了解您的 Novus 系統

Novus

NetApp
November 03, 2026

This PDF was generated from https://docs.netapp.com/zh-tw/novus/get-started/learn-novus.html on November 03, 2026. Always check docs.netapp.com for the latest.

目錄

了解您的 Novus 系統 . . . 1
- 了解 Novus 儲存系統 . . . 1
  - 概況 . . . 1
  - 系統設計原則 . . . 1
  - 網路檔案系統 (NFS) 實作 . . . 2
  - 其他技術和概念 . . . 3
  - 相關資訊 . . . 3
- Novus 儲存系統架構 . . . 3
  - ONTAP 資料伺服器 . . . 3
  - Novus Data Director . . . 4
  - Novus 部署 . . . 4
- Novus 儲存系統環境及處理 . . . 5
  - Novus Data Director 與 ONTAP 資料伺服器內部通訊 . . . 5
  - NDD 叢集的高可用度與彈性 . . . 6

了解您的 **Novus** 系統

了解 **Novus** 儲存系統

NetApp Novus 是一款功能強大的儲存系統，專為滿足大規模 AI 的需求而打造。它將 NetApp 在企業儲存設備的經驗擴展到訓練和推論所需的高效能基礎架構。了解其管理結構和組態選項對於有效管理和最佳效能至關重要。

概況

NetApp Novus 是一款橫向擴充的檔案儲存設備系統解決方案，專為滿足現代 AI 工作負載所需的處理量和規模而設計。它將許多獨立的 ONTAP 儲存設備叢集結合到命名空間中，用戶端將其視為一個檔案系統。Novus 建構於平行 NFS（pNFS）之上，可將中繼資料與檔案資料分隔。專用的中繼資料服務負責管理命名空間，而儲存設備叢集則平行地直接向用戶端提供檔案資料。這消除了單一儲存設備叢集的處理量上限，不再限制系統交付資料的速度。

Novus 部署可以透過新增儲存設備叢集輕鬆擴充。每個新叢集都會將其全部頻寬和容量新增至命名空間，因此隨著系統規模的成長，總處理量也會增加，而無需重新掛載用戶端或進行中斷性的重新平衡。中繼資料服務與資料伺服器獨立擴充，因此您可以擴充命名空間大小和中繼資料效能以符合工作負載。Novus 的設計是讓每個儲存節點以全速運作，並在大規模環境下將元件故障視為例行且不中斷的情況。主要目標是即使系統規模成長至非常龐大，仍能保持資料流向運算。

用戶端使用標準的 NFS 用戶端，這使得 Novus 能夠輕鬆適應大型 GPU 和運算叢集。此外，Novus 具備多租戶特性，因此一個部署可以主機多個隔離的客戶或工作負載，每個都擁有自己的分隔命名空間。它也建構在 ONTAP 獲證實的資料管理員和保護之上。對客戶而言，這意味著儲存設備能讓 GPU 保持忙碌，透過從每個節點獲得更多效能來降低基礎架構成本，並在命名空間中支援非常大的資料集。

Novus 可以協助大規模執行 AI 和其他資料密集型工作負載的組織。這通常包括新型雲端、GPU 即服務供應商、主權雲端、超大規模雲端供應商，以及執行 AI 訓練和推論的大型企業和研究機構。Novus 可為任何使用高效能運算或媒體處理、需要每秒數 TB 處理量，並希望在運算的同時擴充儲存設備的客戶帶來優點。

系統設計原則

Novus 的設計和建造基於幾個核心原則。這些底層架構選擇驅動了其效能、擴充性和彈性。這有助於需要高效能儲存設備、能夠充分利用 GPU 資源並降低基礎架構成本的客戶。

支援百億億次級運算的儲存設備

Novus 的設計旨在成為支援 Exascale 的軟體定義儲存基礎堆疊。它將水平擴充的 pNFS 中繼資料伺服器與分隔且解構的資料伺服器結合。

處理量超出傳統 **ONTAP** 叢集限制

傳統上，單一 ONTAP 叢集的處理量上限是固定的，由其控制器數量和互連頻寬決定。AI 訓練和推論所需的處理量是單一 ONTAP 叢集無法達到的。

透過新增叢集來擴充

Novus 將多個獨立的 ONTAP 叢集（每個叢集最小可包含一個 HA 配對）合併到一個命名空間。每個新叢集都會貢獻其全部頻寬，因此總處理量呈線性成長。用戶端無需重新掛載，機組也無需重新平衡。

可獨立擴充的中繼資料

中繼資料服務階層和分片中繼資料儲存均可從 1 個成長到超過 1,000 個執行個體。中繼資料容量和請求速率與使用的資料節點數量無關。

線速資料節點

每個資料節點的讀取應以 NIC 全速運作，寫入應以 NVMe 全速運作。任何未使用的每節點處理量只會變成額外的硬體和成本。

彈性是主要的設計限制

在擁有 1,000 個或更多儲存節點的規模下，故障並不罕見。彈性是首要的設計限制，而不是營運上的事後考量。

網路檔案系統 (NFS) 實作

Novus 是網路檔案系統（NFS）協定的軟體定義實作。它依賴 NFSv4（特別是平行 NFS（pNFS）擴充）來分割中繼資料和資料路徑。用戶端可以使用標準 NFS 用戶端透過 NFSv4 存取 Novus Data Director（NDD），同時使用 NFSv3 直接存取 ONTAP 資料伺服器上的資料。

平行網路檔案系統（**pNFS**）

並行 NFS (pNFS) 是 NFSv4.1 中引入的一項功能，它將檔案中繼資料管理與檔案資料存取分隔。NFS 伺服器處理中繼資料操作，並向用戶端提供檔案資料儲存位置的資訊。用戶端隨後可以直接存取專用儲存設備和資料伺服器上的資料，而無需將所有讀取和寫入操作都傳送到單一伺服器。

NFS 伺服器提供一個佈局，用於描述每個檔案的資料排列方式以及用戶端如何存取這些資料。根據佈局和儲存組態，用戶端可以並行與一個或多個資料伺服器通訊。這可提高資料傳輸效能，並降低中繼資料伺服器的資料處理負載。

pNFS 支援多種佈局類型，包括檔案、區塊、物件和靈活檔案。每種類型都描述了將檔案資料對應到基礎儲存設備的不同方式，以及用戶端存取資料的相應方式。Novus 使用靈活檔案佈局。

pNFS 靈活檔案佈局是最近新增的佈局選項。它提供靈活的以檔案為基礎的佈局，可描述儲存在多個資料伺服器上的資料，並支援鏡射複本。用戶端使用佈局資訊來選擇適當的資料伺服器路徑以進行存取。這可讓部署根據儲存組態和工作負載要求來調整資料放置與存取行為。

主要元件

Novus 有兩個主要元件，這些元件是建構在 pNFS 標準與 NetApp 的企業儲存設備專業知識之上。

ONTAP 資料伺服器

ONTAP 資料伺服器實作了 Novus 資料平面。它儲存檔案資料並將其提供給 NFS 用戶端。一個 Novus 命名空間可以將其檔案資料分散在多個 ONTAP 叢集上，而一個 ONTAP 叢集可以服務多個命名空間。

Novus Data Director（NDD）

Novus Data Director 提供 Novus 中繼資料平面。它管理命名空間中繼資料，處理用戶端的中繼資料操作要求，並為用戶端提供描述檔案資料在 ONTAP 叢集中儲存位置的配置。用戶端使用具備 pNFS 的 NFSv4.2 與 NDD 互動，而實際的檔案資料則直接從 ONTAP 存取。

請檢閱 "Novus 架構" 以取得有關這些元件及其如何協同運作的詳細資料。

其他技術和概念

還有幾個與 Novus 及其底層技術相關的概念。

命名空間

命名空間是一種邏輯結構，用於組織和識別資源，允許在定義的範圍內一致地參照這些資源。在 NFSv4 中，命名空間是呈現給用戶端的檔案和目錄階層結構；它提供了匯出檔案系統的統一檢視，用戶端可以使用路徑進行導覽。每個 NDD 命名空間都有自己的中繼資料伺服器執行個體。

聯邦

聯盟是一種將獨立系統連接起來的技術，使它們能夠在保持各自獨立代管的同時協同工作；它可以應用於身分、應用程式服務、網路、資料庫和其他資源。當與資料供應商結合使用時，聯盟可讓使用者或應用程式透過統一的介面存取來自多個資料來源的資料，同時供應商保留對其資料的控制——這支援了互通性，並減少了集中化或複製資訊的需求。

分片

分片是一種將資料庫資料分割到多個伺服器上的技術，每個伺服器儲存和處理一部分資料。它透過分散儲存設備和查詢負載來提高容量和效能。

邊界閘道協定

邊界閘道協定 (BGP) 是一種路由協定，用於在自治系統（即構成網際網路的網路）之間交換可達性資訊，並使用路由原則和屬性選擇路徑。Novus 依賴 BGP 來支援其多租戶架構。

易擴充的分散式中繼資料儲存庫

易擴充的分散式中繼資料儲存系統是管理跨多個節點之中繼資料的系統，可確保高可用度、容錯能力和有效率的存取。在 Novus 中，Novus Data Director（NDD）實作了易擴充的分散式中繼資料儲存系統（或簡稱中繼資料儲存系統），以處理命名空間中繼資料，並為用戶端提供存取檔案資料的配置。

相關資訊

- "RFC 8435：平行 NFS （pNFS） 靈活檔案佈局"
- "RFC 8178：NFSv4 擴充和次要版本規則"
- "RFC 8881：網路檔案系統（NFS）版本 4 次版本 1 協定"
- "RFC 7862：網路檔案系統（NFS）版本 4 次版本 2 協定"

Novus 儲存系統架構

Novus 架構由 Novus Data Director（NDD）和 ONTAP 資料伺服器的執行個體組成。ONTAP 資料伺服器支援資料路徑，而 NDD 則透過 Parallel NFS（pNFS）管理中繼資料和用戶端存取。

ONTAP 資料伺服器

資料伺服器是專為 Novus 建置的 ONTAP 叢集。它們實作 Novus 資料平面，儲存檔案資料並將其提供給 NFS 用戶端。一個 Novus 命名空間可以將其檔案資料分佈在多個 ONTAP 叢集上；一個叢集可以服務多個不同的命名空間。這種靈活度可藉由視需要新增資料伺服器，來增加容量與效能。

用戶端使用標準的 NFS 用戶端存取 ONTAP 資料伺服器，無需自訂的 NetApp 軟體。在用戶端從 Novus Data

Director 取得檔案配置後，它可以透過 NFSv3（並可選擇透過 RDMA）直接在 ONTAP 資料伺服器上讀取和寫入檔案資料。大量資料絕不會通過 Novus Data Director，因此資料流量會在資料伺服器之間平行傳輸。

資料伺服器不參與管理命名空間。它們將每個檔案的資料作為後端檔案儲存在 ONTAP 磁碟區中。ONTAP 資料伺服器不維護 Novus 命名空間中繼資料，例如目錄樹或檔案在資料伺服器之間的條帶化方式。

適用於 NetApp Novus 的 AFF A90 是專為執行 ONTAP 的 Novus 所打造的 AFF A90 版本。其特點包括：

- 額外的雙 NVRAM 卡
- 高頻寬網路介面卡
- 設定為標準 ONTAP HA 配對
- 使用 RAID-DP 或 RAID-TEC 保護資料。

標準 ONTAP 叢集無法作為 Novus 資料伺服器；僅支援專為 Novus 建置的平台。

Novus Data Director

Novus Data Director (NDD) 是 pNFS 架構中的中繼資料伺服器 (MDS) 角色的一種實作，並提供 Novus 中繼資料平面。*MDS* 主機 是一部執行 Linux 以及 MDS 協定服務和中繼資料儲存區的實體伺服器。

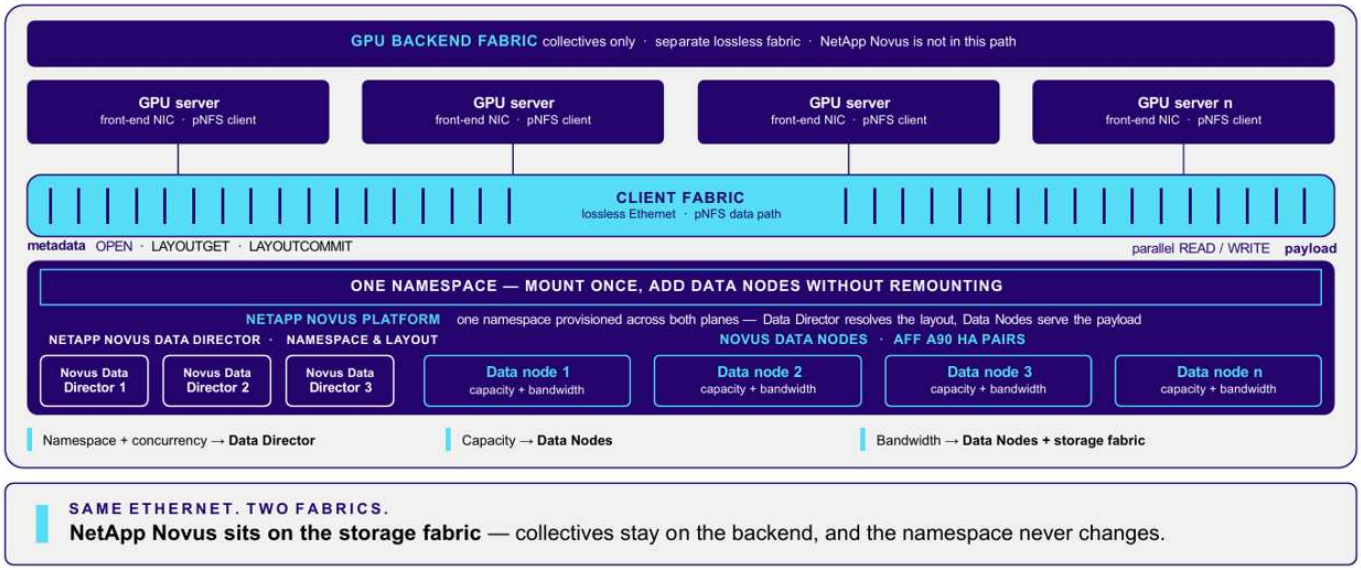
NDD 將每個命名空間作為單一檔案系統呈現給用戶端。它擁有命名空間中繼資料：檔案和目錄、檔案屬性、每個檔案資料儲存位置的對應，以及協調用戶端存取所需的協定狀態。用戶端透過其虛擬 IP 位址裝載命名空間，並使用具有平行 NFS（pNFS）的 NFSv4.2 與 NDD 通訊。NDD 處理所有中繼資料操作，例如查閱、檔案建立和屬性變更。當用戶端開啟檔案時，NDD 會為其提供 pNFS 靈活檔案配置，告知用戶端哪些資料伺服器保存檔案資料。然後，用戶端可以直接與 ONTAP 資料伺服器進行讀取和寫入；NDD 位於中繼資料路徑中，但不在資料路徑中。

NDD 實作為軟體，在至少三台專用伺服器的叢集上執行。它將中繼資料伺服器程序與複寫的分散式資料庫結合，該資料庫將命名空間中繼資料儲存在加密的本機儲存設備上。每個命名空間都在其中幾台伺服器上執行自己的中繼資料伺服器程序。所有運作正常的伺服器都會發布命名空間的虛擬 IP 位址，這會分散用戶端負載，並讓用戶端進行容錯移轉而無需重新裝載。由於權威中繼資料儲存在複寫的資料庫中，因此伺服器即使發生故障並重新啟動，也不會遺失命名空間狀態。

ONTAP 叢集上的資料放置是由 NDD 所控制。它會維護資料伺服器的登錄檔，使用輕量級 NFS 探測來確保每台資料伺服器都可用，並追蹤每台伺服器上的可用容量。建立檔案時，NDD 會選擇資料伺服器，透過 NFS 建立後端檔案，並記錄其位置。如果資料伺服器變得無法使用，NDD 會停止在該處放置新資料，並撤回指向該處的用戶端配置。

Novus 部署

下圖從高階層級展示了典型的 Novus 部署。值得注意的方面包括具有多個 Novus Data Directors 和 ONTAP 資料伺服器節點的命名空間。此架構提供高可用度和低延遲，以支援後端 GPU 伺服器。



Novus 儲存系統環境及處理

您應該了解 Novus 環境及其處理過程的幾個重要方面。

Novus Data Director 與 ONTAP 資料伺服器內部通訊

Novus Data Directors 與 ONTAP 資料伺服器通訊，以支援檔案存取和管理。其中大部分通訊由 NDD 發起，而不是由 ONTAP 伺服器發起。

ONTAP 資料伺服器探索

Novus Data Directors 需要探索有哪些資料伺服器可用。ONTAP 從不主動與 Novus Data Director 進行通訊。營運者或工具會透過在啟動時使用的 `mds.conf` 組態檔來註冊每個資料伺服器；也可以使用命令 `mds-admin ds add` 手動新增資料伺服器。每個資料伺服器在分散式中繼資料儲存區中都以一個資料列表示，包含其主機、匯出、連接埠、狀態和容量。

ONTAP 資料伺服器可用度

有幾種程序可以用來確定每個 ONTAP 資料伺服器的可用度。

心跳

MDS 透過 TCP 向每個線上資料伺服器發送 NFSv3 NULL RPC 呼叫。預設情況下，它每 5 秒使用心跳訊息執行一次此操作。

離線

連續發生一定次數的故障後，資料伺服器會被標記為離線。MDS 會停止將新檔案放置於其上，並通知用戶端歸還指向它的佈局。

回來

在連續多次成功的探測後，它會恢復線上。冷卻期會阻止它來回切換。

用戶端錯誤不算在內

當用戶端報告錯誤時，只會記錄下來。ONTAP 資料伺服器健全狀況僅由探測決定。

ONTAP 容錯移轉是不可見的

如果 ONTAP HA 配對發生接管，MDS 只能透過偵測逾時來察覺。

ONTAP 資料伺服器資料內容

NDD 不會詢問 ONTAP 它儲存了什麼。它已經知道了，因為它建立了每個後備檔案：

- 當建立檔案時，MDS 會選擇資料伺服器，在這些伺服器上建立後端檔案，並將結果儲存到分散式中繼資料儲存區的等量磁碟區對應表中。等量磁碟區對應表記錄了每個等量磁碟區和鏡射分別由哪個資料伺服器和 NFS 檔案控制代碼保存。
- ONTAP 上的後端檔案在 NDD 中具有特定的名稱格式。
- 容量是 NDD 針對每個已裝載的匯出，使用 statvfs() 從 ONTAP 讀取的唯一操作值。
- 檔案大小未根據 ONTAP 進行驗證。它是由用戶端的佈局提交所決定。

NDD 叢集的高可用度與彈性

以 Novus Data Directors 為核心的中繼資料平面確保了系統的高可用度和彈性。

叢集設計

Novus 使用多主機叢集來確保 NDD 服務的高可用度和容錯能力。

三台主機，三個故障網域

最小的 NDD 叢集是由部署在分隔伺服器上的三台實體主機或虛擬機器（VM）組成。每台主機執行一個 MDS 程序和一個中繼資料儲存節點。此叢集的設計可承受一台主機的遺失。

所有副本均提供服務

每個 MDS 複本同時處理用戶端，沒有任何複本被指定為待命模式。所有健全的複本都會透過 BGP 通告相同的 VIP。由於單靠 BGP 本身的計時器並非完全可接受的容錯移轉機制，因此需要 BFD（快速連結故障偵測）。複本僅在其 NFS 服務、閘道和中繼資料儲存連線均已準備就緒，可處理用戶端時，才會通告 VIP。

中繼資料複寫

分散式中繼資料儲存區會保留三個同步複本。仲裁器會選擇變成無法使用，而不是讓雙方同時寫入，藉此防止發生分割腦情況。

重新啟動是安全的

MDS 流程本身不保存任何狀態；分散式中繼資料儲存區保存權威中繼資料。

限制與注意事項

您還應感知以下限制和注意事項：

- 子樹所有權不會自動容錯移轉。命名空間可以分割為多個子樹，每個子樹由一個 MDS 擁有。如果某個所有者發生故障，即使 VIP 已經遷移，其子樹也會一直處於掛起狀態，直到所有者恢復為止。
- 一台主機故障會影響該主機上的所有租戶，因為每個租戶在每個主機都有一個副本。

- 用戶端必須裝載 VIP，而不是主機 IP。否則，用戶端無法容錯移轉。
- 硬裝載在恢復期間可能會阻塞。阻塞的裝載並不代表資料遺失。
- 務必一次對一個複本執行維護。先將複本從 VIP 中移出並將其排空。

版權資訊

Copyright © 2026 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 保留隨時變更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。