



NVIDIA DGX SuperPOD與NetApp - 設計指南

NetApp artificial intelligence solutions

NetApp
August 18, 2025

目錄

NVIDIA DGX SuperPOD與NetApp - 設計指南	1
執行摘要	1
解決方案概述	1
解決方案技術	1
用例摘要	3
技術要求	3
硬體需求	3
軟體需求	4
解決方案驗證	4
結論	4
在哪裡可以找到更多信息	4

NVIDIA DGX SuperPOD與NetApp - 設計指南

此NetApp驗證架構描述了具有NetApp BeeGFS 構建塊的NVIDIA DGX SuperPOD的設計。此解決方案是一個全端資料中心平台，在NVIDIA的專用驗收集群上進行了驗證。



NetApp 的Amine Bennani、Christian Whiteside、David Arnette 和 Sathish Thyagarajan

執行摘要

在當今快速發展的技術格局中，人工智慧正在徹底改變消費者體驗並推動各行各業的創新。然而，這也給 IT 部門帶來了巨大的挑戰，他們面臨著部署能夠處理 AI 工作負載的強烈需求的高效能運算 (HPC) 解決方案的壓力。隨著各組織競相利用人工智慧的力量，對易於部署、擴展和管理的解決方案的需求也日益迫切。

NVIDIA DGX SuperPOD是一個 AI 資料中心基礎架構平台，作為 IT 的交鑰匙解決方案提供，以支援當今企業面臨的最複雜的 AI 工作負載。任何精確的深度學習 (DL) 模型的核心都是大量數據，需要能夠有效地提供和重新提供這些數據的高吞吐量儲存解決方案。NetApp BeeGFS 解決方案由具有 BeeGFS 平行檔案系統的NetApp EF600 儲存陣列組成，使NVIDIA DGX SuperPOD能夠充分發揮其功能。NetApp BeeGFS 解決方案已通過NVIDIA驗證，可與 SuperPOD 架構整合和擴充。其結果是簡化了 AI 資料中心的部署和管理，同時提供了幾乎無限的效能和容量可擴展性。

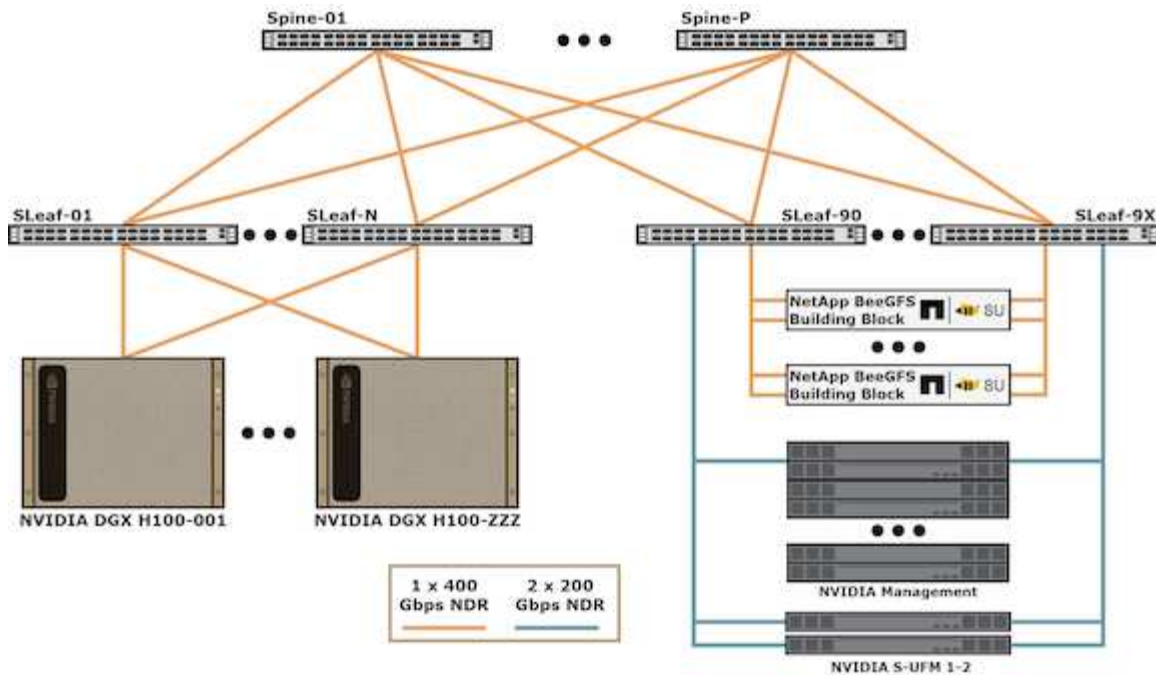
解決方案概述

NetApp BeeGFS 解決方案由高效能NetApp EF600 NVMe 儲存系統和可擴展的 BeeGFS 平行檔案系統提供支援，為要求苛刻的 AI 工作負載提供了強大而高效的儲存基礎。其共享磁碟架構確保高可用性，即使面臨系統挑戰也能保持一致的效能和可存取性。該解決方案提供了可擴展且靈活的架構，可客製化以滿足不同的儲存需求。客戶可以透過整合額外的儲存構建塊來輕鬆擴展其儲存效能和容量，以處理最苛刻的工作負載。

解決方案技術

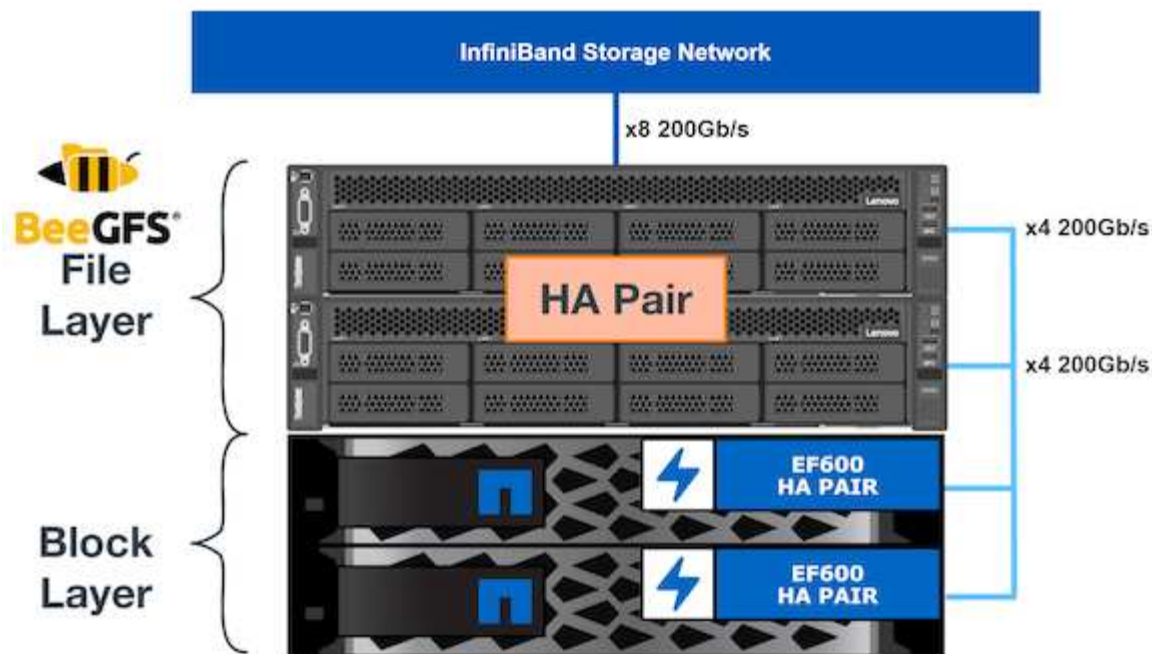
- NVIDIA DGX SuperPOD利用 DGX H100 和 H200 系統以及經過驗證的外部連接共享儲存：
 - 每個 DGX SuperPOD 可擴充單元 (SU) 由 32 個 DGX 系統組成，能夠以 FP8 精度達到 640 petaFLOPS 的 AI 效能。NetApp建議為單一 DGX SuperPOD 設定使用至少 2 個建置區塊來調整NetApp BeeGFS 儲存解決方案的大小。

解決方案的高層視圖



- NetApp BeeGFS 建置區塊由兩個NetApp EF600 陣列和兩台 x86 伺服器組成：
 - 透過以NVIDIA DGX SuperPOD為基礎的NetApp EF600 全快閃陣列，客戶可以獲得可靠的儲存基礎，並享有 6 個 9 的正常運作時間。
 - NetApp EF600 和NVIDIA DGX 系統之間的檔案系統層是 BeeGFS 平行檔案系統。BeeGFS 由德國弗勞恩霍夫高效能運算中心創建，旨在解決傳統平行檔案系統的痛點。其結果是一個具有現代用戶空間架構的檔案系統，現在由 ThinkParQ 開發和交付，並被許多超級運算環境使用。
 - NetApp對 BeeGFS 的支援使 NetApp 優秀的支援組織與客戶對效能和正常運作時間的要求保持一致。客戶可以獲得優質的支援資源、提前獲得 BeeGFS 版本，以及使用部分 BeeGFS 企業功能，例如配額實施和高可用性 (HA)。
- NVIDIA SuperPOD SU 和NetApp BeeGFS 構建塊的結合提供了一種敏捷的 AI 解決方案，其中計算或儲存可以輕鬆無縫地擴展。

NetApp BeeGFS 構建塊



用例摘要

此解決方案適用於以下用例：

- 人工智慧（AI）包括機器學習（ML）、深度學習（DL）、自然語言處理（NLP）、自然語言理解（NLU）和生成人工智慧（GenAI）。
- 中大規模人工智慧訓練
- 電腦視覺、語音、音訊和語言模型
- HPC，包括透過訊息傳遞介面（MPI）和其他分散式運算技術加速的應用程式
- 應用程式工作負載具有以下特點：
 - 讀取或寫入大於 1GB 的文件
 - 多個客戶端（10 個、100 個和 1000 個）讀取或寫入相同文件
- 多 TB 或多 PB 資料集
- 需要針對大檔案和小檔案混合進行最佳化的單一儲存命名空間的環境

技術要求

本節介紹採用NetApp解決方案的NVIDIA DGX SuperPOD的技術需求。

硬體需求

下表 1 列出了單一 SU 實施解決方案所需的硬體元件。解決方案規模從 32 個NVIDIA DGX H100 系統和兩個或三個NetApp BeeGFS 構建塊開始。單一NetApp BeeGFS 建置區塊由兩個NetApp EF600 陣列和兩台 x86 伺服器組成。隨著部署規模的增加，客戶可以添加額外的建置區塊。有關詳細信息，請參閱 ["NVIDIA DGX H100 SuperPOD 參考架構"](#)和 ["NVA-1164-DESIGN： NetApp NVA 上的 BeeGFS 設計"](#)。

硬體	數量
NVIDIA DGX H100 或 H200	32
NVIDIA Quantum QM9700 交換機	8 片葉子，4 根脊柱
NetApp BeeGFS 構建塊	3

軟體需求

下表 2 列出了實施此解決方案所需的軟體元件。解決方案的任何特定實施中使用的軟體元件可能會根據客戶要求而有所不同。

軟體
NVIDIA DGX 軟體堆疊
NVIDIA 基礎指令管理器
ThinkParQ BeeGFS 平行檔案系統

解決方案驗證

NVIDIA DGX SuperPOD 與 NetApp 透過使用 NetApp BeeGFS 構建塊在 NVIDIA 的專用驗收集群上進行了驗證。驗收標準是基於 NVIDIA 執行的一系列應用程式、效能和壓力測試。有關詳細信息，請參閱 ["NVIDIA DGX SuperPOD：NetApp EF600 和 BeeGFS 參考架構"](#)。

結論

NetApp 和 NVIDIA 有著長期的合作，致力於為市場提供一系列 AI 解決方案。NVIDIA DGX SuperPOD 與 NetApp EF600 全快閃陣列結合，是經過驗證的解決方案，客戶可以放心部署。這種完全整合的交鑰匙架構消除了部署風險，使任何人都可以走上贏得人工智慧領導地位的道路。

在哪裡可以找到更多信息

要了解有關本文檔中描述的信息的更多信息，請查看以下文檔和/或網站：

- ["NVIDIA DGX SuperPOD 參考架構"](#)
- ["NVIDIA DGX SuperPOD 資料中心設計參考指南"](#)
- ["NVIDIA DGX SuperPOD：NetApp EF600 和 BeeGFS"](#)

版權資訊

Copyright © 2025 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。