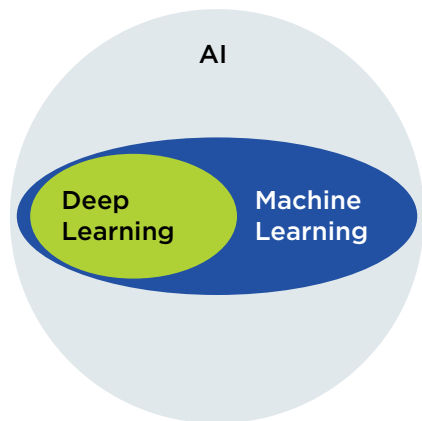


Rationalisez l'intelligence artificielle grâce au partenariat technologique entre Nutanix, NVIDIA et Mellanox



« Même si vous disposez des meilleurs data scientists, des meilleurs ingénieurs en machine learning et des meilleurs outils, vous n'obtiendrez pas des résultats optimaux sur de grands ensembles de données si vous n'êtes pas équipé de la bonne infrastructure. »

- InformationWeek

Donnez un nouvel élan à vos initiatives d'IA grâce à une architecture puissante et évolutive

Les entreprises de divers secteurs, dont la production industrielle, les services financiers, le retail et la santé, se tournent vers la transformation numérique et l'intelligence artificielle (IA) pour mieux comprendre leurs données et ainsi accroître l'engagement client, améliorer l'efficacité des processus métier, optimiser la supply chain et créer de nouvelles opportunités commerciales. Le succès de l'IA passe souvent par l'identification d'idées novatrices et la mise à l'échelle rapide des projets, de la preuve de concept à la production, en passant par le projet pilote.

L'IA englobe toutes les techniques utilisées pour permettre aux ordinateurs d'effectuer des tâches qui, autrement, nécessiteraient l'intelligence humaine, y compris les technologies de machine learning et de deep learning qui sont actuellement à la une de la presse grand public et spécialisée.

Le machine learning (ML), ou apprentissage automatique, repose sur les statistiques et l'optimisation mathématique et s'applique à la prédiction, à l'analyse et au data mining. Les méthodes de machine learning peuvent inclure des approches mathématiques et statistiques telles que les arbres de décision et les algorithmes de clustering en plus du deep learning.

Le deep learning (DL), ou apprentissage profond, est le domaine de la ML qui utilise différentes architectures de réseaux de neurones artificiels à forte intensité de calcul. Les entreprises appliquent le deep learning pour détecter des tendances, comprendre les comportements et faire des prédictions sur la base de grandes quantités de données telles que les données recueillies à partir de l'IoT.

Le grand volume et la variété des données requises, combinés à l'intensité de calcul inhérente au deep learning, créent des besoins particuliers en infrastructure qui peuvent constituer un défi pour les équipes chargées des projets d'intelligence artificielle. L'architecture d'une solution d'infrastructure d'IA, le déploiement du matériel et des logiciels nécessaires, ainsi que l'identification et la suppression des goulots d'étranglement peuvent être un processus coûteux, long et sujet aux erreurs.

Afin de permettre à votre équipe de se concentrer sur ses objectifs d'IA plutôt que sur la complexité de son infrastructure, Nutanix s'est associée à NVIDIA et à Mellanox pour concevoir, tester et valider une architecture de référence capable de résoudre les problèmes de deep learning les plus complexes. Cette solution est rapide à déployer, simple à mettre à l'échelle et fournit les services de données nécessaires pour répondre aux besoins des développeurs de logiciels et des data scientists.

COMPOSANTS D'ARCHITECTURE IA

Appliance Nutanix 100 % Flash

Mise à l'échelle prévisible à partir d'un système distribué, défini par logiciel, conçu pour s'auto-guérir en cas de panne.

Le Nutanix 3060 offre un mélange de performances de traitement et de capacité de stockage. Une empreinte 2RU fournit :

- 4 nœuds de calcul indépendants.
- Jusqu'à 6 disques SSD à connexion directe par nœud, 24 périphériques par appliance.
- Jusqu'à 20 cœurs/socket et 768 Go de mémoire par nœud.
- Des solutions Edge sont disponibles pour l'inférence.

NVIDIA DGX-1

La plateforme d'intelligence artificielle leader du marché, DGX-1, est optimisée pour les problèmes d'intelligence artificielle les plus complexes. Une empreinte 2U fournit :

- 8 GPU connectés par NVlink.
- NVIDIA GPU Cloud (NGC), avec un catalogue complet d'outils logiciels d'intelligence artificielle optimisés.
- Registre de conteneurs : conteneurs optimisés par NVIDIA pour les principaux algorithmes.

LES AVANTAGES D'UNE ARCHITECTURE DE POINTE

L'approche de Nutanix en matière d'intelligence artificielle combine la technologie éprouvée d'infrastructure hyperconvergée de Nutanix Enterprise Cloud avec les processeurs GPU de NVIDIA et la commutation Ethernet 100 Gbit/s à faible latence de Mellanox. Le résultat est une solution de pointe équilibrée conçue pour éliminer les goulots d'étranglement. L'architecture prend en charge des ensembles de données volumineux et variés, délivrant efficacement les données aux GPU NVIDIA et évoluant en fonction de la croissance de vos besoins.

Le deep learning nécessite de grandes capacités de calcul. L'équipe IT peut avoir besoin des capacités de traitement combinées de plusieurs GPU lors de la formation des modèles d'inférence. Un grand ensemble de données peut être nécessaire pour la formation initiale d'un modèle. Le modèle est continuellement formé à mesure que de nouvelles données arrivent, afin de l'affiner. Le processus de formation est donc itératif et la charge de travail totale de l'IA s'accroît avec le temps. Les avantages de la solution Nutanix incluent :

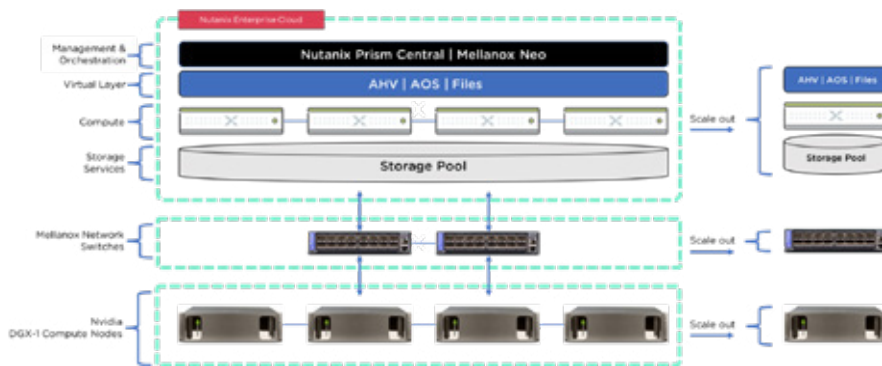
- **La facilité de déploiement et de gestion.** Un ensemble complet d'outils simplifie le déploiement et garantit que les mises à niveau de micrologiciel, de correctifs et de logiciels sont transparentes. La protection des données et la reprise après sinistre sont faciles à gérer et protègent vos opérations d'IA.
- **Des performances et une évolutivité exceptionnelles.** Les déploiements d'IA peuvent prendre rapidement de l'ampleur lorsque vous passez à la production. Les ensembles de données s'agrandissent, de nouvelles sources de données sont ajoutées et les algorithmes deviennent de plus en plus complexes. Avec Nutanix, vous pouvez commencer à petite échelle et évoluer en fonction de la croissance de vos besoins sans avoir à vous soucier des goulots d'étranglement ni à reconfigurer votre architecture.
- **La facilité du processus DevOps.** La formation en IA est de nature itérative. Les applications doivent être mises à jour régulièrement à mesure que le modèle se forme. Nutanix Calm simplifie la gestion du cycle de vie des applications et permet d'effectuer régulièrement des modifications sans avoir recours aux processus manuels sujets aux erreurs.
- **La sécurité des données intégrée.** Pour assurer la sécurité des données sensibles, de nombreux architectes estiment n'avoir d'autre choix que de déployer une infrastructure dédiée à l'IA. Le logiciel Nutanix intègre des fonctions de sécurité, dont l'authentification à deux facteurs et le chiffrement des données au repos, dans un cadre de sécurité renforcé qui a été certifié pour assurer la conformité aux normes les plus strictes.

CARACTÉRISTIQUES DE L'ARCHITECTURE

L'architecture de référence combine les éléments suivants, comme l'illustre la figure ci-dessous :

- Les serveurs Nutanix 3060-G6 fournissent des services de stockage et de données de haute performance, ainsi que des capacités de calcul CPU.
- NVIDIA DGX-1 fournit les capacités de calcul GPU nécessaires aux algorithmes d'IA complexes.
- Les commutateurs réseau Mellanox Ethernet 100 Gbit/s offrent un haut débit et une faible latence pour acheminer efficacement les données entre Nutanix et NVIDIA.

Tous les composants sont évolutifs, ce qui garantit que les CPU, la capacité de stockage et les E/S de données, la connectivité réseau et les GPU peuvent tous être étendus rapidement en fonction des besoins.



Mellanox

Mellanox offre une commutation Ethernet à faible latence et haut débit. Le Mellanox SN2100 présente une empreinte compacte 1RU avec :

- 16 ports (fonctionnant à 25, 40, 50 ou 100 Gbit/s Ethernet)
- Un débit de 3,2 To/s
- Une capacité de traitement de 2,38 Bpps
- Montée en charge jusqu'à 128 ports

PREMIERS PAS

Pour en savoir plus sur l'approche Nutanix en matière d'IA, téléchargez la présentation détaillée de l'architecture de référence sur www.nutanix.com/ai, contactez Nutanix par e-mail à l'adresse contact-france@nutanix.com, suivez-nous sur [Twitter @NutanixFrance](https://twitter.com/NutanixFrance) ou envoyez-nous une demande sur www.nutanix.com/demo pour organiser votre propre briefing personnalisé.

NUTANIX
YOUR ENTERPRISE CLOUD

T. +33 (0)1 82 88 15 90

contact-france@nutanix.com | www.nutanix.fr | [@NutanixFrance](https://twitter.com/NutanixFrance)

© 2019 Nutanix, Inc. Tous droits réservés. Nutanix, le logo Nutanix et tous les noms de produits et services mentionnés dans le présent document sont des marques déposées ou des marques de commerce de Nutanix inc. aux États-Unis et dans d'autres pays. Toutes les autres marques de commerce mentionnées dans le présent document le sont à des fins d'identification seulement et peuvent être des marques de commerce de leurs détenteurs respectifs.