



利用 Dell EMC AI 產品組合 應對 AI 工作負載的挑戰

Dell AI 產品組合與 HPE 類似產品的 比較

毫無疑問，人工智慧 (AI) 已經改變了商業局面，使各種規模的行業和組織能夠從資料中獲得更深入的見解，實現業務流程自動化，提供量身定製的客戶和使用者體驗，並提升行業競爭力。為了有效利用 AI 的力量，組織需要的基礎結構供應商必須能夠提供涵蓋整個 AI 生命週期的全方位整合式解決方案組合。

諸如 Dell Technologies 和 Hewlett Packard Enterprise (HPE) 等基礎結構供應商能幫助客戶滿足不斷增長的 AI 需求並駕馭其固有的複雜性。這些供應商提供了各種支援 AI 的產品組合，以及不同等級的 AI 解決方案，結合了高效能內部部署和雲端基礎結構解決方案，與策略性合作關係和一系列支援與諮詢服務。

本報告檢視有關 Dell 和 HPE AI 產品組合*的公開資訊，旨在歸納出客戶選擇 Dell Technologies 來滿足其 AI 需求，而可能在架構、效能和支援等方面享有的具體優勢。我們比較了 Dell 為支援 AI 部署而設計的伺服器細節，並參考了 ML Commons® 的產業基準測試結果。我們還探索了其他軟體和服務方案，這些產品可在客戶的 AI 之旅的每個階段為其提供支援。

*注意：PT 的所有研究完成於 2023 年 12 月 5 日或之前，因此本文不會反映 Dell 或 HPE 在該日期之後發布的產品或變更。

採用 AI 的挑戰

採用 AI 策略為資料中心及其 IT 人員帶來了許多新挑戰，這些挑戰包括：

- 透過內部 AI 培訓或外部招聘來解決現有員工的技能落差。
- 瞭解 AI 的資料準備需求，包括業務資料的品質、量、位置和目前狀態。
- 評估具體的業務 AI 目標，以確定哪種 AI 模型和實施方式能帶來益處。
- 評估規劃的 AI 系統在運算、網路和儲存方面的需求，並制定採購計畫。

這些只是公司的資料中心在導入 AI 並希望從中獲益的路上，即將面臨的許多重大障礙中的幾個例子。

Dell AI 產品組合能透過專業和諮詢服務來幫助客戶應對這些挑戰，協助客戶建構實施路線圖並為 AI 模型準備資料。¹ 這些產品組合還包含涵蓋機器學習 (ML) 概念和其他教育主題的培訓課程，並提供經過驗證的 AI 設計，有助確保實施成功。² 此外，Dell 還與第三方合作，為客戶提供其他 AI 工具，例如 Hugging Face 社群中的自訂 Dell 入口網站，其中包含可用於部署開源 AI 模型³及 Meta Llama 2 大型語言模型 (LLM) 的專用容器和指令碼。⁴ 除了豐富的運算和 PC 產品 (從行動工作站到支援多達 8 個高端 NVIDIA GPU 的伺服器) 外，Dell EMC 還提供一系列的高效能檔案和物件儲存陣列產品，能滿足 AI 領域的非結構化資料儲存需求。這些儲存產品 (包括 Dell PowerScale、ObjectScale、ECS 和板載儲存) 可以處理 AI 工作負載經常使用的非結構化資料。⁵ Dell 還與 Snowflake 合作，為 Dell 客戶提供混合雲儲存解決方案。⁶ 根據 Dell 的分析，截至 2023 年 8 月，他們提供了「最廣泛的生成式 AI 產品組合」，這些產品組合能夠在整個 AI 實施過程中提供資源，功能超越了傳統定義的伺服器和儲存裝置。⁷

AI 效能和加速運算選項：Dell 與 HPE 的比較

AI 工作負載可以根據其大小或類型，使用 CPU 和/或 GPU 做為運算資源。一些 CPU 提供 AI 專用加速器，例如最新系列 Intel Xeon 可擴充處理器中的 Intel Advanced Matrix Extensions (Intel AMX)。⁸ GPU 通常更適合更大和/或要求更高的工作負載，但 GPU 外形尺寸會影響效能水準。例如，部分 NVIDIA A100 和 H100 型號的 GPU 採用通用 PCIe 或專有 SXM 外形尺寸，後者使用更高效能的 NVIDIA SXM 架構。⁹ 大記憶體容量和伺服器設計功能 (如冷卻架構和電源效率) 也會影響效能。大多數資料中心仍然使用氣冷技術，這表示高效能運算 (HPC) 工作負載需要能夠盡可能有效使用氣冷技術的伺服器。下面，我們將重點介紹 PowerEdge 伺服器產品在元件、冷卻選項等方面的特點，以及其發布的 MLCommons® MLPerf® 分數。

AI 模型指標效能：MLPerf 結果比較

MLPerf® 是一個用於測試 AI 訓練和推論效能的基準套件。要發布 MLPerf® 正式結果的組織，其結果必須符合基準開發方 MLCommons® 設定的具體條件。¹⁰ 這些條件準則提供了效能比較的便利標準。MLPerf® 使用 Datacenter、Edge、Mobile 和 Tiny 資料集進行推論測試，並在報告中詳列 AI 分數和測試期間的耗電量。推論基準套件包含對許多常見 AI、ML 和 DL 模型的測試；請參閱表 1。

表 1：MLPerf® 測試中包含的 AI、ML 和 DL 模型，以及每個模型的典型使用案例。資料來源：Principled Technologies。

常見的 AI 模型	典型使用案例
ResNet	一種圖像分類模型，可幫助電腦學習、記住和識別不同的圖像，可應用於醫學成像、社交媒體內容審查和臉部識別等使用案例
RetinaNet	與 ResNet 相比，是一種可以處理更複雜任務的物件偵測模型。此模型可協助電腦識別和定位圖像或影像畫格中的物件，並可以按重要性對物件進行分類。可應用於自動駕駛、車輛自動輔助技術、監控、臉部辨識等領域
3D-UNet	專門用於醫學影像分割
RNN-T	可應用於自動語言翻譯等使用案例的語音辨識模型
BERT	可應用於文字摘要、語言翻譯和任務自動完成等使用案例的自然語言處理
DLRM-v2-99.9	可應用於定向廣告和個人化產品推薦等使用案例的推薦模型
GPTJ-99 與 99.9	自然語言處理的大型語言模式，可用於聊天機器人和聊天式 AI 工具等使用案例中的文字生成，並且表現優異

MLPerf

除了 AI 模型本身的表現外，MLPerf® 結果還包含多項參數，便於在單個圖表或表格中解析大量資料。以下是這些參數的簡單介紹：

- 99.0 和 99.9：這些數值是指模型訓練的準確率。您需要的輸出越準確，模型就越複雜，處理資料所需的時間就越長。
- 離線樣本/秒：基準在測試開始時傳送所有查詢以模擬系統上現有資料的模式。
- 伺服器查詢/秒：基準在整個測試時間持續傳送查詢以模擬分析即時資料串流的模式。

有關 MLCommons® 和 MLPerf® 結果的更多資訊，請參閱 <https://mlcommons.org/benchmarks/inference-datacenter/>。



本報告中的結果來自 2023 年 11 月在 MLCommons® 網站上發布的 MLPerf® v3.1 推論資料中心結果。¹¹ 這些結果包含由技術製造商和雲端服務供應商提交、涵蓋各種組態的測試結果。與 HPE 提交的公開結果相比，Dell 伺服器在某些 AI 模型中的效能表現較佳。(注意：伺服器的不同 GPU 組態可能會增加直接比較的困難度。)詳情請參閱表 2。

表 1：截至 2023 年 11 月 29 日發布的 MLCommons® MLPerf® 3.1 結果中所包含的 Dell 和 HPE 伺服器。
資料來源：Principled Technologies。

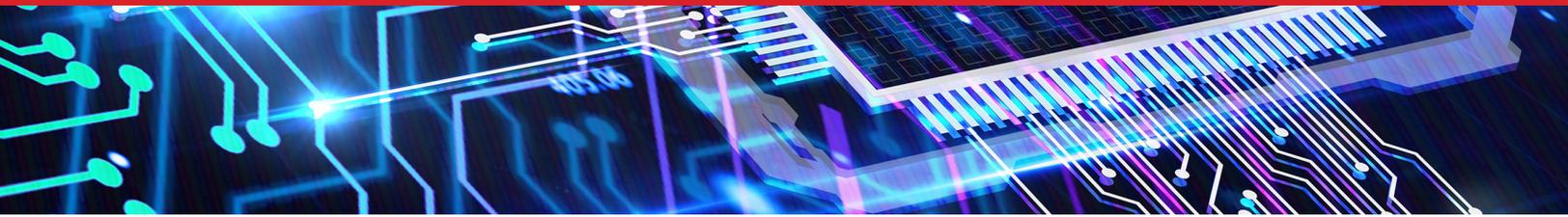
提交者	伺服器型號	GPU 型號數量	說明
Dell ¹²	PowerEdge XE9680	8 個 NVIDIA H100 SXM	適用於大型工作負載 (如大型語言模型) 的 AI 訓練和推論
	PowerEdge XE9640	4 個 NVIDIA H100 SXM	適合訓練高密度和液冷資料中心的大型 AI 模型
	PowerEdge XE8640	4 個 NVIDIA H100 SXM	適用於氣冷資料中心 4U 規格裝置中的傳統 AI 訓練、HPC 和資料分析應用程式
	PowerEdge R760xa	4 個 NVIDIA H100 PCIe	適用於不需要頂級效能 GPU 的各種高運算工作負載，包括 AI-ML/DL 訓練和推論
HPE ^{13,14}	ProLiant XL675d Gen10 Plus	8 個 NVIDIA A100 SXM	適用於高效能運算和 AI
	ProLiant DL380a Gen11	4 個 NVIDIA H100 PCIe	適用於中等運算 AI 工作負載的 2U 伺服器

Dell 和 HPE 伺服器的直接比較

雖然一個全方位的 AI 策略要考量的不只是硬體需求，但強大的硬體效能是 AI 工作負載流暢運行的關鍵要素之一。隨著更新型 GPU 和其他技術的出現，AI 工作負載能力也在不斷增強。在首次公布的 MLPerf® v3.1 結果中，效能最佳的 NVIDIA GPU 是 H100 Tensor Core，Dell 在其幾款 PCIe 和 SXM5 規格的伺服器中搭載此款 GPU 並發布了 MLPerf® 結果。¹⁵ HPE 發布的結果僅列出一個 H100 結果，且規格僅含 PCIe。我們的研究顯示，僅有少數幾款搭載 GPU 的 HPE 伺服器可支援以 SXM5 H100 尺寸規格使 NVIDIA GPU 發揮最佳效能，且沒有任何一款 HPE ProLiant 伺服器能做到。¹⁶ 正如下方所示，GPU 性能越佳，AI 工作負載的效能越高。

8 個 GPU MLPerf 結果

Dell PowerEdge XE9680 可支援多達 8 個 NVIDIA H100 SXM5 GPU (用於 AI 加速) 和多達 2 個第 4 代 Intel® Xeon® 可擴充處理器。PowerEdge XE 產品系列採用模組化架構，可支援 SXM4 或 SXM5 NVIDIA GPU 以及 AMD 的開放運算專案加速器模組 (OAM) GPU 組件，與標準 PCIe GPU 相比，效能更高。¹⁷ PowerEdge XE9680 是一款搭載 8 路 NVIDIA H100 SXM5 的緊湊型伺服器，僅佔用 6U 機架空間。最新的 HPE ProLiant Gen11 伺服器目前不支援 H100 SXM 尺寸規格，¹⁸ 目前僅有幾款 HPE Cray 超級運算伺服器能支援。¹⁹ 由於 HPE 沒有提交任何 Cray 伺服器的 MLPerf® 結果，並且僅在其 AI 產品組合頁面上列出其 ProLiant 伺服器，因此本文的重點將放在 ProLiant 伺服器上。(請參閱圖 1.)



Featured AI products and services

PRODUCT

HPE Ezmeral Unified Analytics Software

Unlock data and insights faster by helping you develop and deploy data and analytic workloads. Provides fully managed, secure, enterprise-grade versions of the most popular open-source frameworks with a consistent SaaS experience.

[Explore more →](#)

PRODUCT

HPE Machine Learning Development Environment

Uncover hidden insights from your data by helping engineers and data scientists collaborate, build more accurate ML models and train them faster.

[Explore more →](#)

PRODUCT

HPE Machine Learning Data Management Software

Uncover hidden insights with a data pipelining and versioning solution that automates data pipelines and accelerates time to ML model production by processing petabyte-scale workloads.

[Explore more →](#)

PRODUCT

HPE ProLiant Servers

Speed time to value with systems that are optimized for computer vision inference, generative visual AI, and end-to-end natural language processing.

[Explore more →](#)

1 <https://www.hpe.com/us/en/solutions/ai-artificial-intelligence.html> 上的重點 AI 產品和服務螢幕擷取畫面列出了截至 2023 年 12 月 5 日的 HPE ProLiant 伺服器。

在 2023 年 11 月首次發布的 8 GPU 伺服器 MLPerf® v3.1 結果中，配備 NVIDIA SXM5 H100 GPU 的 Dell PowerEdge XE9680 效能比配備 NVIDIA SXM4 A100 GPU 的 HPE ProLiant XL675d Gen10 Plus 高出 4.25 倍 (見圖 2)。

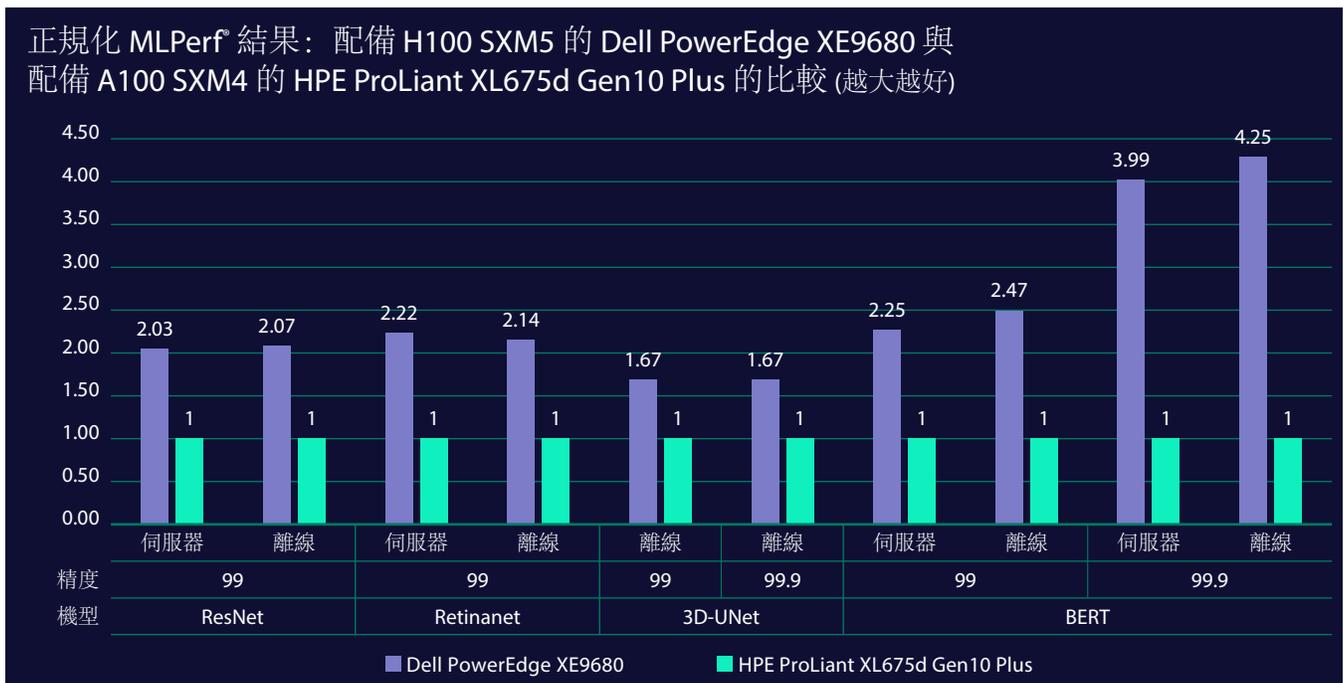


圖 2：已公布的 Dell PowerEdge XE9680 和 HPE ProLiant XL675d Gen10 Plus 的 MLPerf® 測試結果 (截至 2023 年 11 月 29 日的結果)。Dell 系統使用 NVIDIA H100 GPU，而 HPE 系統使用老一代的 GPU。資料來源：Principled Technologies 使用來自 MLCcommons® 的資料。20、21

為了便於比較，我們對圖 2 到圖 5 中的測試結果進行了標準化。也就是說，我們將數值 1 指派給每個 HPE ProLiant DL380a Gen 11 結果，並顯示與之相關的相應 Dell PowerEdge R760xa 結果。正如這些結果所示，即使 GPU 型號只有一代差異，也會對您在眾多 AI 工作負載中看到的效能產生重大影響。

4 個 GPU MLPerf 結果

若將節省資料中心耗電或空間列為關鍵考量要素時，2U Dell PowerEdge XE9640 可以提供解方。PowerEdge XE9640 配備多達 4 個 NVIDIA H100 SXM GPU，GPU 運算能力為 XE9680 的一半，而空間卻減少了三分之二。²² 密集封裝的 Dell PowerEdge XE9640 採用 Dell Smart Cooling 技術，提供一系列散熱技術，包括為 CPU 和 GPU 配備直接液體冷卻技術。²³

PowerEdge XE9640 的 2U 機箱改良了氣流設計，配備更大的風扇和散熱器，有助於冷卻如 PCIe 卡和記憶體等其他重要元件。²⁴ PowerEdge XE9640 是目前 Dell 或 HPE 當中唯一一款配備 2U 四路 HGX H100 GPU 的產品。HPE AI 產品組合提供 1U 和 2U Gen11 ProLiant 伺服器，但僅限於 PCIe 規格的 GPU。²⁵

Dell PowerEdge XE9640 伺服器還支援 Intel Max GPU 系列 1550 OAM GPU，這些 GPU 提供低功耗、高密度的 GPU 選項，包括 PCIe 卡和 OpenCompute 加速器模組 (OAM)。²⁶ 雖然我們無法確定截至 2023 年 12 月 5 日，HPE 是否提供了配備這些 GPU 的伺服器，但他們確實提供了配備 Intel 資料中心 GPU Max 380 GPU 的 HPE ProLiant DL380 Gen11 和 DL11a Gen1100。²⁷ 這表示 Dell PowerEdge XE9640 可能是目前唯一一款在 2U 伺服器中配備四個 Intel Max 1550 OAM GPU 的產品。對於重視資料中心空間和電源效率的公司，配備四個 Intel Max 1550 GPU 的 2U 伺服器提供了一種解決方案，可以在不犧牲資料中心空間的情況下同時享有高效能運算和能源效率。

在已公布的 MLPerf[®] 3.1 結果中，配備四個 HGX H100 GPU 的 Dell PowerEdge XE9640，效能比配備四個 PCIe H100 GPU 的 HPE ProLiant DL380a 高出 1.99 倍 (見圖 3)。

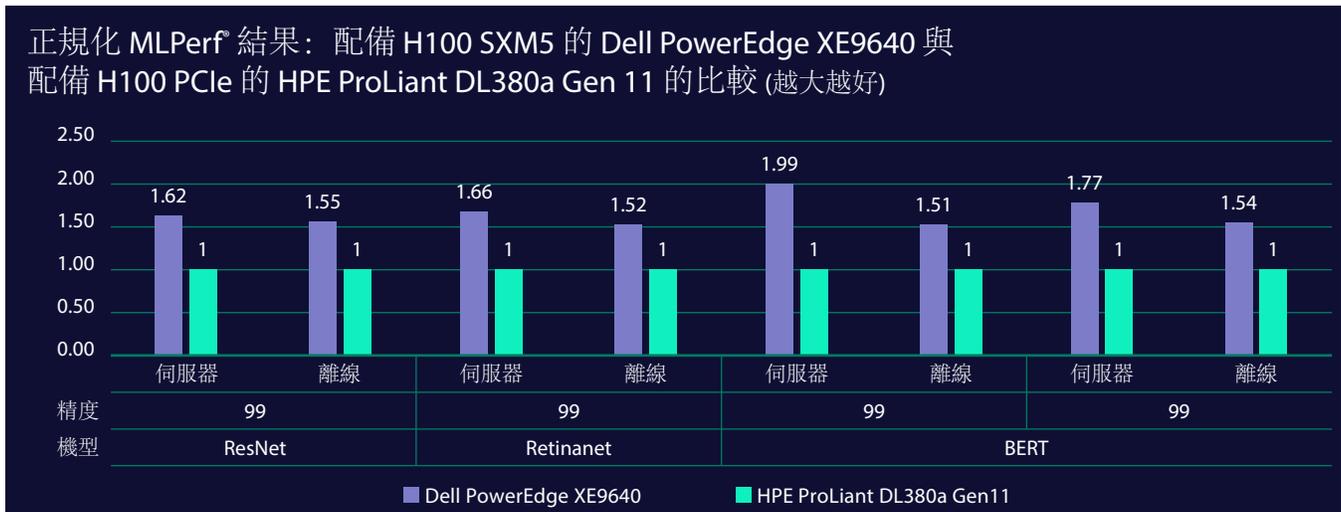


圖 3：已公布的 Dell PowerEdge XE9680 和 HPE ProLiant XL675d Gen10 Plus 的 MLPerf[®] 測試結果 (截至 2023 年 11 月 29 日的結果)。Dell 系統使用 NVIDIA H100 GPU，而 HPE 系統使用老一代的 GPU。資料來源：Principled Technologies 使用來自 MLCommons[®] 的資料。^{28,29}

PowerEdge XE8640 提供四路 GPU 組態，為處理器提供氣冷功能，為 GPU 提供液體輔助空氣冷卻散熱器，無需將設施供水管連接到機架。對於不使用或不能使用外部冷卻技術的公司，³⁰4U Dell PowerEdge XE8640 支援四個 NVIDIA H100 SXM5 GPU，可提供與 PowerEdge XE9640 相等的運算能力，而且無需使用直接液體冷卻技術。³¹

Dell PowerEdge XE8640 採用最新的第 4 代 Intel Xeon 可擴充處理器和高達 4 TB 的記憶體，³²能夠處理 AI 和資料分析中常見的大型資料集和複雜運算。HPE 確實在 HPE Cray 系統中提供了 NVIDIA H100 SXM5 GPU，但 HPE ProLiant GPU 的伺服器則並不支援。

與 2023 年 11 月公布的 MLPerf[®] 資料相比，配備四個 NVIDIA H100 SXM5 GPU 的 PowerEdge XE8640 伺服器在 9 個不同類別的所有 4 GPU 中，AI 輸送量表現最佳。如圖 4 所示，與 HPE ProLiant DL380a 伺服器相比，其得分高出 2.07 倍。

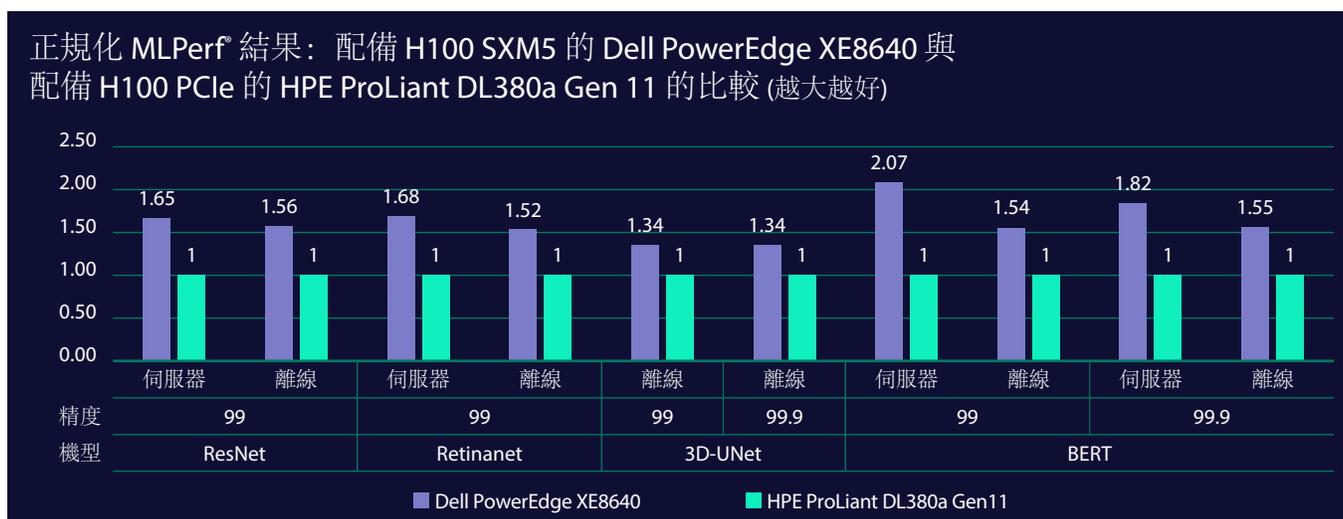


圖 4：已公布的 Dell PowerEdge XE8640 和 HPE ProLiant DL380a Gen11 的 MLPerf[®] 測試結果 (截至 2023 年 11 月 29 日的結果)。Dell 系統使用 NVIDIA H100 SXM 規格，而 HPE 系統使用效能較低的 PCIe 規格。資料來源：Principled Technologies 使用來自 MLCommons[®] 的資料。³³、³⁴

最後，對於希望從小規模著手，然後根據需要擴充的組織，2U Dell PowerEdge R760xa 伺服器可容納 NVIDIA、AMD 和 Intel 的一系列 GPU，並可支援多達 4 個雙寬度 PCIe Gen 5 GPU 或 12 個單寬度 PCIe GPU。³⁵ 此款伺服器配備 32 個 DIMM 插槽，一個 8 槽 2.5 吋磁碟機箱和 12 個 PCIe 插槽，提供可擴充的儲存空間，可隨著 AI 資料需求的增加而擴充，並可支援多達 12 個單寬度 PCIe GPU 或 4 個雙寬度 PCIe GPU，如 NVIDIA H100 或 L40S。³⁶ 如此的可擴充能力代表伺服器能適應不斷發展的 AI 任務，從機器學習模型訓練到高階資料處理都不成問題。

PowerEdge R760xa 的氣冷系統支援高密度運算環境，可容納熱設計功耗 (TDP) 高達 350W 的加速器，³⁷因此能夠在密集運算負載下保持優異效能。在截至 2023 年 11 月使用「伺服器」模式公布的 MLPerf[®] ResNet、RetinaNet 和 BERT 伺服器測試結果中，配備四個 NVIDIA H100 PCIe GPU 的 PowerEdge R760xa，效能優於同樣配備四個 H100 PCIe GPU 的 HPE ProLiant DL380a Gen 11 (見圖 5)。

正規化 MLPerf® 結果：配備 H100 PCIe 的 Dell PowerEdge R760xa 與 配備 H100 PCIe 的 HPE ProLiant DL380a Gen 11 的比較 (越大越好)

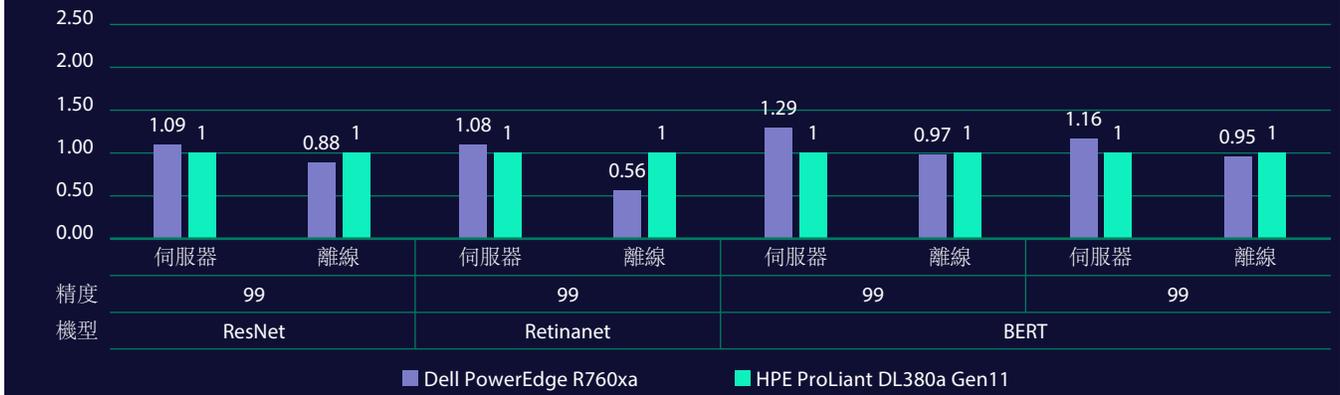


圖 5：已公布的 Dell PowerEdge R760xa 和 HPE ProLiant DL380a Gen11 的 MLPerf® 測試結果 (截至 2023 年 11 月 29 日的結果)。兩個系統都使用 NVIDIA H100 GPU 的 PCIe 規格。資料來源：Principled Technologies 使用來自 MLCommons® 的資料。^{38,39}

總而言之，MLPerf® 結果表明，伺服器元件的效能差異很大，因此選擇正確的選項來支援您的工作負載和效能需求至關重要。可支援 AI 工作負載的 Dell PowerEdge 伺服器提供了多種冷卻和密度選項，能滿足公司的各種資料中心需求，同時提供強大的 MLPerf® 效能。

Dell AI 產品組合的更詳細介紹

雖然運算效能至關重要，但運算效能只是規劃 AI 工作負載時的一個考慮因素。在開始導入 AI 時，您還必須考慮供應商的 AI 產品組合中可提供的其他產品。下面我們將討論這些 AI 產品組合的其他重要考量因素，包括用戶端工作站、雲端原生產品和儲存裝置等等。我們還重點列出了與 HPE 相比，Dell 產品可能具有優勢的領域。

工作站

Dell Precision Data Science 工作站為 AI 開發人員和資料科學家提供了 NVIDIA RTX® GPU 和 Intel Xeon® CPU 以及一整套資料科學工具。⁴⁰ 這些系統能運用多種專業級運算選項，包括具有 100 多種專業應用程式認證的 NVIDIA GPU⁴¹，以及 Intel Xeon 可擴充處理器加速器 (如 Intel DL Boost)。⁴² Precision 工作站提供行動、立式和機架式三種外型規格，可滿足從大型動態資料分析到行動科學現場建模的各種需求。

HPE 工作站產品範圍較窄，主要產品為配備 NVIDIA L4 的單一立式工作站；HPE 未提供行動工作站選項。⁴³ 雖然足以完成許多任務，但 HPE 立式工作站的靈活性和工作負載覆蓋範圍均無法與陣容更豐富的 Dell 媲美。Dell Precision 工作站提供眾多尺寸和便攜性選項，可為客戶量身定製解決方案，滿足實驗室、辦公室和現場作業等環境的不同需求。

儲存裝置

談到執行 AI 工作負載，儲存能力可能與運算能力一樣重要。較多的資料可以提高 AI 模型的準確性，但巨量資料集的儲存和管理作業，會對許多資料中心的能力構成挑戰。此外，由於模型通常使用非結構化資料進行訓練，因此支援 AI 的儲存系統必須能夠輕鬆處理許多不同類型的資料。⁴⁴ 為了提供 AI、ML 和 DL 資料集的容量和擴充能力，Dell EMC 提供了 PowerScale[®] 儲存產品系列，和適合用於儲存物件的 Elastic Cloud Storage (ECS) 或軟體定義版的 ObjectScale。

PowerScale 全快閃記憶體 NAS 產品組合提供每個節點 3.84 TB 到 720 TB 原始儲存容量的選項，叢集全快閃記憶體容量可達 186 PB 原始儲存容量。PowerScale 的靈活性和擴充能力可支援各種客戶和 AI 使用案例。⁴⁵ 建立叢集後，PowerScale F900 的總原始儲存容量最高可達 186PB。⁴⁶ 三種全快閃記憶體 PowerScale 型號 (F200、F600 和 F900) 都具有內嵌資料壓縮和重複資料刪除功能，以提高儲存效率。⁴⁷ 這裡的每個 PowerScale 儲存型號都使用 Dell OneFS[™] 檔案系統，該系統運用原則將儲存進行分層，最重要的資料優先排列在最快速的儲存層中，以達成工作負載最佳化。⁴⁸ Dell 還在 AWS Marketplace 中提供 OneFS 軟體 APEX File Storage for AWS。客戶可將 OneFS 用於 AWS 運算例項，透過使用與內部部署 OneFS 陣列相同的功能，提供一致的使用者體驗。⁴⁹ 雖然 HPE 提供公有雲整合以達成混合儲存解決方案，但我們在其產品中沒有找到像 Dell APEX File Storage for AWS 這樣的雲端原生選項。

Dell 的物件儲存選項包括 Dell Enterprise Object Storage (ECS)，這是「專門打造來以公有雲規模儲存非結構化資料」的。⁵⁰ 除了內建可與 Amazon S3 物件儲存技術相容的特性而可用於混合雲外，ECS 儲存節點還可提供每個機架高達 14PB 的容量。⁵¹ HPE 對於非結構化儲存還提供檔案和物件儲存選項，只不過其物件儲存產品是透過與 Scality 合作的方式提供。客戶可以向 HPE 購買適用於 Scality 的 HPE 解決方案。⁵²

專業服務

Dell EMC 提供廣泛的專業服務，包括諮詢、資料準備、部署、支援和受管理服務等，以支援 AI 部署。對於尋求經驗證的架構與解決方案的組織，Dell 針對特定的使用案例提供 Validated Design for AI，讓組織在設計和部署 AI 資源時不需要猜測。這些經過 Dell 驗證的 AI 解決方案包含硬體和軟體套件、對話式 AI 模型和機器學習操作等。透過將預先設定的專用解決方案與 AI 相關服務結合，Dell EMC 提供了涵蓋各種 AI 需求的全方位 AI 解決方案。與臨時拼湊的解決方案相比，這些產品可讓客戶更快速順利的邁向 AI 成功之路。

從諮詢到實作，在您的 AI 之旅中，Dell 服務還可以全程提供指引。Dell ProConsult Advisory 服務可協助客戶確定使用者能透過採用 GenAI 流程獲得哪些益處，並協助客戶制定涵蓋所需解決方案和 IT 技能的完整路線圖。Dell 服務可以為大型語言模型整合準備資料，並對 IT 團隊進行 AI 知識培訓。為了全面採用 GenAI，Dell 團隊會審視您的使用案例，確定最適合您需求的 AI 模型，並完成相關部署和設定作業。HPE 也提供可支援公司 AI 技術導入的專業服務。⁵³ ⁵⁴

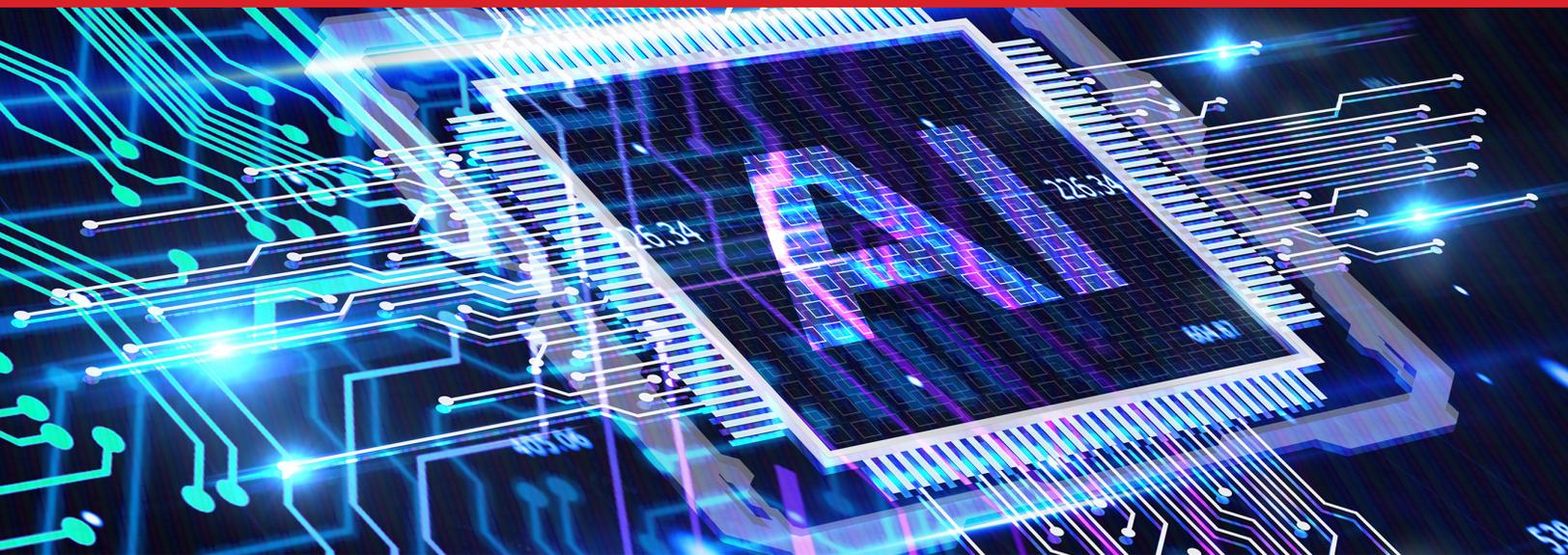
管理面考量

伺服器需要持續不斷進行管理，這會佔用系統管理員的許多時間。韌體、軟體和驅動程式需要定期更新，IT 人員必須執行效能和溫度最佳化與維護等繁瑣的任務。在之前的 Principled Technologies (PT) 測試中，我們評估了採用 Integrated Dell Remote Access Controller 9 (iDRAC9) 的 Dell 伺服器的管理功能。⁵⁵ 系統管理員可以放心依靠 iDRAC9 OpenManage Enterprise (OME) 的自動線上更新功能安排更新時程，讓伺服器隨時保持最新狀態，並且在工作負載增加時，使用設定檔快速便捷地加入新的伺服器。與使用 HPE OneView 和 HPE iLO 相比，藉助 iDRAC 和 OME，Dell 客戶可以更輕鬆地存取更多遠端管理功能、更輕鬆地部署伺服器和更新韌體。Dell PowerEdge 伺服器提供的 Dell 管理和服務，可協助組織「減少耗費在監測系統健全狀況或更新韌體的時間和精力」，將省下來的時間和精力用於創新和其他任務。⁵⁶

表 2：節錄自 2022 年 11 月的 PT 報告的 Dell 和 HPE 管理工具比較摘要。⁵⁷

資料來源：Principled Technologies。

	Dell 管理工具的與眾不同之處	贏過多少
更多遠端管理功能 iDRAC 與 iLO 的比較	更多 HTML5 主控台和 BIOS 組態功能，可在 iDRAC 中提供更多遠端功能	2.5 倍的 HTML5 主控台功能和 13 倍的 BIOS 功能
伺服器部署更輕鬆 OME 與 OneView 的比較	使用 OME 進行一對多設定檔部署	與 OneView 相比，部署伺服器的時間減少了 52%
韌體更新更輕鬆 OME 與 OneView 的比較	使用 OME 進行自動線上更新	透過連線到 Dell.com 同時更新多個伺服器，省下使用 OneView 手動上傳套件來更新伺服器的時間
更便捷的警示功能 OME 與 OneView 的比較	在 OME 中設定警示原則並根據警示執行自動化動作	與每次在 OneView 中收到警示時手動執行動作相比，自動執行此流程可以節省時間並減少出錯的可能性
易於使用的安全性功能 (系統鎖定和動態 USB) iDRAC 與 iLO 的比較	使用 iDRAC 的步驟更少，時間更短且無需重新開機	只需 1/4 的步驟，系統鎖定時間減少 91%
更強大的分析功能 CloudIQ for PowerEdge 與 InfoSight 的比較	CloudIQ for PowerEdge 提供可自訂的報告和更多健全狀況指標，可讓系統管理員更有效地進行控管	與 InfoSight 相比，可供選擇的指標多了 15 倍以上



結論

利用 AI 的強大功能來簡化和改善業務營運是一項具有挑戰性的任務，並且會對業務營運帶來重大影響。技術的演進從未如此快速，與合適的 AI 供應商合作是成功致勝的關鍵。Dell 不僅可以提供全方位的 AI 產品組合，還可以提供完善的規劃、準備、實施和管理服務，選擇與 Dell 這樣的公司合作，客戶就能夠更從容地面對這些挑戰。MLPerf® 基準測試顯示，Dell AI 產品組合可為 AI 工作負載提供強大穩定的效能。憑藉高效能和靈活的伺服器選項，以及多種儲存選擇、經過驗證的解決方案和專為 AI 量身定製的專業服務，Dell 可以協助企業擁抱 AI 及其優勢。

1. Dell · “Increasing Your Data Value with Dell Generative AI Solutions”(使用 Dell 生成式 AI 解決方案提高您的資料價值) · 2023 年 12 月 19 日存取 · <https://www.dell.com/en-us/blog/increasing-your-data-value-with-dell-generative-ai-solutions/>。
2. Dell · “Dell AI solutions”(Dell AI 解決方案) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>。
3. Dell · “Dell Technologies and Hugging Face to Simplify Generative AI with On-Premises IT”(Dell Technologies 和 Hugging Face 透過內部部署 IT 簡化生成式 AI) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/corporate/newsroom/announcements/detailpage.press-releases~usa~2023~11~20231114-dell-technologies-and-hugging-face-to-simplify-generative-ai-with-on-premises-it.htm#/lter-on/Country:en-us>。
4. Dell · “Dell and Meta Collaborate to Drive Generative AI Innovation”(Dell 和 Meta 合作推動生成式 AI 創新) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/blog/dell-and-meta-collaborate-to-drive-generative-ai-innovation/>。
5. Dell · “Dell AI-Ready Data Platform”(Dell 的 AI 就緒資料平台) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/storage-for-ai.htm?hve=explore+unstructured+storage#tab0=0>。
6. Dell · “Snowflake and Dell Partnership Gains Momentum”(Snowflake 和 Dell 合作，如虎添翼) · 2023 年 12 月 19 日存取 · <https://www.dell.com/en-us/blog/snowflake-and-dell-partnership-gains-momentum/>。
7. Robert McNeal · “Dell, VMware and NVIDIA Bring AI to Your Data”(Dell、VMware 和 NVIDIA 將 AI 技術應用於您的資料) · 2024 年 1 月 17 日存取 · <https://www.dell.com/en-us/blog/dell-vmware-and-nvidia-bring-ai-to-your-data/>。根據上面的連結：「根據 2023 年 8 月的 Dell 分析，Dell Technologies 提供的解決方案可支援從工作站電腦 (行動和固定式) 到伺服器的 AI 工作負載，適用於高效能運算、資料儲存、雲端原生軟體定義式基礎架構、網路交換器、資料保護、HCI 及各種服務。」
8. Intel · “Accelerate Artificial Intelligence (AI) Workloads with Intel Advanced Matrix Extensions (Intel AMX)”(使用 Intel Advanced Matrix Extensions (Intel AMX) 加速人工智慧 (AI) 工作負載) · 2023 年 12 月 12 日存取 · <https://www.intel.com/content/www/us/en/content-details/785250/accelerate-artificial-intelligence-ai-workloads-with-intel-advanced-matrix-extensions-intel-amx.html>。

9. Viperia · “NVIDIA’s H100 and A100 GPU Cards: Exploring the Intricacies of SXM and PCI-E Connections”(NVIDIA 的 H100 和 A100 GPU 卡：探索 SXM 和 PCI-E 連接的複雜性) · 2023 年 12 月 12 日存取 · <https://www.viperatech.com/unraveling-the-mysteries-sxm-vs-pci-e-connections-in-nvidias-high-end-h100-and-a100-gpus/>。
10. GitHub · “MLPerf® Results Messaging Guidelines”(MLPerf® 結果訊息傳遞指南) · 2024 年 1 月 16 日存取 · https://github.com/mlcommons/policies/blob/master/MLPerf_Results_Messaging_Guidelines.adoc。
11. MLCommons® · “MLPerf® Inference: Datacenter Benchmark Suite Results”(MLPerf® 推論：資料中心基準套件結果) · 2023 年 12 月 12 日存取 · <https://mlcommons.org/en/inference-datacenter-31/>。
12. Dell · “PowerEdge XE Servers”(PowerEdge XE 伺服器) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/servers/specialty-servers/poweredge-xe-servers.htm?hve=explore+poweredge+xe#tab0=0>。
13. HPE · “HPE ProLiant XL675d Gen10 Plus Configure-to-order Server”(HPE ProLiant XL675d Gen10 Plus 客製化產品規格訂購伺服器) · 2023 年 12 月 12 日存取 · <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.1013142988.html>。
14. HPE · “HPE ProLiant DL380a Gen11” · 2023 年 12 月 12 日存取 · <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.proliant-dl380-server.1014696168.html>。
15. MLCommons® · “MLPerf® Inference: Datacenter Benchmark Suite Results v 3.1”(MLPerf® 推論：資料中心基準套件結果 3.1 版) · 2023 年 12 月 12 日存取 · <https://mlcommons.org/benchmarks/inference-datacenter/>。
16. HPE · “NVIDIA Accelerators for HPE ProLiant Servers”(適用於 HPE ProLiant 伺服器的 NVIDIA 加速器) · 2023 年 12 月 12 日存取 · https://www.hpe.com/psnow/doc/c04123180.html?jumpid=in_pdp-psnow-qs。
17. Dell · “PowerEdge XE9680 Specification Sheet”(PowerEdge XE9680 規格表) · 2024 年 1 月 19 日存取 · <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9680-spec-sheet.pdf>。
18. HPE · “HPE & NVIDIA Financial services solution sets new records in performance”(HPE 和 NVIDIA 金融服務解決方案在效能方面創下新記錄) · 2023 年 12 月 12 日存取 · <https://community.hpe.com/t5/alliances/hpe-amp-nvidia-financial-services-solution-sets-new-records-in/ba-p/7197388>。
19. HPE · “QuickSpecs: HPE Cray Supercomputing XD670”(規格概述：HPE Cray 超級運算 XD670) · 2023 年 12 月 12 日存取 · <https://www.hpe.com/psnow/doc/a50004292enw>。
20. 已驗證的 MLPerf® v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0069 條。MLPerf® 名稱和標誌是 MLCommons® Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
21. 已驗證的 MLPerf® v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0085 條。MLPerf® 名稱和標誌是 MLCommons® Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
22. Dell · “PowerEdge XE9640 Rack Server”(PowerEdge XE9640 機架式伺服器) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>。
23. Accelsius · “Enabling the AI Revolution with Liquid Cooling”(利用液冷技術實現 AI 革新) · 2023 年 12 月 12 日存取 · <https://www.accelsius.com/blog/enabling-the-ai-revolution-with-liquid-cooling>。
24. Dell · “Dell PowerEdge XE9640 Technical Guide”(Dell PowerEdge XE9640 技術指南) · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9640-technical-guide.pdf>。
25. HPE · “HPE ProLiant DL380a Gen11” · 2023 年 12 月 12 日存取 · https://www.hpe.com/psnow/doc/PSN1014696168WWEN.pdf?jumpid=in_pdp-psnow-dds。
26. Intel · “Intel® Data Center GPU Max Series Technical Overview”(Intel® 資料中心 GPU Max 系列技術概覽) · 2023 年 12 月 12 日存取 · <https://www.intel.com/content/www/us/en/developer/articles/technical/intel-data-center-gpu-max-series-overview.html#gs.088741>。
27. HPE · “Intel Data Center GPU Max 1100 48GB Accelerator for HPE Data sheet”(適用於 HPE 的 Intel 資料中心 GPU Max 1100 48GB 加速器資料表) · 2023 年 12 月 12 日存取 · <https://www.hpe.com/psnow/doc/PSN1014779728WWEN>。
28. 已驗證的 MLPerf® v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0066 條。MLPerf® 名稱和標誌是 MLCommons® Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
29. 已驗證的 MLPerf® v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0084 條。MLPerf® 名稱和標誌是 MLCommons® Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。

30. Dell · “AI and HPC —With Air or Liquid Cooling”(AI 和 HPC — 採用氣冷或液冷技術) · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/servers/briefs-summaries/poweredge-xe9640-and-xe8640-infographic.pdf>。
31. Dell · “PowerEdge XE8640 : Drive AI, HPC modeling and simulation workloads with superior performance”(PowerEdge XE8640 : 以卓越的效能驅動 AI、HPC 建模和模擬工作負載) · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe8640-spec-sheet.pdf>。
32. Dell · “PowerEdge XE8640 Rack Server”(PowerEdge XE8640 機架式伺服器) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>。
33. 已驗證的 MLPerf[®] v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0067 條。MLPerf[®] 名稱和標誌是 MLCommons[®] Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
34. 已驗證的 MLPerf[®] v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0084 條。MLPerf[®] 名稱和標誌是 MLCommons[®] Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
35. Dell · “PowerEdge R760xa Rack Server”(PowerEdge R760xa 機架式伺服器) · 2023 年 12 月 12 日存取 · https://www.dell.com/en-us/shop/dell-poweredge-servers/poweredge-r760xa-rack-server/spd/poweredge-r760xa/pe_r760xa_16902_v1_vp#features_section。
36. SANStorageWorks · “Dell EMC PowerEdge R760xa: Powerful and scalable for GPU workloads”(Dell EMC PowerEdge R760xa : 功能強大且可擴充，適用於 GPU 工作負載) · 2023 年 12 月 12 日存取 · <https://www.sanstorageworks.com/PowerEdge-R760xa.asp>。
37. Dell · “Dell PowerEdge Servers and NVIDIA GPUs”(Dell PowerEdge 伺服器和 NVIDIA GPU) · 2023 年 12 月 12 日存取 · <https://infohub.delltechnologies.com//design-guide-generative-ai-in-the-enterprise-inferencing/dell-poweredge-servers-and-nvidia-gpus-1/>。
38. 已驗證的 MLPerf[®] v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0064 條。MLPerf[®] 名稱和標誌是 MLCommons[®] Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
39. 已驗證的 MLPerf[®] v3.1 Inference Closed 分數。檢索自 <https://mlcommons.org/benchmarks/inference-datacenter/> 2023 年 12 月 5 日 · 3.1-0084 條。MLPerf[®] 名稱和標誌是 MLCommons[®] Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。
40. Dell · “Workstations for AI”(AI 專用工作站) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/ai-technologies/index.htm?hve=explore+dell+precision+for+ai#pdf-overlay=//www.delltechnologies.com/asset/en-us/products/workstations/briefs-summaries/ai-industry-brochure.pdf>。
41. NVIDIA · “NVIDIA RTX in Professional Workstations”(專業工作站中的 NVIDIA RTX) · 2023 年 12 月 12 日存取 · <https://www.nvidia.com/en-us/design-visualization/desktop-graphics/>。
42. Intel · “Intel[®] Deep Learning Boost (Intel[®] DL Boost)” · 2023 年 12 月 12 日存取 · <https://www.intel.com/content/www/us/en/artificial-intelligence/deep-learning-boost.html>。
43. HPE · “HPE ProLiant ML350 Gen11” · 2023 年 12 月 12 日存取 · <https://buy.hpe.com/us/en/compute/tower-servers/proliant-ml300-servers/proliant-ml350-server/hpe-proliant-ml350-gen11/p/1014696172>。
44. ComputerWeekly.com · “Storage requirements for AI, ML and analytics in 2022”(2022 年 AI、ML 和分析的儲存要求) · 2023 年 12 月 12 日存取 · <https://www.computerweekly.com/feature/Storage-requirements-for-AI-ML-and-analytics-in-2022>。
45. Dell · “PowerScale AI-Ready Data Platform”(PowerScale 的 AI 就緒資料平台) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale>。
46. Dell · “Compare PowerScale”(比較 PowerScale) · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale#compare-module>。
47. Dell · “Dell PowerScale All-Flash”(Dell PowerScale 全快閃記憶體) · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h15963-ss-powerscale-all-flash-nodes.pdf>。
48. Dell · “Dell PowerScale OneFS Software Features”(Dell PowerScale OneFS 軟體功能) · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h18275-onefs-software-features-data-sheet.pdf>。

49. Dell · “Dell APEX File Storage for AWS” · 2023 年 12 月 12 日存取 · <https://www.delltechnologies.com/asset/en-us/products/storage/briefs-summaries/h19575-so-apex-file-storage-for-aws.pdf>。
50. Dell · “Dell ECS Enterprise Object Storage” · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/storage/ecs/index.htm?hve=explore+ecs#tab0=0&tab1=0>。
51. Dell · “Dell ECS Enterprise Object Storage” · 2023 年 12 月 12 日存取 · <https://www.dell.com/en-us/dt/storage/ecs/index.htm#tab0=0&tab1=0&accordion0>。
52. HPE · “Storage Solutions for Scality”(適用於 Scality 的儲存解決方案) · 2023 年 12 月 12 日存取 · <https://www.hpe.com/us/en/storage/file-object/scality.html>。
53. HPE · “Make AI work for you”(讓 AI 為您工作) · 2024 年 1 月 16 日存取 · <https://www.hpe.com/us/en/solutions/artificial-intelligence.html>。
54. HPE · “HPE AI Services | Generative AI Implementation”(HPE AI 服務 | 生成式 AI 實施) · 2024 年 1 月 16 日存取 · <https://www.hpe.com/us/en/services/generative-ai-implementation-service.html>。
55. Principled Technologies · “Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE”(使用 Dell 管理產品組合中的工具與 HPE 的類似工具簡化管理員任務並改善安全性和健全狀況監控) · 2023 年 12 月 12 日存取 · <https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>。
56. Principled Technologies · “Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE”(使用 Dell 管理產品組合中的工具與 HPE 的類似工具簡化管理員任務並改善安全性和健全狀況監控) · 2023 年 12 月 12 日存取 · <https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>。
57. Principled Technologies · “Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE”(與 HPE 的同類工具相比，使用 Dell 管理產品組合中的工具簡化管理員任務並改善安全性和健全狀況監控)。

MLPerf 名稱和標誌是 MLCommons Association 在美國和其他國家/地區的註冊和未註冊商標。保留所有權利。未經授權不得使用。如需詳細資訊，請參閱 www.mlcommons.org。

► 檢視本報告的原始英文版本，網址為 <https://facts.pt/zPmSx4c>

此專案是由 Dell Technologies 委託執行。



Facts matter.®

Principled Technologies 為 Principled Technologies, Inc. 的註冊商標。
所有其他產品名稱皆為各自所有人之商標。

免責聲明；賠償責任限制：

Principled Technologies, Inc. 已盡合理努力確保其測試之正確性及有效性，然而，Principled Technologies, Inc. 特此排除任何與測試結果及分析相關、指涉其正確性、完整性或品質之明示或默示保證，其中包含適合任何特定用途之默示擔保。以任何測試程序或結果存在任何可能之錯誤或瑕疵為由，並據此所提之任何損失或損害訴訟，Principled Technologies, Inc. 及其員工與承包商概不負責，運用任何測試結果之所有個人或機構均需自負風險，不得異議。

在任何情況下，Principled Technologies, Inc. 一概不為其測試所產生之間接、特殊、意外或連帶性損害負責，即使已於事前告知此等損害之可能性者亦同。在任何情況下，Principled Technologies, Inc. 之賠償責任，包括直接損害在內，不得大於 Principled Technologies, Inc. 所收取之測試相關費用金額。此處所記載者為客戶唯一的具排他性救濟措施。