

快速回應用戶需求 主動創造商機與營收

搭載Intel® Xeon®處理器的HP ProLiant伺服器讓企業發揮巨量資料分析最大效益

聯合報系(UDN)透過搭載Intel® Xeon® E5/E7處理器之HP ProLiant系列伺服器的導入，讓該公司IT團隊得以善用並擁抱當前最新、最快的巨量資料分析技術

聯合報系為了對旗下電子商務、聯合新聞網、售票網與文創平台等系統日誌、活動狀況與客戶行為進行深入分析，進而開發出更符合消費者與市場需求之精準商品、服務與內容的開發、製作、銷售及推薦，該集團因而決定導入巨量資料(Big Data)技術與分析平台，在搭載效能出眾Intel® Xeon® E5/E7處理器之HP ProLiant系列伺服器的穩定運作基礎下，該公司能夠心無旁騖不斷採用更新、更快的分散式架構巨量資料分析技術，加快市場需求回應速度，並提升企業整體競爭力。

打造穩定Data Lake巨量分析架構

聯合報系一開始採用甲骨文(Oracle)資料庫、SPSS資料探礦、ETL資料轉換工具等傳統BI工具建置資料倉儲(Data Warehouse, DW)，再結合運用巨量資料技術之Hadoop推薦引擎，對來自電子商務「UDN買東西購物網中心」、聯合新聞網、「UDN售票網」與「有設計uDesign」文創設計電子商務平台等網站蒐集的日誌進行分析與推薦。

聯合報系電子商務部資訊研發中心總監黃俊賢指出：「透過Hadoop推薦引擎，進行客戶瀏覽行為、購物車採購行為的紀錄與分析，然後再根據不同商品、服務及內容項目，對個別使用者提供專屬的精準推薦。除此之外，未來會更進一步地將EDM發送系統中客戶開信與點擊紀錄、客服系統之話務客服紀錄、APP瀏覽行為等資料一併納入數據分析框架中，以建構更全面、更完整的營運視野。」

但上述架構卻出現兩處效能瓶頸，他進一步表示：「我們嘗試將Hadoop資料倒進Oracle資料庫中，進行多面向的客戶行為分析，卻發生因資料量體過大導致系統無法運作的情形，同時也在SPSS建模部分出現效能瓶頸，如此一來，我們完全無法做更深度的跨產業、跨載具的客戶行為分析。」為了改善傳統資料倉儲因巨量資料所產生的效能問題，該公司決定在搭載Intel® Xeon® E5/E7處理器之HP ProLiant系列伺服器的穩固基礎與強大I/O效能基礎下，全面朝向MPP巨量資料分析架構的目標邁進。

系統穩定性高於一切

為了解決效能瓶頸問題，聯合報系耗費3個月的

時間針對市面上支援大規模平行處理(MPP)架構的巨量資料分析引擎方案進行評估測試，最終選擇同時具備欄位導向(Column-oriented)與列導向(Row-oriented)儲存能力並支援強大統計分析、資料探勘與平行運算能力之開放源碼R語言的EDW方案，而成功解決上述兩個效能瓶頸的問題。

一開始為了快速擁抱巨量資料技術，並展現其強大效益，所以聯合報系一開始選擇能夠快速又方便部署的設備型Hadoop推薦引擎平台。「隨著該平台的確發揮受到老闆肯定的推薦效益之後，為了讓基礎架構更具彈性調動能力並發揮更大效益，如今Hadoop平台全面採用搭載效能與穩定性兼具之Intel® Xeon® E5/E7處理器的HP ProLiant系列伺服器，同時每台並內建兼具成本效益與I/O效能的1-2TB SATA硬碟與96GB記憶體。」黃俊賢表示。

「系統穩定性高於一切，我們不會拿不確定的產品去開玩笑，」黃俊賢斬釘截鐵地指出：「這也是我們為何選擇搭載Intel® Xeon® 處理器的伺服器之原因，Intel長年居於市場領導地位，以及在市占率、效能及穩定性上的市場口碑與表現，更堅定我們採用的決心。」

在巨量資料系統運作過程中難免會出現一些效能瓶頸狀況，對此，黃俊賢表示，所幸經銷商大綜電腦系統能不時提供許多可以解決問題的新技術或新產品訊息，再加上隨傳隨到的服務品質，讓問題能夠立即獲得圓滿的解決。此外，在巨量資料分析引擎的導入過程中，經銷商也能提供許多便利測試的足夠資訊與數據，讓新產品與技術的測試更加深入、順暢。

為提升巨量資料分析戰力奠立基礎

儘管一開始Hadoop推薦引擎的資料無法倒進原有的資料倉儲中進行處理與分析，但在Hadoop推薦引擎的加持下，兩年來「UDN買東西購物網中心」的轉換率提升了50%，而且客戶停留時間也增加了50%，對於成交機率的提升有很大的幫助。黃俊賢強調指出，正因為Hadoop推薦平台成功為公司創造營收，因而能說服老闆更進一步地導入巨量資料分析技術，以打造能進行深度分析、建模與預測的Data Lake分析架構。



聯合報系(UDN)電子商務部資訊研發中心總監 黃俊賢

對於聯合報系而言，推薦與分析兩大巨量資料系統的穩定性非常重要，令人讚賞的是，隨著搭載Intel® Xeon® E5/E7處理器之HP ProLiant系列伺服器的導入，在長達四年24小時全天候運轉下的穩定性與效能表現亮麗，讓IT人員可以無後顧之憂地將心思全放在能發揮關鍵效益的巨量資料系統上。值得一提的是，HP ProLiant系列伺服器零組件同時具備強大互通性，因而能發揮更具彈性的共用效益。更重要的是，內含Intel® Xeon®處理器的伺服器具備強大效能與可擴充性，因而能有效因應不斷快速成長的巨大資料量。

在搭載Intel® Xeon®處理器的伺服器兼具穩定、效能與可擴充性的堅實基礎下，聯合報系能夠儘快完成EDW分析引擎的全面導入，同時也能不斷尋求更新、更快的巨量資料技術，並開發更有價值的應用。例如該公司最近想推行讓在移動過程中透過手持裝置瀏覽商品的使用者，能在回家後的PC上繼續瀏覽並且成交的跨載具推薦系統。

「巨量資料可說是當前投資報酬率最高的關鍵應用，更是今後企業存活的一大關鍵，在推薦效益的初步發揮與基礎運算平台的穩定無虞下，讓我們可以為下一步全面提升聯合報系IT人員巨量資料分析戰力而努力。」黃俊賢總結表示。

更多資訊請洽大綜電腦02-55821008 #225

蕭世偉先生

