NVIDIA DGX SuperPOD avec NetApp - Guide de conception

NetApp artificial intelligence solutions

NetApp
August 18, 2025

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

Sommaire

NVIDIA DGX SuperPOD avec NetApp - Guide de conception . . . 1
Résumé exécutif . . . 1
Présentation de la solution . . . 1
Technologie des solutions . . . 1
Résumé du cas d’utilisation . . . 3
Exigences technologiques . . . 3
Configuration matérielle requise . . . 3
Configuration logicielle requise . . . 4
Vérification de la solution . . . 4
Conclusion . . . 4
Où trouver des informations supplémentaires . . . 4

NVIDIA DGX SuperPOD avec NetApp - Guide de conception

Cette architecture vérifiée NetApp décrit la conception du NVIDIA DGX SuperPOD avec les blocs de construction NetApp BeeGFS. Cette solution est une plateforme de centre de données full-stack validée sur un cluster d'acceptation dédié chez NVIDIA.

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

Résumé exécutif

Dans le paysage technologique actuel en évolution rapide, l'IA révolutionne les expériences des consommateurs et stimule l'innovation dans tous les secteurs. Cependant, cela présente également des défis importants pour les services informatiques, qui sont sous pression pour déployer des solutions de calcul haute performance (HPC) capables de gérer les demandes intenses des charges de travail de l'IA. Alors que les organisations se précipitent pour exploiter la puissance de l'IA, l'urgence d'une solution facile à déployer, à faire évoluer et à gérer augmente.

NVIDIA DGX SuperPOD est une plate-forme d'infrastructure de centre de données IA fournie sous forme de solution clé en main pour l'informatique afin de prendre en charge les charges de travail IA les plus complexes auxquelles sont confrontées les entreprises d'aujourd'hui. Au cœur de tout modèle d'apprentissage profond (DL) précis se trouvent de grands volumes de données, nécessitant une solution de stockage à haut débit capable de servir et de réutiliser efficacement ces données. La solution NetApp BeeGFS, composée de baies de stockage NetApp EF600 avec le système de fichiers parallèle BeeGFS, permet au NVIDIA DGX SuperPOD de libérer toutes ses capacités. La solution NetApp BeeGFS a été validée par NVIDIA pour s'intégrer et évoluer avec l'architecture SuperPOD. Le résultat est un déploiement et une gestion simplifiés du centre de données IA tout en offrant une évolutivité pratiquement illimitée en termes de performances et de capacité.

Présentation de la solution

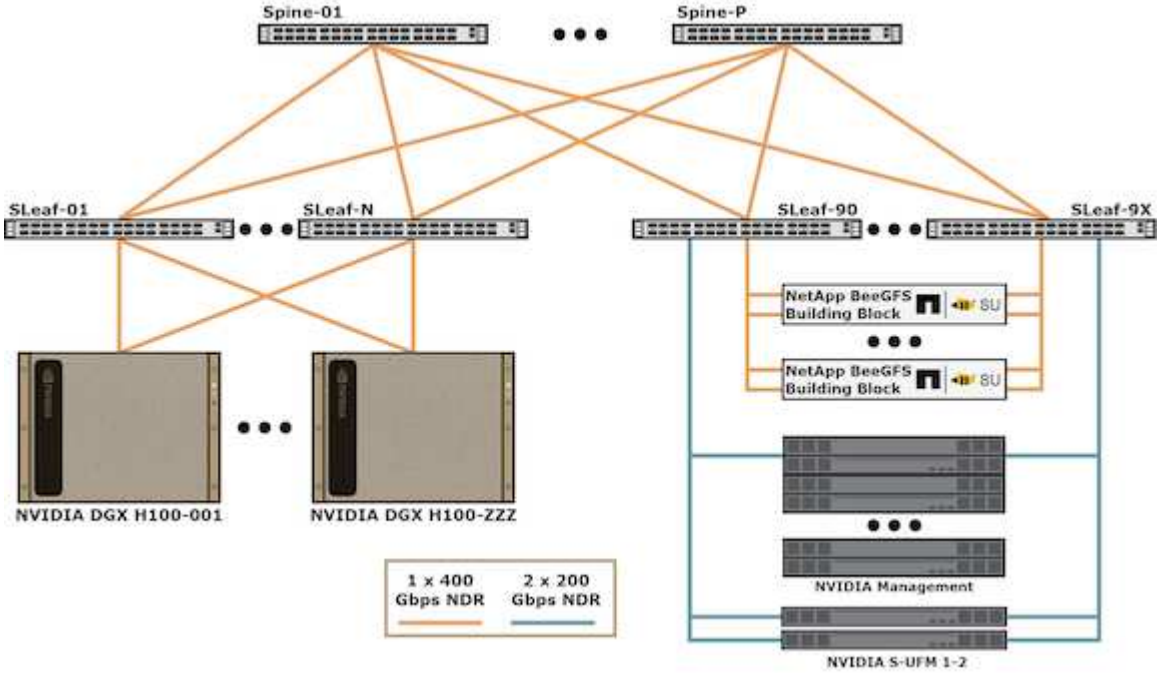
La solution NetApp BeeGFS, optimisée par les systèmes de stockage NVMe NetApp EF600 hautes performances et le système de fichiers parallèles évolutif BeeGFS, offre une base de stockage robuste et efficace pour les charges de travail d'IA exigeantes. Son architecture de disque partagé garantit une haute disponibilité, en maintenant des performances et une accessibilité constantes, même face aux défis du système. Cette solution fournit une architecture évolutive et flexible qui peut être personnalisée pour répondre à diverses exigences de stockage. Les clients peuvent facilement étendre leurs performances et leur capacité de stockage en intégrant des blocs de construction de stockage supplémentaires pour gérer même les charges de travail les plus exigeantes.

Technologie des solutions

- NVIDIA DGX SuperPOD exploite les systèmes DGX H100 et H200 avec un stockage partagé externe validé :

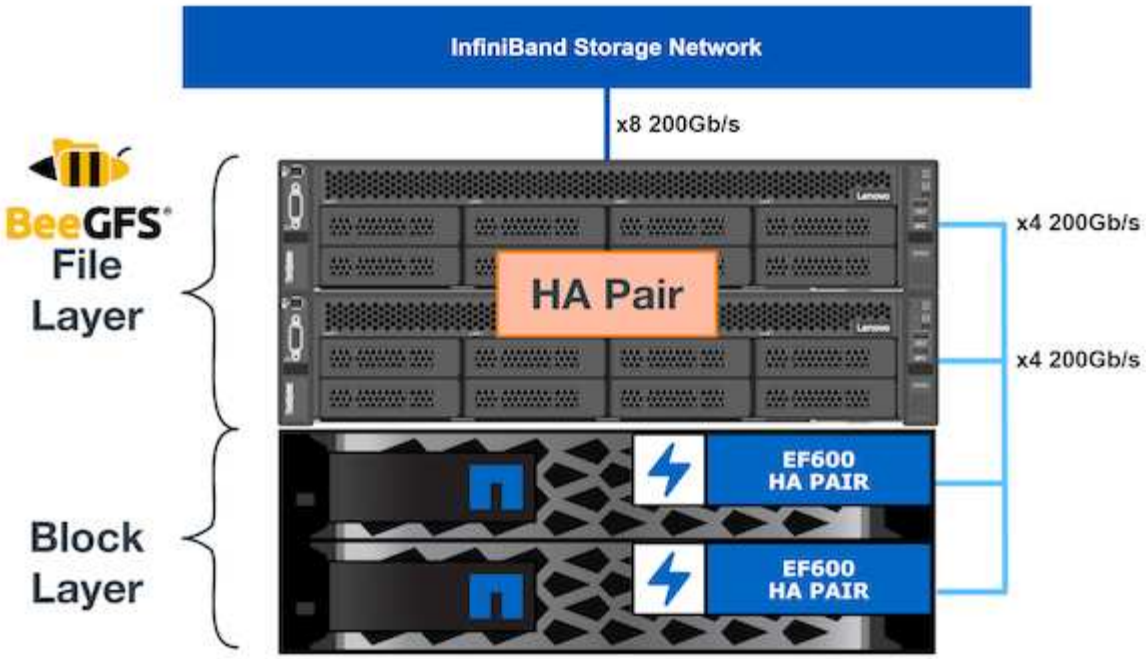
- Chaque unité évolutive (SU) DGX SuperPOD se compose de 32 systèmes DGX et est capable de 640 pétaFLOPS de performances d'IA avec une précision FP8. NetApp recommande de dimensionner la solution de stockage NetApp BeeGFS avec au moins 2 blocs de construction pour une seule configuration DGX SuperPOD.

Une vue d'ensemble de la solution



- Les blocs de construction NetApp BeeGFS se composent de deux baies NetApp EF600 et de deux serveurs x86 :
  - Avec les baies 100 % flash NetApp EF600 à la base de NVIDIA DGX SuperPOD, les clients bénéficient d'une base de stockage fiable soutenue par six 9 de disponibilité.
  - La couche du système de fichiers entre les systèmes NetApp EF600 et NVIDIA DGX est le système de fichiers parallèle BeeGFS. BeeGFS a été créé par le Centre Fraunhofer pour le calcul haute performance en Allemagne pour résoudre les problèmes des systèmes de fichiers parallèles hérités. Le résultat est un système de fichiers avec une architecture d'espace utilisateur moderne qui est désormais développé et fourni par ThinkParQ et utilisé par de nombreux environnements de supercalcul.
  - Le support NetApp pour BeeGFS aligne l'excellente organisation de support de NetApp sur les exigences des clients en matière de performances et de disponibilité. Les clients ont accès à des ressources d'assistance supérieures, à un accès anticipé aux versions de BeeGFS et à l'accès à certaines fonctionnalités d'entreprise de BeeGFS telles que l'application des quotas et la haute disponibilité (HA).
- La combinaison des SU NVIDIA SuperPOD et des blocs de construction NetApp BeeGFS fournit une solution d'IA agile dans laquelle le calcul ou le stockage évolue facilement et de manière transparente.

Bloc de construction NetApp BeeGFS



Résumé du cas d'utilisation

Cette solution s'applique aux cas d'utilisation suivants :

- Intelligence artificielle (IA) comprenant l'apprentissage automatique (ML), l'apprentissage profond (DL), le traitement du langage naturel (NLP), la compréhension du langage naturel (NLU) et l'IA générative (GenAI).
- Formation en IA à moyenne et grande échelle
- Vision par ordinateur, parole, audio et modèles de langage
- HPC incluant les applications accélérées par l'interface de passage de messages (MPI) et d'autres techniques de calcul distribué
- Charges de travail des applications caractérisées par les éléments suivants :
  - Lecture ou écriture dans des fichiers de plus de 1 Go
  - Lecture ou écriture dans le même fichier par plusieurs clients (10, 100 et 1000)
- Ensembles de données multitéraoctets ou multipétaoctets
- Environnements nécessitant un espace de stockage unique optimisable pour un mélange de fichiers volumineux et petits

Exigences technologiques

Cette section couvre les exigences technologiques pour la solution NVIDIA DGX SuperPOD avec NetApp .

Configuration matérielle requise

Le tableau 1 ci-dessous répertorie les composants matériels nécessaires à la mise en œuvre de la solution pour un seul SU. Le dimensionnement de la solution commence avec 32 systèmes NVIDIA DGX H100 et deux ou trois blocs de construction NetApp BeeGFS. Un seul bloc de construction NetApp BeeGFS se compose de deux baies NetApp EF600 et de deux serveurs x86. Les clients peuvent ajouter des blocs de construction

supplémentaires à mesure que la taille du déploiement augmente. Pour plus d'informations, consultez le "Architecture de référence NVIDIA DGX H100 SuperPOD" et "NVA-1164-DESIGN : Conception de NVA BeeGFS sur NetApp" .

Matériel	Quantité
NVIDIA DGX H100 ou H200	32
Commutateurs NVIDIA Quantum QM9700	8 feuilles, 4 épines
Blocs de construction NetApp BeeGFS	3

Configuration logicielle requise

Le tableau 2 ci-dessous répertorie les composants logiciels nécessaires à la mise en œuvre de la solution. Les composants logiciels utilisés dans une implémentation particulière de la solution peuvent varier en fonction des exigences du client.

Logiciels
Pile logicielle NVIDIA DGX
Gestionnaire de commandes de base NVIDIA
Système de fichiers parallèle ThinkParQ BeeGFS

Vérification de la solution

NVIDIA DGX SuperPOD avec NetApp a été validé sur un cluster d'acceptation dédié chez NVIDIA en utilisant les blocs de construction NetApp BeeGFS. Les critères d'acceptation étaient basés sur une série de tests d'application, de performances et de stress effectués par NVIDIA. Pour plus d'informations, consultez le "NVIDIA DGX SuperPOD: architecture de référence NetApp EF600 et BeeGFS" .

Conclusion

NetApp et NVIDIA ont une longue histoire de collaboration pour proposer un portefeuille de solutions d'IA sur le marché. NVIDIA DGX SuperPOD avec la baie entièrement flash NetApp EF600 est une solution éprouvée et validée que les clients peuvent déployer en toute confiance. Cette architecture clé en main entièrement intégrée élimine les risques liés au déploiement et met tout le monde sur la voie de la victoire dans la course au leadership de l'IA.

Où trouver des informations supplémentaires

Pour en savoir plus sur les informations décrites dans ce document, consultez les documents et/ou sites Web suivants :

- "Architecture de référence NVIDIA DGX SuperPOD"
- "Guide de référence de conception du centre de données NVIDIA DGX SuperPOD"
- "NVIDIA DGX SuperPOD: NetApp EF600 et BeeGFS"

Informations sur le copyright

Copyright © 2025 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site http://www.netapp.com/TM sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.