



DELL Technologies

AMD
together we advance_

Générer de puissantes informations avec l'IA dans un écosystème ouvert



Serveur Dell PowerEdge XE9680 avec accélérateur AMD Instinct™
MI300X et plateforme logicielle ouverte ROCm™ 6

Avec la popularité croissante de ChatGPT et d'autres applications d'IA génératives (GenAI), l'intelligence artificielle (IA) se généralise rapidement, et de plus en plus d'organisations veulent en profiter. Une récente étude IDC montre que :



77 % des organisations

explorent les possibilités d'utilisation de l'IA ou investissent déjà de manière significative dans des initiatives.¹

Les organisations qui cherchent à tirer parti d'un accès plus rapide aux informations et d'autres avantages de l'IA doivent agir rapidement, si ce n'est pas déjà fait. Le moment est venu, en particulier avec un écosystème ouvert fortement impliqué dans l'avancement de l'IA et une nouvelle solution prête à stimuler l'innovation en matière d'IA, ainsi que votre activité.

¹ IDC, Future Enterprise Resiliency and Spending Survey, Wave 6, juillet 2023.



La voie est ouverte

Avec la progression de GenAI et d'autres applications avancées, la technologie de puce devient plus puissante et sophistiquée. Mais la base des systèmes propriétaires ont créé des obstacles pour les développeurs en première ligne.

Cette situation est sur le point de changer. Dell Technologies et AMD collaborent pour améliorer vos capacités avec des solutions d'infrastructure basées sur des technologies Open Source. L'utilisation de l'accélérateur AMD Instinct™ MI300X et de la plateforme logicielle ouverte AMD Radeon Open Compute (ROCm™) sur le serveur Dell PowerEdge XE9680 vous offre un écosystème puissant, une perspective sur l'avenir, pour susciter des opportunités et libérer votre potentiel.

Augmenter les performances et la flexibilité avec Dell Technologies et AMD

Dell Technologies et AMD aident les entreprises du monde entier à améliorer leurs résultats grâce à un système qui permettra aux développeurs de réussir.



Le serveur Dell PowerEdge XE9680 est une centrale de traitement des données spécialement conçue pour les tâches d'IA. Il prend en charge huit accélérateurs, ce qui est idéal pour les charges applicatives de formation et d'inférence en apprentissage automatique (ML) / Deep Learning (DL), en particulier pour ceux qui forment de grands modèles de langage (LLM).

L'accélérateur AMD Instinct MI300X, alimenté par l'architecture AMD CDNA™ (Compute DNA) de 3e génération, offre la plus grande capacité de mémoire à bande passante élevée sur le marché des processeurs graphiques, à savoir 1,5 téraoctet de mémoire d'accélérateur de processeur graphique cohérente par serveur pour la formation des plus grands modèles d'IA. Il permet également d'augmenter la capacité d'inférence en utilisant moins de serveurs.

Libérer l'accélération avec ROCm 6

ROCm, pour sa part, libère la puissance de calcul d'AMD Instinct. Plateforme logicielle ouverte, flexible et évolutive de calcul accéléré, ROCm est équipé de pilotes, d'API et d'autres outils de processeur graphique qui prennent en charge l'optimisation du code pour les accélérateurs AMD Instinct et simplifient la portabilité du code accéléré du processeur graphique existant.

Il simplifie également radicalement votre expérience. ROCm est optimisé au niveau de l'appareil avec une base de code librement partagée qui peut évoluer à mesure que les systèmes et leurs pièces évoluent. Cette flexibilité ouverte permet également un portage sans effort et un codage presque nul.

Un écosystème sans frontières

Fondés sur des normes ouvertes et l'interopérabilité, l'accélérateur AMD Instinct MI300X et ROCm sont pris en charge de manière native par un grand nombre de cadres et de bibliothèques d'IA. Cela offre un plus grand choix et une plus grande flexibilité à ceux qui souhaitent exploiter la puissance d'AMD tout en conservant la compatibilité avec les systèmes logiciels existants.

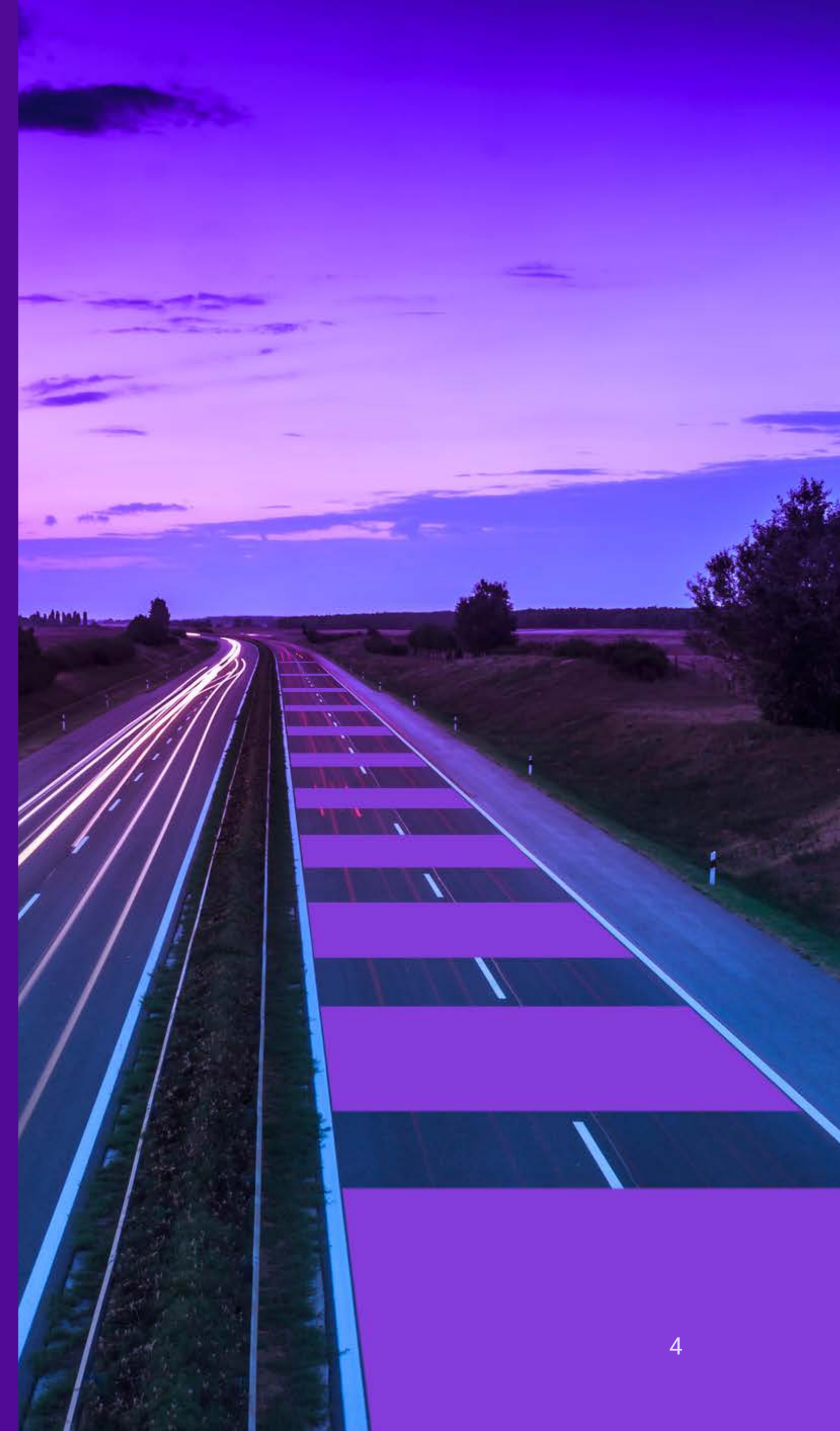


AMD ouvre la voie et vous donne la liberté de faire avancer vos projets à toute vitesse.

Avantages de ROCm

- Libère la puissance du processeur graphique pour accélérer les tâches de calcul
- Optimisé pour les charges applicatives ML/DL/GenAI à grande échelle
- Fournit une prise en charge immédiate des principaux cadres et modèles (y compris PyTorch[®], TensorFlow[™], etc.)
- Permet l'adoption rapide de nouvelles innovations (y compris la prise en charge native dans JAX, Open Neural Network Xchange (ONNX), OpenAI Triton, etc.)
- Favorise l'innovation, la différenciation et la collaboration avec la communauté de développement OSS
- Fournit une base flexible pour la portabilité ouverte

En savoir plus sur les [Nouveautés de ROCm](#).



Porter et optimiser facilement les applications pour AMD Instinct avec ROCm

- Permet le portage de logiciels à faible friction vers AMD. HIP permet aux développeurs de créer des applications portables pour AMD Instinct MI300X à partir d'un code source unique.
- Offre toutes les fonctionnalités avec les bibliothèques et opérateurs couramment utilisés
- Garantit le portage rapide et l'optimisation des performances de la plupart des applications

Prise en charge croissante en amont

ROCm fournit une prise en charge en amont pour les cadres actuels les plus populaires, y compris, mais sans s'y limiter :

- PyTorch
- TensorFlow
- Open Neural Network Xchange (ONNX)
- JAX



Gagner un temps précieux grâce à la portabilité ouverte

ROCm apporte une tranquillité d'esprit aux développeurs qui pourraient se méfier d'un nouvel écosystème et de son impact éventuel sur le code existant. L'un des principaux avantages de la plateforme ROCm est la portabilité des applications en apportant des modifications minimales au code existant. Le portage est simple et vous gagnez énormément de temps.

Quatre méthodes de portage des applications existantes sur l'accélérateur AMD Mi300X :

1 Exploiter les cadres d'IA existants qui prennent en charge ROCm en amont.

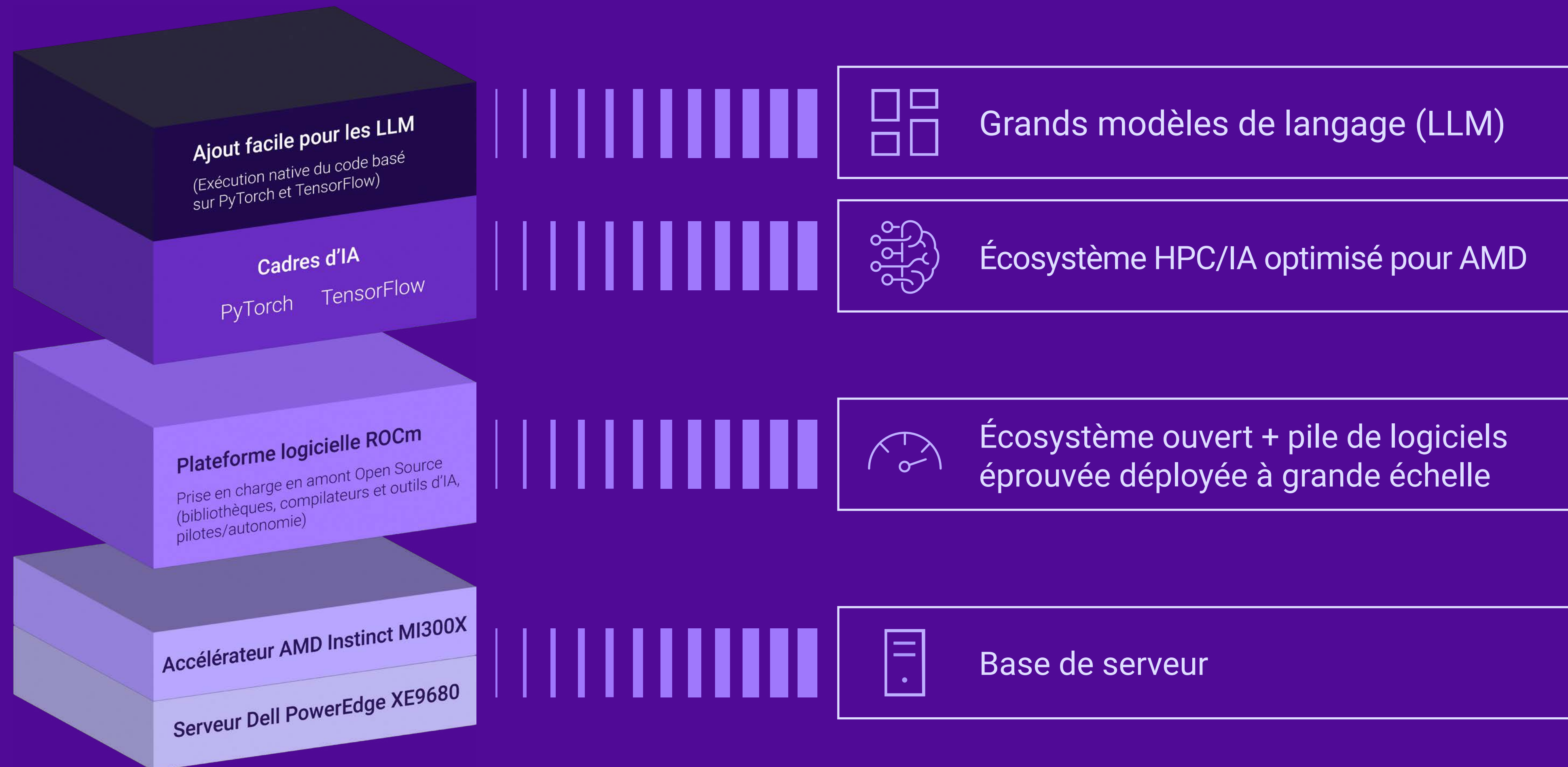
2 Utiliser des noyaux ML drop-in optimisés pour AMD Instinct MI300X. Il est également possible de tirer parti de l'outil de conversion AMD HIP, qui permet d'optimiser les applications CUDA sur des processeurs graphiques ROCm.

3 Utiliser des versions en miroir des bibliothèques CUDA existantes, qui sont optimisées et gérées pour ROCm.

4 Accéder aux conteneurs et applications d'IA préconfigurés AMD.

Lancez-vous

Accélérez l'IA existante et les nouveaux projets GenAI avec un cadre ouvert de Dell Technologies et AMD. Optimisez vos charges applicatives sur une base Dell PowerEdge éprouvée qui prend en charge les derniers accélérateurs AMD, y compris l'accélérateur AMD Instinct MI300X. [Rendez-vous sur le hub de développeur ROCm.](#)

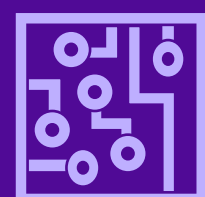


L'accélérateur AMD Instinct MI300X est prêt à être déployé

Doté de performances exceptionnelles, l'accélérateur AMD Instinct MI300X fournit un écosystème ouvert, rapide, évolutif et flexible, répondant aux besoins actuels des entreprises, en particulier pour faire face aux défis liés aux ressources, à la complexité, à la latence et à l'architecture. L'accélérateur AMD Instinct MI300X offre un avantage notable en termes de :



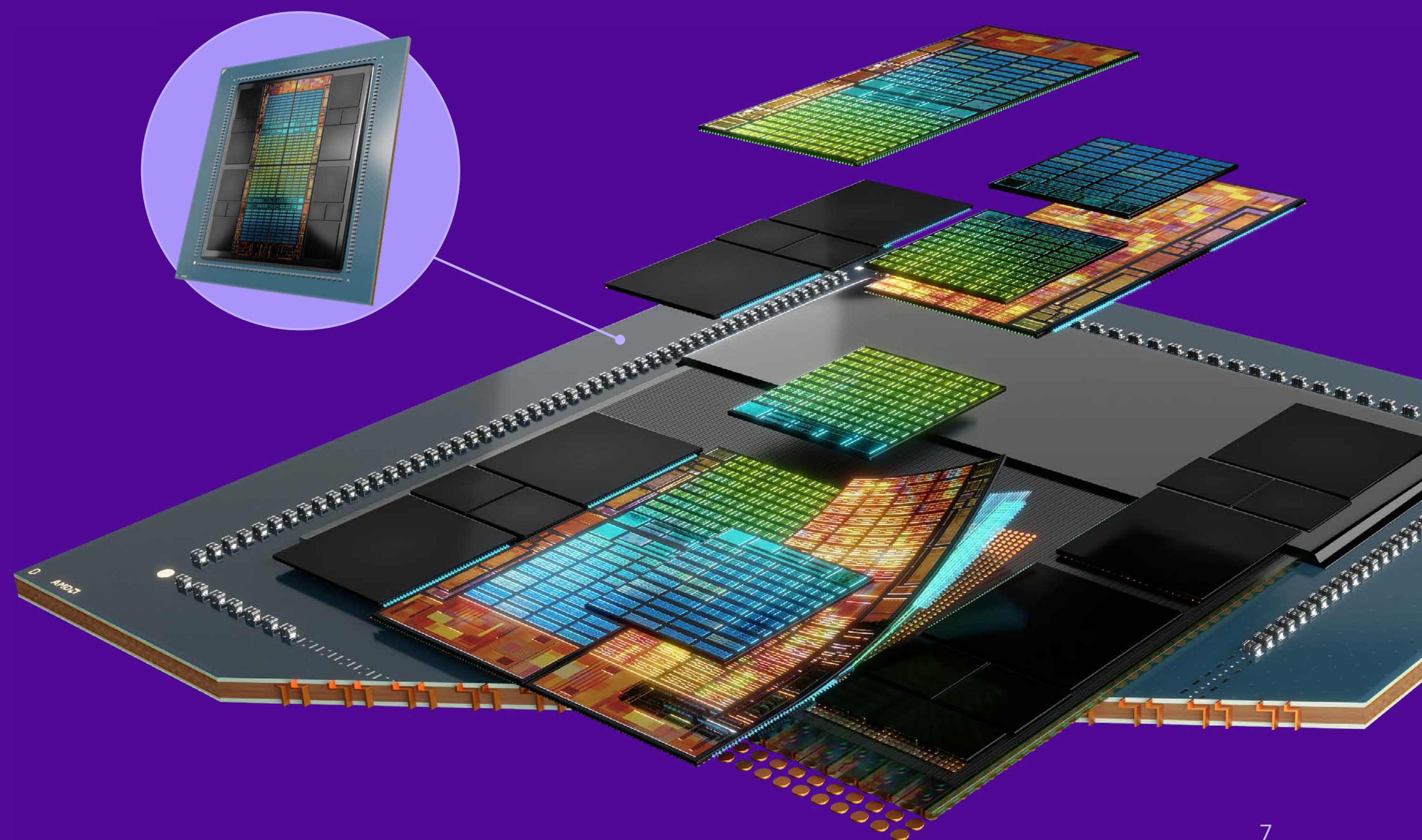
Performance : mettez en place de nouveaux workflows d'IA avec une mémoire HBM3 empilée sans précédent de 192 Go par processeur graphique fonctionnant à 5,3 To par seconde.



Densité : réduisez les déplacements de données grâce aux technologies de pointe d'empilement de puces qui permettent d'augmenter la capacité de calcul et de mémoire par serveur.



Ouvert : facilitez la recherche de résultats avec un écosystème conçu pour l'accessibilité et l'adaptabilité.



Libérer le potentiel de l'IA avec le serveur Dell PowerEdge XE9680

Obtenez des résultats plus rapides et plus intelligents avec une plateforme de serveur spécialement conçue pour l'IA. Améliorez les activités basées sur les données, obtenez des résultats visuels et optimisez les opérations d'IA à l'échelle de l'entreprise grâce à une solution qui accélère l'information, fournit une IA fiable et simplifie les opérations.



IA de confiance

Réduisez les risques et gagnez du temps en accélérant le cycle de vie de votre IA avec une infrastructure de solutions fiables et de haute qualité. PowerEdge XE9680 offre une protection intégrée de type Silicon Root of Trust jusqu'au retrait des actifs.



Opérations simplifiées

Contrôlez et gérez efficacement votre infrastructure d'IA et vos charges applicatives de performances où que vous soyez grâce à l'automatisation de l'infrastructure IA. Dell Technologies fournit des systèmes intelligents qui fonctionnent ensemble et indépendamment pour accélérer les tâches et simplifier votre expérience.



Accès plus rapide aux informations

Stimulez les performances tout au long du cycle de vie de l'IA et optimisez la prestation d'opérations d'IA au rythme de l'activité. Dell Technologies propose des calculs innovants pour vous aider à exploiter l'IA avec des performances maximales.

Stimulez l'innovation

Dell Technologies et AMD s'efforcent de faire progresser l'innovation au profit de nos clients privilégiés. Ensemble, nous apportons au marché la diversité du silicium dans une solution viable qui donne aux développeurs les moyens d'agir. Notre solution commune vous permet :



- **D'accélérer la commercialisation** en offrant davantage d'options et en réduisant les obstacles.
- **D'obtenir des informations plus rapidement** grâce à un écosystème ouvert qui ne vous empêche pas d'avancer.
- **De choisir la solution la mieux adaptée à vos besoins.** Profitez de l'avantage de la chaîne logistique Dell pour commercialiser des technologies innovantes au moment où elles sont le plus nécessaires.

Avancer à plein régime avec Dell Technologies et AMD

L'IA est une réalité aujourd'hui, et son influence ne fera que s'étendre et s'accélérer. Pour les entreprises qui cherchent à exploiter ses capacités, les possibilités semblent infinies.

C'est pourquoi vous ne souhaitez pas que votre potentiel soit limité par les obstacles technologiques qui peuvent parfois se dresser sur votre chemin. Au lieu de cela, profitez d'ensemble de solutions puissantes dédiées aux charges applicatives de l'IA.

L'accélérateur AMD Instinct MI300X et le logiciel ROCm sur le serveur Dell PowerEdge XE9680 sont la solution. Accélérez la transition vers de nouvelles applications et la commercialisation, et vous commencerez à obtenir des résultats significatifs pour votre entreprise et le monde entier.

Pour en savoir plus, consultez la page Dell.com/GPU

Copyright © 2023 Dell Inc. ou ses filiales. Tous droits réservés. Dell et les autres marques citées sont des marques commerciales de Dell Inc. ou de ses filiales. AMD, Instinct™, ROCm™ et leurs combinaisons sont des marques commerciales d'Advanced Micro Devices, Inc. PyTorch® est une marque commerciale ou déposée de PyTorch ou des concédants de licence de PyTorch. TensorFlow™ et toutes les marques associées sont des marques commerciales de Google Inc. CUDA® est une marque commerciale et/ou déposée de NVIDIA Corporation aux États-Unis et dans d'autres pays. Les autres marques peuvent être la propriété de leurs détenteurs respectifs. e-book publié en France en décembre 2023.

Dell Technologies estime que les informations figurant dans ce document sont exactes à la date de publication. Ces informations peuvent faire l'objet de modifications sans préavis.