

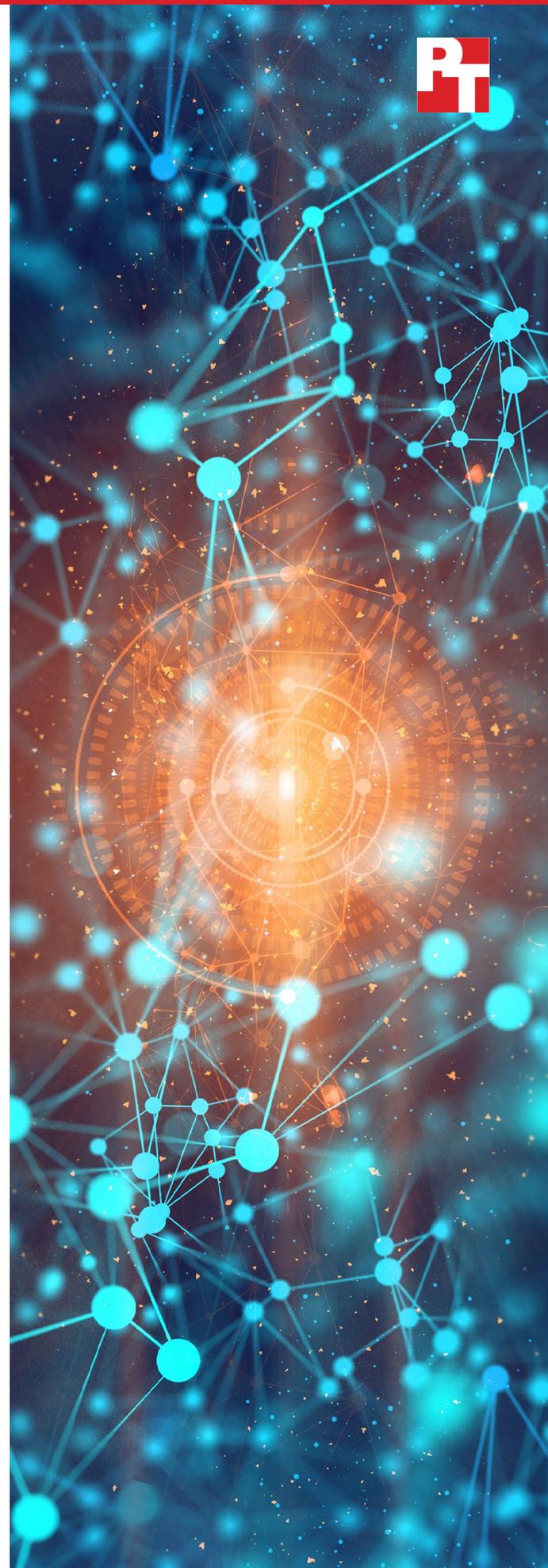


# Trouver la voie de la réussite en matière d'IA avec la gamme Dell de solutions d'IA

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

L'intelligence artificielle (IA) ouvre de nouveaux horizons et est en passe de transformer les activités métier dans tous les secteurs. Alors que les organisations explorent les moyens de tirer parti de l'IA pour améliorer leurs activités, il est important de rappeler qu'implémenter l'IA et en récolter les bénéfices ne se fait pas du jour au lendemain. Chaque entreprise est unique. Il incombe donc à chacune d'évaluer ses données et ses objectifs métier pour voir comment l'utilisation de l'IA peut produire les résultats spécifiques souhaités. Faire appel à une société comme Dell qui propose une gamme complète de solutions d'IA, notamment pour la planification, la préparation des données, la sélection du matériel approprié, la conception du modèle d'IA, les tests de validation technique, les architectures de référence et le support de bout en bout peut se traduire par des projets couronnés de succès.

Compte tenu de la myriade d'options disponibles sur le marché, trouver un partenaire capable d'aider dans toutes ces décisions peut faire la différence entre une mise en œuvre réussie de l'IA et une erreur coûteuse. Dans ce rapport, nous nous intéressons de plus près aux gammes de solutions d'IA de Dell et de Supermicro dans le but d'informer les lecteurs sur les avantages que Dell peut offrir au client tout au long de son parcours vers l'adoption de l'IA. Nous nous concentrerons d'abord sur les options de serveur et de calcul, un segment dans lequel chaque société propose une palette d'offres riches et variées. Nous verrons ensuite comment Dell va au-delà des simples considérations matérielles pour répondre aux besoins des sociétés en matière de formation, de services de planification, d'écosystèmes de partenaires, etc.



# Serveurs et performances pour les charges applicatives d'IA

Les serveurs, l'infrastructure informatique de base qui alimente les charges applicatives d'IA, peuvent utiliser comme ressources de calcul des processeurs, des processeurs graphiques, ou les deux, en fonction de la taille ou du type de charge applicative. Pour les charges applicatives plus importantes ou plus exigeantes comme les charges HPC ou d'IA, les processeurs graphiques offrent des performances optimales. Ceux-ci se présentent sous différents formats : PCIe universel, module Open Compute Project Accelerator Module (OAM) et architecture propriétaire NVIDIA SXM, qui offre actuellement des performances optimales.<sup>1</sup> Les capacités étendues de mémoire et les caractéristiques de conception des serveurs, comme l'architecture de refroidissement et l'efficacité énergétique, jouent également sur les performances. La plupart des datacenters sont encore refroidis par air, ce qui signifie que les charges applicatives d'IA ont besoin de serveurs conçus pour assurer un refroidissement par air aussi efficace que possible. Nous présentons ci-dessous des offres de serveurs Dell PowerEdge avec leurs caractéristiques (composants et options de refroidissement, entre autres), ainsi que les scores MLPerf<sup>®</sup> publiés par MLCommons<sup>®</sup>.

## Résultats des tests

MLPerf<sup>®</sup> est une suite d'analyse comparative de points de référence qui teste les performances de l'IA à la fois en matière d'entraînement et d'inférence. Pour qu'une organisation puisse publier les résultats MLPerf<sup>®</sup> officiels, ceux-ci doivent être conformes aux conditions spécifiques définies par le développeur de l'outil d'analyse comparative des points de référence, MLCommons<sup>®</sup>.<sup>2</sup> Ces directives de conformité fournissent des normes qui facilitent la comparaison des performances. Pour les tests d'inférence, MLPerf<sup>®</sup> utilise des jeux de données provenant du datacenter, de la périphérie, de données mobiles et de microdonnées, et indique les scores de l'IA et l'énergie consommée en watts pendant les tests. Cette suite d'analyse comparative de points de référence d'inférence comprend des tests de nombreux modèles courants d'IA, d'apprentissage automatique (ML) et de Deep Learning (DL) (voir Tableau 1).

Tableau 1 : Modèles d'IA, de ML et de DL pris en charge par MLPerf<sup>®</sup>, avec tests et cas d'utilisation types pour chacun. Source : Principled Technologies.

Modèles d'IA courants	Exemples d'utilisation classiques
ResNet	Modèle de classification d'images qui aide les ordinateurs à apprendre, à mémoriser et à identifier différentes images pour des cas d'utilisation tels que l'imagerie médicale, la modération de contenus sur les réseaux sociaux et la reconnaissance faciale.
RetinaNet	Type de détection d'objets capable de gérer une plus grande complexité que ResNet. Ce modèle aide les ordinateurs à identifier et à localiser des objets dans des images ou dans des vidéos, et peut les classer par ordre d'importance. Il est utilisé pour des applications telles que les véhicules autonomes, les technologies d'aide à la conduite, la surveillance et la reconnaissance faciale.
3D-UNet	Modèle de segmentation spécifique utilisé dans le domaine de l'imagerie médicale.
RNN-T	Reconnaissance vocale pour des cas d'utilisation comme la traduction automatique de langues.
BERT	Traitement du langage naturel pour des cas d'utilisation comme la rédaction de synthèses de texte, la traduction de langues et l'exécution automatique de tâches.
DLRM-v2-99.9	Modèle de recommandation dans des cas d'utilisation tels que les publicités ciblées et les recommandations de produits personnalisés.
GPTJ-99 et 99.9	LLM pour le traitement du langage naturel qui excelle dans la génération de texte pour des cas d'utilisation comme les chatbots et les outils d'IA basés sur le chat



## À propos de MLPerf

Les résultats MLPerf® prennent en compte plusieurs paramètres en plus des modèles d'IA eux-mêmes, ce qui peut représenter beaucoup de données à analyser dans un seul graphique ou tableau. Voici une brève explication de ces paramètres :

- 99,0 et 99,9 : ces mesures indiquent la précision d'entraînement du modèle. Plus vous avez besoin que le résultat soit précis, plus le modèle sera complexe, et plus le traitement des données sera long.
- Échantillons hors ligne par seconde : mode dans lequel l'outil d'analyse comparative des points de référence envoie toutes les requêtes au début du test, simulant les données déjà présentes sur le système.
- Requêtes du serveur par seconde : mode dans lequel l'outil d'analyse comparative des points de référence envoie des requêtes pendant toute la durée du test, simulant l'analyse d'un flux de données en direct.

Pour en savoir plus sur MLCommons® et les résultats MLPerf®, consultez <https://mlcommons.org/benchmarks/inference-datacenter/>.

Les résultats présentés dans ce rapport proviennent des résultats de MLPerf® v3.1 Inference Datacenter publiés sur le site Web de MLCommons® en novembre 2023.<sup>3</sup> Ces résultats se fondent sur des informations soumises par des fabricants de technologies et des prestataires de services Cloud et couvrent un large éventail de configurations. En comparaison avec les informations soumises accessibles au public fournies par Supermicro, les serveurs Dell PowerEdge ont produit des résultats comparables. Le Tableau 2 répertorie des informations détaillées sur les serveurs.

Tableau 2 : Serveurs Dell et Supermicro évalués pour produire les résultats MLPerf® 3.1 publiés par MLCommons® en novembre 2023.  
Source : Principled Technologies.

Fournisseur	Modèle du serveur	Référence et modèle de processeurs graphiques	Description
Dell <sup>4</sup>	PowerEdge XE9680	8 NVIDIA H100 SXM	Pour l'entraînement de l'IA et l'inférence avec d'importantes charges applicatives, par exemple les grands modèles de langage
	PowerEdge XE9640	4 NVIDIA H100 SXM	Pour l'entraînement de grands modèles d'IA dans des datacenters haute densité et à refroidissement liquide
	PowerEdge XE8640	4 NVIDIA H100 SXM	Pour piloter l'entraînement traditionnel de l'IA, les applications HPC et d'analytique des données dans un format 4U pour les datacenters refroidis par air
Supermicro <sup>5</sup>	**AS-8125GS-TNHR	8 NVIDIA H100 SXM	Pour l'entraînement de l'IA à grande échelle et les charges applicatives HPC avec processeurs AMD
	**SYS-821GE-TNHR	8 NVIDIA H100 SXM	Pour l'entraînement de l'IA à grande échelle et les charges applicatives HPC avec processeurs Intel
	**SYS-421GU-TNXR	4 NVIDIA H100 SXM	Conception modulaire pour la flexibilité de prise en charge des charges applicatives HPC et IA

Dell et Supermicro ayant tous deux présenté des résultats avec des configurations de processeurs graphiques identiques, la comparaison des performances se fait aisément. Comme le montrent les Figures 1 et 2, la similitude de configuration des deux fournisseurs produit des résultats sensiblement équivalents. Pour d'autres configurations, par exemple celles présentées aux Figures 3 et 4, Dell a obtenu de meilleures performances que Supermicro au test de la configuration à 4 processeurs graphiques sur le modèle GPTJ-99.9. Il convient de noter que Dell a présenté des résultats pour tous les modèles disponibles avec les trois serveurs, ce qui n'est pas le cas de Supermicro. Nous comparons uniquement les modèles pour lesquels les deux serveurs ont produit des résultats. Pour consulter l'ensemble des résultats de Dell, reportez-vous aux résultats MLPerf® MLCommons®.

## Résultats des serveurs équipés de huit processeurs graphiques

Le serveur Dell PowerEdge XE9680 prend en charge jusqu'à huit processeurs graphiques NVIDIA H100 SXM5 pour l'accélération de l'IA et jusqu'à deux processeurs Intel® Xeon® Scalable de 4<sup>e</sup> génération. La famille de produits PowerEdge XE présente une architecture modulaire qui prend en charge des processeurs graphiques NVIDIA SXM4 ou SXM5 ou des assemblages de processeurs graphiques Open Compute Project Accelerator Module (OAM), qui peuvent optimiser les performances par rapport à celles d'un processeur graphique PCIe standard. Le serveur Dell PowerEdge XE9680 propose également l'accélérateur AMD Instinct™ MI300X.<sup>6</sup> Occupant seulement 6U de l'espace rack, le système PowerEdge XE9680 est un serveur NVIDIA H100 SXM5 huit voies compact.

En comparaison, le serveur Supermicro équipé de 8 processeurs graphiques occupe 33 % plus d'espace rack avec 8U alors qu'il propose le même format SXM pour les processeurs graphiques NVIDIA. Autrement dit, un rack peut accueillir sept serveurs Dell PowerEdge XE9680, contre seulement cinq serveurs Supermicro. Dans les Figures 1 et 2, nous comparons les résultats du serveur Dell PowerEdge XE9680 par rapport à ceux de deux configurations du serveur Supermicro équipé de 8 processeurs graphiques : le modèle SYS-821GE-TNHR équipé de processeurs Intel et le modèle AS-8125GS-TNHR équipé de processeurs AMD. Il convient de noter que Supermicro n'a pas présenté de résultats pour RNN-T sur le modèle SYS-821GE-TNHR, c'est pourquoi nous excluons ce modèle du tableau de la Figure 1.

### Résultats MLPerf® normalisés : Dell PowerEdge XE9680 avec H100 SXM5 par rapport à Supermicro SYS-821GE-TNHR avec H100 (plus les valeurs sont élevées, mieux c'est)

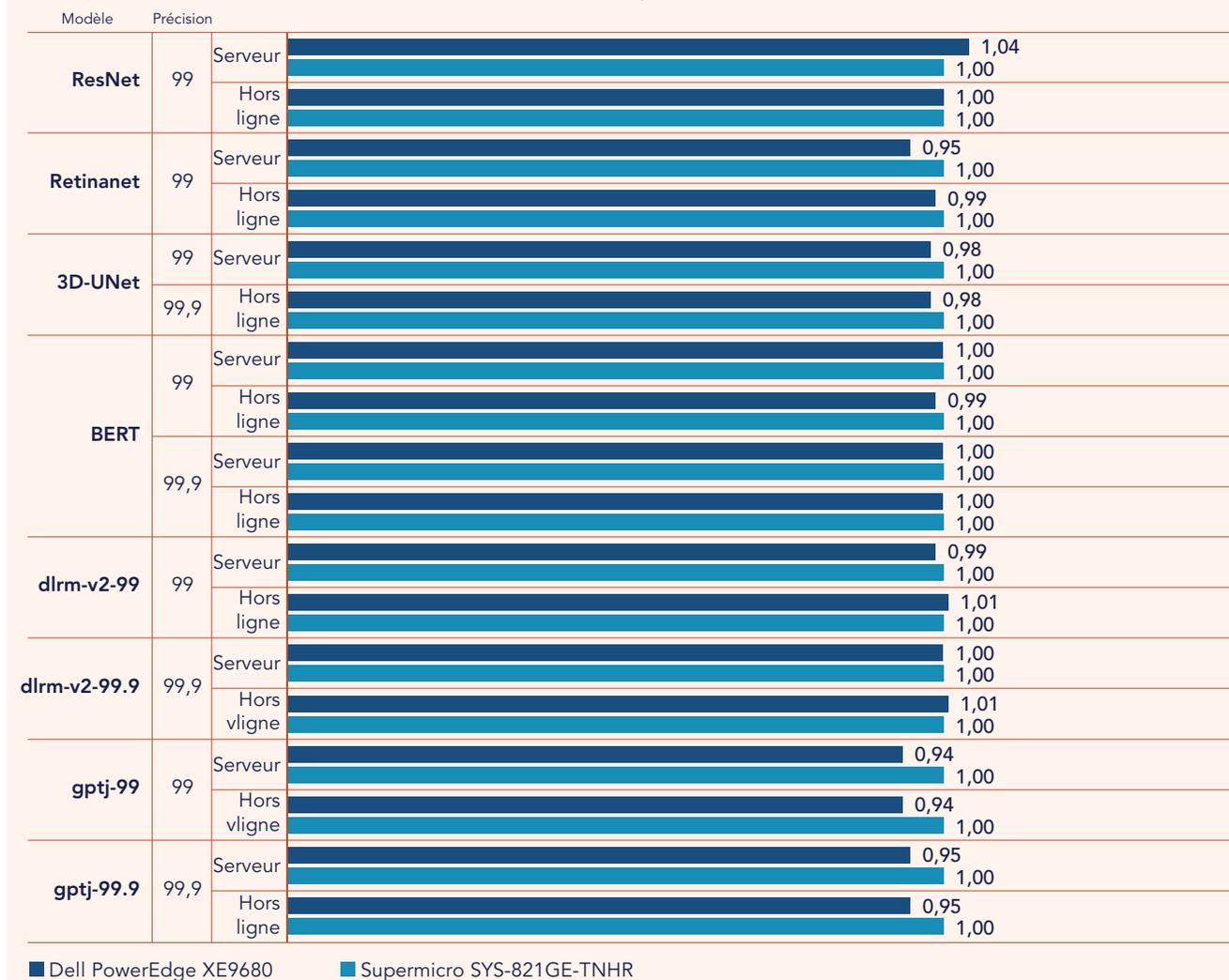


Figure 1 : Résultats MLPerf® publiés pour les serveurs Dell PowerEdge XE9680 et Supermicro SYS-821GE-TNHR en date du 29 novembre 2023. Les deux systèmes utilisent le format SXM du processeur graphique NVIDIA H100. Source : Principled Technologies d'après des données de MLCommons<sup>7,8</sup>



## Résultats MLPerf® normalisés : Dell PowerEdge XE9680 avec H100 SXM5 par rapport à Supermicro AS-8125GS-TNHR avec H100 SXM5 (plus les valeurs sont élevées, mieux c'est)

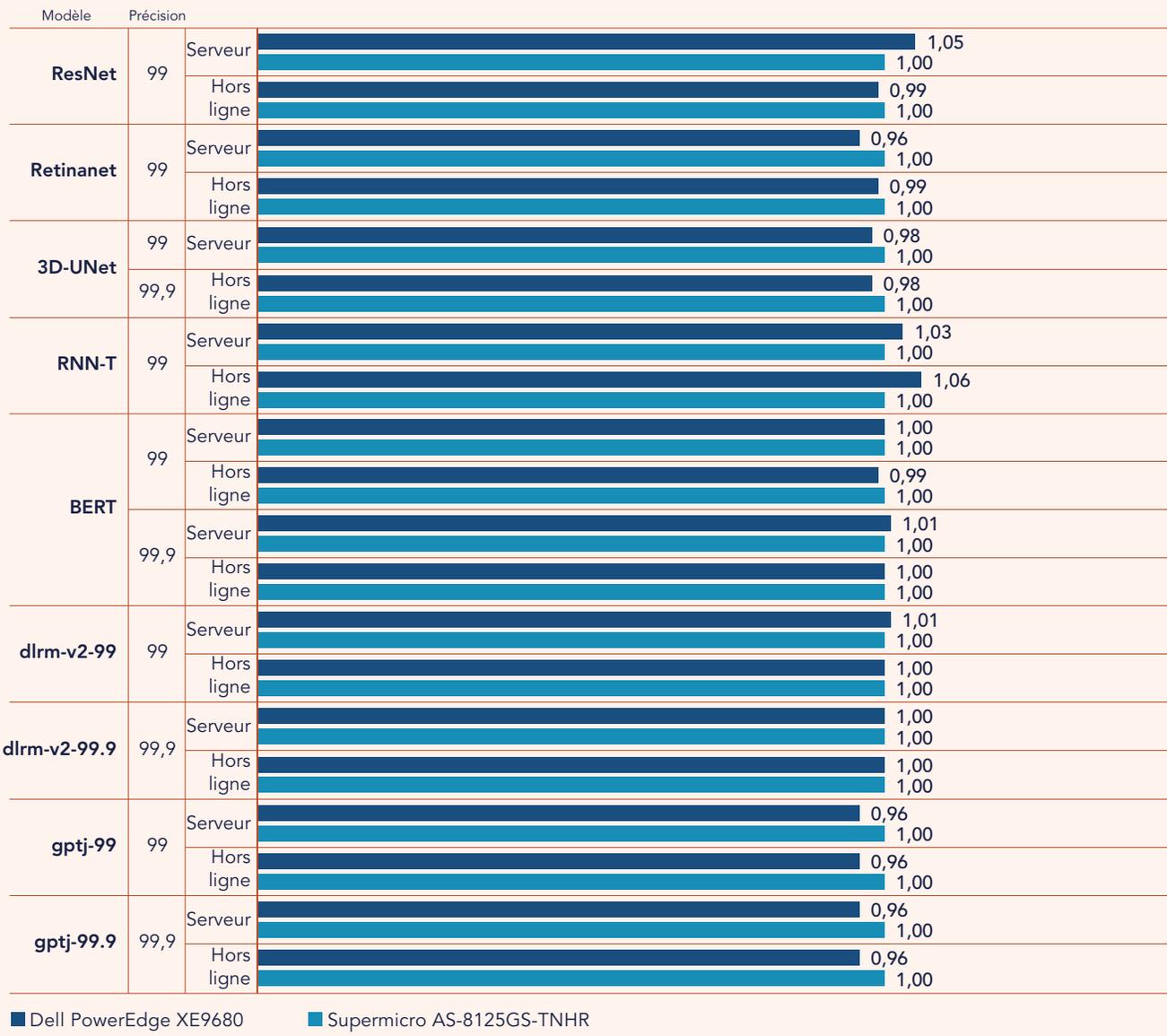


Figure 2 : Résultats MLPerf® publiés pour les serveurs Dell PowerEdge XE9680 et Supermicro AS-8125GS-TNHR en date du 29 novembre 2023. Les deux systèmes utilisent le format SXM du processeur graphique NVIDIA H100. Source : Principled Technologies d'après des données de MLCommons®<sup>9,10</sup>

Comme le montrent ces résultats, le choix du serveur Dell PowerEdge XE9680 vous permet de bénéficier de performances similaires sur une myriade de charges applicatives d'inférence de l'IA tout en occupant moins d'espace dans le datacenter.



## Résultats des serveurs équipés de quatre processeurs graphiques

Lorsque la réduction de la consommation d'énergie ou le gain d'espace dans le datacenter est une priorité absolue, le serveur Dell PowerEdge XE9640 2U pourrait être la solution la plus adaptée. Pouvant être équipé de quatre processeurs graphiques NVIDIA H100 SXM, le serveur PowerEdge XE9640 offre la moitié de la puissance de calcul du processeur graphique du serveur PowerEdge XE9680 dans un tiers de l'espace.<sup>11</sup> Le serveur Dell PowerEdge XE9640, haute densité, intègre la technologie Dell Smart Cooling, qui fournit un large éventail de technologies thermiques, dont le refroidissement liquide direct pour les processeurs et les processeurs graphiques.<sup>12</sup> Le boîtier 2U du serveur PowerEdge XE9640 intègre des mécanismes de circulation d'air améliorés, notamment des ventilateurs et des dissipateurs de chaleur de plus grande taille, pour aider à refroidir les autres composants essentiels, comme les cartes PCIe et la mémoire.<sup>13</sup>

Supermicro propose un serveur SYS-220GQ-TNAR+ plus ancien, équipé de quatre processeurs graphiques NVIDIA A100 HGX dans un format 2U, mais nous n'avons pas pu trouver de serveurs Supermicro 2U équipés de quatre processeurs graphiques H100 HGX plus récents qui puissent correspondre au PowerEdge XE9640.<sup>14</sup> Le serveur à 4 processeurs graphiques équipé de 4 processeurs graphiques NVIDIA HGX H100 de Supermicro évalué par MLPerf<sup>®</sup> est le serveur SYS-421GU-TNXR, qui est un serveur 4U. Comme nous l'avons indiqué précédemment, Supermicro a soumis des résultats MLPerf<sup>®</sup> 3.1 uniquement pour le serveur SYS-421GU-TNXR sur le modèle d'IA GPTJ-99.9. Nous ne pouvons donc pas le comparer au serveur PowerEdge XE9640 sur les autres modèles. Cependant, dans les résultats publiés, 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 3).

### Résultats MLPerf<sup>®</sup> 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)

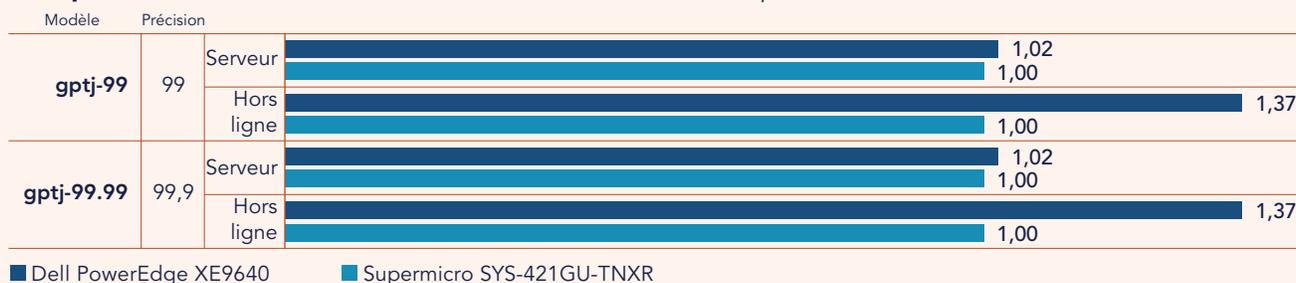


Figure 3 : Résultats MLPerf<sup>®</sup> publiés pour les serveurs Dell PowerEdge XE9640 et Supermicro SYS-421GU-TNXR en date du 29 novembre 2023. Les deux systèmes utilisent le format SXM du processeur graphique NVIDIA H100. Source : Principled Technologies d'après des données de MLCommons<sup>®15,16</sup>

Nous pouvons également comparer les résultats du serveur Supermicro SYS-421GU-TNXR à ceux du serveur Dell PowerEdge XE8640, un serveur 4U équipé de 4 processeurs graphiques qui prend également en charge des processeurs graphiques NVIDIA H100 HGX. Plus grand que le serveur PowerEdge XE9640, le modèle PowerEdge XE8640 ne nécessite pas de refroidissement liquide direct, ce qui en fait un bon compromis entre densité et technologies de refroidissement pour les datacenters qui n'ont pas accès au refroidissement par eau. Le serveur PowerEdge XE8640 dispose d'un système de refroidissement par air pour les processeurs et d'un radiateur de refroidissement par air assisté par liquide pour les processeurs graphiques. Le rack n'a donc pas besoin d'être raccordé à une arrivée d'eau.<sup>17</sup> Le serveur Dell PowerEdge XE8640 est équipé des tout derniers processeurs Intel Xeon Scalable de 4<sup>e</sup> 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.<sup>18</sup> 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 obtient de meilleurs scores GPTJ-99 dans les tests hors ligne que le serveur Supermicro (Figure 4).

## 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 4 : 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<sup>19,20</sup>

Comme nous l'avons vu, les résultats des performances MLPerf® des offres de serveurs Supermicro et Dell équipés de processeurs graphiques sont similaires dans l'ensemble, avec un net avantage pour les serveurs Dell PowerEdge XE8640 et PowerEdge XE9640 sur un modèle d'IA.

Étant donné que les performances ne constituent qu'un aspect du parcours d'implémentation de l'IA, nous avons également examiné le reste des offres de Dell et de Supermicro en matière d'IA, depuis les stations de travail, les solutions de stockage et de gestion de réseau, jusqu'aux services en passant par le support, la formation, etc. Nous avons constaté que la gamme Dell AI propose un plus large éventail que les offres de Supermicro et qu'elle permet de surmonter de nombreux obstacles auxquels les entreprises sont confrontées, au-delà de celui des performances de calcul.

## Stations de travail clientes et offres de stockage

### Autres options de calcul avec les stations de travail

Certains cas d'utilisation de l'IA nécessitent une approche différente du calcul, et les utilisateurs qui ont besoin d'accéder à du matériel compatible avec l'IA ne peuvent pas tous rester rattachés au datacenter. Les scientifiques qui travaillent dans des laboratoires n'ont pas toujours la place d'installer un rack de serveurs, et il n'est pas pratique pour les utilisateurs qui travaillent en périphérie de transporter un système de bureau lourd et volumineux. C'est là que les stations de travail entrent en jeu. La gamme de solutions Dell AI comprend plusieurs stations de travail Precision prêtes pour l'IA, y compris des configurations au format tour, mobiles, et au format rack, pour répondre aux différents besoins.<sup>21</sup> Supermicro, en revanche, ne propose pas de stations de travail mobiles, facilement accessibles en déplacement. Ses offres de stations de travail équipées de processeurs graphiques sont constituées de plusieurs configurations de tours différentes, dont certaines peuvent, selon Supermicro, être montées en rack.<sup>22</sup> Lors de nos recherches, nous avons identifié deux stations de travail montables en rack, mais il s'avère qu'elles sont plus anciennes, disponibles uniquement avec des processeurs graphiques NVIDIA A100, et elles ne sont d'ailleurs peut-être plus disponibles à la vente.<sup>23</sup> Si votre entreprise a besoin de flexibilité en matière de type et de mobilité des stations de travail avec processeurs graphiques qu'elle déploie, la gamme de solutions Dell AI est mieux adaptée à ces besoins.

### Considérations relatives au stockage

Le stockage peut être tout aussi crucial que le calcul lors de l'exécution de charges applicatives d'IA. Disposer d'une plus grande quantité de données améliore la précision des modèles d'IA, mais le stockage et la gestion d'énormes jeux de données peuvent mettre à l'épreuve les capacités de nombreux datacenters. En outre, étant donné que l'entraînement des modèles se fait généralement à l'aide de données non structurées, les systèmes de stockage prêts pour l'IA doivent pouvoir gérer facilement de nombreux types de données différents.<sup>24</sup> Pour fournir la capacité et la mise à l'échelle des jeux de données IA, ML et DL, Dell propose la série PowerScale™ pour le stockage en mode fichier et un système Elastic Cloud Storage (ECS) ou ObjectScale software-defined pour le stockage en mode objet.

La gamme NAS All-Flash Dell PowerScale offre des options allant de 3,84 To à 720 To de capacité brute par nœud, avec des capacités All-Flash en cluster atteignant 186 Po de capacité brute. La flexibilité et l'évolutivité de PowerScale peuvent prendre en charge une grande variété de clients et de cas d'utilisation d'IA.<sup>25</sup> Les trois modèles PowerScale All-Flash (F200, F600 et F900) incluent la compression et la déduplication des données inline pour améliorer l'efficacité du stockage.<sup>26</sup> Chaque modèle de stockage PowerScale utilise le système de fichiers Dell OneFS™, qui tire parti de stratégies de hiérarchisation avancées pour garantir que les données les plus fréquemment consultées résident sur les niveaux de stockage les plus performants.<sup>27</sup> Dell propose également le logiciel OneFS sur le marché Amazon Web Services (AWS) avec APEX File Storage for AWS. Les clients peuvent utiliser OneFS avec leurs instances de calcul AWS pour une expérience utilisateur cohérente avec les mêmes fonctionnalités que celles disponibles dans les baies OneFS sur site.<sup>28</sup>

Les offres de stockage de Supermicro consistent en des serveurs de stockage (montés en rack, à haute densité de stockage), disponibles dans différentes tailles et densités.<sup>29</sup> Pour obtenir le stockage en mode fichier de Supermicro, les clients doivent choisir parmi diverses offres de stockage software-defined tierces, par exemple WekaIO, Scality RING ou OSNEXUS.<sup>30</sup> Scality RING et OSNEXUS incluent certes des options de stockage en mode fichier dans la description de leur plateforme, mais WekaIO semble être le choix numéro un pour les clients à la recherche d'un stockage en mode fichier de base. Supermicro propose plusieurs architectures de référence couvrant un large éventail de cas d'utilisation, mais les clients doivent disposer d'un abonnement ou d'une licence pour le logiciel WekaIO, ce qui peut augmenter le coût global de la solution.<sup>31</sup>

Les options de stockage en mode objet de Dell comprennent la solution de stockage en mode objet d'entreprise Dell ECS qui est « spécialement conçue pour stocker les données non structurées à l'échelle du Cloud public ».<sup>32</sup> Outre la compatibilité intégrée avec la fonctionnalité de stockage en mode objet Amazon S3 pour le Cloud hybride, les nœuds de stockage ECS offrent une capacité allant jusqu'à 14 Po par rack.<sup>33</sup> Comme pour le stockage en mode fichier, les offres de stockage en mode objet de Supermicro nécessitent des configurations tierces. La plateforme OSNEXUS est une plateforme de stockage combinée pour le stockage en mode fichier, en mode bloc et en mode objet, tandis que la solution Scality RING combine le stockage en mode fichier et en mode objet ; toutes deux nécessitent des licences auprès de fournisseurs tiers<sup>34,35</sup>. En ce qui concerne le stockage en mode objet uniquement, les clients peuvent acheter la solution Supermicro pour Quantum ActiveScale à des fins de stockage en mode objet dans le Cloud privé avec un abonnement au logiciel Quantum.<sup>36</sup> Au moment de la rédaction de ce rapport, nous n'avons pas trouvé d'options de consommation/dépenses d'exploitation flexibles proposées par Supermicro.

Étant donné que les offres de stockage de Supermicro nécessitent un engagement auprès de fournisseurs de logiciels tiers, les clients risquent de devoir payer des frais de licence et d'abonnement supplémentaires et de rencontrer des difficultés en matière de support, de dépannage, etc. La gamme de solutions Dell AI offre aux clients Dell Storage une solution de services et de support fiable et unique couvrant tous les aspects de leur solution de stockage.

## À propos de Dell APEX

Pour les clients qui souhaitent utiliser leur stockage en mode fichier as-a-service, Dell propose APEX Data Storage Services, une solution qui comprend le stockage en mode fichier, en mode bloc et le stockage de sauvegarde. Grâce à Dell APEX Console, les clients peuvent commander de nouveaux abonnements, ajuster et surveiller la capacité de stockage, etc. Selon Dell, cette solution permet de « profiter de la simplicité et de l'agilité de l'expérience Cloud avec plus de contrôle sur vos applications et vos données »<sup>37</sup>.

Pour en savoir plus sur Dell APEX, consultez <https://www.dell.com/en-us/dt/apex/storage/data-storage-services/index.htm>



## Options de mise en réseau

La gestion de réseau est un autre élément essentiel de l'infrastructure d'IA. De nombreuses charges applicatives d'IA s'exécutent sur de grands clusters de serveurs qui nécessitent une communication constante entre eux et avec le système de stockage. Les charges de travail d'IA ont donc besoin d'une gestion de réseau robuste pour éviter les goulots d'étranglement. Si votre réseau n'est pas suffisant pour la charge applicative d'IA, l'entraînement et l'inférence prendront plus de temps, ce qui ralentira le traitement des données et le délai d'obtention des informations. Dell propose des commutateurs PowerSwitch top-of-rack (ToR) pour le datacenter et des modules d'E/S PowerEdge MX pour les réseaux Ethernet et de fabric.<sup>38</sup> Les offres PowerSwitch délivrent des débits allant de 1 GbE à 400 GbE pour répondre à un large éventail de besoins. En outre, les commutateurs Dell PowerSwitch série Z offrent des connexions 100 GbE et 400 GbE optimisées pour les fabrics leaf/spine.<sup>39</sup>

Supermicro propose également des commutateurs équipés de ports Ethernet délivrant des débits jusqu'à 400 GbE pour les applications top-of-rack (ToR) et d'autres applications comme l'architecture spine-leaf du datacenter.<sup>40</sup> Cependant, les services de réseau Dell offrent plusieurs avantages en matière de facilité d'utilisation et de flexibilité que Supermicro ne propose pas. Des services tels que le Dell Fabric Design Center peuvent contribuer à éviter les incompatibilités, les lacunes ou les inefficacités du réseau, en aidant les clients à planifier et à déployer des fabrics réseau de manière automatisée.<sup>41</sup> Pour les environnements spécifiques comme VMware VxRail, VMware ESXi et les configurations Dell PowerStore, Dell propose des services SmartFabric Services, qui permettent le déploiement d'une infrastructure software-defined et la gestion du cycle de vie. Avec PowerStore, les services SmartFabric Services peuvent automatiser « jusqu'à 99 % des tâches de connectivité LAN avec un fabric plug-and-play ». <sup>42</sup> Les services de ce type, qui fournissent automatisation, conseils et autres pour la conception et l'implémentation du réseau, aident les clients dans leur parcours vers l'adoption de l'IA.

## Des services, des formations, et bien plus encore

Le principal défi autre que matériel dans le cadre de l'implémentation de l'IA, c'est la nécessité de disposer d'une expertise interne pour la stratégie, la planification, la préparation et la gestion des données. Gérer et assurer la maintenance des charges applicatives d'IA nécessite des connaissances uniques, comprenant à la fois une expertise matérielle plus traditionnelle, mais aussi en apprentissage automatique et en science des données. Les personnes chargées de concevoir et d'implémenter la stratégie d'IA doivent également avoir une connaissance approfondie des objectifs opérationnels propres à l'entreprise afin de s'assurer que les nouvelles charges applicatives d'IA permettent de les atteindre.<sup>43</sup>

Autre obstacle de taille : l'intégration transparente de l'IA dans les systèmes opérationnels existants. Cette intégration nécessite un alignement stratégique des nouvelles technologies d'IA sur les processus métier actuels, afin de s'assurer que l'introduction de l'IA ne perturbe pas les workflows établis. Établir un partenariat avec une société comme Dell, qui propose plusieurs architectures de référence de solutions optimisées et validées, des cours de formation, des options de gestion et un vaste écosystème de partenaires, peut faciliter votre parcours vers l'adoption de l'IA.

## Services professionnels pour l'IA

Pour répondre aux besoins de formation et de planification, Dell offre une variété de services spécifiquement conçus pour l'IA.<sup>44</sup> Les services Dell assurant l'implémentation de l'IA comprennent des services de conseil, la préparation des données, le déploiement, des services de support et de formation, chacun ciblant des aspects spécifiques de l'adoption de l'IA. Les services de conseil Dell pour l'IA générative aident les clients à créer une feuille de route pour identifier les cas d'utilisation et accompagner les sociétés dans la rationalisation de leurs processus.<sup>45</sup> De même, les services d'adoption pour l'IA générative proposent des ateliers avec des professionnels de Dell pour étudier vos besoins et vos défis uniques afin de déterminer un modèle préentraîné pour votre entreprise et de mener des sessions de transfert de connaissances dans l'objectif de former votre personnel IT.<sup>46</sup> Dell propose également des services d'implémentation, de mise à l'échelle et de services managés pour l'IA générative qui offrent différents niveaux de support et de formation, jusqu'à produire une infrastructure d'IA entièrement gérée, qui permet à votre personnel IT de se concentrer sur les modèles et les données tandis que Dell gère le matériel.<sup>47</sup> Les services ProSupport Services garantissent des performances système optimales, fournissent une assistance matérielle et logicielle (essentielle à l'exécution des opérations d'IA en continu) et résolvent les problèmes techniques.<sup>48</sup>

Les services de formation font partie intégrante du développement des compétences et des connaissances nécessaires à l'utilisation de l'IA. Les offres de formation de Dell comprennent des programmes de formation complets en science des données, des certifications en analytique avancée et des ateliers sur des technologies IA spécifiques, comme l'apprentissage automatique.<sup>49</sup>

Les services de Supermicro, quant à eux, se limitent principalement au dépannage, aux manuels, aux autorisations de retour de marchandises (RMA) et aux garanties.<sup>50</sup> Nous n'avons pas trouvé de services de conception, d'implémentation, de gestion ou de formation dans la gamme IA de Supermicro. Pour les sociétés à la recherche d'un partenaire de formation qui les aide à s'y retrouver dans les méandres du parcours d'adoption de l'IA, Dell est le choix qui s'impose.

## Partenariats avec des tiers pour les charges applicatives d'IA

Dell Technologies et NVIDIA travaillent en collaboration pour proposer des solutions Dell Validated Designs, qui visent à fournir une solution complète pour l'IA générative dans les environnements professionnels. Ce projet crée une infrastructure évolutive et hautes performances basée sur les technologies et les logiciels Dell et NVIDIA, avec un cadre de modèle d'IA qui permet aux entreprises de construire et d'exécuter des modèles d'IA personnalisés. Cette solution permet aux clients d'obtenir rapidement des charges applicatives d'IA opérationnelles.<sup>51</sup> Pour en savoir plus, lisez la section Dell Validated Designs ci-dessous.

Dell a établi des partenariats avec plusieurs sociétés pour améliorer les applications des technologies d'IA. Avec Hugging Face, Dell facilite la configuration de grands modèles de langage (LLM) sur site. Ce partenariat associe l'expertise en IA de Hugging Face avec les serveurs et les systèmes de stockage de Dell. Un portail Hugging Face spécifique pour Dell fournira des outils pour un déploiement simple et sécurisé des modèles d'IA Open Source de Hugging Face. L'objectif est de continuer à améliorer ces modèles pour les systèmes Dell, afin d'optimiser les performances et de prendre en charge de nouvelles applications de l'IA.<sup>52</sup>

Dell et Starburst travaillent sur un Data Lake hautes performances et évolutif qui intègre l'analytique de Starburst dans les technologies de stockage et de calcul de Dell, cherchant ainsi à offrir un point d'accès unique à toutes les sources de données pour les outils IA et ML. Les clients pourront profiter de ce partenariat pour aider à éliminer les silos de données.<sup>53</sup>

D'après nos recherches, Supermicro a des partenariats beaucoup plus limités pour l'IA. SiMa.ai et Supermicro ont codéveloppé le système Supermicro SYS-E300-13AD, un serveur Edge ML compact conçu pour le traitement analytique vidéo multiflux. Ce serveur, équipé du pipeline ML de SiMa.ai sur une puce, traite efficacement plusieurs canaux vidéo, réduit le coût total de possession et améliore la fiabilité et la sécurité. Il offre un environnement de calcul conçu pour le traitement et l'analyse de nombreux flux vidéo, offrant une intelligence de pointe adaptée à diverses applications d'entreprise.<sup>54</sup>

## Dell Validated Designs

Afin d'éviter les approximations des solutions matérielles d'IA, Dell propose des architectures de référence validées en laboratoire et optimisées pour plusieurs charges applicatives d'IA, entre autres. Ces conceptions validées comprennent des concepts architecturaux, des présentations de solutions complètes, des validations de performances et d'autres validations en laboratoire prouvant les capacités de la solution sur la charge applicative pour laquelle elle a été conçue. Ces charges applicatives comprennent les environnements virtualisés, les outils MLOps, l'apprentissage automatique, l'IA conversationnelle, l'inférence de l'IA générative, le réglage des modèles d'IA générative, NVIDIA Fleet Command et OpenShift AI.<sup>55</sup>

Par exemple, la conception validée d'IA pour les environnements virtualisés combine l'IA basée sur VMware avec NVIDIA AI Enterprise sur l'infrastructure Dell, optimisant ainsi l'IA dans les paramètres virtuels.<sup>56</sup> Le guide Validated Design inclut des résultats de performances montrant l'entraînement du modèle ResNet, ce qui prouve aux clients que la conception fonctionne et présente le type de performance auquel ils peuvent s'attendre.<sup>57</sup> Ces validations offrent aux clients une valeur qui va au-delà d'une simple liste de systèmes matériels fonctionnant ensemble. Elles expliquent les concepts, recommandent des configurations et guident les clients dans leurs réflexions et leurs attentes en matière de performances.<sup>58</sup>

Supermicro propose certes des solutions basées sur les cas d'utilisation, mais n'atteint pas le niveau d'architecture de référence observé avec Dell Validated Designs. Plutôt que de recommander une solution spécialement conçue pour des charges applicatives spécifiques, Supermicro organise ses serveurs et ses processeurs graphiques en catégories, par exemple l'inférence et l'entraînement de l'IA, les solutions HPC/IA, la visualisation et la conception, etc.<sup>59</sup> Dans ses brochures et fiches techniques, Supermicro recommande quelques serveurs et processeurs graphiques les mieux adaptés pour une tâche, présente plusieurs cas d'utilisation, répertorie les technologies clés à l'œuvre et suggère des logiciels.<sup>60</sup> Contrairement aux solutions Dell Validated Designs, Supermicro ne semble pas inclure d'architectures réseau, ni de données de validation ou sur les performances. Supermicro propose également plusieurs conceptions de référence qui offrent une architecture de référence plus détaillée pour certaines solutions d'IA, comme un rapport sur l'entraînement de l'IA à grande échelle avec une solution de refroidissement liquide publié par Supermicro en collaboration avec NVIDIA.<sup>61</sup> Les clients trouveront peut-être une architecture de référence plus approfondie pour un scénario spécifique, mais au moment de notre étude, nous n'en avons trouvé que trois : l'architecture à refroidissement liquide susmentionnée, une architecture de station de travail IA et une architecture RedHat OpenShift.<sup>62</sup>

Dans l'ensemble, nous avons constaté que les Dell Validated Designs couvraient davantage de charges applicatives d'IA et offraient des conseils plus détaillés que les offres de Supermicro.

## Services de gestion et iDRAC

D'après un [rapport d'avril 2023](#) publié par Principled Technologies, le contrôleur iDRAC (Integrated Dell Remote Access Controller) offre plusieurs fonctions avancées par rapport à l'interface IPMI (Intelligent Platform Management Interface), notamment en matière d'automatisation, de sécurité et de configuration.<sup>63</sup> Le Tableau 3 présente une comparaison des fonctions de gestion de Dell et de Supermicro, tirée de ce rapport, qui montre comment le contrôleur iDRAC pourrait faciliter le déploiement et la mise à jour de firmware et offrir davantage de fonctions de sécurité que l'interface IPMI de Supermicro. Il convient de noter que certaines conclusions peuvent avoir changé depuis la publication initiale.

Tableau 3 : Résumé du rapport d'avril 2023 publié par Principled Technologies comparant les outils de gestion de Dell et de Supermicro. Certaines conclusions peuvent avoir changé depuis la publication. Source : Principled Technologies <https://facts.pt/V5fDf06>.

	Ce qui est différent avec les outils de gestion Dell	À quel point c'est mieux
<p><b>Mises à jour de firmware plus faciles</b> iDRAC9 par rapport à IPMI de Supermicro OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>Mises à jour en ligne automatisées avec iDRAC9, avec options de programmation</li> <li>OME permet de créer des référentiels de firmware personnalisés et de mettre à jour le firmware du BIOS, du BMC et d'autres composants du serveur sans outils ou agents supplémentaires</li> </ul>	<ul style="list-style-type: none"> <li>Nous avons configuré des mises à jour automatiques dans le contrôleur iDRAC en <b>seulement 74 secondes</b></li> <li>L'interface IPMI de Supermicro <b>n'a aucune fonction de mise à jour automatique</b>, les administrateurs doivent donc procéder aux mises à jour manuellement</li> <li>SSM prend uniquement en charge les mises à jour de firmware du BIOS et du BMC et nécessite le programme SUM pour mettre à jour les autres composants</li> </ul>
<p><b>Plus de fonctionnalités de sécurité</b> iDRAC9 par rapport à IPMI de Supermicro OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>iDRAC9 propose l'authentification multifactor et la désactivation dynamique des ports USB sans interruption de service du système</li> <li>OME offre à la fois un contrôle d'accès basé sur les rôles (RBAC) et un contrôle d'accès basé sur le périmètre (SBAC) pour limiter la gestion des appareils à un sous-ensemble de groupes d'appareils</li> </ul>	<ul style="list-style-type: none"> <li>L'interface IPMI de Supermicro n'offre pas de <b>fonctionnalités d'authentification multifactor</b></li> <li>L'interface IPMI nécessite un redémarrage du système et une configuration du BIOS pour désactiver les ports USB</li> <li>Le programme SSM de Supermicro offre un contrôle d'accès RBAC, mais pas le contrôle d'accès plus restrictif SBAC</li> </ul>
<p><b>Gestion du cycle de vie simplifiée</b> OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>Gestion complète du cycle de vie sans agent via OME pour faciliter la gestion et la surveillance</li> </ul>	<ul style="list-style-type: none"> <li>SSM a besoin de l'agent SuperDoctor5 pour obtenir des mesures détaillées de l'intégrité du système local et pour que Supermicro Update Manager (SUM) mette à jour des composants supplémentaires</li> </ul>
<p><b>Déploiement de serveur plus facile</b> iDRAC9 par rapport à IPMI de Supermicro</p>	<ul style="list-style-type: none"> <li>Importation d'un profil complet de serveur Dell en seulement 12 étapes avec iDRAC9</li> <li>Options de configuration BIOS robustes avec iDRAC9 offrant 52 fonctions BIOS et prise en charge de la configuration de composants comme la carte NIC RAID et iDRAC</li> </ul>	<ul style="list-style-type: none"> <li>L'interface IPMI de Supermicro nous a permis de sauvegarder et de restaurer uniquement la configuration IPMI plutôt que l'ensemble du profil du serveur</li> <li>iDRAC9 offre 52 fonctionnalités de BIOS, tandis que l'interface IPMI n'offre <b>aucune option de configuration du BIOS</b></li> </ul>
<p><b>Davantage d'options de création de rapport et d'analytique</b> iDRAC9 par rapport à IPMI de Supermicro OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>iDRAC9 offre un flux de télémétrie qui permet aux utilisateurs d'envoyer facilement les données du serveur à des outils d'analytique comme Splunk</li> <li>OME envoie des données de télémétrie directement à CloudIQ pour faciliter la surveillance</li> </ul>	<ul style="list-style-type: none"> <li>L'interface IPMI offre uniquement une fonctionnalité SYSLOG que les administrateurs peuvent utiliser pour envoyer des messages à des fins d'agrégation et d'analyse éventuelle</li> <li>SSM n'a pas de solution de gestion Cloud équivalant à Dell CloudIQ</li> </ul>
<p><b>Davantage de fonctionnalités favorisant les pratiques durables</b> OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>Davantage de métriques de surveillance dans OME Power Manager, y compris des données sur l'empreinte carbone</li> </ul>	<ul style="list-style-type: none"> <li>SSM dispose de métriques d'utilisation moins robustes et <b>n'a aucun moyen de suivre l'empreinte carbone</b></li> </ul>
<p><b>Davantage de méthodes de surveillance</b> OME par rapport à SSM de Supermicro</p>	<ul style="list-style-type: none"> <li>Possibilité de gérer les serveurs Dell de n'importe où via l'application mobile OpenManage</li> <li>Surveillance des appareils tiers avec OME grâce aux adresses IP et aux informations d'identification des serveurs avec la possibilité d'importer des MIBS SNMP tiers</li> </ul>	<ul style="list-style-type: none"> <li>SSM n'a pas <b>d'application mobile</b></li> <li>SSM <b>n'autorise pas</b> la surveillance d'appareils tiers avec les adresses IP des serveurs</li> </ul>



## Conclusion

Lorsqu'il s'agit de concevoir des solutions d'IA dans votre entreprise, de les implémenter, de les gérer et d'en assurer la maintenance, de nombreux facteurs sont à prendre en compte. Pour vous aider à investir judicieusement et à tirer le meilleur parti de votre solution d'IA, vous recherchez plus qu'un simple fournisseur de matériel. Nos recherches indiquent que Dell offre des services qui peuvent vous aider en tant que partenaire tout au long de votre parcours. Alors, envisagez d'investir dans des solutions Dell lorsque vous vous lancez dans un parcours d'adoption de l'IA.

1. Vipera, « NVIDIA's H100 and A100 GPU Cards: Exploring the Intricacies of SXM and PCI-E Connections », consulté le 5 janvier 2024, <https://www.viperatech.com/unraveling-the-mysteries-sxm-vs-pci-e-connections-in-nvidias-high-end-h100-and-a100-gpus/>.
2. MLCommons, « MLPerf Inference: Datacenter Benchmark Suite Results », consulté le 5 janvier 2024, <https://mlcommons.org/en/inference-datacenter-31/>.
3. MLCommons, « MLPerf Inference: Datacenter Benchmark Suite Results. »
4. Dell, « Serveurs PowerEdge XE », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/servers/specialty-servers/powerededge-xe-servers.htm>.
5. Supermicro, « Next Leap of AI Infrastructure is Here », consulté le 5 janvier 2024, <https://www.supermicro.com/en/accelerators/nvidia>.
6. Dell, « PowerEdge XE9680 », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/powerededge-xe9680-spec-sheet.pdf>.
7. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0069. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
8. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0135. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
9. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0069. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
10. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0132. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).

11. Dell, « Serveur au format rack PowerEdge XE9640 », consulté le 5 janvier 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
12. Accelsius, « Enabling the AI Revolution with Liquid Cooling », consulté le 5 janvier 2024, <https://www.accelsius.com/blog/enabling-the-ai-revolution-with-liquid-cooling>.
13. Dell, « Guide technique PowerEdge XE9640 », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9640-technical-guide.pdf>.
14. Supermicro, « GPU Server Systems », consulté le 5 janvier 2024, [https://www.supermicro.com/en/products/gpu?pro=pl\\_grp\\_type%3D1](https://www.supermicro.com/en/products/gpu?pro=pl_grp_type%3D1).
15. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0067. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
16. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0133. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
17. Dell, « PowerEdge XE8640 », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe8640-spec-sheet.pdf>.
18. Dell, « Serveur au format rack PowerEdge XE8640 », consulté le 5 janvier 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.
19. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0066. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
20. Score vérifié MLPerf® v3.1 Inférence fermée. Extrait de <https://mlcommons.org/benchmarks/inference-datacenter/> 5 décembre 2023, entrée 3.1-0133. Le nom et le logo MLPerf sont des marques ou 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, reportez-vous au site [www.mlcommons.org](http://www.mlcommons.org).
21. Dell, « Technologies d'intelligence artificielle (IA) optimisées par les stations de travail Dell Precision », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/ai-technologies/index.htm?hve=explore+dell+precision+for+ai#tab0=0>.
22. Supermicro, « Super Workstations », consulté le 5 janvier 2024, <https://www.supermicro.com/en/products/superworkstation>.
23. Supermicro, « Rackmount Workstations », consulté le 5 janvier 2024, <https://www.supermicro.com/en/products/rackmount-workstations>.
24. Stephen Pritchard, « Storage requirements for AI, ML and analytics in 2022 », consulté le 5 janvier 2024, <https://www.computerweekly.com/feature/Storage-requirements-for-AI-ML-and-analytics-in-2022>.
25. Dell, « PowerScale Plateforme de données prête pour l'IA », consulté le 5 janvier 2024, <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale>.
26. Dell, « Dell PowerScale All-Flash », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h15963-ss-powerscale-all-flash-nodes.pdf>.
27. Dell, « Fonctionnalités du logiciel Dell EMC PowerScale OneFS », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h18275-onefs-software-features-data-sheet.pdf>.
28. Dell, « Stockage en mode objet d'entreprise Dell ECS », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/storage/ecs/>.
29. Supermicro, « Accelerating AI Data Pipelines », consulté le 5 janvier 2024, <https://www.supermicro.com/en/products/storage>.
30. Supermicro, « Supermicro Software-Defined Storage and Memory Solutions », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/software-defined-storage>.
31. Supermicro, « Supermicro WEKA Distributed Storage Solution », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/wekaio>.
32. Dell, « Stockage en mode objet d'entreprise Dell ECS », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/storage/ecs/>.

- 
33. Dell, « Stockage en mode objet d'entreprise Dell ECS »
  34. Supermicro, « Supermicro OSNEXUS Software-Defined Storage Solution », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/osnexus>.
  35. ASBIS, « Supermicro Solution for Scality RING », consulté le 5 janvier 2024, <https://news.asbis.com/news/suppliers/supermicro-renewed-the-line-of-scality-ring-solution/>.
  36. Supermicro, « Supermicro solution for Quantum ActiveScale », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/activescale>.
  37. Dell, « Stockage as-a-service évolutif et élastique », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/apex/storage/data-storage-services/>.
  38. Dell, « Optez pour une gestion de réseau ouverte avec PowerSwitch », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/networking/>.
  39. Dell, « Commutateurs de datacenter Dell PowerSwitch », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/networking/data-center-switches/>.
  40. Supermicro, « SSE-T7132S - 400Gb Ethernet Switch », consulté le 5 janvier 2024, <https://www.supermicro.com/en/products/accessories/Networking/SSE-T7132SR.php>.
  41. Dell, « Dell EMC Networking SmartFabric Services Deployment with VxRail 4.7—Fabric Design Center », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/dell-emc-networking-smartfabric-services-deployment-with-vxrail-4-7-1/fabric-design-center-26/>.
  42. Dell, « Dell SmartFabric Services », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/networking/smartfabric/>.
  43. Penny Madsen, « Scaling Skills for AI: Lessons from Early Adopters », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/industry-market/idc-brief-importance-of-skills-for-ai-dell.pdf>.
  44. Dell, « Design Guide—Generative AI in the Enterprise – Model Customization—Overview », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/overview-5381/>.
  45. Dell, « Design Guide—Generative AI in the Enterprise – Model Customization—Advisory Services for Generative AI », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/advisory-services-for-generative-ai-1/>.
  46. Dell, « Design Guide—Generative AI in the Enterprise – Model Customization—Adoption Services for Generative AI », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-model-customization/adoption-services-for-generative-ai-1/>.
  47. Dell, « Design Guide—Generative AI in the Enterprise – Model Customization—Adoption Services for Generative AI ».
  48. Dell, « Artificial Intelligence (AI) Ready Solution Services », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/services/solutions/artificial-intelligence-services.htm>.
  49. Dell, « Comprehensive AI Training Modules Tailored for You », consulté le 5 janvier 2024, <https://education.dell.com/content/emc/en-us/home/training/aiml.html>.
  50. Supermicro, « Services and Support », consulté le 5 janvier 2024, <https://www.supermicro.com/en/support>.
  51. Travis Vigil, « Dell and NVIDIA: Bringing Generative AI to the Enterprise », consulté le 5 janvier 2024, <https://www.dell.com/en-us/blog/dell-and-nvidia-bringing-generative-ai-to-the-enterprise/>.
  52. Dell, « Dell Technologies and Hugging Face to Simplify Generative AI with On-Premises IT », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/corporate/newsroom/announcements/detailpage.press-releases~usa~2023~11~20231114-dell-technologies-and-hugging-face-to-simplify-generative-ai-with-on-premises-it.htm>.
  53. Richard DeMare, « Starburst and Dell expand partnership to accelerate AI efforts with more intelligent data collection », consulté le 5 janvier 2024, <https://www.starburst.io/blog/starburst-and-dell-expand-partnership-to-accelerate-ai-efforts-with-more-intelligent-data-collection/>.
  54. Business Wire, « SiMa.ai and Supermicro Announce Partnership to Accelerate Power Efficient ML at the Edge », consulté le 5 janvier 2024, <https://www.businesswire.com/news/home/20231129609794/en/SiMa.ai-and-Supermicro-Announce-Partnership-to-Accelerate-Power-Efficient-ML-at-the-Edge>.
  55. Dell, « Solutions Dell AI », consulté le 5 janvier 2024, <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>

56. Dell, « Unlock the power of AI in virtualized environments », consulté le 5 janvier 2024, <https://www.delltechnologies.com/asset/en-us/products/ready-solutions/briefs-summaries/ai-vxrail-powerscale-brief.pdf>.
57. Dell, « Design Guide—Virtualizing GPUs for AI with VMware and NVIDIA Based on Dell Infrastructure—Performance Results », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/design-guide-virtualizing-gpus-for-ai-with-vmware-and-nvidia-based-on-dell-infrastructure-1/performance-results-15/>.
58. Dell, « Design Guide—Virtualizing GPUs for AI with VMware and NVIDIA Based on Dell Infrastructure—Design Considerations », consulté le 5 janvier 2024, <https://infohub.delltechnologies.com/l/design-guide-virtualizing-gpus-for-ai-with-vmware-and-nvidia-based-on-dell-infrastructure-1/design-considerations-105/>.
59. Supermicro, « Accelerate Every Workload », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/ai-deep-learning>.
60. Supermicro, « Supermicro Enterprise AI Inference & Training », consulté le 5 janvier 2024, [https://www.supermicro.com/datasheet/Datasheet\\_AI-Workloads\\_Enterprise\\_AI\\_Inferencing\\_and\\_Training.pdf](https://www.supermicro.com/datasheet/Datasheet_AI-Workloads_Enterprise_AI_Inferencing_and_Training.pdf).
61. Supermicro, « SUPERMICRO RACK SCALE SOLUTIONS: LARGE SCALE AI TRAINING WITH LIQUID COOLING », consulté le 5 janvier 2024, [https://www.supermicro.com/solutions/Solution-Brief\\_Rack\\_Scale\\_AI.pdf](https://www.supermicro.com/solutions/Solution-Brief_Rack_Scale_AI.pdf).
62. Supermicro, « Accelerate Every Workload », consulté le 5 janvier 2024, <https://www.supermicro.com/en/solutions/ai-deep-learning>.
63. Principled Technologies, « Dell management tools made server deployment and updates easier, offered more comprehensive security, and provided more robust infrastructure analytics », consulté le 5 janvier 2024, <https://www.principledtechnologies.com/Dell/Management-tools-vs-Supermicro-0423.pdf>.

► Consultez la version d'origine en anglais de ce rapport à l'adresse <https://facts.pt/q9p46K9>

Ce projet a été réalisé à la demande de Dell Technologies.



**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.

**EXCLUSION DE GARANTIE, LIMITATION DE RESPONSABILITÉ :**

Principled Technologies, Inc. a pris toutes les mesures raisonnables pour garantir la précision et la validité de ses tests. Toutefois, Principled Technologies, Inc. décline spécifiquement toute garantie, expresse ou implicite, relative aux résultats et à l'analyse des tests, à leur précision, à leur exhaustivité ou à leur qualité. Cela inclut toute garantie implicite d'adéquation à un usage particulier. Toute personne ou entité s'appuyant sur les résultats d'un de ces tests le fait à son propre risque et accepte que Principled Technologies, Inc., ses collaborateurs et ses sous-traitants ne soient en aucun cas responsables de toute perte ou tout préjudice causés par une erreur ou un défaut éventuels dans le cadre d'une procédure ou d'un résultat de test.

Principled Technologies, Inc. ne peut en aucun cas être tenu responsable des dommages indirects, spéciaux, fortuits ou consécutifs résultant de ses tests, même si la société a été informée de la possibilité de tels dommages. La responsabilité de Principled Technologies, Inc. ne peut en aucun cas, notamment en cas de dommages directs, excéder les montants versés en relation avec les tests de Principled Technologies, Inc. Les recours uniques et exclusifs du client sont définis dans le présent document.