



加快研究開發的決策速度、強化基因組分析作為新藥研發基礎 實現靈活自由分析的基因組學資料大數據儲存

以成為「在癌症方面最先進國際藥劑研發公司」為願景的第一三共股份有限公司（以下稱為第一三共），正積極致力於基因組分析為主的資料分析研究。為了超越競爭對手，以及持續的推動研究開發，必須建置能儲存海量基因序列及分析資料，並迅速解析的分析基礎架構。為解決這個問題，第一三共採用 IBM Software Defined Storage (軟體定義儲存) 的分層型儲存系統，以及搭載 GPU 的 IBM Power Systems 的組合。這幾項解決方案將大大地開啟新藥研發的未來。

[導入產品介紹及服務]

- IBM Elastic Storage Server (ESS)
- IBM TS4500 磁帶庫
- IBM Power System S822LC for HPC (Minsky)



課題

- 因使用新一代排序器進行基因組分析，資料量急劇增加。進而使得儲存空間不足，也需要花費更長的時間進行資料分析

解決方案

- 導入高速儲存系統基礎架構「ESS」和磁帶庫「TS4500」建構的分層型儲存系統，以及為進行人工智慧導入與深度學習的高速伺服器基礎架構「Minsky」

效果

- 由於能夠在高效率兼具擴展性的資料分析系統基礎架構上進行高速處理，現在可從各種角度分析全基因組

【客戶課題】

每週需手動將資料轉傳到磁帶，以解決儲存空間緊張的問題

在新藥開發上，基因和蛋白質等基因組學資料分析的重要性正在迅速增長。第一三共目標成為「在癌症方面具有優勢的最先進國際藥劑研發公司」，也致力於分析基因組學資料。

企業集團中的第一三共 RD Novare 股份有限公司（以下，稱第一三共 RD Novare）在此研究開發中擔任重要角色。該公司運用以開發突破性醫藥為目的的技術基礎架構，致力建構出高品質的研發平台。

第一三共 RD Novare 的轉化研究部門部長菅野道裕表示，「今後最重要的是，專注於如何將基礎研究結果投入新藥研發的轉化研究 (Translational Research)，以及能闡明醫藥對患者影響的生物標記研究這兩個面向。在此我們需要運用電腦運算，將資訊用科學的方式呈現。」強調 IT 在當前新藥研發過程中的重要性。

這些研究活動的起點，是使用能快速讀取 DNA 核苷酸序列的排序器進行分析。第一三共 RD Novare 於 2014 年導入了新一代的排序器，並一直致力於基因組分析。但是，這個計畫也帶來了新的課題。同公司的轉化研究部門主任研究員小野祥正表示，「比起舊式排序器，新一代的排序器能一口氣讀取多 1 億倍的資料，如果繼續採用往常的資料儲存方法，以目前儲存系統的基礎架構來看，很明顯地很快就會達到極限。」

他更進一步提到，「以前都是把分析的工作委託給外面執行，但是隨著公司導入排序器後，變得可以每天在公司內進行分析的工作，使得排序器的工作量暴增，產生出的資料量也超出了我們原本的預期。」（小野）

在一次的分析當中，所生成的資料量約為 100GB，如果要進行處理的話，則需要 2 倍的容量。當初我們假設所需容量為 30TB，但是實際上每月都持續不斷增加 2TB，因此磁碟空間很快就不敷使用。

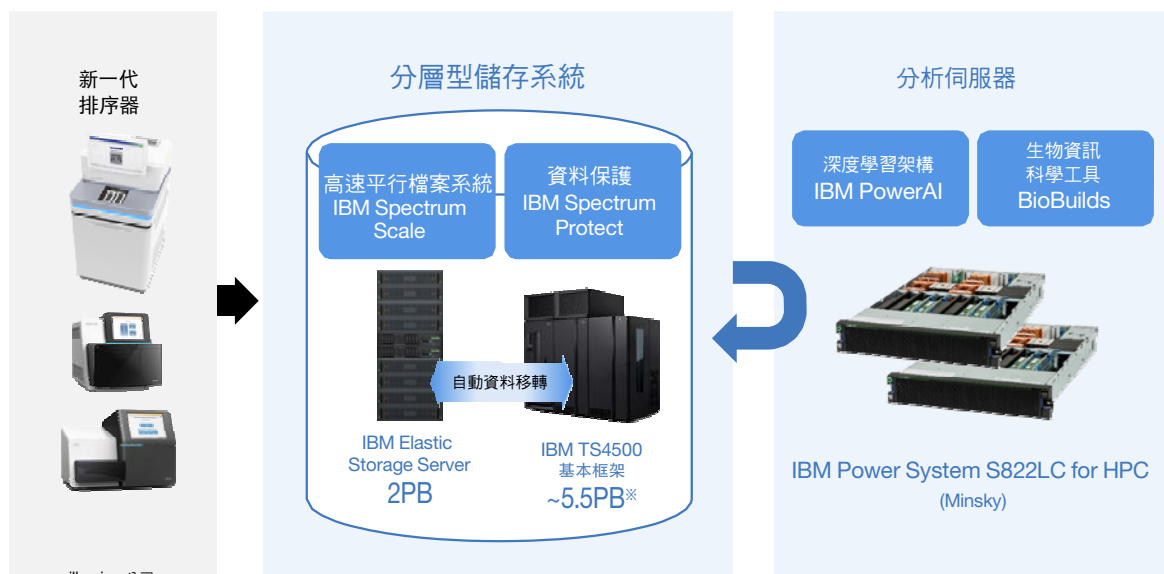
但是還沒有辦法可以應對這種磁碟空間不足的狀況，因此只要資料累積過多，我們都會用手動的方式歸檔到磁帶。

IBM 解決方案所提供的新型分析基礎架構，功能十分強大。我們期待，未來新藥研發的速度將不斷加速。



第一三共 RD Novare 股份有限公司
轉化研究部門
部長 菅野 道裕

新型基因組分析系統基礎架構的示意圖



* 每增設 1 個擴充框架，最多可以追加 10PB

【解決方案】

藉由使用分層型儲存系統和高速伺服器 構建高效能且高擴展性的分析系統基礎架構

針對這些課題，轉化研究部門立即與第一三共的 IT 企劃部門共享資訊，並著手檢討如何進行分析系統基礎架構的更新。「為了開發治療癌症的藥物，未來也會繼續增加樣品。該部門決定要優先解決分析系統基礎架構的問題，2017 年 6 月開始拜訪了幾家公司，向他們提出 RFP（委託提案書）」，第一三共 IT 企劃部門企劃小組負責人的中野暢也娓娓道來。

該公司整體 IT 環境是將「雲端優先」納入公司政策當中，但是由於基因組分析資料的容量龐大，對公司內部網路的影響甚鉅，以及考量提升研究速度，最後決定在公司的內部部署設定下使用。對於新型分析系統基礎架構，最重視的是資料傳輸的容量和速度，以及轉換到分析階段的便利性。

針對了第一三共的需求，幾家公司提出他們的提案，最後第一三共的選擇是 IBM 的分層型儲存系統「ESS」和磁帶庫「TS4500」，以及搭載 GPU 的高速伺服器基礎架構「Minsky」。中野表示，「在選擇合作夥伴時，我們不僅著重硬體，也將重點放在生命科學知識上。IBM 的解決方案，具有全面的高品質，還兼具卓越的擴充性，可以在公司內部部署環境和雲端環境之間，很透明地處理資訊。」

IBM Elastic Storage Server，是在 HPC 領域中已經發展 20 年以上的分散式檔案管理軟體 IBM Spectrum Scale（前身為 GPFS: General Parallel File System），實現更高速、靈活以及高擴展性的儲存系統基礎架構。這回系統導入的結構，擁有超過 2PB 的大容量，傳輸速度超過 23GB/s，實現了快速且大量汲取組學資料的需求。此外，在歸檔分析結果時，搭載企業磁帶驅動器的 IBM TS4500 磁帶庫之間，透明地執行並有計畫的自動分層管理，提供高效經濟的保存資料。

愛用者的小野十分重視系統的導入實績。「IBM 的解決方案已導入東京大學醫學科學研究所和許多其他海外基因組研究機關當中，這些實證案例證明了系統對高速處理大量資料的能力。實際上，我也認為這就是我們的唯一最佳選擇。」（小野）

IBM POWER 處理器和 NVIDIA GPU 藉由 NVLink 互連整合在 Minsky 內，不但可支援該公司的分析軟體外，還可以高速連接既有的分析伺服器，以及裝置能靈活運用 AI 的架構，這些成績都受到客戶的高度肯定。「雖然現在採用資料型系統，但未來可預期由 AI 來推動系統作業。我們也認為這方面在未來深具潛力。」中野如此描述。

新型分析系統基礎架構的建設進行得非常迅速。2017 年 9 月即開始進行專案導入，10 月上旬開始架設硬體，且所有分析資料都移轉至 ESS，11 月中旬時即已開始正式使用。從導入到正式啟用的這段期間僅花了 2 個月。「正因為是以日本 IBM 為中心團隊合作，所以才能如此順利導入系統」，中野如此回顧這次的專案。

【成果/未來展望】

專注於基因組分析的環境已經完備， 提升新藥研發的速度將指日可待

新型分析系統基礎架構的成果很快就展現出來。將資料存儲在 ESS 中，通過 Minsky 進行分析，分析速度最快可以提升至 10 倍，平均下來也可達到 3 倍到 4 倍之多。菅野表示，「IBM 基礎系統十分強大。我們期待未來新藥研發的研究速度能更加提升。」他如此肯定。

從原本苦無足夠空間的壓力中解放，也是此專案的一大成就。小野提到，「專案的第一步，即構建具有可擴展性系統來架構整個分析系統基礎架構，因此終於可以從資料容量限制的壓力當中解放。除了減輕心理跟工作負擔之外，還帶來具備卓越擴充能力的系統，協助我們嘗試新專案。」

藉由構建具有擴展性的系統來架構整個分析系統基礎架構，可以解決資料儲存空間不足的煩惱，此外，這個系統不僅具備擴充能力，還能協助我們嘗試新專案。



第一三共 RD Novare 股份有限公司
轉化研究部門
主任研究員 小野 祥正

IBM 的解決方案，具有全面高品質，還有其高擴展性，可以在公司內部部署環境和雲端環境之間，透明清楚地處理資訊。



第一三共股份有限公司
IT 企劃部門
企画小組
負責人 中野 暢也

由於癌症這種疾病的性質，新藥研發必須要使用整體基因組來進行全基因組分析，藉由新型分析系統基礎架構，現在可以自由地選擇分析方法，使用全部或是一部分來進行分析。「現在也可運用多個應用程序，我認為分析環境已規劃十分完備。」（中野）

中野也指出，「這次分析系統基礎架構藉由軟體版本的升級，未來也期望能作為驗證系統來驗證科學依據。」所有的紀錄檔 (log) 都保留在分層型儲存系統的龐大空間，所以能確保資料正確性，將來也可直接引用作為新藥申請和臨床報告的參考資料，功能十分強大。

「在這麼多的安全性保障下，活用此分析基礎架構變成很重要的任務。升級軟體版本強化安全性的同時，我們也計畫提升員工的 IT 素養，辦理一系列的教育訓練。」（中野）

另外，在未來，我們還將藉由擴大此分析基礎架構，來擴張更多研究開發的可能性。中野表示，「今後希望能引進外部知識來改善公司的使用環境，繼續引發出新型分析系統基礎架構的優勢。對於研發通用量子電腦保持領導地位的 IBM，十分期待今後的發展。」

為了讓公司成為「在癌症方面具有優勢的最先進國際藥劑研發公司」，架構出基因組分析基礎的新藥開發體系十分關鍵。本專案的新型分析基礎架構，將能對該公司的 2025 年願景提供重大貢獻。



Daiichi-Sankyo

第一三共股份有限公司

〒103-8426 東京都中央區日本橋本町三丁目 5 番 1 號

<https://www.daiichisankyo.co.jp/>

該公司由三共與第一製藥營運合併，於 2005 年成立。企業理念為「致力於持續創造革新的藥品，並且提供能滿足多樣化醫療需求的藥品，為世人的健康、豐富其生活作出貢獻。」2025 年的願景為「在癌症方面具有優勢的最先進國際藥劑研發公司」，透過引進先進技術加快研發速度，致力於不斷創造革新性藥品。



©Copyright IBM Japan, Ltd. 2018

〒103-8510 東京都中央區日本橋箱崎町 19-21

本目錄中的資訊截至 2018 年 5 月為止。規格可能在沒有預告的情況下進行變更。所記載的案例內容與特定客戶相關，並不代表在所有情況下都能獲得同樣成效。效果會因客戶的使用環境以及其他因素而有不同。有關產品服務等更多資訊，請聯繫我們或事業合作夥伴的銷售代表。IBM、IBMLOGO、ibm.com、GPFS、IBM Spectrum Protect、IBM Spectrum Scale、Power、POWER 以及 Power Systems，是 International Business Machines Corp. 在世界各國的註冊商標。其他產品名及服務名稱等，可能是 IBM 或是各自其他公司的商標。關於目前 IBM 商標的完整清單，請前往 www.ibm.com/legal/copytrade.shtml 參閱。