NVIDIA DGX SuperPOD con NetApp - Guida alla progettazione

NetApp artificial intelligence solutions

NetApp
August 18, 2025

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

Sommario

NVIDIA DGX SuperPOD con NetApp - Guida alla progettazione . . . 1
Sintesi . . . 1
Panoramica della soluzione . . . 1
Tecnologia delle soluzioni . . . 1
Riepilogo del caso d’uso . . . 3
Requisiti tecnologici . . . 3
Requisiti hardware . . . 3
Requisiti software . . . 4
Verifica della soluzione . . . 4
Conclusione . . . 4
Dove trovare ulteriori informazioni . . . 4

NVIDIA DGX SuperPOD con NetApp - Guida alla progettazione

Questa architettura verificata da NetApp descrive la progettazione di NVIDIA DGX SuperPOD con i componenti di base NetApp BeeGFS. Questa soluzione è una piattaforma di data center full-stack convalidata su un cluster di accettazione dedicato presso NVIDIA.

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

Sintesi

Nell'attuale panorama tecnologico in rapida evoluzione, l'intelligenza artificiale sta rivoluzionando le esperienze dei consumatori e stimolando l'innovazione in tutti i settori. Tuttavia, presenta anche sfide significative per i reparti IT, che sono sotto pressione per implementare soluzioni di calcolo ad alte prestazioni (HPC) in grado di gestire le intense richieste dei carichi di lavoro dell'intelligenza artificiale. Mentre le organizzazioni si affrettano a sfruttare la potenza dell'intelligenza artificiale, cresce l'urgenza di una soluzione che sia facile da implementare, scalare e gestire.

NVIDIA DGX SuperPOD è una piattaforma infrastrutturale AI per data center, fornita come soluzione chiavi in mano per l'IT, per supportare i carichi di lavoro AI più complessi a cui devono far fronte le aziende odierne. Alla base di qualsiasi modello di deep learning (DL) accurato ci sono grandi volumi di dati, che richiedono una soluzione di archiviazione ad alta capacità in grado di gestire e rigestire in modo efficiente questi dati. La soluzione NetApp BeeGFS, composta da array di storage NetApp EF600 con file system parallelo BeeGFS, consente a NVIDIA DGX SuperPOD di sfruttare appieno le sue potenzialità. La soluzione NetApp BeeGFS è stata convalidata da NVIDIA per integrarsi e scalare con l'architettura SuperPOD. Il risultato è una distribuzione e una gestione semplificate dei data center basati sull'intelligenza artificiale, offrendo al contempo una scalabilità praticamente illimitata in termini di prestazioni e capacità.

Panoramica della soluzione

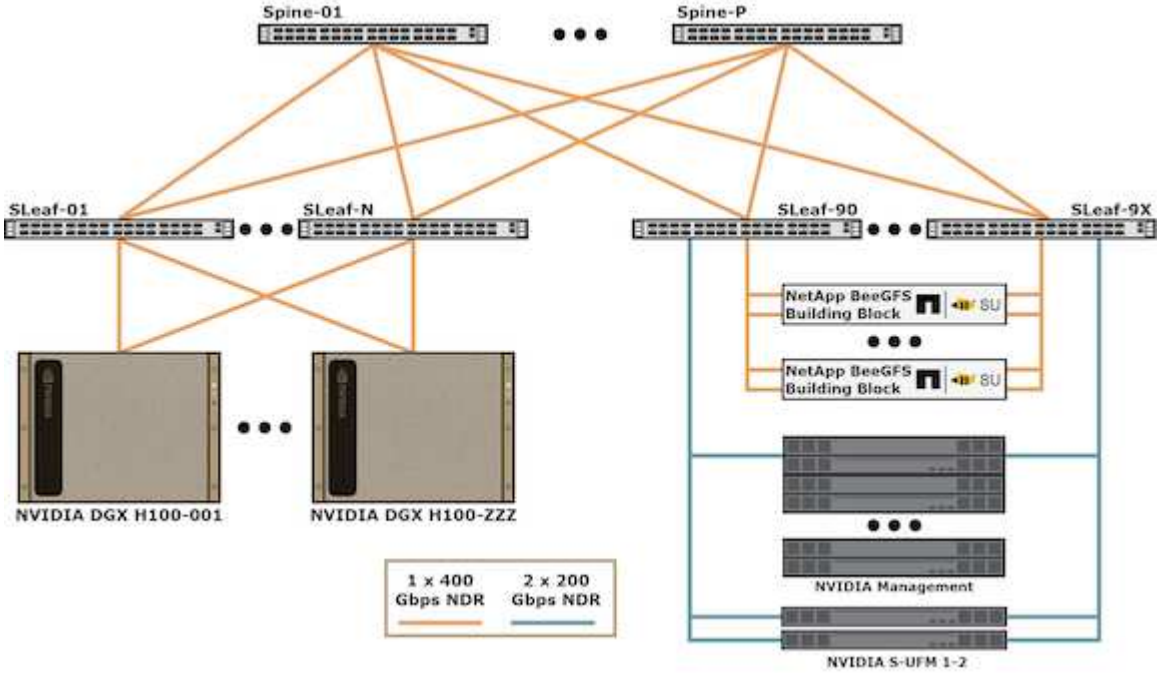
La soluzione NetApp BeeGFS, basata sui sistemi di storage NetApp EF600 NVMe ad alte prestazioni e sul file system parallelo scalabile BeeGFS, offre una base di storage solida ed efficiente per carichi di lavoro di intelligenza artificiale impegnativi. La sua architettura a disco condiviso garantisce un'elevata disponibilità, mantenendo prestazioni e accessibilità costanti, anche di fronte a sfide di sistema. Questa soluzione fornisce un'architettura scalabile e flessibile che può essere personalizzata per soddisfare diverse esigenze di archiviazione. I clienti possono espandere facilmente le prestazioni e la capacità di archiviazione integrando ulteriori blocchi di archiviazione per gestire anche i carichi di lavoro più impegnativi.

Tecnologia delle soluzioni

- NVIDIA DGX SuperPOD sfrutta i sistemi DGX H100 e H200 con un archivio condiviso validato collegato esternamente:

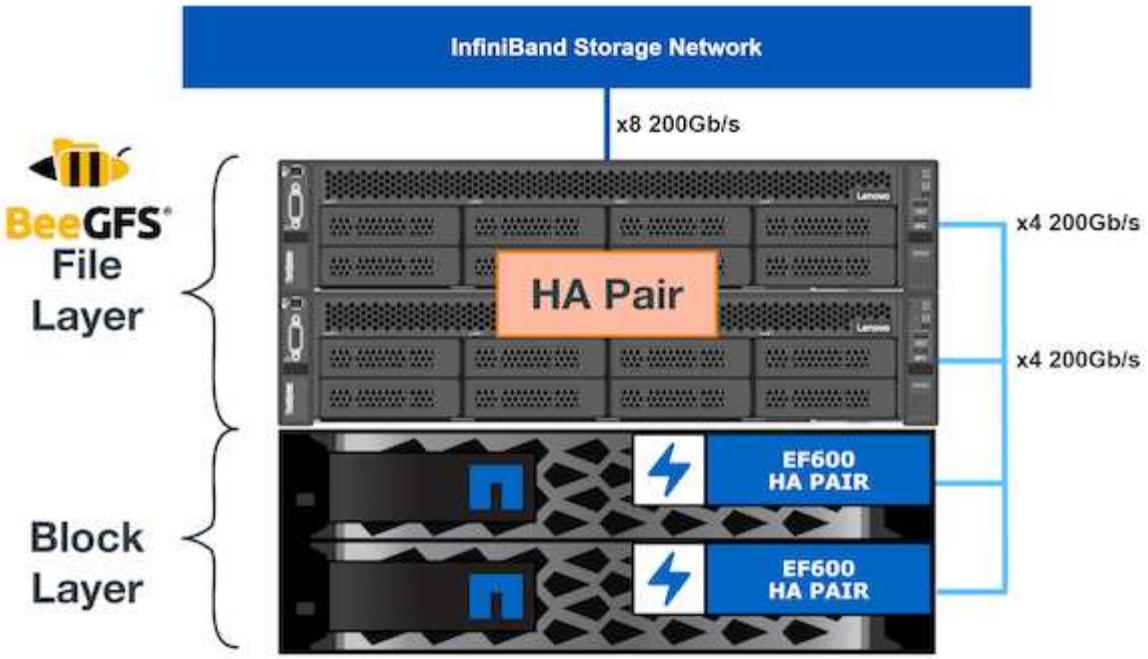
- Ogni unità scalabile (SU) DGX SuperPOD è composta da 32 sistemi DGX ed è in grado di raggiungere prestazioni di intelligenza artificiale pari a 640 petaFLOPS con precisione FP8. NetApp consiglia di dimensionare la soluzione di storage NetApp BeeGFS con almeno 2 blocchi di base per una singola configurazione DGX SuperPOD.

Una visione d'insieme della soluzione



- I blocchi di costruzione NetApp BeeGFS sono costituiti da due array NetApp EF600 e due server x86:
  - Grazie agli array all-flash NetApp EF600 alla base di NVIDIA DGX SuperPOD, i clienti ottengono una base di storage affidabile supportata da un uptime di sei 9.
  - Il livello del file system tra i sistemi NetApp EF600 e NVIDIA DGX è il file system parallelo BeeGFS. BeeGFS è stato creato dal Fraunhofer Center for High-Performance Computing in Germania per risolvere i problemi dei file system paralleli legacy. Il risultato è un file system con un'architettura moderna e user space, ora sviluppato e distribuito da ThinkParQ e utilizzato da molti ambienti di supercalcolo.
  - Il supporto NetApp per BeeGFS allinea l'eccellente organizzazione di supporto di NetApp alle esigenze dei clienti in termini di prestazioni e tempi di attività. I clienti hanno accesso a risorse di supporto di livello superiore, accesso anticipato alle versioni di BeeGFS e accesso a funzionalità aziendali selezionate di BeeGFS, come l'applicazione delle quote e l'alta disponibilità (HA).
- La combinazione di NVIDIA SuperPOD SU e dei componenti di base NetApp BeeGFS fornisce una soluzione di intelligenza artificiale agile in cui l'elaborazione o l'archiviazione sono scalabili in modo semplice e fluido.

Blocco di costruzione NetApp BeeGFS



Riepilogo del caso d'uso

Questa soluzione si applica ai seguenti casi d'uso:

- Intelligenza artificiale (IA), tra cui apprendimento automatico (ML), apprendimento profondo (DL), elaborazione del linguaggio naturale (NLP), comprensione del linguaggio naturale (NLU) e IA generativa (GenAI).
- Formazione AI su media e larga scala
- Modelli di visione artificiale, parlato, audio e linguaggio
- HPC, comprese le applicazioni accelerate dall'interfaccia di passaggio dei messaggi (MPI) e altre tecniche di elaborazione distribuita
- Carichi di lavoro applicativi caratterizzati da quanto segue:
  - Lettura o scrittura di file più grandi di 1 GB
  - Lettura o scrittura sullo stesso file da parte di più client (decine, centinaia e migliaia)
- Set di dati multiterabyte o multipetabyte
- Ambienti che necessitano di un singolo namespace di archiviazione ottimizzabile per un mix di file di grandi e piccole dimensioni

Requisiti tecnologici

Questa sezione illustra i requisiti tecnologici per la soluzione NVIDIA DGX SuperPOD con NetApp .

Requisiti hardware

Nella tabella 1 sottostante sono elencati i componenti hardware necessari per implementare la soluzione per una singola SU. Il dimensionamento della soluzione inizia con 32 sistemi NVIDIA DGX H100 e due o tre blocchi di costruzione NetApp BeeGFS. Un singolo blocco di costruzione NetApp BeeGFS è costituito da due array NetApp EF600 e due server x86. I clienti possono aggiungere ulteriori componenti man mano che

aumentano le dimensioni dell'implementazione. Per maggiori informazioni, vedere il "Architettura di riferimento NVIDIA DGX H100 SuperPOD" E "NVA-1164-DESIGN: BeeGFS su NetApp NVA Design" .

Hardware	Quantità
NVIDIA DGX H100 o H200	32
Switch NVIDIA Quantum QM9700	8 fogli, 4 spine
Blocchi di costruzione NetApp BeeGFS	3

Requisiti software

Nella tabella 2 sottostante sono elencati i componenti software necessari per implementare la soluzione. I componenti software utilizzati in una particolare implementazione della soluzione potrebbero variare in base alle esigenze del cliente.

Software
Stack software NVIDIA DGX
Gestore dei comandi di base NVIDIA
Sistema di file parallelo ThinkParQ BeeGFS

Verifica della soluzione

NVIDIA DGX SuperPOD con NetApp è stato convalidato su un cluster di accettazione dedicato presso NVIDIA utilizzando i blocchi di costruzione NetApp BeeGFS. I criteri di accettazione si basavano su una serie di test applicativi, prestazionali e di stress eseguiti da NVIDIA. Per maggiori informazioni, vedere il "NVIDIA DGX SuperPOD: architettura di riferimento NetApp EF600 e BeeGFS" .

Conclusione

NetApp e NVIDIA collaborano da lungo tempo per offrire al mercato un portafoglio di soluzioni di intelligenza artificiale. NVIDIA DGX SuperPOD con l'array all-flash NetApp EF600 è una soluzione collaudata e convalidata che i clienti possono implementare con sicurezza. Questa architettura completamente integrata e chiavi in mano elimina i rischi derivanti dall'implementazione e consente a chiunque di vincere la corsa alla leadership dell'intelligenza artificiale.

Dove trovare ulteriori informazioni

Per saperne di più sulle informazioni descritte nel presente documento, consultare i seguenti documenti e/o siti web:

- "Architettura di riferimento NVIDIA DGX SuperPOD"
- "Guida di riferimento alla progettazione del data center NVIDIA DGX SuperPOD"
- "NVIDIA DGX SuperPOD: NetApp EF600 e BeeGFS"

Informazioni sul copyright

Copyright © 2025 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEQUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina http://www.netapp.com/TM sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.