使 bypass 寫入提前終止，現有資料也不會遺失。當 C 的寫入完成時，儲存 c 的 band 被標記為空閒 band [X]。
- 步驟 3：按照與步驟 2 相同的方式，依序寫入 D、E、F 和 G。
- 步驟 4：MC 中的 B 轉移到 [S]，並且 [X] 被標記為備用 band [S]。這意味著 [X] 已同步刪除。

接下來，我們要檢視同一個 band group 如何進行多次 bypass 寫入 (如上圖 4 (b) 所示)。在此範例中，同步執行兩個 bypass 寫入串流 (B-C-D 和 G-H-I)。此流程使用將 B 和 G 寫入 MC 時產生的兩個空閒 band [X] 和 [Y]，以便對 C 和 H 同步進行 bypass 寫入。

將資料 band 寫入 MC 時可以產生任意數量的空閒 band，因此可以同時進行三個以上的 bypass 寫入串流。bypass 寫入是逐 band 執行。由於 band 大小與監控攝影機系統的 Cluster 大小無關，視訊資料的前端和尾端不足以填充 band。因此，這些零散的資料元素在寫入 band 之前被組合，其方式與第 4.1 節所述的方法 (a) 相同。

如此，S300 系列硬碟利用空閒 band 進行 bypass 寫入，以適應監控攝影機系統的資料流特性。

5. 以 CMR 格式記錄系統資料

除了視訊資料之外，監控攝影機系統還同步記錄管理和分析資料 (以下合稱「系統資料」)。系統資料由許多小資料區塊組成。系統資料本質上是隨機的，因此不適合 DM-SMR。然而，由於監控攝影機系統的效能不斷提升，系統資料量也隨之持續增長。

為有效記錄監控攝影機系統的系統資料，S300 系列提供了 CMR 區域，該區域位於固定的 LBA 範圍內。雖然隨機資料可以直接寫入 CMR 區域，但保證系統資料的完整性比單一視訊資料的完整性更重要。為此，監控攝影機系統經常使用 Flush Cache 和 Flush Cache EXT 命令，請求硬碟對所有未寫入的資料進行 flush-write，刷新所有未完成寫入操作的寫入資料。由於典型的硬碟在 DRAM 緩衝區中進行寫入資料快取，因此大部分寫入資料保留在 DRAM 緩衝區中，等待寫入磁碟。由於需要進行 flush-write 的資料量很大，因此刷新寫入必須高速進行，以確保視訊資料錄製的穩定性。

S300 系列會回應 flush-write 請求，立即將所有未寫入的資料寫入磁碟。此時，由於寫入資料已在 MC 中進行快取，因此可以在短時間內完成要放置在 SMR 區域中的寫入操作。相較之下，對 CMR 區域的寫入需要大量時間，因為目的地為 CMR 區域的資料將導致許多尋軌操作和旋轉延遲。即使這些寫入按照寫入操作的順序進行最佳化，前述情況也是不可避免。為解決這個問題，S300 系列硬碟提供了一個可以暫時儲存 CMR 區域資料的區域 (如下圖 5 所示)。

為回應 flush-write 的工作要求，此資料會寫入暫時儲存區域，以縮短操作時間。

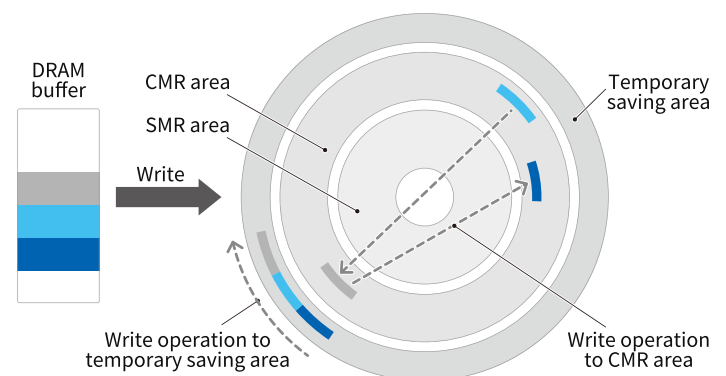


圖 5. CMR 和磁碟上的暫時儲存區域 - CMR 區域適合記錄系統資料。S300 系列提供了一個暫時儲存區域，以加快 flush-write 的速度。

6. 結論

適用於監控攝影機系統的 S300 系列 3.5 吋 SMR 硬碟 (4TB 和 6TB)，為客戶提供每磁片 2TB 的高密度儲存空間。此外，這些型號還結合了針對多重串流監控環境進行最佳化的韌體，最高可支援同時連線至 64 台攝影機。隨著 4K (3,840×2,160 像素) 和 8K (7,680×4,320 像素) 超高畫質 (UHD) 攝影機日益普及，監控系統產生的資料量預計將會持續增加，資料儲存資源必須能跟上更多容量的需求。

監控攝影機系統也融入了各種創新功能，以充分發揮 SMR 硬碟的特性。因此，監控攝影機系統對 SMR 硬碟的需求預計將會增加。為因應監控硬碟市場的增長，Toshiba 將繼續提高監控型硬碟的容量，以滿足客戶的需求。

參考資料

Toshiba Corporation. “Large-Capacity, High-Performance 3.5-inch HDDs for Surveillance Camera Systems Applying SMR Technologies” Toshiba Review: 76 (6), November 2021

Chrome-extension://efaidnbmninnibpcjpcglclefindmkaj/https://-toshiba.semicon-storage.com/content/dam/toshiba-ss-v3/master/en/company/technical-review/pdf/3_5-inch-hdd-smr_202303_en.pdf

(1) Shimomura, K. 2019. “Large-Capacity HDDs Applying SMR Technology for Data Centers.” Toshiba Review: 74(6): 12–16. Accessed August 2, 2021.

https://www.global.toshiba/content/dam/toshiba/migration/corp/techReviewAssets/tech/review/2019/06/74_06pdf/a04.pdf.

(2) Toshiba Corporation, Haga, T., Tanaka, H. 2016. “Disk storage device and method.” Japanese Patent No. 5951472. Published July 13, 2016.