

中央氣象局「大量資料儲存系統」簡介

壹、前言

中央氣象局大量資料儲存系統（Massive Storage System，MSS）係因應本局數值天氣預報模式系統及其他資訊系統所產製之大量資料儲存所需而建置。由於氣象局之主要作業系統運作型態具有即時性，並且作業環境必須維持 7 天 24 小時的高可用性，而儲存媒體更必須滿足網際網路上大量資料的流通與配合高速電腦之資料處理能力，因此本局之大量資料儲存系統需具備以下之特性：

- 高容量：提供足夠的儲存空間，但須兼顧經濟規模。
- 高可靠：資料需有一份以上之備份，並維持資料在存、取之正確性。
- 高效率：資料的存取效能需達一定水準，以配合作業或研發所需，而不會降低系統的執行效率。
- 高可用：系統要能穩定，同時配合方便易用的服務軟體，以維持系統的可用性。

基於以上之需求，本局大量資料儲存系統採用階層式儲存管理（Hierarchical Storage Management，以下簡稱 HSM）架構，將磁碟系統和自動磁帶櫃系統之不同存取特性的儲存設備，整合成一個階層式超大容量的儲存體，在各階層的儲存體間，具有自動資料轉移的特性。HSM 架構在提供資訊系統大量資料儲存空間的同時，並能兼顧到資料存取效能和經濟規模。

然而由於整個大量資料儲存系統由不同層次的軟、硬體，包括伺服器主機、磁碟陣列、磁帶館和各式磁帶機與磁帶媒體組合而成，軟體部分則涵蓋管理軟體和資料庫軟體，系統架構複雜且各牽涉不同的專業領域，每個環節的風險均可對系統產生重大影響，例如因系統之 single point of failure 導致作業停頓、因作業容量不足而影響系統穩定性、因不同層次之軟、硬體維護之協調不良而導致不當維護等，這些風險一旦產生作用，勢必對本局之各式資訊系統產生重大影響，因此本局之大量資料儲存系統不惜投入大量經費，各個環節均採用最新的資訊科技，以期能維持整個系統的高效能和高可用性。是故，該大量資料儲存系統係本局之重要資產，平時即由專人負責系統之維運，並隨時監控系統的作業效能；除此之外，我們也隨時關心最新資訊科技的發展，以使本系統能保持在最佳狀況，以提供局內各式作業最好的儲存後援。

貳、中央氣象局大量資料儲存系統演進

中央氣象局大量資料儲存系統於民國 80 年隨著 CRAY 超級電腦採購案開始建置，初期買進 STK4410 磁帶館，該磁帶館可容納約 5000 捲匣式磁帶，此外，該系統亦包括 8 台 STK 4480 磁帶機，STK 4480 磁帶機可讀取的磁帶容量為 200 MB，所以當時大量資料儲存系統的最大總容量只有 1 TB。

隨著時間的推進，氣象局所需存放的資料量越來越多，因此大量資料儲存系統亦隨之升級。首先，於民國 85 年採購一部檔案伺服器主機（convex），並使用 uni-tree 軟體提供 HSM 服務。此外，也將 STK 4410 磁帶館升級至 STK 9310 磁帶館，另外將 4 台 STK4480 磁帶機升級成 TimberLine 9490，其磁帶容量由 200 MB 提升為 800 MB，磁帶機速度由 3 M/sec 提昇到 6 M/sec。同時購進 3 台 Red Wood SD-3 磁帶機，其磁帶容量有 10 GB、25 GB 和 50 GB，於是整個大量資料儲存系統的最大容量提升至 250 TB。

為擴充大量資料儲存系統之處理容量，於民國 87 年隨著 Fujitsu 超級電腦採購案建置另一部檔案伺服器主機（Sun E5500，記憶體 4 GB）、磁碟陣列一部（容量為 450 GB）以及 ASM 階層式檔案管理系統一套。至此，本局大量資料儲存系統的系統架構便依據當時的設計而沿用至今。

民國 88 年更進一步採購 3 座 STK9740 磁帶館，約增加 1300 捲磁帶存放容量，另外採購 5 台 STK9840 磁帶機，其可讀寫的磁帶容量為 20 GB，存取速度為 19 MB/sec，此時整個大量資料儲存系統的最大容量達 126 TB。

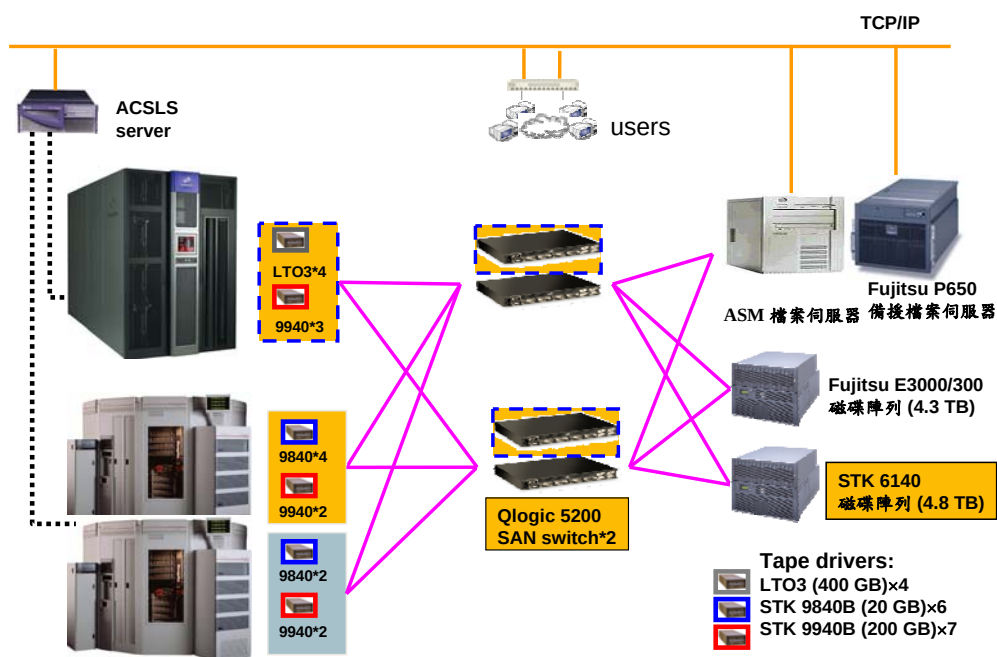
為因應本局各式資料源增加趨勢，同時也提供給局內各個資訊系統儲存相關資料之用，並配合局內異地備援機制之建置與發展，大量資料儲存系統於民國 92 年進一步擴充，包括新建置一部 STK9310 磁帶館、擴充檔案伺服器之磁碟陣列（HP StorageWorks MA8000，容量為 800 GB）、另採購 STK9940 磁帶機 5 部，其可讀寫的磁帶容量為 200 GB，存取速度為 30 MB/sec，至此局本部之大量資料儲存系統儲存容量達 2400 TB。在此同時將局本部 3 部 STK9740 磁帶館和 2 部 STK9940、1 部 STK9840 磁帶機移往南區氣象中心，並新採購一部 Sun V480（記憶體 4 GB）做為檔案伺服器，一部 Sun Blade150 做為磁帶館控制主機、磁碟陣列一部（容量為 750 GB）以及 ASM 階層式檔案管理系統一套，至此，南區氣象中心建置之獨立的大量資料儲存系統，其最大容量可達 260 TB，此一大量資料儲存系統除作為局本部資料異地備援之設備，同時也提供南區氣象中心資料儲存之用。

在另一方面，由於大量資料儲存系統所服務的容量與時俱增，雖然磁帶儲存系統的容量不斷增加，但檔案伺服機的作業容量並未隨之擴充，特別是隨著各年度計畫的推動及各相關應用系統陸續上線作業，原先 Sun E5500 及其磁碟陣列早已不敷所需，並直接影響到整個大量資料儲存系統的可用性，而 HP StorageWorks MA8000 磁碟陣列與 Sun E5500 檔案伺服器系統產生相容性問題，並無法實際發揮 HSM 之磁

碟快取的功能。因此，於民國 94 年汰換原有的 Sun E5500 系統，並採購 Fujitsu Primepower 650 伺服器一部(含 SPARC64 V 1.89 GHz CPU 4 顆、記憶體 16 GB)、Fujitsu ENTERNUS 3000 磁碟陣列一部(容量達 4.3 TB)以及 STK9840 磁帶機一部，此部磁帶機建置於南區氣象中心，此時南區氣象中心的大量資料儲存系統擁有 2 部 STK9840、2 部 STK9940 磁帶機，可有效提高磁帶機的可用性。在新的檔案伺服系統建置完成後，本局之大量資料儲存系統藉由最新資訊科技輔助而具體提升其可用性，以作為局內各式應用系統資料儲存的最佳後盾。

近年來氣候監測與氣候變遷成為舉世關注的焦點，同時災變天氣對台灣經濟與民眾身家安全所造成的影響與日遽增，因此本局為提升氣候監測與預報，以及改善劇烈天氣預報，積極推動「氣候變異及劇烈天氣監測預報系統發展計畫」，也因此對 MSS 產生更多的需求。為此，本局在民國 96 年進一步擴充 MSS 系統，以滿足各式作業所需的儲存空間，以下將進一步介紹本局目前 MSS 的系統架構與作業現況。

參、中央氣象局大量資料儲存系統架構簡介



圖一：中央氣象局大量資料儲存系統架構圖

圖一是本局大量資料儲存系統架構圖，圖中顯示，該系統主要由線上儲存設備(包括檔案伺服主機、磁碟陣列)以及近線儲存設備(包括磁帶館、磁帶機和 ACSLS 伺服主機)等整合在一個儲域網路(Storage Area Network, SAN)的環境中，茲將各個系統組成成分詳述於下。

一、檔案伺服主機

大量資料儲存系統之檔案伺服主機為 Sun Fire V890，共有 4 顆 1.8GHz UltraSPARC IV+ 中央處理器和 16 GB 記憶體，此外由 Fujitsu Primepower 650（含 SPARC64 V 1.89 GHz CPU 4 顆、記憶體 16 GB）作為備援主機，以達到高可用性的目的。大量資料儲存系統在檔案伺服主機上執行階層式儲存管理軟體（Sun-FS ASM），提供磁碟檔案系統至磁帶系統間的檔案自動移轉（遷入及遷出）功能，並可由管理者依據磁碟空間、檔案使用頻率、檔案大小等參數設定移轉規則。ASM 支援以下功能：

- Archiving—自動將檔案拷貝至磁帶，但不清除磁碟中的資料。
- Releasing—自動將已經上磁帶的檔案自磁碟刪除並釋放磁碟空間。
- Staging—自動將檔案媒體從磁帶回存至磁碟。
- Recycling—將到期的檔案（expired）從磁帶中清除。

ASM 授權數量以儲存容量計，目前總計為 1200 TB。

二、 磁碟陣列

大量資料儲存系統擁有兩部磁碟陣列，作為 HSM 之磁碟快取的角色，一部是 STK 6140，容量為 4,8 TB，另一部為 Fujitsu ENTERNUS 3000 磁碟，容量達 4.3 TB，兩部均支援光纖介面。

兩部磁碟陣列均受 ASM 軟體管轄，具有 2 組光纖磁碟陣列控制器，支援動態負載平衡，且具有 HA 備援功能，不會因單一光纖磁碟陣列控制器故障而使系統無法正常運作，其上的檔案系統能將檔案自動在磁碟和磁帶系統之間轉移。

三、 光纖交換器

本系統使用 4 部 16 埠光纖交換器（Qlogic 5200）連結伺服主機、兩部磁碟陣列和光纖介面之磁帶機，以形成一個儲域網路環境，4 部光纖交換器彼此互相串連，磁碟陣列和主機均支援動態負載平衡，且具有 HA 備援功能，因此充分發揮儲域網路環境之高速頻寬與高可用性之功能。

四、 ACSLS 伺服主機

ACSLS 伺服主機為一部 Sun ultra 25 主機，其上安裝有 StorageTek 自動化卡匣系統磁帶櫃軟體（Automated Cartridge System Library Software，ACSLS），藉由單一中央控管、效率式管理磁帶櫃資源。其主要任務為提供主機對磁帶要求的服務，並監控磁帶館各設備的狀況，包括操控磁帶館機器手臂、磁帶館內之磁帶管理（auditing, reporting, event-log recording, querying）以及自動清潔磁帶機等。目前 ACSLS 授權數目（即可管理之磁帶數目總計為 16,000 捲）

五、 磁帶館

大量資料儲存系統包括一部 STK8500 和 2 部 STK 9310 磁帶館，後者兩館互相串連，磁帶可以在兩館之間自動交換。上述磁帶館系統支援 STK 9840、STK 9940、STK 9490 和 LTO 3 等各式磁帶機。STK 8500 最大可容納 7500 捲磁帶，STK 9310 則為 6000 捲磁帶，因此本局大量資料儲存系統最大可儲存容量高達 5,400 TB。

六、 磁帶機

本局大量資料儲存系統總計有 17 部磁帶機，其中包括 4 部 LTO-3 磁帶機、7 部 STK 9940 磁帶機和 6 部 STK 9840 磁帶機。由於不同磁帶機的效能、儲存容量等均不相同，因此可以依據不同使用者之使用行為來區分磁帶機的使用策略，以使整體效益達到最高。

肆、 結論與展望

本局大量資料儲存系統係因應本局數值天氣預報模式系統及其他資訊系統所產製之大量資料儲存所需而建置。本局大量資料儲存系統實際使用量在民國 93 年約 134 TB，94、95 年則劇升為 202 TB 與 329 TB，年增加率超過 60 %，顯見本局資料量呈現爆炸性的成長。本局大量資料儲存系統採用階層是檔案管理架構，由不同層次的硬體，包括伺服主機、磁碟陣列、磁帶館和各式磁帶機與磁帶媒體組合而成，軟體部分則涵蓋管理軟體和資料庫軟體。針對該系統，本局不惜投入大量經費，各個環節均採用最新的資訊科技，以期能維持整個系統的高效能和高可用性。在此同時，我們也隨時監控各式資源的使用分析，以量化提供系統擴充與採購的決策依據，以使充分達到最大的投資效益，並進而提升整體系統的可用性與安全性，以及據以提供局內更有效率、更安全之資料儲存服務。

（本文由中央氣象局氣象資訊中心洪景山 提供）