



Dell propose une gamme plus étendue de solutions, de services et d'options de consommation pour prendre en charge les déploiements d'IA des clients



Jusqu'à 37 % de performances en plus dans la moitié de l'espace rack pour le traitement du langage naturel (GPT-J) selon les données MLperf



Les conceptions validées de Dell comprennent plus d'architectures de référence axées sur l'IA que les solutions validées par Supermicro

Trouver la voie de la réussite en matière d'IA avec la gamme Dell de solutions d'IA

Comparaison de la gamme Dell AI par rapport aux offres similaires de Supermicro

Pour implémenter l'intelligence artificielle (IA) et en tirer parti, vous devez prendre en compte tous les aspects d'une solution. Il est alors essentiel de trouver un partenaire capable d'assurer une prise en charge de bout en bout, de la planification à la préparation des données, en passant par la sélection du matériel, la conception des modèles d'IA, les tests de validation de principe, les architectures de référence et le support complet. Dans cette synthèse, nous mettons en évidence les avantages des serveurs des gammes IA de Dell™ en matière de performances et d'espace par rapport à ceux de Supermicro, en nous basant sur les résultats des tests d'analyse comparative de MLPerf® accessibles au public. MLPerf® teste les performances d'entraînement et d'inférence sur plusieurs modèles d'IA. Les données de cette synthèse proviennent des résultats MLPerf® v3.1 Inference Datacenter publiés sur le site Web MLCommons® depuis novembre 2023.¹ Nous comparons également les informations accessibles au public sur la gamme de produits et de solutions proposés par chaque fournisseur pour prendre en charge les déploiements d'IA.

Serveurs à quatre processeurs graphiques : les serveurs Dell PowerEdge offrent de meilleures performances MLPerf que les serveurs Supermicro

Pouvant être équipé de quatre processeurs graphiques NVIDIA H100 SXM, le serveur Dell PowerEdge XE9640 offre la moitié de la puissance de calcul du processeur graphique du serveur PowerEdge XE9680 dans un tiers de l'espace.² Le serveur à quatre processeurs graphiques équipé de processeurs graphiques NVIDIA HGX H100 de Supermicro évalué par MLPerf® est le serveur SYS-421GU-TNXR, qui est un serveur 4U. Dans les résultats publiés pour le modèle d'IA gptj-99.9, le serveur PowerEdge XE9640 a obtenu de meilleures performances que le serveur Supermicro dans les tests hors ligne, atteignant jusqu'à 1,37 fois le score (voir Figure 1).

Nous avons également comparé les résultats du serveur Supermicro SYS-421GU-TNXR à ceux du serveur Dell PowerEdge XE8640, un serveur 4U équipé de quatre processeurs graphiques qui prend également en charge les processeurs graphiques NVIDIA H100 HGX. Le serveur Dell PowerEdge XE8640 est équipé des tout derniers processeurs Intel Xeon Scalable de 4e génération et d'une mémoire pouvant atteindre 4 To pour gérer les jeux de données volumineux et les calculs complexes fréquents en IA et en analytique des données.³ Avec son format 4U, le serveur PowerEdge XE8640 est similaire au serveur Supermicro SYS-421GU-TNXR à la fois en matière de densité et de capacités des processeurs graphiques. Cependant, comme nous l'avons vu avec le serveur PowerEdge XE9640, le modèle Dell PowerEdge XE8640 a obtenu de meilleurs scores gptj-99 que le serveur Supermicro dans les tests hors ligne (voir Figure 2).

Résultats MLPerf® normalisés : Dell PowerEdge XE9640 avec H100 SXM5 par rapport à Supermicro

SYS-421GU-TNXR avec H100 SXM5
(plus les valeurs sont élevées, mieux c'est)

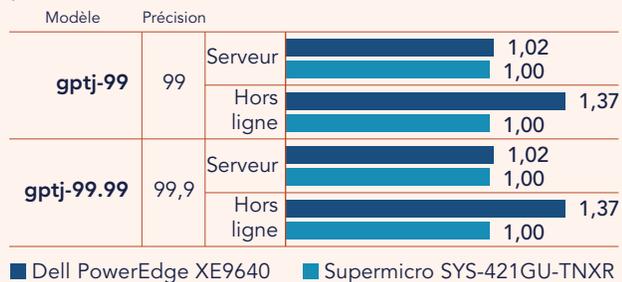


Figure 1 : Résultats MLPerf® publiés pour les serveurs Dell PowerEdge XE9640 et Supermicro SYS-421GU-TNXR en date du 29 novembre 2023. Les deux systèmes utilisent le format SXM du processeur graphique NVIDIA H100. Source : Principled Technologies d'après des données de MLCommons®.*

Résultats MLPerf® normalisés : Dell PowerEdge XE8640 avec H100 SXM5 par rapport à Supermicro

SYS-421GU-TNXR avec H100 SXM5
(plus les valeurs sont élevées, mieux c'est)

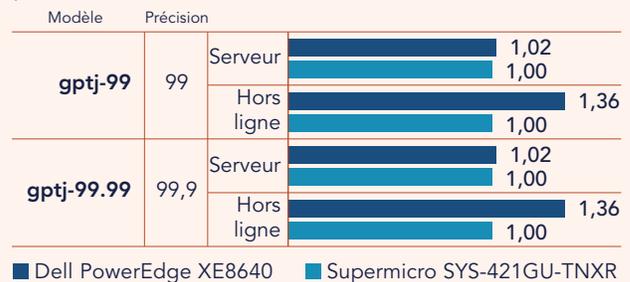


Figure 2 : Résultats MLPerf® publiés pour les serveurs Dell PowerEdge XE8640 et Supermicro SYS-421GU-TNXR en date du 29 novembre 2023. Les deux systèmes utilisent le format SXM du processeur graphique NVIDIA H100. Source : Principled Technologies d'après des données de MLCommons®.*

Au-delà du matériel : avoir une vue d'ensemble

Les organisations à la recherche de flexibilité dans les offres de stations de travail équipées de processeurs graphiques peuvent trouver une gamme d'options plus large dans le portefeuille IA de Dell. Plus que du matériel, la gamme d'IA de Dell offre aux clients une solution unique de services et de support pour chaque étape de leur déploiement. En revanche, la gamme d'IA de Supermicro offre des services limités aux manuels, au dépannage et aux retours de produits. Dans le cadre de nos recherches, nous n'avons pas pu trouver de services de conception, d'implémentation, de gestion ou de formation dans la gamme d'IA de Supermicro. L'adoption de l'IA étant très complexe, les services proposés par Dell en font un partenaire de formation plus prometteur pour le déploiement de ces charges applicatives. De plus, nous avons constaté que les conceptions validées de Dell incluent des architectures de référence plus centrées sur l'IA que les solutions validées par Supermicro et qu'elles proposent des conseils plus détaillés.

Lisez notre rapport complet, qui explore en détail les avantages offerts par la gamme d'IA de Dell en matière d'options de calcul, de stockage et de gestion de réseau, de services professionnels pour l'IA, de partenariats tiers pour les charges applicatives d'IA, de conceptions validées Dell qui aident à éliminer les conjectures liées aux solutions matérielles d'IA, ainsi que de services de gestion et d'iDRAC.

*Score vérifié MLPerf v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrées 3.1-0133, 3.1-0066 et 3.1-0067. Le nom et le logo MLPerf sont des marques et des marques déposées de MLCommons Association aux États-Unis et dans d'autres pays. Tous droits réservés. Utilisation non autorisée strictement interdite. Pour plus d'informations, voir le site www.mlcommons.org.

1. MLCommons, « MLPerf Inference: Datacenter Benchmark Suite Results », consulté le 7 février 2024, <https://mlcommons.org/en/inference-datacenter-31/>.
2. Dell, « Serveur au format rack PowerEdge XE9640 », consulté le 5 janvier 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
3. Dell, « Serveur au format rack PowerEdge XE8640 », consulté le 5 janvier 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.

► Consultez la version d'origine en anglais de cette synthèse à l'adresse <https://facts.pt/kRTwky3>

Cliquez ici pour lire le rapport ►



Facts matter.®

Principled Technologies est une marque déposée de Principled Technologies, Inc. Tous les autres noms de produit sont des marques déposées par leurs propriétaires respectifs. Pour plus d'informations, consultez le rapport.