

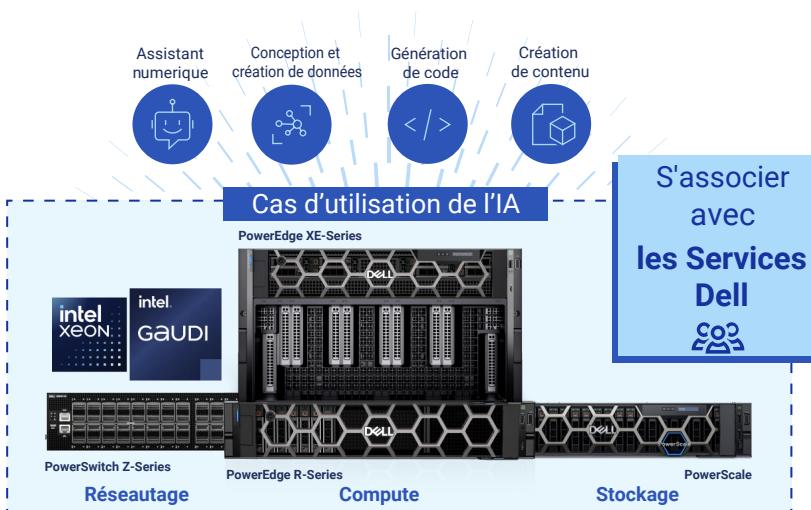
Plate-forme Dell AI

Dans le domaine de l'intelligence artificielle et de l'informatique avancée, la compétitivité exige des solutions robustes, fiables et performantes. Pour répondre à ces besoins, Dell Technologies a lancé la Dell AI Factory : une initiative conçue pour accélérer l'innovation en IA au sein des organisations de toutes tailles, grâce à un portefeuille de produits, de solutions et de services optimisés pour les charges de travail en IA. Elle regroupe ces technologies optimisées par l'IA avec un écosystème ouvert de partenaires, des solutions validées et intégrées, des services d'experts, ainsi que les meilleures pratiques, afin d'aider les clients à obtenir des résultats en matière d'IA plus rapidement.

L'usine Dell AI inclut désormais la plate-forme Dell AI avec Intel. Optimisée par les accélérateurs Intel® Gaudi® 3 AI, cette solution de pointe offre des performances élevées à grande échelle et la flexibilité de choisir le logiciel, l'Ethernet et le modèle de mise en œuvre qui répondent le mieux à vos besoins.

Des solutions conjointes pour les flux de travail de l'IA

La plate-forme d'IA Dell avec Intel combine la puissance d'Intel Gaudi 3 AI avec des serveurs, du stockage, des réseaux, des services et des logiciels Dell pour fournir des solutions conçues, testées et validées conjointement afin de garantir une expérience de déploiement transparente, le tout selon vos conditions. Éprouvée avec Dell Validated Designs, cette solution peut vous aider à démarrer votre parcours avec des cas d'utilisation AI qui vous aident à accélérer vos résultats commerciaux. Pour en savoir plus sur la façon dont le Dell PowerEdge XE9680 et l'Intel Gaudi 3 s'associent pour alimenter les flux de travail GenAI, cliquez ici.



Services professionnels conçus pour l'IA

Dell dispose d'un vaste portefeuille de services professionnels pour accélérer des résultats commerciaux à chaque étape de votre parcours AI. Notre équipe de services d'experts peut vous aider à développer, mettre en œuvre et faire évoluer votre AI solution, de l'élaboration de la stratégie à la préparation des données, plateforme déploiement, modèle aligné sur les cas d'utilisation, réglage du fonctionnement à l'échelle pour garantir L'organisation est prête à réussir dans l'avenir.

Principaux avantages



Performance

Hautes performances avec un TCO réduit



Scalabilité

Des architectures évolutives et une facilité d'utilisation combinées pour offrir un chemin simple vers la croissance.



Flexibilité

Élargit le choix en silicium, logiciels, Ethernet et options d'implémentation.

Faits rapides sur Intel Gaudi 3

- Universal Baseboard (HLB-325)
- 14.6 PFLOPS (FP8)
- >1TB HBM2e
- 29.6TB/S HBM bande passante
- 64 Moteurs de multiplication matricielle
- 192 200gbE RDMA NICs
- 9.6TB/s Mise en réseau bidirectionnelle



En savoir plus sur [Intel Gaudi 3](#)

Accélérateurs Intel® Gaudi® 3 AI

Les accélérateurs d'IA Intel Gaudi 3 offrent des avantages transformateurs aux entreprises qui déploient des solutions d'IA génératives, combinant des performances exceptionnelles, une évolutivité et une facilité d'utilisation, le tout avec un retour sur investissement et une efficacité améliorés par rapport aux solutions concurrentes. Son architecture avancée prend en charge les applications AI exigeantes charges de travail, offrant un rapport prix/performances supérieur à celui des concurrents. La mise en réseau Ethernet standard ouverte assure une intégration transparente et la flexibilité, en éliminant la dépendance vis-à-vis d'un fournisseur et en permettant aux entreprises de Choisir parmi une variété de solutions de mise en réseau standard de l'industrie.

Intégration avec les principaux frameworks et plateformes AI tels que Pytorch et Hugging Face améliore la productivité des développeurs et simplifie les migrations de modèles, ce qui facilite l'exploitation d'un large éventail d'applications de l'aluminium. Avec Intel Gaudi 3 accélérateurs d'IA, les organisations peuvent obtenir des informations plus rapidement et maintenir un avantage concurrentiel dans le paysage de l'IA en évolution rapide.maintain a competitive edge in the rapidly evolving AI landscape.

Une approche logicielle ouverte pour un maximum de flexibilité

La plate-forme d'IA Dell avec Intel comprend une pile complète de logiciels open source qui ont été testés et validés avec du matériel Dell et des accélérateurs d'IA Intel Gaudi 3. Cela permet aux organisations d'utiliser les outils dont elles ont besoin et d'éviter la dépendance vis-à-vis d'un fournisseur sans conjecture.

Pile logicielle de nœud de calcul

Outils et modèles	Hugging Face	Meta Llama 3
Cadre de ML	PyTorch	
Outils ML	jupyter	Kubeflow
Diffusion de modèles ML	vLLM	TGI TEI
Observabilité	Prometheus	Metrics: Intel Gaudi 3, Kubernetes, iDRAC, OneFS
Bibliothèques optimisées		
Environnement de développement	Intel® Gaudi® Software	
Modèles de programmation	Optimum	
Pilotes et durée d'exécution		
Gestionnaire de charge de travail	kubernetes	
Système d'exploitation	ubuntu	
Dell PowerEdge (iDRAC, BIOS, Firmware) Configuration		



Prêt à commencer?

Découvrez les accélérateurs d'IA Intel Gaudi 3 sur [l'Intel® Tiber™ AI Cloud](#)
Commencez à tester vos cas d'utilisation cibles avec des fonctions d'IA réelles sur [le catalogue de logiciels Intel AI](#)



En savoir plus
sur l'IA générative



Contactez un Expert
Dell Technologies



Joignez-vous à la
conversation

À propos de Dell Generative AI

Dell Technologies propose le plus large portefeuille de solutions d'IA générative au monde, des ordinateurs de bureau aux centres de données en passant par le Cloud, le tout en un seul endroit.

Vous pouvez accélérer votre transition vers l'IA grâce aux services Dell Technologies et à un vaste écosystème de partenaires qui inclut la prise en charge de nombreux modèles d'IA générative open source. Qu'un membre de l'équipe travaille avec une infrastructure Dell spécialement conçue pour alimenter des applications d'IA générative à l'échelle de l'entreprise, ou qu'il invite simplement l'IA à partir d'un ordinateur portable Dell, il travaille avec le portefeuille de solutions d'IA le plus large et le plus avancé du secteur.

Notre approche consiste à apporter de l'IA à vos données. Cela signifie répondre à des exigences strictes en matière de souveraineté des données et de conformité, tout en dimensionnant correctement l'infrastructure d'IA pour réduire les coûts, soutenir les objectifs de durabilité et vous permettre de garder le contrôle du modèle et de vos données.

Prochaines étapes

Commencez par un atelier d'accélération gratuit pour l'IA générative, qui peut vous aider à obtenir un consensus entre vos parties prenantes commerciales et techniques sur votre solution et les cas d'utilisation prioritaires.

Pour en savoir plus

[Atelier d'accélération pour l'IA générative](#)