NVIDIA DGX SuperPOD con NetApp : Guía de diseño

NetApp artificial intelligence solutions

NetApp
August 18, 2025

This PDF was generated from https://docs.netapp.com/es-es/netapp-solutions-ai/infra/ai-eseries-superpod.html on August 18, 2025. Always check docs.netapp.com for the latest.

Tabla de contenidos

NVIDIA DGX SuperPOD con NetApp : Guía de diseño . . . 1
- Resumen ejecutivo . . . 1
- Descripción general de la solución . . . 1
  - Tecnología de soluciones . . . 1
  - Resumen del caso de uso . . . 3
- Requisitos tecnológicos . . . 3
  - Requisitos de hardware . . . 3
  - Requisitos de software . . . 4
- Verificación de la solución . . . 4
- Conclusión . . . 4
- Dónde encontrar información adicional . . . 4

NVIDIA DGX SuperPOD con NetApp : Guía de diseño

Esta arquitectura verificada de NetApp describe el diseño del NVIDIA DGX SuperPOD con bloques de construcción NetApp BeeGFS. Esta solución es una plataforma de centro de datos de pila completa validada en un clúster de aceptación dedicado en NVIDIA.

Amine Bennani, Christian Whiteside, David Arnette y Sathish Thyagarajan, NetApp

Resumen ejecutivo

En el panorama tecnológico actual en rápida evolución, la IA está revolucionando las experiencias de los consumidores e impulsando la innovación en todas las industrias. Sin embargo, también presenta desafíos importantes para los departamentos de TI, que están bajo presión para implementar soluciones de computación de alto rendimiento (HPC) capaces de manejar las intensas demandas de las cargas de trabajo de IA. A medida que las organizaciones compiten por aprovechar el poder de la IA, crece la urgencia de contar con una solución que sea fácil de implementar, escalar y administrar.

NVIDIA DGX SuperPOD es una plataforma de infraestructura de centro de datos de IA entregada como una solución llave en mano para TI para respaldar las cargas de trabajo de IA más complejas que enfrentan las empresas actuales. En el núcleo de cualquier modelo de aprendizaje profundo (DL) preciso hay grandes volúmenes de datos, lo que requiere una solución de almacenamiento de alto rendimiento que pueda servir y re-servir estos datos de manera eficiente. La solución NetApp BeeGFS, que consta de matrices de almacenamiento NetApp EF600 con el sistema de archivos paralelo BeeGFS, permite que NVIDIA DGX SuperPOD libere toda su capacidad. La solución NetApp BeeGFS ha sido validada por NVIDIA para integrarse y escalar con la arquitectura SuperPOD. El resultado es una implementación y gestión simplificadas del centro de datos de IA al tiempo que ofrece una escalabilidad prácticamente ilimitada en términos de rendimiento y capacidad.

Descripción general de la solución

La solución NetApp BeeGFS, impulsada por los sistemas de almacenamiento NetApp EF600 NVMe de alto rendimiento y el sistema de archivos paralelos BeeGFS escalable, ofrece una base de almacenamiento sólida y eficiente para cargas de trabajo de IA exigentes. Su arquitectura de disco compartido garantiza una alta disponibilidad, manteniendo un rendimiento y una accesibilidad consistentes, incluso ante desafíos del sistema. Esta solución proporciona una arquitectura escalable y flexible que se puede personalizar para satisfacer diversos requisitos de almacenamiento. Los clientes pueden ampliar fácilmente su rendimiento y capacidad de almacenamiento integrando bloques de almacenamiento adicionales para manejar incluso las cargas de trabajo más exigentes.

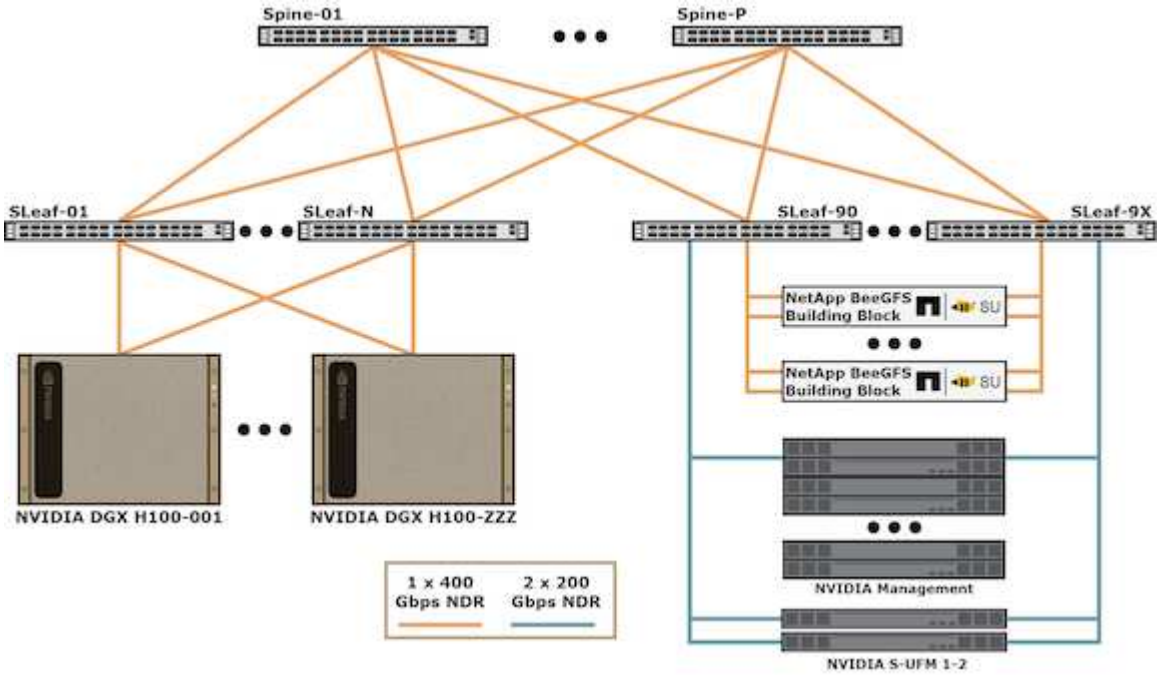
Tecnología de soluciones

- NVIDIA DGX SuperPOD aprovecha los sistemas DGX H100 y H200 con un almacenamiento compartido

externo validado:

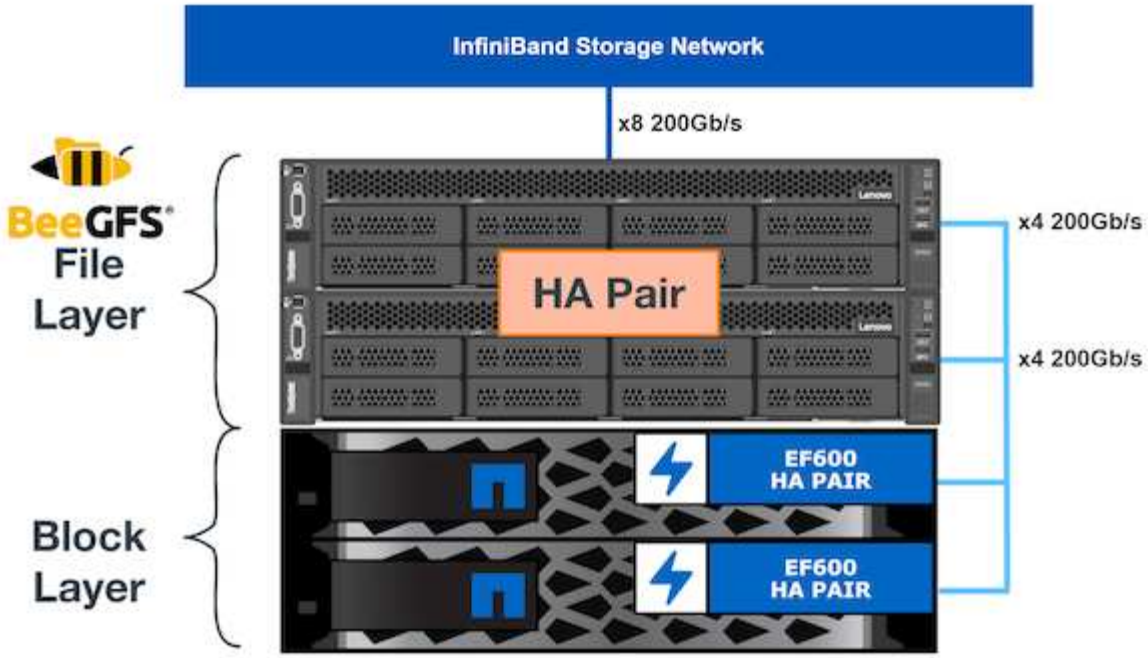
- Cada unidad escalable (SU) DGX SuperPOD consta de 32 sistemas DGX y es capaz de alcanzar 640 petaFLOPS de rendimiento de IA con precisión FP8. NetApp recomienda dimensionar la solución de almacenamiento NetApp BeeGFS con al menos 2 bloques de construcción para una única configuración DGX SuperPOD.

Una visión de alto nivel de la solución



- Los bloques de construcción NetApp BeeGFS constan de dos matrices NetApp EF600 y dos servidores x86:
  - Con las matrices all-flash NetApp EF600 en la base de NVIDIA DGX SuperPOD, los clientes obtienen una base de almacenamiento confiable respaldada por seis nueves de tiempo de actividad.
  - La capa del sistema de archivos entre los sistemas NetApp EF600 y NVIDIA DGX es el sistema de archivos paralelo BeeGFS. BeeGFS fue creado por el Centro Fraunhofer de Computación de Alto Rendimiento en Alemania para resolver los problemas de los sistemas de archivos paralelos heredados. El resultado es un sistema de archivos con una arquitectura de espacio de usuario moderna que ahora desarrolla y distribuye ThinkParQ y que utilizan muchos entornos de supercomputación.
  - El soporte de NetApp para BeeGFS alinea la excelente organización de soporte de NetApp con los requisitos del cliente en cuanto a rendimiento y tiempo de actividad. Los clientes obtienen acceso a recursos de soporte superiores, acceso anticipado a las versiones de BeeGFS y acceso a funciones empresariales seleccionadas de BeeGFS, como cumplimiento de cuotas y alta disponibilidad (HA).
- La combinación de las SU NVIDIA SuperPOD y los bloques de construcción BeeGFS de NetApp proporciona una solución de IA ágil en la que el cómputo o el almacenamiento se escalan de manera fácil y sin inconvenientes.

Bloque de construcción BeeGFS de NetApp



Resumen del caso de uso

Esta solución se aplica a los siguientes casos de uso:

- Inteligencia artificial (IA), que incluye aprendizaje automático (ML), aprendizaje profundo (DL), procesamiento del lenguaje natural (NLP), comprensión del lenguaje natural (NLU) e IA generativa (GenAI).
- Entrenamiento de IA a escala media y grande
- Modelos de visión artificial, habla, audio y lenguaje
- HPC incluye aplicaciones aceleradas por la interfaz de paso de mensajes (MPI) y otras técnicas de computación distribuida
- Cargas de trabajo de aplicaciones caracterizadas por lo siguiente:
  - Leer o escribir en archivos de más de 1 GB
  - Lectura o escritura en el mismo archivo por varios clientes (decenas, centenas y millares)
- Conjuntos de datos multiterabyte o multipetabyte
- Entornos que necesitan un único espacio de almacenamiento optimizable para una combinación de archivos grandes y pequeños

Requisitos tecnológicos

Esta sección cubre los requisitos tecnológicos para la solución NVIDIA DGX SuperPOD con NetApp .

Requisitos de hardware

En la Tabla 1 a continuación se enumeran los componentes de hardware necesarios para implementar la solución para una sola SU. El dimensionamiento de la solución comienza con 32 sistemas NVIDIA DGX H100 y dos o tres bloques de construcción NetApp BeeGFS. Un solo bloque de construcción NetApp BeeGFS consta de dos matrices NetApp EF600 y dos servidores x86. Los clientes pueden agregar bloques de

construcción adicionales a medida que aumenta el tamaño de la implementación. Para obtener más información, consulte la "Arquitectura de referencia NVIDIA DGX H100 SuperPOD" y "NVA-1164-DISEÑO: BeeGFS en el diseño de NVA de NetApp" .

Hardware	Cantidad
NVIDIA DGX H100 o H200	32
Conmutadores NVIDIA Quantum QM9700	8 hojas, 4 lomos
Bloques de construcción BeeGFS de NetApp	3

Requisitos de software

En la Tabla 2 a continuación se enumeran los componentes de software necesarios para implementar la solución. Los componentes de software que se utilizan en cualquier implementación particular de la solución pueden variar según los requisitos del cliente.

Software
Pila de software NVIDIA DGX
Administrador de comandos base de NVIDIA
Sistema de archivos paralelo ThinkParQ BeeGFS

Verificación de la solución

NVIDIA DGX SuperPOD con NetApp se validó en un clúster de aceptación dedicado en NVIDIA mediante el uso de bloques de construcción BeeGFS de NetApp . Los criterios de aceptación se basaron en una serie de pruebas de aplicación, rendimiento y estrés realizadas por NVIDIA. Para obtener más información, consulte la "NVIDIA DGX SuperPOD: Arquitectura de referencia NetApp EF600 y BeeGFS" .

Conclusión

NetApp y NVIDIA tienen una larga trayectoria de colaboración para ofrecer una cartera de soluciones de IA al mercado. NVIDIA DGX SuperPOD con la matriz all-flash NetApp EF600 es una solución probada y validada que los clientes pueden implementar con confianza. Esta arquitectura llave en mano totalmente integrada elimina el riesgo de la implementación y pone a cualquiera en el camino para ganar la carrera hacia el liderazgo en IA.

Dónde encontrar información adicional

Para obtener más información sobre la información que se describe en este documento, revise los siguientes documentos y/o sitios web:

- "Arquitectura de referencia NVIDIA DGX SuperPOD"
- "Guía de referencia de diseño de centros de datos NVIDIA DGX SuperPOD"
- "NVIDIA DGX SuperPOD: NetApp EF600 y BeeGFS"

Información de copyright

Copyright © 2025 NetApp, Inc. Todos los derechos reservados. Imprimido en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPCIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en http://www.netapp.com/TM son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.