

過度供應 對三星的好處 數據中心固態硬盤

預留空間是一項功能，專門為要從 SSD 中擦除的數據提供額外容量，而不會中斷系統性能。

專用的預留空間可以根據用戶的喜好進行調整，帶來更快的速度和更長的 SSD 壽命等好處。

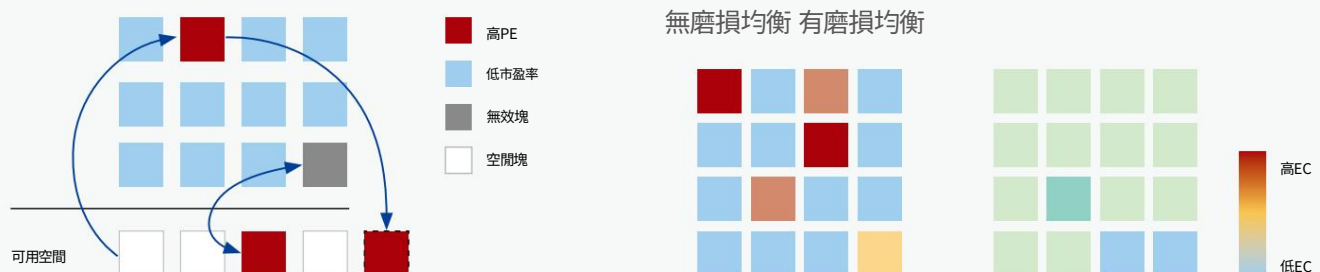
本白皮書提供了有關預留空間的深入信息，以及有關如何調整預留空間的說明，以及調整前的注意事項。

背景與硬盤驅動

器 (HDD) 不同，固態驅動器 (SSD) 在寫入數據時將電子存儲在 NAND 單元上。使用 NAND 閃存，存儲或擦除新數據時不會覆蓋已存儲的數據。由於固態硬盤的寫入和擦除操作 (Program/Erase) 是在不同的單元中進行的，分別稱為頁面和塊，因此在寫入和管理數據時不可避免地需要多次循環編程和擦除。隨著更多此類循環的重複，一些電子會被困在電池之間，隨著時間的推移，這些電池會達到其使用壽命的盡頭並遇到耐久性問題。這種現象稱為單元磨損，是造成 NAND 壽命物理極限的原因。

因此，正確管理 NAND 對於延長 SSD 的使用壽命至關重要。當在某個區域重複寫入數據時，相應的單元會很快磨損，因此應避免重複寫入相同的單元。Wear-leveling 是一種防止對特定區域重複寫入操作的功能，通過將暴露於大量 P/E 週期的塊與空間塊交換，使單元能夠被均勻地利用，從而允許用戶在給定的條件下使用更長時間的 SSD 狀況。

圖 1：磨損均衡

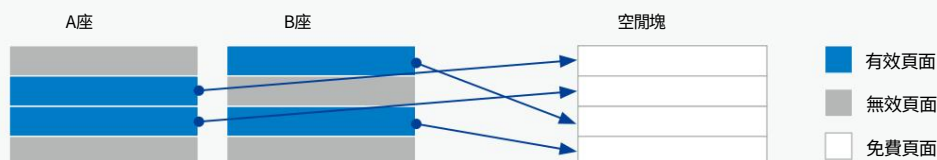


由於 NAND 閃存無法覆蓋，因此必須先擦除現有數據才能將新數據寫入該單元，這會降低 SSD 的整體寫入性能。通常，擦除數據比寫入數據需要更長的時間，因為如前所述，寫入操作是在頁中執行的，而擦除操作是在塊中執行的。為了緩解寫入性能的下降，實施了稱為垃圾收集 (GC) 的過程以在 SSD 中創建空間塊。該技術通過將有效頁面收集到一個位置並擦除由無效頁面組成的塊來保護空間塊。但是，在垃圾收集干擾主機寫入的意外情況下，這有時也可能導致性能下降。

因此，需要 SSD 中的可用空間才能使固件 (FW) 功能順利運行。分配額外空間的過程稱為超額配置 (OP)。

圖 2 :垃圾收集

步驟1
收集有效頁面

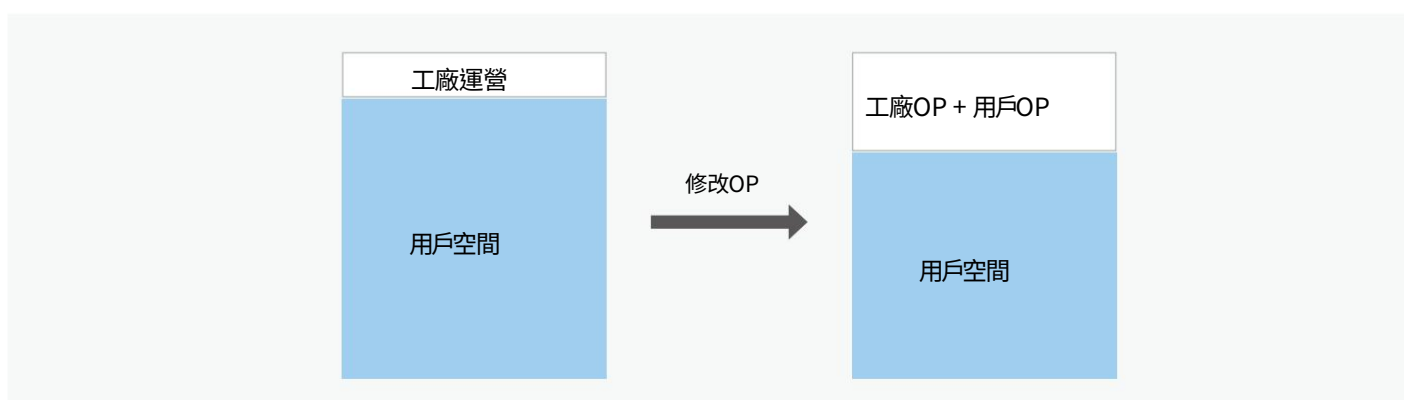


第2步
擦除由無效頁面組成的塊



什麼是 OP（過度配置）？

如前所述，預留空間是指通過將一定數量的 SSD 的 NAND 閃存分配到預留空間來確保額外空間以允許高效使用 SSD 的功能。此空間只能由 SSD 的控制器訪問，而不能由主機訪問。OP 區域僅由空間塊組成，有助於在進行磨損均衡或垃圾收集時有效地交付空間塊，並有助於提高 SSD 的性能和使用壽命。通常，三星 DC SSD 默認設置為 OP 提供 6.7% 的容量，但如果用戶需要額外的 OP，則可以根據用戶環境手動調整空間大小。



如何計算 OP 比率？

OP比例公式 : $OP(\%) = ((Physical\ Capacity - User\ Capacity) / User\ Capacity) * 100$ Ex)

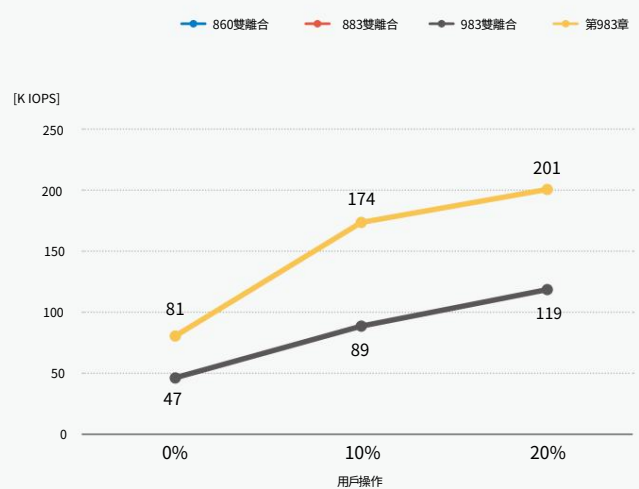
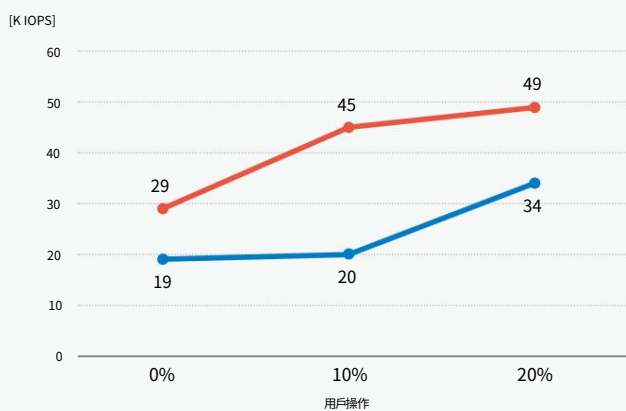
當使用128GB SSD中的120GB作為用戶容量，8GB分配給OP時，OP (%) 是 $((128 - 120) / 120) * 100 = 6.7\%$ 。

增加OP有什麼好處？

儘管開箱即用 (FOB) NAND 的順序寫入和隨機寫入性能沒有差異，但是一旦數據被寫入整個空間，隨機寫入的性能不如順序寫入與非門。

隨機寫，比順序寫更小，在塊中混合有效和無效頁面，這會導致頻繁的 GC 並導致性能下降。如果增加 OP，則可以確保主機無法訪問的更多空間，由此產生的 GC 效率有助於提高性能。持續性能以相同的方式得到改善。

圖 4. SATA/NVMe DC SSD 每個用戶 OP 的隨機寫入性能 (QD32)



從產品的生命週期來看，GC等內部操作導致NAND寫入次數變得大於主機寫入次數，這也導致WAF (Write Amplification Factor)增加，定義為比率主機寫入到 NAND 寫入。WAF 值的增加表明意外的 NAND 使用正在增長，並且在達到寫入的總字節數 (TBW) 之前，產品的使用壽命可能會縮短。足夠的 OP 空間通過提高內部 NAND 操作的效率來減少 NAND 的使用。它具有增加保修期內可用的每日工作量 (DWPD) 的優勢。

三星 DCT 系列示例

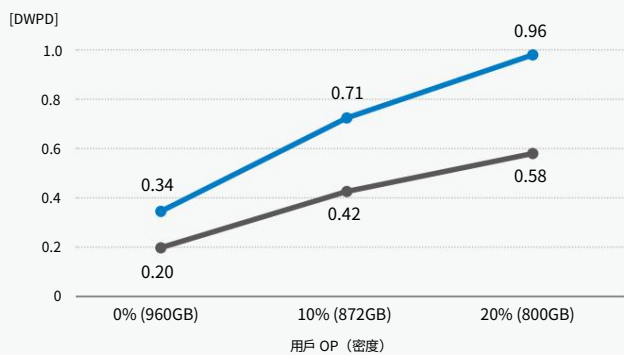
下圖通過為三星數據中心產品（860/883/983 DCT、983 ZET）分配額外的用戶 OP，顯示了每個保修的估計 DWPD。圖表中的值是使用以下公式計算的，是每個 SSD 的計算值，不是保證值。DWPD 隨著 OP 率的增加而增加。用戶可以通過 SMART 屬性找到下面計算所需的數字，這些在下一節“使用 SMART 屬性估算 SSD 的壽命”中有更詳細的解釋。

$$\text{WAF} = \frac{\text{物理寫入量}}{\text{主機寫入量}}$$

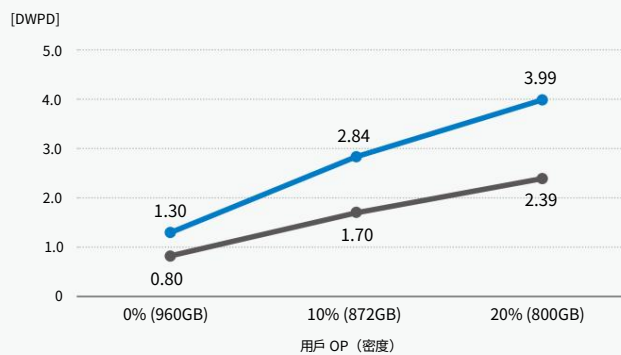
$$\text{DWPD} = \frac{\text{NAND PE 週期} \times \text{原始密度}}{\text{邏輯密度} \times 365 \times \text{質保年限} \times \text{WAF}}$$

圖 5. 每個用戶 OP 的 DWPD

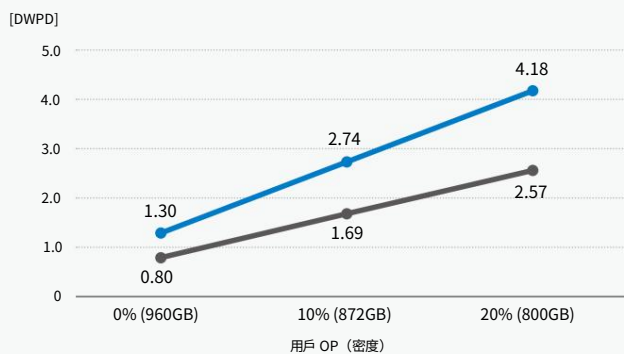
適用於860 DCT 960GB



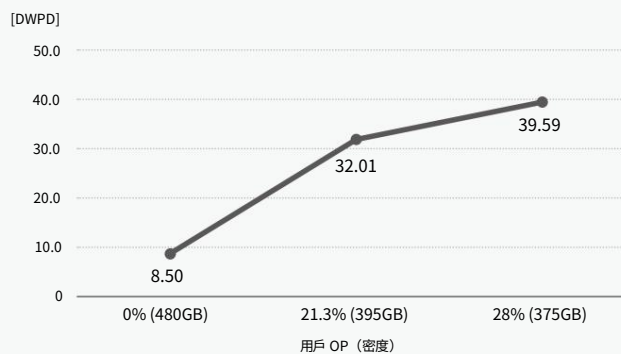
適用於883 DCT 960GB



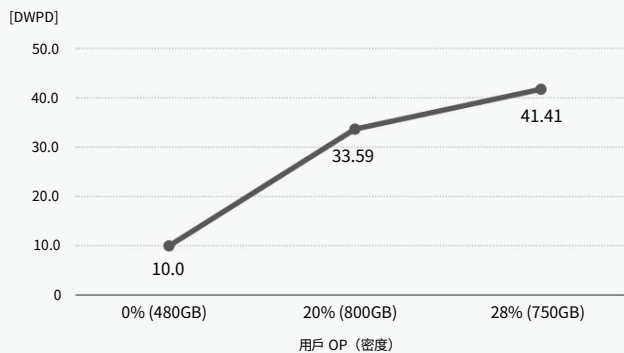
適用於983 DCT 960GB



適用於983 DCT 480GB



適用於983 ZET 960GB



如何估算具有 SMART 屬性的 SSD 的壽命？

- 查看 SMART 屬性表可以讓用戶了解他們的 SSD 在特定工作負載和時間段內的磨損情況。或者，這些屬性允許用戶推斷其 SSD 的使用壽命。請注意以下屬性僅適用於 SATA SSD。
- ID 241和ID 251分別表示host和NAND的寫入量，用戶可以計算自己SSD的WAF有了這些價值觀。
- ID 177 表示磨損均衡操作的次數，也可以解釋為總體平均編程/擦除週期。該值與 WAF 值一起允許用戶找出 DWPD。
- ID 247* 表示自工作負載計時器啟動以來SSD 已運行的時間（以秒為單位）。（用戶可以自行決定啟動/停止所述計時器或讓它持續運行。它通過他們的 SSD 軟件工具進行控制）。
* 請注意，對於 860 DCT，ID 247 以分鐘錶示。
- ID 246 顯示自工作負載計時器 (ID 247) 啟動以來讀取命令的 I/O 操作份額，並以百分比表示。（相反，寫 I/O 操作的份額可以通過從 100 中減去給定的智能屬性讀數來確定）。
- ID 245 衡量SSD 在給定工作負載(ID 246) 和這些工作負載持續的時間段(ID 247) 下的磨損情況。它顯示為 SSD 在其使用壽命內總磨損的千分之一讀數（即讀數為 1000 意味著 SSD 在給定的時間和使用模式下已經磨損）。
- 適用於 883 DCT 的示例
一位用戶目睹了他的 SSD 的使用模式最近從 80% 減少到 70% 的讀取 I/O 操作，他想了解這一變化對其 SSD 的使用壽命有何影響。他決定進行為期 1 週的測試。在他的測試運行結束時，SMART 屬性顯示如下：ID 245 :4、ID 246 :70 和 ID 247 :604,800（7 天 x 24 小時 x 60 分鐘 x 60 秒）

要根據上述讀數找到估計的使用壽命結束時間，用戶需要進行以下計算：首先，用戶想要了解在 SSD 完全耗盡之前，在給定的測試場景下還可以運行多少個週期。因此，計算得出： $1000 / 4 = 250$ 個週期。

其次，用戶隨後將該數字乘以測試運行的持續時間，以得出 SSD 的總預期壽命（以秒為單位）。此計算得出 $250 \times 604,800 = 151,200,000$ 秒。

最後，考慮到以秒錶示的大數的相對抽象性質，用戶隨後希望以年、月或週來表示計算出的生命週期。如果用戶選擇以年表示壽命，則需要進行如下計算： $151,200,000 / (365 \times 24 \times 60 \times 60) = 4.79$ 年。

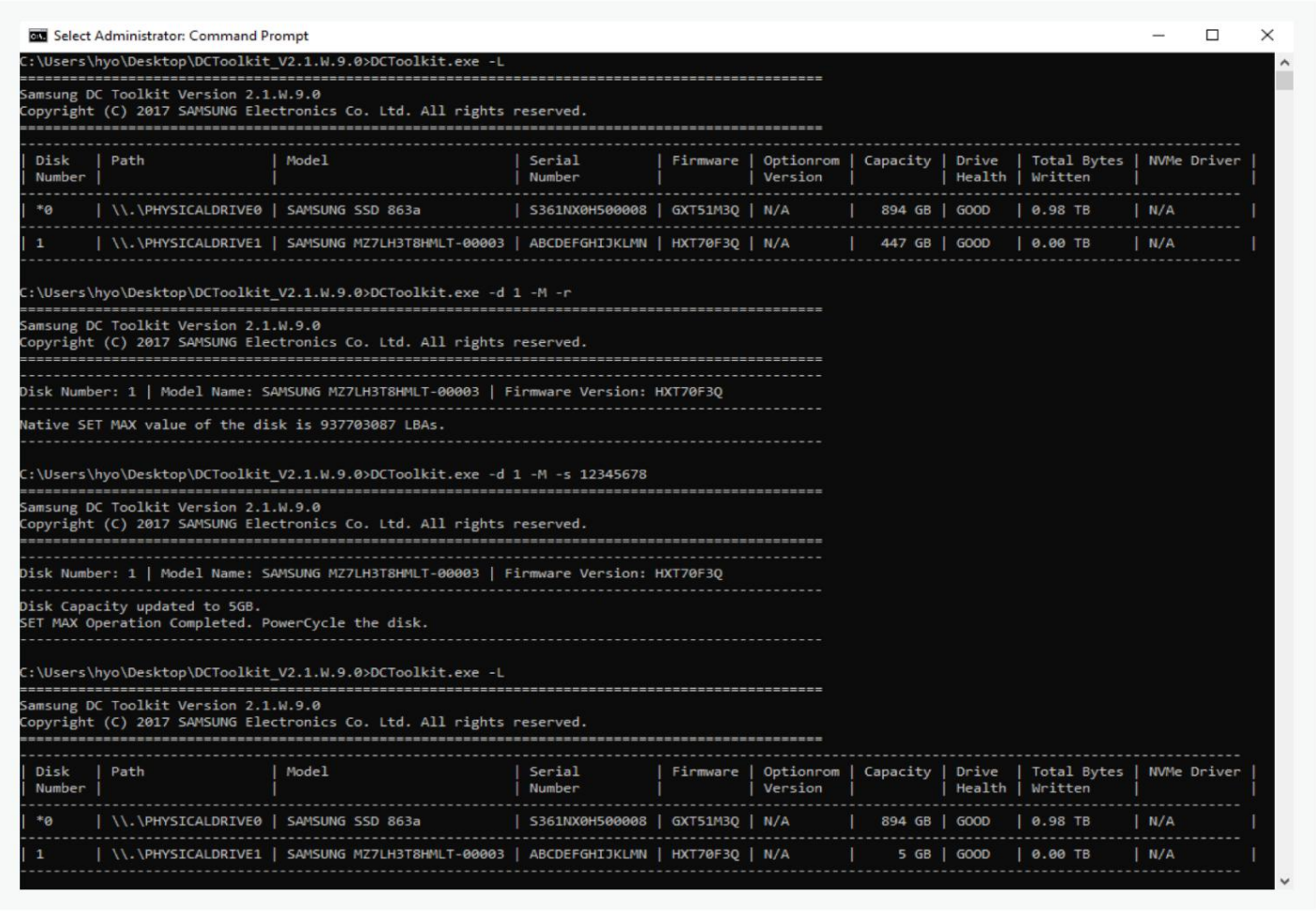
ID	屬性名稱	狀態標誌	臨界點 (%)
58	重置扇區計數	110011	10
9	開機時間	110010	-
12	開機計數	110010	-
177	磨損均衡計數	010011	58
179	使用的保留塊計數 (總計)	010011	10
180	未使用的保留塊計數 (總計)	010011	10
181	程序失敗計數 (總計)	110010	10
182	擦除失敗計數 (總計)	110010	10
183	運行時錯誤計數 (總計)	010011	10
184	端到端錯誤數據路徑錯誤計數	110011	97
187	不可糾正的錯誤計數	110010	-
190	氣流溫度	110010	-
194	溫度	100010	-
195	ECC錯誤率	011010	-
197	待定扇區計數	110010	-
199	CRC 錯誤計數	111110	-
202	SSD 模式狀態	110011	10
235	POR 恢復計數	010010	-
241	編寫的 LBA 總數	110010	-
242	讀取的 LBA 總數	110010	-
243	SATA 降檔控制	110010	-
244	熱節流狀態	110010	-
245	定時工作負載介質磨損	110010	-
246	定時工作負載主機讀/寫比率	110010	-
247	定時工作負載定時器	110010	-
251	NAND 寫入	110010	-

如何調整 OP ？

用戶可以使用 DC Toolkit 或 Linux HDParm 磁盤管理來調整 OP 的可用空間。此調整僅限於未使用的空間。如果用戶希望增加 OP 的可用空間，那麼他或她必須清理一些已經在使用的空間。

如何使用 DC Toolkit 調整 OP 設置

步驟說明	命令指令
1. 識別連接到系統的設備。	DCToolkit.exe -L
2. 檢查 MAX ADDRESS 設置可用範圍。	DCToolkit.exe -d 1 -M -r
3. 設置最大地址值。	DCToolkit.exe -d 1 -M -s 12345678
4. 確認設定值。	DCToolkit.exe -L



結論

雖然增加 OP 具有提高用戶 SSD 的性能和使用壽命的優勢，但它也會減少主機的可用空間。這意味著在調整 OP 時，必須考慮為用戶的工作負載提供適當的空間，而不是過度增加它。如果在 SSD 為 FOB 時設置首選項，用戶可以體驗到 OP 的最大好處。因此，建議在設置 OP 設置時，用戶的 SSD 產品為 FOB。

有關三星 SSD 的更多信息，請訪問 www.samsungssd.com。

版權所有 © 2019 三星電子有限公司。保留所有權利。Samsung 是 Samsung Electronics Co., Ltd. 的註冊商標。規格和設計如有更改，恕不另行通知。非公制重量和尺寸均為近似值。所有數據在創建時都被認為是正確的。三星對錯誤或遺漏概不負責。所有品牌、產品、服務名稱和徽標均為其各自所有者的商標和/或註冊商標，特此予以承認。