

## Livre blanc

# Accélérer la modernisation de l'IA grâce à l'infrastructure de données

Sponsorisé par : IBM Corporation

Ashish Nadkarni                      Sriram Subramanian                      Matt Leib  
Février 2021

## SYNTHÈSE

---

Les fonctionnalités de l'intelligence artificielle (IA), du machine learning (ML) et, souvent, du deep learning (DL) sont désormais des composantes essentielles des projets de transformation numérique (DX). Les opportunités business qui peuvent être réalisées grâce à l'IA sont exceptionnellement prometteuses. Les entreprises sont de plus en plus conscientes que le fait de ne pas exploiter l'IA peut être un désastre business, si les concurrents obtiennent une multitude d'informations auparavant indisponibles pour développer et satisfaire leur clientèle. Très peu d'entreprises pensent aujourd'hui que "l'IA n'est pas pour nous" ou que "l'IA est surtout un phénomène de mode". Au contraire, des réalisations sérieuses en matière d'IA sont entreprises au niveau mondial, dans tous les secteurs et dans les sociétés de toutes tailles.

Les entreprises cherchent à développer une transformation et une modernisation axée sur l'IA, qui implique de passer les phases d'essais et de parvenir à générer de la valeur business grâce à leurs investissements dans l'IA. Le succès des investissements des entreprises dans la transformation numérique basée sur l'IA est directement lié à l'ampleur de l'expertise requise pour développer, mettre en œuvre et faire évoluer les solutions d'IA. Les secteurs d'activité (LOB), le personnel IT, les data scientists et les développeurs de nombreuses entreprises ont travaillé pour découvrir l'IA, comprendre les cas d'usage, définir une stratégie IA pour leur entreprise, lancer des projets d'IA, mais aussi développer et tester les applications d'IA qui offrent de nouvelles perspectives et fonctionnalités à l'aide d'algorithmes de machine learning, en particulier de deep learning (DL).

Au fur et à mesure que les projets se développent, de nouvelles questions émergent. Les entreprises savent, et en ont peut-être fait l'expérience, qu'elles ne peuvent pas utiliser l'informatique standard et l'infrastructure de stockage existante. Elles se rendent également compte que l'entraînement de l'IA (l'entraînement du modèle d'IA) et l'inférence de l'IA (l'utilisation du modèle entraîné pour comprendre ou prédire un événement) nécessitent différents types de calcul évolutif avec une infrastructure de stockage tout aussi évolutive.

IDC constate que les entreprises qui tentent d'utiliser leur infrastructure existante sans en moderniser une partie ou la totalité ont plus de risques d'échouer. Même en termes de modernisation, les objectifs varient d'une entreprise à l'autre. Même si les entreprises maîtrisent bien les fonctions de calcul, elles sous-estiment souvent la valeur du stockage dans l'IA. De plus, les applications d'IA, et notamment les systèmes de deep learning, qui analysent des quantités de données à croissance

**IDC constate que les entreprises qui tentent d'utiliser leur infrastructure existante sans en moderniser une partie ou la totalité ont plus de risques d'échouer. Même au cours de la phase de modernisation, les objectifs varient d'une entreprise à l'autre. Si les entreprises maîtrisent bien les fonctions de calcul, elles sous-estiment souvent la valeur du stockage dans l'IA.**

exponentielle, sont extrêmement exigeantes et nécessitent de puissantes capacités de traitement basées sur un grand nombre de cœurs de calcul, et les systèmes de stockage standard ne permettent pas de réaliser ces tâches d'IA. Enfin, il est important que les projets liés à l'IA soient pris en compte dans les efforts de modernisation des applications qui incluent Kubernetes et/ou des conteneurs et l'intégration avec un ou plusieurs services de cloud public par le biais d'une architecture de cloud hybride.

Les recherches d'IDC montrent qu'en termes d'infrastructure de stockage, une attention inadéquate ou insuffisante concernant les détails, peut rapidement faire dérailler les projets de transformation via l'IA. Pour combler cette lacune, les entreprises qui utilisent leur infrastructure existante et sont maintenant prêtes à la mettre en production, doivent remanier cette infrastructure afin d'obtenir les performances de traitement parallèle requises. Pour cela, elles doivent investir dans des solutions de stockage plus modernes, capables d'évoluer à grande échelle et intégrées au cloud, des conteneurs et des outils de calcul intensif pour le déploiement global et l'accès aux données. Les solutions telles que IBM Spectrum Scale et IBM Elastic Storage System (ESS) fournissent les composants nécessaires à une architecture d'IA. Ces solutions sont adaptées aux applications d'IA, au déploiement conteneurisé et à un déploiement de cloud hybride spécifiquement axé sur les applications d'IA.

## APERÇU DE LA SITUATION

---

### IA et ML, ici et maintenant

Les entreprises du monde entier utilisent activement les nouvelles opportunités offertes par les investissements dans l'IA pour accélérer leurs projets de transformation numérique. L'intelligence artificielle est un ensemble de technologies qui utilisent le traitement automatique du langage naturel (NLP), l'analyse d'images/vidéos, le machine learning, les graphes de connaissances et autres technologies pour répondre à des questions, découvrir des informations et fournir des recommandations. Ces systèmes émettent des hypothèses et formulent des réponses possibles sur la base des preuves disponibles, peuvent être entraînés par l'importation de grandes quantités de contenu, et s'adaptent et apprennent de leurs erreurs et de leurs échecs grâce à un réentraînement ou à une supervision humaine. IDC prévoit que d'ici 2022, au moins 60% de ces cas d'usage centrés sur l'IA seront déployés dans au moins 65% des entreprises du Global 2000, soit une croissance de 34% par rapport à 2019.

L'intelligence artificielle devient rapidement omniprésente dans les entreprises pour l'automatisation des processus et des workflows. En 2019, IDC a examiné 176 cas d'usage de transformation numérique dans huit domaines fonctionnels, notamment l'expérience client (CX), la stratégie juridique et d'entreprise, les installations et les achats. IDC en a conclu qu'environ 26% de ces cas d'usage étaient à la fois dépendants de l'IA et actuellement déployés dans la plupart des entreprises.

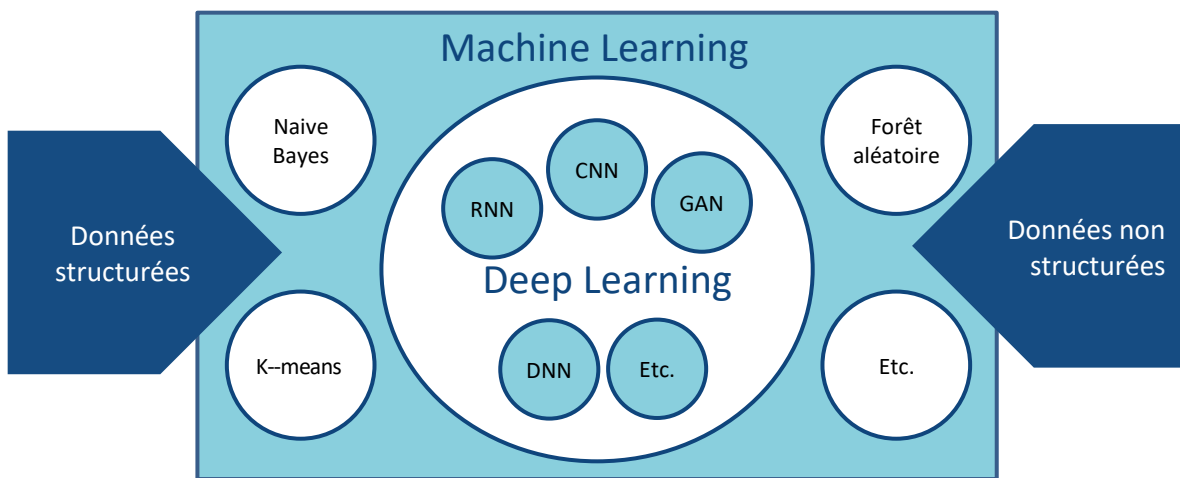
Cela signifie que, bientôt, la plupart des grandes entreprises exploiteront les technologies d'IA, telles que le traitement automatique du langage naturel, le machine learning, le deep learning et la reconnaissance vocale, dans l'ensemble de l'entreprise, afin de développer les opérations, de donner un sens aux données non structurées et de fournir des informations métiers intelligentes. Cependant, les entreprises qui n'ont pas encore trouvé comment faire passer les cas d'usage d'IA d'un POC à la production prendront encore plus de retard, creusant ainsi le fossé numérique.

**IDC prévoit que d'ici 2022, au moins 60% de ces cas d'usage centrés sur l'IA seront déployés dans au moins 65% des entreprises du Global 2000, soit une croissance de 34% par rapport à 2019.**

Le machine learning est un sous-ensemble de techniques d'IA qui permet aux systèmes informatiques d'apprendre et d'améliorer leur comportement pour une tâche donnée sans devoir être programmés par un humain. Les modèles de machine learning sont des algorithmes qui peuvent s'améliorer au fil du temps en se testant de manière répétée à l'aide de grandes quantités de données structurées et/ou non structurées jusqu'à ce qu'ils soient considérés comme ayant "appris" une tâche (par exemple, la reconnaissance d'un visage humain). La figure 1 illustre le fait que le deep learning est un sous-ensemble du machine learning. Les architectures DL typiques sont les réseaux neuronaux profonds (DNN), les réseaux neuronaux convolutifs (CNN), les réseaux neuronaux récurrents (RNN), les réseaux adversaires génératifs (GAN), et bien d'autres encore.

**FIGURE 1**

**Applications de machine learning et de deep learning**



Source : IDC, 2021

Les plateformes d'IA comprennent les logiciels d'IA conversationnelle (par exemple, les assistants numériques) ; l'analyse prédictive pour découvrir les relations cachées entre les données et faire des prédictions ; l'analyse de texte et le langage naturel pour reconnaître, comprendre et extraire la valeur d'un texte ; l'analyse vocale/de la parole pour reconnaître, identifier et extraire des informations issues de l'audio, de la voix et du discours ; et l'analyse d'images et de vidéos pour reconnaître, identifier et extraire des informations à partir d'images et de vidéos, en particulier la reconnaissance des formes, des objets, des couleurs et autres éléments tels que les personnes, les visages, les émotions, les voitures et les paysages, pour n'en citer que quelques-uns.

De nombreuses entreprises sont bien avancées dans les projets d'IA et ont atteint un stade où elles sont prêtes à commencer le déploiement de l'IA à grande échelle. D'autres sont en train d'expérimenter l'IA, tandis qu'un troisième groupe en est au stade de l'évaluation de ce que les applications d'IA peuvent apporter à leur entreprise.

En ce qui concerne le premier groupe (prêt pour le déploiement), IDC observe une série de cas d'usage de l'IA qui ont commencé à être mis en œuvre par les entreprises, les gouvernements et autres organisations. Les cinq cas d'usage les plus courants aujourd'hui, classés en fonction du montant que les entreprises y consacrent en termes de matériel, de logiciels et de services, sont les suivants :

- **Agents de service client automatisés.** Dans le secteur bancaire, par exemple, ces applications d'IA fournissent un service client via un programme d'apprentissage qui comprend les besoins et les problèmes des clients et aide une banque à réduire le temps et les ressources nécessaires pour résoudre les problèmes des clients. Ces agents sont de plus en plus utilisés dans tous les types de secteurs.
- **Recommandation et automatisation des processus de vente.** Utilisées dans divers secteurs, il s'agit d'applications d'IA qui fonctionnent avec les systèmes de gestion de la relation client (CRM) pour comprendre le contexte du client en temps réel et recommander des actions pertinentes aux agents commerciaux.
- **Systèmes automatisés de veille et de prévention des menaces.** Devenant un élément essentiel de la prévention des menaces au sein des gouvernements et des entreprises, ces applications d'IA traitent les rapports de veille, en extraient des informations, établissent des relations entre les divers éléments, puis identifient les menaces pesant sur les bases de données, les systèmes, les sites web, etc.
- **Analyse et enquête sur les fraudes.** Dans le secteur des assurances, mais aussi dans d'autres secteurs, ces applications d'IA utilisent les règles de fonctionnement pour identifier les transactions frauduleuses, et elles apprennent automatiquement à identifier les différents systèmes de fraude liés aux assurances.
- **Maintenance prédictive automatisée.** Dans le secteur industriel, ces applications d'IA sont basées sur des algorithmes de machine learning qui construisent un modèle prédictif précis des défaillances potentielles des machines dans les usines, réduisant ainsi les temps d'arrêt et les coûts de maintenance.

Les autres cas d'usage de l'IA qui ont gagné du terrain dans les entreprises (classés par ordre de dépenses en matériel, logiciels et services) sont les suivants :

- Conseillers en programmes et systèmes de recommandation
- Systèmes de diagnostic et de traitement
- Automatisation intelligente des processus
- Systèmes d'enquête et de recommandation en matière de gestion de la qualité
- Automatisation informatique et assistants numériques pour les collaborateurs des entreprises
- Conseillers d'achat experts et recommandations produits
- Approvisionnement, logistique et veille réglementaire
- Gestion des actifs/flottes et traitement automatisé des réclamations
- Jumeaux numériques et simulations numériques avancées
- Sécurité publique et intervention d'urgence
- Apprentissage adaptatif
- Mise en réseau intelligente
- Gestion de flotte, du fret et des actifs
- Recherche et découverte pharmaceutique

## La transformation par l'IA nécessite une base de données solide

Les projets ML et DL reposent en grande partie sur une combinaison de diverses données structurées et non structurées (reportez-vous à la figure 1). Les projets d'IA comportent de multiples étapes, notamment la collecte de données, la configuration, le nettoyage, la vérification, la construction de modèles, l'entraînement, les tests, l'inférence et la suppression des données. Chaque étape ayant des exigences différentes, allant d'un accès plus rapide, d'un stockage à faible latence à un stockage d'archives à faible coût, les entreprises doivent choisir une infrastructure de stockage appropriée et rentable pour ces étapes. Elles doivent également utiliser un ensemble d'outils pour gérer les données d'IA tout au long de leur cycle de vie.

Avec le déploiement de la 5G et la forte croissance des capteurs basés sur l'IoT, davantage de données sont générées et consommées à la périphérie. Les caméras et les assistants intelligents sont déployés à la périphérie, sur l'appareil, afin de fournir une réponse plus rapide et une meilleure expérience utilisateur. Le besoin d'une IA locale, traitée à la périphérie du réseau ou sur les points d'extrémité, augmente. Les applications d'IA sensibles aux temps de latence et fonctionnant sur des devices périphériques, avec une connectivité limitée, nécessiteront une grande évolutivité. Les cas d'usage du edge computing basé sur l'IA seront adoptés.

## Une infrastructure de données adaptée à l'IA accélère la transformation par l'IA

Lorsque les entreprises investissent dans l'IA, plusieurs facteurs liés à la conception et au déploiement de l'infrastructure deviennent capitaux. S'inspirant des fournisseurs de services hyperscale et cloud, les entreprises abordent les besoins en infrastructure par la création d'une infrastructure de données unifiée. Une infrastructure de données constitue une base commune pour les projets d'IA, avec notamment une couche de calcul et de stockage hautement efficace et évolutive. Au lieu d'aborder les applications d'IA de façon homogène, une infrastructure de données les traite de manière composite et connecte une partie de cette application à la bonne couche de calcul alimentée par une couche de stockage appropriée en fonction d'un mélange de données structurées et non structurées. Cette approche composite accélère la transformation par l'IA selon trois axes :

- **L'évolutivité.** Cette dimension d'échelle décrit le niveau auquel l'application fonctionne. Les sous-dimensions fondamentales (calcul, mise en réseau et persistance des données (stockage)) sont toutes liées au matériel. Il est crucial que les sous-dimensions liées au logiciel, telles que l'orchestration, gagnent en équité pour maintenir l'équilibre avec l'augmentation de la taille et de la complexité de la pile.
- **La portabilité.** Il s'agit de la capacité de l'application à être déployée entre le cœur, la périphérie et les points d'extrémité. Aujourd'hui, nombre de ces applications sont statiques par nature (c'est-à-dire conçues pour être exécutées en un seul déploiement). De plus en plus, les entreprises envisagent de développer des applications dans un environnement (par exemple, le cloud public) puis de les installer (en production) dans un autre (par exemple, en périphérie). Ce modèle est analogue au modèle actuel de développement et de déploiement des applications mobiles.
- **La durée.** Cette dimension concerne la continuité temporelle de l'application elle-même. De nombreuses applications d'IA empruntent leur conception à l'informatique haute performance ou au déploiement du big data et de l'analytique. Elles sont conçues pour être de nature discontinue. De plus en plus, et cela grâce à la forte croissance des accélérateurs haute performance, les applications d'IA peuvent analyser des données en continu en temps réel ou quasi réel.

### Les défis des projets d'IA

Les recherches d'IDC montrent que les entreprises citent souvent les défis suivants concernant leurs projets d'IA du point de vue de la gestion des données :

- Des cycles d'acquisition et de préparation des données consommateurs de temps
- Des silos d'infrastructures pour les différents cas d'usage de l'analytique
- Les multiples copies des mêmes données sans une seule source unique de données de référence
- La nécessité de gérer et de protéger en toute sécurité la provenance des données pour assurer leur réplication
- Un besoin d'accessibilité (cloud hybride) et de collaboration globales
- L'intégrité des données une fois qu'elles ont été récupérées et nettoyées

## Les mythes et les incontournables de l'infrastructure de données de l'IA

Une base essentielle, bien que souvent négligée, est l'infrastructure de stockage. Le déploiement de l'IA à grande échelle impose souvent des besoins plus élevés de l'infrastructure de stockage en termes de capacité (croissance) et de performances (IOPS et bande passante). Les entreprises supposent souvent que le stockage interne sur serveur ou le stockage d'entreprise utilisé pour d'autres applications est suffisant pour exécuter des applications d'IA. Et une fois l'infrastructure mise en place, elles se rendent compte que le stockage est le maillon faible de la chaîne. Chacune de ces applications d'IA présente des besoins différents et, par conséquent, de défis pour le service IT. Les acheteurs et les éditeurs doivent donc éviter d'adopter un raisonnement limité où chaque problème peut être résolu par une même solution miracle.

Une approche plus prudente consiste à adopter une approche globale de l'infrastructure des données. Même si les mécanismes d'accès et de mise à l'échelle de la persistance des données sont des enjeux majeurs, les entreprises doivent élargir leur champ d'action pour inclure la mise en réseau et l'intégration entre les fonctions de calcul, les logiciels de stockage et les systèmes. Les entreprises doivent adopter une approche cohérente de l'infrastructure de données de bout en bout, et pas seulement une approche consistant à "rajouter" d'autres systèmes de stockage. IDC estime que les exigences en matière d'infrastructure de données, autour du stockage, peuvent être ramenées aux domaines clés abordés dans les sections qui suivent.

### *Intégration du calcul*

On suppose souvent, et à tort, que toutes les applications d'IA sont conteneurisées. Ce n'est pas le cas. Elles sont nombreuses à être exécutées sur des serveurs bare metal ou sont même virtualisées. De nombreuses applications d'IA sont optimisées pour utiliser des accélérateurs. Cela ne signifie pas que toutes les applications d'IA sont plus performantes sur un ordinateur accéléré. Mais l'ordinateur accéléré représente un ensemble différent de défis pour les applications en général.

### *Persistance et accès aux données*

Si les besoins de calcul des applications de l'IA varient, il en va de même pour les besoins de persistance des données. La couche de persistance des données est un aspect sous-représenté et mal compris de la pile des applications d'IA. On part souvent du principe, encore une fois erroné, que toutes les applications d'IA nécessitent une grande quantité de stockage haute performance. En fait, les applications d'IA ne sont pas toutes des "ensembles de big data". Elles peuvent échantillonner de nombreux petits ensembles de données simultanément pendant une courte période. De même, les applications bare metal exécutées sur des plateformes ouvertes utilisent souvent l'accès aux blocs ou aux fichiers avec un système par ajout (scale-out). Il n'est pas rare que des applications virtualisées soient exécutées sur une infrastructure hyperconvergente (HCI).

L'accès multiprotocole est donc une évidence, avec un mélange de données structurées et non structurées intégrées dans l'infrastructure de stockage. De nombreux appareils IoT et périphériques communiquent via SMB ou NFS, et quelques-uns utilisent S3. L'accès aux données en continu est donc nécessaire dans certains cas. Et dans certains autres cas, le système natif de fichiers parallèles peut également être utilisé.

### *Mise à l'échelle et hiérarchisation*

La prise en charge des applications d'IA et de ML oblige les systèmes de stockage à offrir des performances évolutives. Pour les données non structurées, il s'agit de systèmes de stockage qui utilisent des systèmes de fichiers parallèles avec accès réseau. Pour les données structurées, on utilise des

systèmes de stockage basés sur la technologie flash. L'évolutivité consiste essentiellement à augmenter ou à diminuer les performances et la capacité indépendamment les unes des autres pour répondre aux exigences des applications d'IA et de ML.

En outre, pour une infrastructure à l'épreuve du temps, le système doit également être capable de traiter simplement et à moindre coût des données plus anciennes ou plus froides sur un stockage objet à faible coût avec une interface de stockage objet connue telle que S3.

### *Stockage software-defined*

L'IA et le ML agissent comme des catalyseurs pour le stockage software-defined. Ils permettent d'utiliser l'infrastructure en tant que code ainsi que l'automatisation via une couche de contrôle logicielle hétérogène au-dessus du matériel. Cela permet une meilleure intégration avec les workflows IA/ML, garantissant ainsi que le stockage évolue de manière transparente en fonction des besoins de l'application.

### *Agilité et flexibilité du déploiement*

Les applications qui répondent aux cas d'usage peuvent être développées sur mesure par une entreprise, être basées sur un logiciel d'IA du marché, ou être fournies en tant que SaaS IA. Le déploiement pour le logiciel du marché ou développé sur mesure, se fait on premise, dans une infrastructure cloud pour le SaaS, ou dans un cloud hybride, dans lequel l'environnement on premise interagit avec un environnement de cloud public via une couche d'automatisation et d'orchestration commune.

Étant donné la nature distribuée de l'IA, on peut supposer qu'il est préférable de rapprocher les outils de calcul de la source des données plutôt que l'inverse. Ces derniers temps, le modèle cœur-périphérie-point d'extrémité est devenu de facto la manière de décrire l'IA (où le cœur comprend le cloud et les points d'extrémité comprennent l'intelligence embarquée). Il est important de noter que les profils d'applications de chaque site varient, et donc les besoins d'infrastructure sous-jacentes aussi.

Pour les différents scénarios de déploiement, des solutions doivent être envisagées pour :

- **Traiter en toute sécurité le volume de données nécessaire à l'entraînement des modèles d'IA avec des performances extrêmement élevées.** Les besoins de performance pour l'entraînement par deep learning nécessitent d'exécuter un processus à parallélisme massif à l'aide de GPU, combiné à une intégration de données à large bande passante.
- **Traiter en toute sécurité le volume de données sur lequel le modèle d'IA effectuera des inférences avec des performances extrêmement élevées.** La performance liée à l'inférence signifie la capacité à traiter les données entrantes par le biais du modèle d'IA entraîné et de fournir des informations ou prendre des décisions d'IA en quasi-temps réel.

Pour les datascientists et les développeurs, il peut parfois être plus facile de lancer un projet d'IA dans le cloud, ce qui leur évite de devoir organiser un calcul on premise qui, pour le deep learning, doit généralement être accéléré. Des instances accélérées d'IA sont disponibles sur la plupart des clouds publics, généralement avec des piles d'IA en code open-source. Bien entendu, dans le cas d'instances de cloud accélérées pour l'entraînement de l'IA, c'est le fournisseur de services de cloud (SP) qui détermine ce qui est disponible pour l'utilisateur final en termes de processeurs, de coprocesseurs, d'interconnexions, de tailles de mémoire, de bande passante d'entrée-sortie, etc. Tous les fournisseurs de services de cloud ne proposent pas de combinaisons optimisées de ces composants, qui déterminent en fin de compte la vitesse et la qualité avec lesquelles les datascientists peuvent

développer des modèles d'entraînement. C'est pourquoi de nombreuses entreprises optent pour le déploiement on premise.

Au cours de leurs expériences en matière d'IA ces dernières années, de nombreuses entreprises se sont retrouvées coincées avec leur infrastructure standard ou avec des instances cloud basiques. L'entraînement des modèles prenait trop de temps, et l'inférence était trop lente. L'étude d'IDC montre que 77,1% des personnes interrogées déclarent avoir rencontré une ou plusieurs limites avec leur infrastructure d'IA on premise et 90,3% déclarent avoir rencontré des limites de calcul dans le cloud.

## ARCHITECTURE GLOBALE EXTENSIBLE D'IA UTILISANT LE STOCKAGE IBM

---

IBM Storage pour les solutions l'IA et de données, permettent aux clients de faire passer de manière transparente des projets d'IA au stade de la production dans des environnements de cloud hybride. IBM est toujours un des leaders du secteur des applications évolutives à haute performance, ainsi que du stockage efficace, sécurisé, évolutif et à haute capacité des solutions d'IA et de big data à haute performance. Le portefeuille de stockage d'IBM permet une un stockage intégré et une gestion des données depuis la périphérie, le datacenter principal et le cloud public, accélérant ainsi la modernisation par l'IA. Il est fourni avec un support large et une intégration des conteneurs Kubernetes et de la plateforme Red Hat OpenShift. Il peut être déployé et accessible dans le cloud public ou pour les applications de datacenter. Les solutions IBM Storage pour l'IA et les données, ont pour but de réduire la complexité et les coûts en offrant une intégration accrue et profonde avec une architecture d'IA pouvant être déployée à l'échelle de toute l'entreprise.

### IBM Spectrum Scale

IBM Spectrum Scale repose sur une architecture IT distribuée, conçue pour toutes les applications à haute performance telles que l'IA/machine learning, la modélisation, la simulation, et l'analytique. C'est un système de fichiers parallèles en cluster qui réduit les coûts d'exploitation grâce à une gestion et une évolutivité simples, et les coûts d'investissement grâce à une politique d'optimisation des données et une gestion transparente de leur cycle de vie. IBM Spectrum Scale permet d'accéder simultanément à un seul système de fichiers ou à un ensemble de systèmes de fichiers à partir de plusieurs nœuds. Les nœuds peuvent être rattachés directement, rattachés au réseau, faire l'objet d'un mix entre les deux, ou se trouver dans une configuration en cluster sans aucun partage. Cette solution par ajout (scale-out) et cette plateforme à haute disponibilité permettent un accès partagé très performant à des ensembles de données communs.

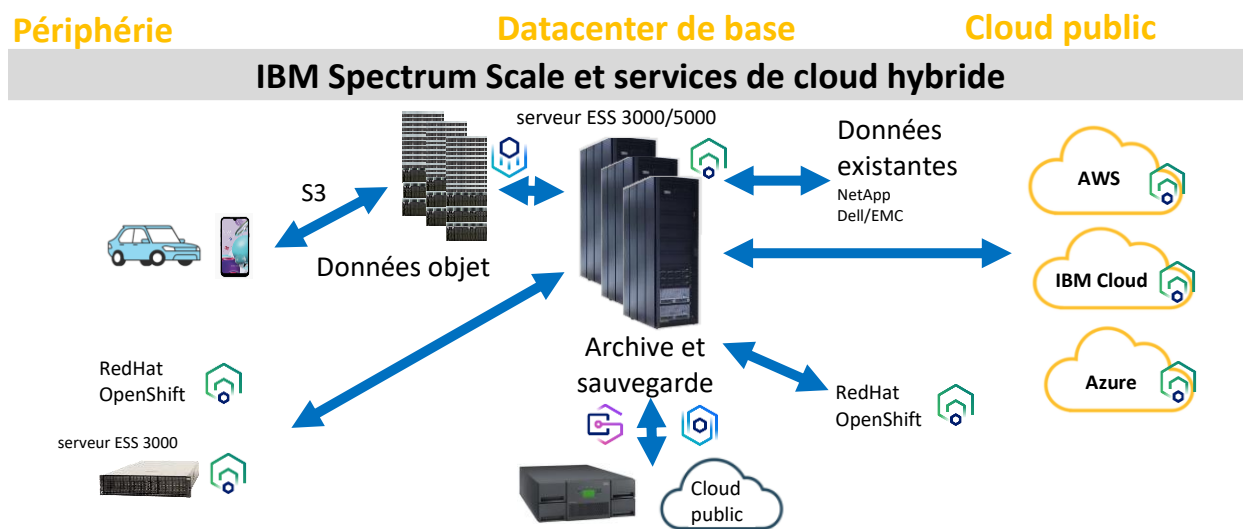
**IBM Spectrum Scale offre une base solide pour l'infrastructure IA, fondée sur des années d'expérience dans le secteur du calcul haute performance.**

IBM Spectrum Scale prend en charge la duplication des données, la gestion du stockage selon des règles précises et les opérations multisites. Les équipes IT peuvent créer un cluster de nœuds de conteneurs Kubernetes, de nœuds IBM AIX, de nœuds IBM Z ou LinuxONE, de nœuds Linux, de nœuds Microsoft Windows Server, ou un mélange des cinq. IBM Spectrum Scale peut fonctionner sur des instances virtualisées ou conteneurisées, fournit ainsi un accès commun aux données partagées dans des environnements, et exploite le partitionnement logique, ou autres hyperviseurs. Plusieurs clusters IBM Spectrum Scale peuvent partager des données au sein d'un même site ou à travers des connexions de réseau étendu (WAN) pour une collaboration et un accès aux données à l'échelle mondiale. IBM Spectrum Scale offre une base solide pour l'infrastructure IA, fondée sur des années d'expérience dans le secteur du calcul haute performance (voir la figure 2).



FIGURE 2

## IBM Spectrum Scale et services de cloud hybride



Source : IDC, 2021

IBM Spectrum Scale offre un espace de noms global, un accès partagé aux systèmes de fichiers entre les clusters IBM Spectrum Scale, un accès simultané aux fichiers à partir de plusieurs nœuds, une grande capacité de récupération et une disponibilité des données grâce à la réplication, la possibilité d'apporter des modifications sur un système de fichiers monté, et une administration simplifiée même dans les grands environnements. Les principales différences d'IBM Spectrum Scale sont les suivantes :

- Un accès partagé au système de fichiers entre les clusters IBM Spectrum Scale, ce qui permet de partager des données entre des clusters distincts sur un même site ou à travers un réseau étendu
- Une amélioration des performances du système grâce à un système de fichiers parallèles breveté
- Une cohérence des fichiers grâce à un accès simultané et précis aux clients dans le cluster en utilisant la gestion des jetons
- Une disponibilité et une fiabilité accrues des données grâce à des fonctions telles que la journalisation de l'audit du système de fichiers et des fonctions personnalisables telles que les montages intelligents pouvant couvrir de longues distances
- Une flexibilité accrue du système permettant l'ajout ou la suppression de ressources de disque ou de serveur pendant le montage du système de fichiers
- Une gestion simplifiée du stockage permettant d'assurer la gestion du cycle de vie des informations (ILM) grâce à une gestion du stockage puissante et automatisée à plusieurs niveaux, allant de la mémoire flash aux disques durs, en passant par le cloud et les bandes, et incluant même la réduction des données en fonction de règles établies
- Une administration simplifiée grâce à de nombreuses interfaces de système de fichiers standard qui peuvent être exécutées directement à partir de la plupart des applications
- Un déploiement de cloud hybride assurant la disponibilité, l'intégrité et la sécurité des données, ainsi qu'un stockage optimisé en mode conteneur et une intégration avec Red Hat OpenShift.

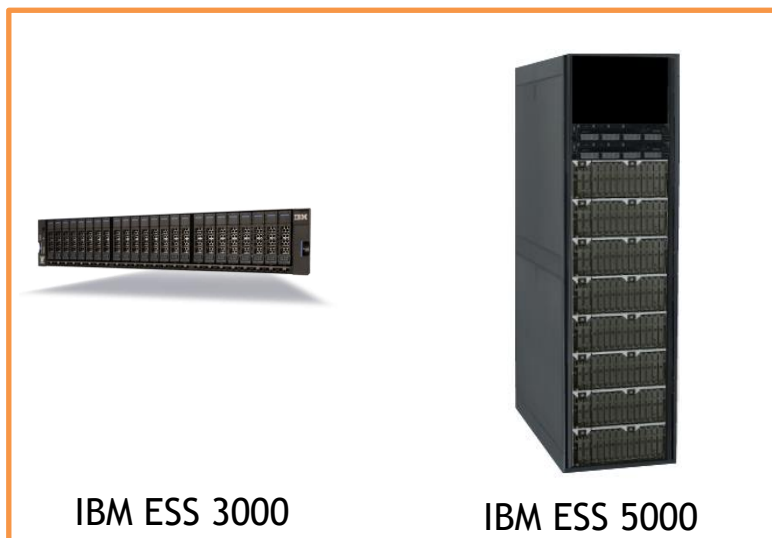
## IBM Elastic Storage System (ESS)

IBM Elastic Storage System est une version moderne du stockage software-defined conçue pour faciliter la configuration et la gestion des blocs de construction, ce qui permet aux sociétés IT de déployer plus rapidement un stockage rapide et hautement évolutif pour les applications à haute performance, telles que les applications d'IA, de big data et d'analytique (voir la figure 3). IBM ESS :

- Est conçu avec le stockage flash NVMe pour offrir une évolutivité vers des niveaux de l'ordre de l'exaoctet et une qualité de service constante dans toute l'infrastructure
- Peut être intégré à la gestion de fichiers et de services de données d'IBM Spectrum Scale pour fournir un système de stockage global fédéré
- Permet aux entreprises de consolider leurs besoins en matière de stockage, de la périphérie aux données centrales, et de les intégrer au cloud public, ce qui réduit l'inefficacité, diminue les coûts d'acquisition, simplifie la gestion du stockage et élimine les silos de données
- Fournit un accès constant et de haute performance à de multiples applications exigeantes déployées en bare metal ou dans des environnements virtualisés (prend en charge l'intégration de Kubernetes et de Red Hat OpenShift, ce qui facilite le déploiement d'applications cloud natives).

### FIGURE 3

#### IBM Elastic Storage System



Source : IDC, 2021

IBM ESS est disponible en deux versions, 3000 et 5000 :

- IBM Elastic Storage System 3000 (ESS 3000) est conçu pour relever le défi de la gestion des données pour l'analytique. Présenté dans un boîtier compact 2U, ESS 3000 accélère la valorisation des applications d'intelligence artificielle/deep learning et de calcul intensif, grâce à son stockage entièrement NVMe et à l'installation et la mise à jour simples et rapides de logiciels conteneurisés. La conception matérielle et logicielle de l'ESS 3000 permet à

l'entreprise d'accéder à des performances de pointe, nécessaires à l'utilisation optimale des ordinateurs gourmands en données.

- IBM ESS 5000 offre des nœuds ayant une capacité de l'ordre du pétaoctet et à haut débit, combinant le stockage software-defined IBM Spectrum Scale avec des serveurs à forte intensité d'E/S basés sur le processeur IBM POWER9. En consolidant les besoins de stockage de l'ensemble de l'entreprise sur IBM ESS 5000 et ESS 3000 basé sur NVMe, les équipes IT peuvent réduire l'inefficacité, diminuer les coûts d'acquisition et prendre en charge les besoins en matière d'IA, de HPC, d'analyse et/ou de stockage à haute capacité qui sont caractéristiques des domaines de la santé, des médias, du secteur public et des services financiers. L'ESS 5000 peut commencer avec des téraoctets et évoluer vers des centaines de pétaoctets, voire des exaoctets, grâce à un espace de noms unique qui élimine les coûteux silos de données. IBM Spectrum Scale est le système de fichiers parallèles situé au cœur d'IBM ESS 5000, ce qui augmente sa capacité de traitement au fur et à mesure de sa croissance. Il permet l'intégration avec les modèles ESS précédents pour la protection des investissements et offre des options moins coûteuses telles que le stockage cloud et IBM Tape. Avec IBM Spectrum Scale, les services IT peuvent éliminer les silos de données et les goulots d'étranglement, simplifier la gestion du stockage et accéder plus rapidement aux données.

## IBM Cloud Object Storage

IBM Cloud Object Storage (COS) est une solution de stockage de pointe, software-defined, hautement évolutive et rentable, qui permet de stocker des données non structurées en périphérie, dans le datacenter principal et dans les clouds privés ou publics. IBM Cloud Object Storage est idéal pour le déploiement ou la modernisation d'infrastructures hautes performances pour l'IA, l'analytique, l'IoT, la vidéo et les référentiels d'images. Il offre également une valeur ajoutée sans précédent en permettant aux entreprises de réduire les coûts de stockage jusqu'à 12%, en utilisant de nouveaux disques à résonance magnétique (SMR) de 18 To, tout en augmentant le débit jusqu'à 55 Go/s dans un cluster de 12 nœuds. L'offre inclut des fonctions de protection des données locales ou géodispersées d'IBM, qui peuvent être personnalisées pour des data lakes exigeants et des systèmes nécessitant de grands besoins (voir la figure 4).

### FIGURE 4

#### IBM Cloud Object Storage



Source : IDC, 2021

IBM Cloud Object Storage est un système fondamental pour l'infrastructure IA avec des avantages tels que :

- **L'évolutivité** : COS prend en charge la croissance exponentielle des données, en faisant évoluer les performances, depuis les téraoctets jusqu'aux exaoctets.
- **La sécurité** : COS dispose d'un système de chiffrement intégré et d'un système de stockage WORM (Write Once Read Many) régi par des règles.
- **La simplicité** : COS peut accéder aux données simultanément à partir de n'importe quel endroit et fournit une reprise en ligne automatique, une régénération des données, une extension automatique et un rééquilibrage.
- **Une efficacité accrue** : COS fournit des données géoprotégées grâce à l'efficacité de l'algorithme IDA (Information Dispersal Algorithm) et est disponible sous forme de logiciel ou de solutions d'appliances entièrement prises en charge.
- **Des fonctionnalités de recherche** : COS offre des analyses et des fonctions de recherche personnalisées pour gagner du temps. Les entreprises peuvent créer des métadonnées personnalisées pour créer de la valeur.
- **L'accès amélioré aux fichiers** : Une nouvelle passerelle logicielle d'accès aux fichiers peut être reliée de manière transparente à tout système de fichiers Windows ou Linux utilisant un accès SMB ou NFS, ce qui permet de connecter facilement des applications basées sur des fichiers au stockage objet.
- **Le transfert à haut débit** : L'option de transfert à grande vitesse d'IBM Aspera facilite le transfert des données, et les niveaux de classe de stockage flexibles permettent de gérer les coûts tout en répondant aux besoins d'accès aux données.

## PERSPECTIVES D'AVENIR

---

En 2022, 65% du PIB mondial sera numérisé, ce qui entraînera des dépenses informatiques de 6 800 milliards de dollars dans le monde pour la période 2020-2023 (voir *IDC FutureScape : Worldwide Digital Transformation 2021 Predictions*, IDC #US46880818, octobre 2020). L'infrastructure numérique ne se limite pas aux services centraux traditionnels des entreprises ni aux datacenters dans le cloud. Elle comprend tous les actifs et les ressources qui permettent le transfert des applications et du code en vue de la transformation. Elle permettra une meilleure expérience client. Elle permet également d'intégrer l'intelligence et l'automatisation dans les opérations business et de soutenir l'innovation permanente jusqu'à la limite numérique de l'entreprise et du secteur. Une stratégie numérique réussie doit transformer l'infrastructure pour éliminer les silos, briser les barrières imposées par la technologie et aller au-delà de la simple prise en charge des outils et applications traditionnels.

IDC estime que l'IA sera la base de l'infrastructure numérique. Le groupe Customer Insights and Analysis d'IDC a récemment mené une enquête pour analyser les dépenses informatiques actuelles et futures, ainsi que les projets d'adoption, dans les entreprises de toutes tailles et de divers secteurs. L'étude a révélé que près de 76% des 3 600 décideurs clés des sociétés IT de divers secteurs d'activité dans le monde ont indiqué que l'intelligence artificielle est ou sera d'ici les deux prochaines années un élément clé de leur stratégie DX. Seuls 22% des répondants ont indiqué que l'IA serait l'élément clé de leur stratégie DX dans les trois à cinq prochaines années. Parmi les secteurs qui s'attendent à ce que l'IA soit un élément clé de leur stratégie DX d'ici un à 2 ans, les télécommunications, les services publics, l'éducation et les sociétés de services professionnels sont les plus susceptibles de se tourner vers l'IA dans leurs efforts de transformation numérique.

La force de l'IA réside dans son "rapport temps/valeur", c'est-à-dire la valeur que les entreprises peuvent obtenir des données dans le délai le plus court possible. Le FutureScape d'IDC sur l'intelligence artificielle estime que "l'intelligence artificielle est l'innovation la plus disruptive de notre vie". L'IA n'est plus seulement "un petit plus à avoir". La pandémie mondiale a accéléré l'adoption de l'IA, qui devient omniprésente dans tous les processus business. Les solutions d'IA alimentées par le machine learning, l'IA conversationnelle et la vision par ordinateur sont à la pointe de la résilience business, de l'innovation accélérée et des expériences clients et employés transformées. Environ 51% des personnes interrogées dans le cadre de l'enquête précédemment évoquée ont indiqué qu'elles envisageaient l'intelligence artificielle ou que cette dernière était déjà en production. Ce chiffre était de 34% en 2019. L'impact le plus important de l'IA est l'aide qu'elle apporte aux salariés dans leur travail. L'adoption de l'IA par les entreprises continuera à se développer à mesure que les avantages des déploiements complets deviendront plus tangibles.

IDC estime que les investissements dans les infrastructures d'IA resteront importants au cours des prochaines années. IDC estime le chiffre d'affaires du matériel d'IA (serveur et stockage confondus) à 13,4 milliards de dollars en 2020, soit une croissance de 10,3% d'une année sur l'autre. Au sein du marché du matériel, la croissance du stockage utilisant l'IA est estimée à 11,4% en 2020. Le marché global du matériel devrait connaître une forte reprise en 2021, avec une croissance de 35,5% par rapport à l'année précédente, tirée par le stockage utilisant l'IA, qui devrait lui-même connaître une croissance de 43,1% par rapport à 2020.

Une grande partie de ce stockage sera déployée dans des environnements de cloud hybride avec une couche de mobilité transparente qui permet de déplacer la couche applicative du cœur vers le cloud ou vers la périphérie et inversement. Le stockage constituera une base essentielle pour le cloud hybride, car il permet de déplacer les fonctions de calcul vers les données tout en fournissant un plan d'accès et de contrôle commun.

Le stockage, et en particulier le stockage dans les environnements de cloud hybride, restera le socle sur lequel les projets d'IA peuvent s'étendre aujourd'hui et à l'avenir. Les investissements dans l'IA et les projets de modernisation des données dues à l'IA entraîneront des investissements dans le stockage de fichiers par ajout (scale-out) et les données non structurées. IDC a récemment interrogé 624 professionnels de l'informatique et équipes d'exploitation dans le monde pour déterminer les tendances en matière d'adoption de l'infrastructure informatique. Dans cette étude, IDC a constaté que plus de 65% des personnes interrogées préfèrent les systèmes de fichiers par ajout (scale-out) accessibles localement ou via NFS pour leurs applications à haute performance telles que l'IA. En fait, pour les applications d'IA en particulier, qui comprennent les applications d'entraînement et d'inférence, la performance est le premier critère en matière de stockage. Viennent ensuite la facilité de déploiement dans le cloud hybride et la qualité du service.

## CONSEILS ESSENTIELS POUR LES ACHETEURS DE TECHNOLOGIES

---

Les acheteurs de technologie sont à juste titre perturbés quant au processus de construction de leur propre infrastructure d'IA. Ils ont défini des cas d'usage, lancé des projets d'IA, et embauché ou formé des datascientists et des développeurs d'applications, mais ils se retrouvent soudain limités par l'infrastructure sur laquelle développer des modèles d'IA. Souvent,

**La croissance du stockage utilisant l'IA est estimée à 11,4% en 2020. Le marché global du matériel devrait connaître une forte reprise en 2021, avec une croissance de 35,5% par rapport à l'année précédente, tirée par le stockage utilisant l'IA, qui devrait lui-même connaître une croissance de 43,1% par rapport à 2020.**

l'infrastructure existante est exploitée brièvement, puis des investissements sont effectués dans une infrastructure accélérée. Les datascientists se retrouvent obligés d'assembler des piles sur l'infrastructure accélérée et d'essayer de les faire fonctionner, ce qui n'est en fait pas leur rôle. Les équipes d'infrastructure IT ne connaissent pas les piles dont les datascientists ont besoin et ne peuvent ni les assembler ni les optimiser. Il en résulte une grave pénurie de compétences que les vendeurs de serveurs et les fournisseurs de services cloud ont tenté de combler avec leurs propres piles, chacun à leur manière. Aujourd'hui, il y a autant de piles que de vendeurs, et elles se chevauchent souvent lorsqu'elles sont développées par plusieurs membres de la même chaîne de valeur.

## Commencer par les résultats de l'entreprise

Les entreprises doivent commencer par établir un lien entre les contraintes de service et les cas d'usage afin d'identifier les résultats business issus d'un investissement dans une infrastructure IA. Elles doivent chercher à quantifier et à mesurer les avantages de ces investissements. Par exemple, lorsqu'on cherche à se différencier de la concurrence, la question est de savoir dans quelle mesure et à quel moment. Ces critères devraient ensuite conduire à la sélection d'une architecture d'application. Les entreprises doivent tenir compte des caractéristiques de l'IA lorsqu'elles cherchent à améliorer leur marque en analysant mieux le sentiment, les désirs et les besoins des clients et en y répondant de manière ciblée, ce qui se traduit par une augmentation des revenus et des bénéfices.

## Adopter une approche globale

Il est important de considérer l'ensemble des données (c'est-à-dire d'adopter une vision globale) lors de la mise en œuvre de tout projet d'IA. Ne considérer qu'un seul problème à la fois peut entraîner la création d'un énième silo ou, pire encore, accroître la complexité de l'environnement en raison d'un manque d'intégration et d'interopérabilité entre plusieurs architectures et solutions. Une infrastructure de données doit être envisagée comme une solution globale, depuis la périphérie jusqu'au cœur du réseau, en passant par le cloud, et par le biais de cas d'usage tels que l'IA et autres applications stratégiques.

## Développer la bonne application et la bonne architecture de données

Le développement d'une application d'IA et d'une architecture de données est une tâche complexe. Il s'agit de convertir les besoins et les résultats business en un workflow déterministe activé par l'IA. Le workflow doit décrire la façon dont les fonctionnalités de l'IA améliorent le comportement de cette application, comment les données sont ingérées et analysées, et comment l'application interagit avec d'autres applications business et avec les utilisateurs. L'accent doit être mis ici sur la façon dont les données sont consommées, produites et analysées par les applications et sur les implications sur le matériel. Lors de la création d'un plan "greenfield", il convient de se concentrer sur le mélange entre les composants logiciels personnalisés (code open source ou propriétaire) et les composants logiciels du marché.

## Choisir la bonne pile de référence

Plusieurs éditeurs et fournisseurs de services ont publié des piles de référence pour la mise en œuvre d'une infrastructure IA. Nombre d'entre elles sont "ouvertes" par nature, ce qui permet une expérience modulaire "prête à l'emploi", et peuvent être consommées sur la base d'un service prépayé pour une mise en œuvre à faible coût. Il s'agit d'une considération importante, car les investissements dans l'infrastructure de l'IA peuvent rapidement devenir coûteux. IDC prévoit de publier son point de vue sur les piles de référence des éditeurs les plus populaires dans un prochain document.

Les avantages à garder à l'esprit lors de l'examen des piles de référence sont la réduction des coûts, la disponibilité des données et des applications, l'utilisation et la consolidation effectives de l'infrastructure et, dans la mesure du possible, une plateforme unique de fourniture d'applications interopérables.

## Créer une architecture pour l'IA

Les entreprises ont besoin d'une stratégie de gestion des données pour fournir un accès souple et organisé à toutes les données, de tout type, quel que soit leur emplacement, et répondre aux préoccupations liées à l'architecture de référence. Un effort de modernisation permettrait de définir et de déployer une architecture fournissant une base ouverte et extensible, offrant choix et flexibilité, et étant capable de communiquer avec d'autres plateformes cloud. La stratégie de gestion des données hybrides d'IBM pour accélérer le parcours vers l'IA est une approche en quatre étapes : collecter, organiser, analyser et intégrer.

- **Collecter** : rendre les données simples et accessibles au bon endroit, à partir de n'importe quelle base de données ou installation de stockage
- **Organiser** : veiller à ce que les données soient fiables, complètes et cohérentes à toutes les étapes du cycle de vie de l'information : profiler, nettoyer et cataloguer les données, assurer la protection et la conformité, et permettre une visibilité, une détection et un reporting selon des règles définies.
- **Analyser** : créer, déployer et gérer des modèles d'IA à l'aide d'outils intégrés pour explorer et analyser des données structurées et non structurées et les déployer en toute sécurité.
- **Intégrer** : assurer la confiance et la transparence dans les décisions recommandées par le modèle, expliquer les décisions, détecter les biais, etc., en utilisant les solutions et services fournis.

Le parcours vers l'IA consiste à faire passer les données de l'acquisition à la compréhension grâce à une architecture qui peut facilement être utilisée dans toute l'entreprise. Il est important que chaque partie de l'échelle d'IA s'intègre à l'ensemble du parcours. Le stockage est généralement mis en œuvre de manière tactique avec des solutions de stockage spécifiques qui créent des silos de données et des solutions qui ne sont pas intégrées ensemble ou avec un ensemble complet de solutions d'infrastructure. Les clients peuvent stocker des données sur un grand système de stockage de fichiers ou d'objets, mais sans avoir de détails sur ces données ou ne pas les utiliser pour obtenir des informations supplémentaires. Les clients peuvent certes commencer ou se concentrer sur un projet sur une partie du parcours, mais chaque projet doit envisager une architecture d'IA globale pour optimiser les ressources et moderniser l'infrastructure en vue de l'extension des applications d'IA.

Les entreprises qui traitent l'IA comme une application composite à haute performance (composée de plusieurs applications interconnectées), en empruntant des éléments cruciaux tels que les systèmes de fichiers par ajout (scale-out), l'informatique hétérogène et les interconnexions à haut débit pour l'accès au calcul et au stockage distribués, sont celles qui parviennent finalement à faire évoluer leur infrastructure d'IA.

## Exploiter les bons partenariats

Pour les acheteurs de technologie, le partenariat avec un fournisseur de solutions de bout en bout est crucial pour le succès à long terme, mais IDC pense qu'actuellement aucun éditeur sur le marché ne propose encore cet environnement de bout en bout, même s'ils travaillent dur pour y parvenir. Néanmoins, en s'engageant avec un partenaire de confiance, les entreprises peuvent mieux utiliser les approches de l'IA pour développer leur business. Elles peuvent devenir agiles et adaptables, et

tirer parti des synergies internes pour devenir rentables. Enfin, elles peuvent devancer les forces disruptives de leur secteur en se réinventant. Le partenaire idéal doit fournir :

- Des solutions éprouvées qui s'adaptent aussi bien aux petits déploiements de test qu'aux déploiements massifs à l'échelle mondiale
- Une expertise en matière de secteurs d'activité sur les segments qui correspondent à son orientation business
- Un accès intégré à un écosystème diversifié d'éditeurs de logiciels et de fournisseurs d'infrastructures
- Un point de vue axé sur les données pour garantir la sécurité et maximiser la valeur à long terme de ses investissements dans de nouvelles sources de données
- Un succès vérifié dans la simplification des aspects matériels, logiciels et sécuritaires de grands projets

## DEFIS ET OPPORTUNITES

---

### Pour les entreprises

Ce livre blanc a abordé une série de défis auxquels les entreprises sont confrontées lorsqu'elles sont prêtes à mettre en production leurs applications d'IA. De la préparation des données au développement de modèles, en passant par les environnements d'exécution, l'entraînement, le déploiement et la gestion des modèles d'IA, les besoins relatifs à l'infrastructure sous-jacente défient les anciens modèles de matériel polyvalent. Les investissements dans une infrastructure conçue pour les applications à forte intensité de données, avec des performances, une mise à l'échelle, un accès aux données et une intégration plus importants, et pouvant se fonder dans un environnement de cloud hybride, offrent une valeur et une qualité de service à long terme. Les entreprises devront décider de remplacer ou de compléter les plateformes de stockage polyvalentes existantes par des systèmes de stockage adaptés aux tâches de traitement spécifiques à l'IA. Ce processus servira de base au développement et à l'exploitation d'applications d'IA de pointe.

### Pour IBM

Le défi pour IBM est toujours celui de la reconnaissance du marché. IBM propose des solutions d'infrastructure d'IA exceptionnelles et complètes, intégrées aux piles d'infrastructure (comme Red Hat OpenShift) et au cloud public, mais que les clients potentiels perçoivent, à tort, comme complexes ou plus coûteuses. La réaction réflexe qui s'ensuit et qui consiste à choisir l'un des grands éditeurs de systèmes de stockage de base pour les applications extrêmement gourmandes en données prive ces entreprises des solutions d'infrastructure d'IA qui pourraient réellement leur être bénéfiques. Les solutions IBM de stockage d'IA et de données avec IBM Spectrum Scale et IBM Elastic Storage System, par exemple, constituent une base du calcul intensif pour bon nombre des plus grands datacenters. Maintenant que les nouvelles applications d'IA commencent à imiter sérieusement les déploiements de supercalculateurs avec des besoins de calcul intensif et à défier l'infrastructure on premise et dans le cloud sur laquelle elles fonctionnent, c'est le moment pour IBM d'entrer en scène et de gagner de nouveaux clients.



## CONCLUSION

---

Au cours des dernières années, IDC a été témoin de la transformation de l'IA dans de nombreuses entreprises et de la façon dont elles ont commencé à développer un large éventail de fonctionnalités d'IA. Initialement lancées à titre expérimental par un personnel relativement inexpérimenté et exécutées sur l'infrastructure disponible, ces projets ont pris de plus en plus d'importance. De nombreuses entreprises ont développé une expertise approfondie en matière d'IA et elles font l'expérience directe de la vitesse à laquelle l'IA devient un aspect essentiel de leur business.

Dans le même temps, les services IT ont également évolué en apprenant à adapter l'infrastructure à utiliser pour l'IA. Aujourd'hui, les besoins en matière d'infrastructure pour l'entraînement le deep learning ou pour l'inférence sont beaucoup plus clairs, de même que la manière de faire évoluer ces environnements vers la production. Il est évident que le deep learning nécessite une infrastructure différente de celle des autres applications. L'entraînement du deep learning nécessite des nœuds en clusters dotés de processeurs et de coprocesseurs puissants, d'interconnexions rapides avec stockage évolutif, d'une large bande passante d'E/S et d'une mémoire abondante.

Aujourd'hui, la plus grande priorité pour les services IT est de concevoir et de déployer au mieux leur infrastructure de données et leurs applications d'IA avec les meilleurs systèmes, puis de les connecter et de les optimiser ensemble.

IDC estime que les solutions de stockage IBM pour les données et l'IA, qui comprennent Spectrum Scale, Elastic Storage System et Cloud Object Storage, offrent une valeur et des performances sans précédent.

**Aujourd'hui, la plus grande priorité pour les services IT est de concevoir et de déployer au mieux leur infrastructure de données et leurs applications d'IA avec les meilleurs systèmes, puis de les connecter et de les optimiser ensemble.**

## A propos d'IDC

International Data Corporation (IDC) est le premier groupe mondial de conseil, d'études et d'événementiels sur les marchés des technologies de l'information, des télécommunications et des technologies grand public. IDC aide les professionnels de l'informatique, les responsables en entreprise et la communauté des investisseurs à prendre des décisions basées sur des données factuelles concernant les achats de technologies et la stratégie métier. Plus de 1100 analystes répartis dans plus de 110 pays apportent une expertise internationale, régionale et locale sur les opportunités technologiques et industrielles. Depuis 50 ans, IDC fournit des informations stratégiques à ses clients afin des aider à atteindre leurs objectifs métier. IDC est une filiale de la société IDG, leader mondial sur le marché des médias de la recherche et des événementiels dédiés aux technologies.

## Siège social international

140 Kendrick Street  
Building B  
Needham, MA 02494  
USA  
508.872.8200  
Twitter: @IDC  
idc-community.com  
www.idc.com

---

### Notice de copyright

Publication externe des informations et des données d'IDC – L'utilisation de toutes les informations d'IDC destinées à être utilisées dans des publicités, des communiqués de presse ou des documents promotionnels nécessite une autorisation écrite préalable de la part du Vice-Président ou du Responsable pays approprié. Une version préliminaire du document proposé doit accompagner cette demande. IDC se réserve le droit de refuser son consentement à un usage externe pour les raisons de son choix.

Copyright 2021 IDC. Toute reproduction sans autorisation écrite préalable est strictement interdite.

