

Le secret pour accélérer votre passage à l'IA

La mise en œuvre de l'IA à grande échelle peut être intimidante. NVIDIA AI Computing by HPE permet de réduire la complexité et d'accélérer l'adoption par l'entreprise

L'intelligence artificielle (IA) figure désormais en bonne place sur la feuille de route technologique de la quasi-totalité des entreprises. Aujourd'hui, 72 % des entreprises ont déployé une forme d'intelligence artificielle (IA), l'IA générative (GenAI) étant de loin la plus courante¹. Cependant, le degré d'adoption de l'IA varie énormément d'une entreprise à une autre.

La plupart en sont encore en effet au stade de la découverte de la GenAI, incitant généralement leurs employés à utiliser des chatbots publics. Un groupe plus restreint d'entreprises réalise des démonstrations de faisabilité à l'aide d'ensembles de données limités afin d'identifier les meilleurs cas d'utilisation de la technologie. Moins d'un tiers des projets de GenAI d'entreprise parviennent à passer de l'expérimentation à la production².

Selon Peter Moser, technologue senior chez HPE, une grande partie de ces premières expérimentations se déroulent dans le cloud public. La raison à cela ? Cette approche facilite le lancement, le développement et les tests d'un modèle simple dans un environnement sandbox. Mais cela peut changer rapidement.

« Dès que vous commencez à faire évoluer le projet et à traiter des volumes de données plus importants, l'IA devient rapidement compliquée et coûteuse », explique P. Moser. « De nombreux clients s'aperçoivent qu'ils ne veulent pas que leurs données soient dans le cloud public, hors de leur contrôle. Ils veulent rapatrier leurs projets sur site. » Cette évolution est déjà bien documentée : aujourd'hui, plus de 60 % des entreprises déclarent qu'elles prévoient d'entraîner et de déployer leurs applications d'IA dans un cloud privé sur site³.

L'un des principaux moteurs du cloud privé est la souveraineté et la sécurité des données, car aucune entreprise ne veut risquer d'exposer ses ressources les plus précieuses dans le cloud public ou d'enfreindre les réglementations de gouvernance des données, qui continuent d'émerger. Le cloud privé offre également davantage de contrôle sur les performances des applications, la latence, l'évolutivité et le coût, selon P. Moser.

Toutefois, le passage d'expériences limitées dans le cloud public à des applications d'IA à l'échelle de la production sur site est un énorme pas en avant, qui explique pourquoi 9 projets pilotes sur 10 n'atteignent jamais la phase de production⁴. « La grande majorité des entreprises ne disposent pas de l'infrastructure technologique et de l'expertise nécessaires pour faire fonctionner l'IA à grande échelle », fait remarquer P. Moser. En d'autres termes, elles ont besoin d'aide.

Vous avez décidé d'adopter un cloud privé pour l'IA. Et maintenant ?

Seules les entreprises les plus avancées au monde, dont beaucoup travaillent avec le machine learning depuis des années, disposent aujourd'hui en interne des outils et de l'expertise nécessaires pour déployer des applications de GenAI sans aide, explique P. Moser. Les autres sont confrontés à une courbe d'apprentissage décourageante, ainsi qu'à des investissements considérables en technologie et en temps.

« Les charges de travail de calcul haute performance comme l'IA utilisent des systèmes de fichiers parallèles », déclare P. Moser. « Elles nécessitent des GPU et des outils complexes dont la plupart des services informatiques ne disposent pas. Tout est nouveau pour elles. »

Pour entraîner ces modèles IA, les entreprises doivent déterminer comment consolider d'énormes volumes de données non structurées stockées dans des formats et des emplacements variés. Elles doivent recruter des spécialistes des données, affiner les compétences de leurs équipes informatiques pour les préparer à l'IA, déployer des garde-fous de gouvernance des données appropriés pour atténuer les risques et intégrer tous ces nouveaux systèmes dans leur infrastructure technologique existante.

Pourtant, elles ne doivent pas trop tarder à réaliser ces objectifs. La pression du marché autour de l'adoption et du déploiement rapides de la GenAI ne fera qu'augmenter.

« La mise en place de ces mesures par l'entreprise elle-même prend énormément de temps et comporte beaucoup de risques », remarque P. Moser. « La plupart des entreprises n'en savent tout simplement pas assez sur l'IA pour y parvenir en temps voulu. »

¹ "The state of AI in early 2024: GenAI adoption spikes and starts to generate value," McKinsey & Co., May 30, 2024

² "Now decides next: Moving from potential to performance," Deloitte, August 2023

³ "Essential Elements for Private Cloud Strategies," IDC, August 2024

⁴ "Reasons Why Generative AI Pilots Fail To Move Into Production," Forbes, Jan 8, 2024

HPE Private Cloud AI : une solution prête à l'emploi

Tous les facteurs décrits précédemment ont poussé Hewlett Packard Enterprise à s'associer à NVIDIA® pour proposer une solution clés en main unique destinée aux entreprises : HPE Private Cloud AI⁵. Cette offre va bien au-delà de la simple infrastructure. L'option de cloud privé évolutive, préconfigurée et complète fait partie du portefeuille NVIDIA AI Computing by HPE. Il s'agit d'une solution d'IA entièrement intégrée englobant le calcul et le stockage, la conception et la planification, l'intégration et l'optimisation, la formation et l'enseignement, le conseil et bien plus encore.

Le processus commence par une conversation approfondie sur les types d'opportunités que l'IA présente pour votre entreprise, une évaluation du degré de préparation de votre entreprise à l'IA, une discussion sur les meilleures pratiques et le développement de cas d'utilisation adaptés à vos besoins. HPE propose des ateliers aux entreprises qui souhaitent se lancer dans l'IA mais ne savent pas par où commencer ni même quelles questions poser, indique P. Moser.

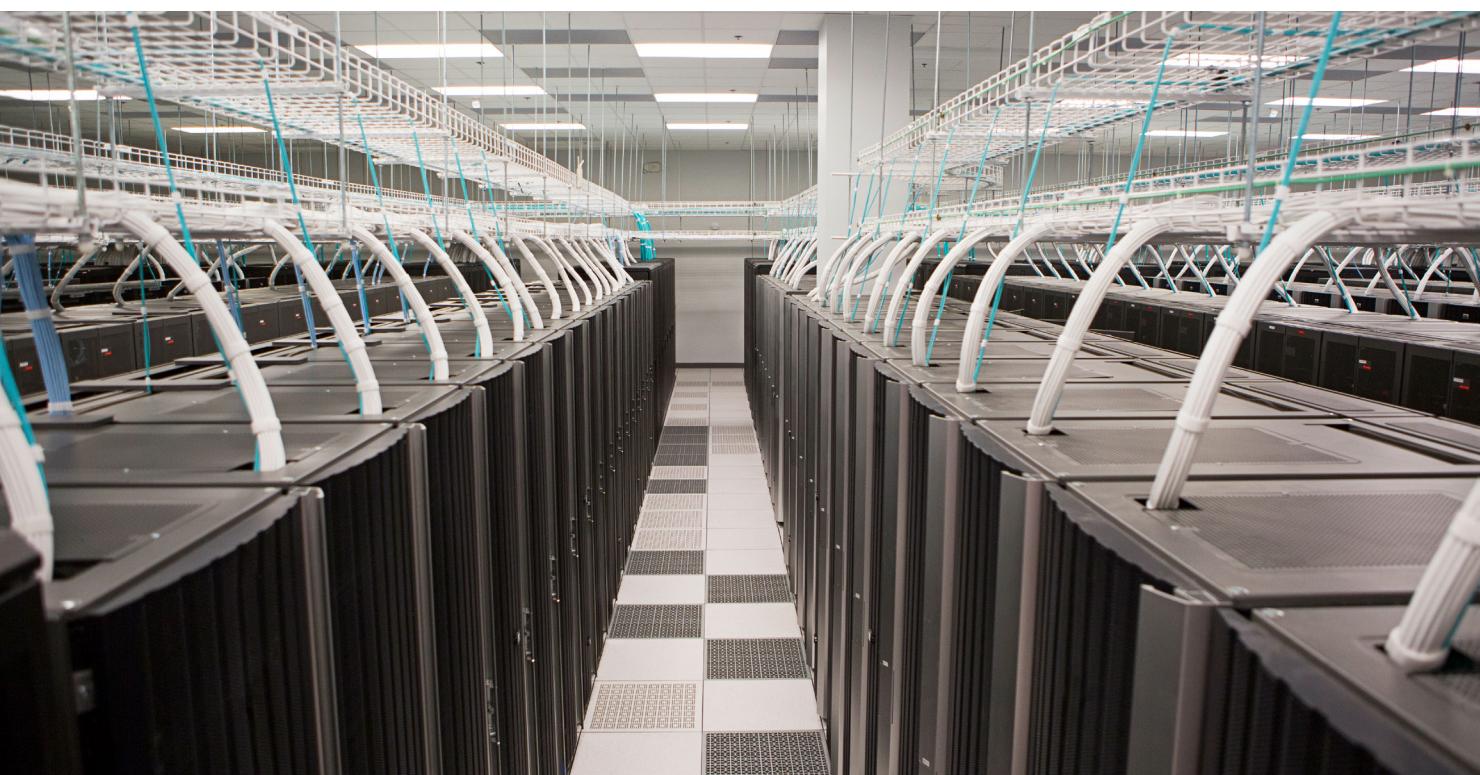
« L'une des grandes erreurs que commettent de nombreuses entreprises est de ne pas s'informer assez sur tout ce que l'IA peut faire », explique-t-il. « Elles réduisent d'emblée les limites du possible. Ces ateliers les aident à se concentrer sur les priorités les plus précieuses pour l'entreprise. Ils font entièrement évoluer le discours. »

L'étape suivante est la phase de conception technique et de planification. Toutes les entreprises n'ont pas les mêmes besoins ou niveaux d'ambition pour leurs projets de GenAI. C'est pourquoi HPE Private Cloud AI est disponible en quatre offres de serveurs préconfigurées, en fonction des besoins de calcul et de stockage de chaque cas d'utilisation et selon le déploiement : apprentissage, ajustement ou inférence.

« Vous obtenez une solution qui a déjà été préconçue et assemblée par les entreprises qui fabriquent tous ces outils et certifient que tout fonctionnera en parfaite harmonie », explique P. Moser. « Le lendemain, nous venons l'installer. En 48 heures, elle est opérationnelle. »

Chaque serveur HPE ProLiant configuré par HPE est certifié pour fonctionner avec 70 grands modèles de langage (LLM) disponibles en téléchargement, ainsi qu'avec des cas d'utilisation courants.

HPE Ezmeral Data Fabric permet aux entreprises de regrouper leurs datastores distribués en un tout cohérent, qui permet aux LLM d'ingérer plus facilement des pétaoctets de données d'entraînement non structurées. Les entreprises doivent simplement choisir le LLM adapté à leurs besoins, lui fournir leurs données compatibles avec l'IA et commencer à découvrir des informations.



⁵ "Hewlett Packard Enterprise introduces one-click-deploy AI applications in HPE Private Cloud AI," HPE, Sep 5, 2024

Soutenir l'ensemble du cycle de vie de l'IA

Lorsque les systèmes entrent en production, les entreprises peuvent choisir de les gérer elles-mêmes ou de faire appel à HPE et à ses partenaires pour obtenir des conseils sur la manière d'intégrer des solutions clés en main HPE et NVIDIA à leurs cadres de sécurité et de gouvernance des données existants. Nous vous proposons notre expertise tout au long du cycle de vie de l'IA, de la mise en œuvre à l'optimisation.

HPE peut également aider les entreprises à résoudre des problèmes épineux concernant la manière d'utiliser les données de manière éthique et transparente, d'éviter les violations de la

confidentialité et du droit d'auteur, et de se conformer à un nombre croissant de réglementations en matière d'IA.

« Il est incroyablement complexe de réussir son projet d'IA », ajoute P. Moser. « Tant qu'on n'a pas réalisé une tâche, on en ignore les tenants et les aboutissants, et ceux qui ont essayé de tout faire eux-mêmes découvrent que c'est beaucoup plus compliqué qu'ils ne le pensaient. »

En savoir plus

HPE.com/ai

Visit HPE.com



[Live Chat](#)

© Copyright 2025 Hewlett Packard Enterprise Development LP. Les informations figurant dans ce document sont susceptibles d'être modifiées sans préavis. Les seules garanties relatives aux produits et services Hewlett Packard Enterprise sont stipulées dans les déclarations de garantie expresses accompagnant ces produits et services. Aucune partie du présent document ne saurait être interprétée comme offrant une garantie supplémentaire. Hewlett Packard Enterprise décline toute responsabilité quant aux éventuelles erreurs ou omissions techniques ou rédactionnelles qui pourraient être constatées dans le présent document.

a50011771FRE, Rév. 2

HEWLETT PACKARD ENTERPRISE

hpe.com