

L'importance de la plateforme de données pour une IA efficace

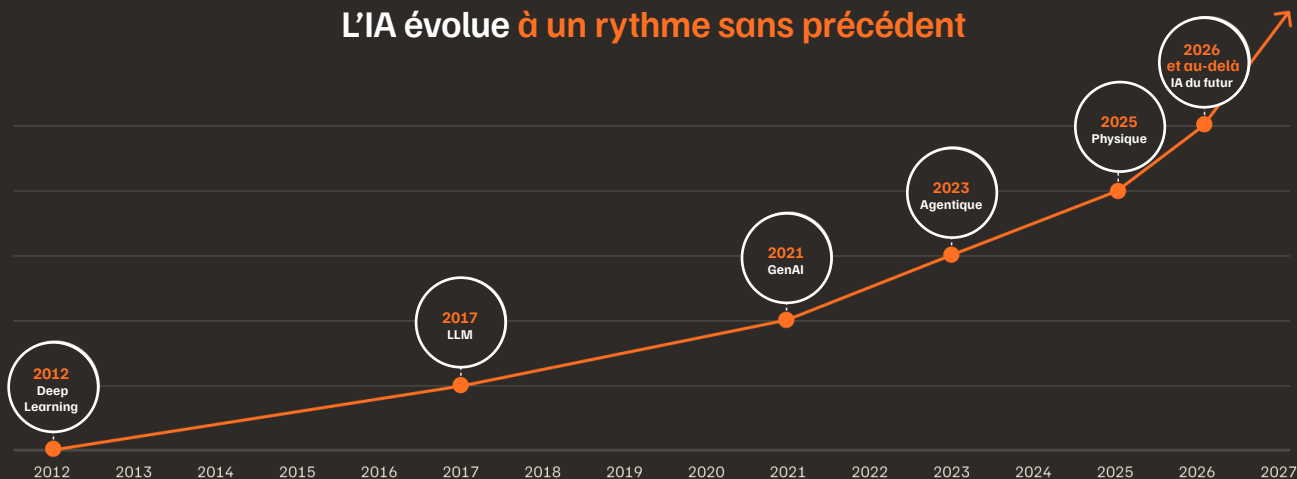
L'IA n'a pas détruit le stockage traditionnel. Elle a mis en lumière ses limites existantes.

Elle oblige les entreprises à enfin remettre en question leur approche en matière de gestion du stockage et des données. Ce qui fonctionnait autrefois, gérer les systèmes de stockage un par un et les applications une par une, peine désormais à faire face au rythme effréné des changements, à l'explosion de la réutilisation des données et à l'augmentation des risques. Dans ce document, vous découvrirez pourquoi, à l'ère de l'IA, il devient impossible d'ignorer ces brèches et comprendrez qu'il n'est pas simplement question de mettre à niveau votre solution de stockage, mais bien de repenser la gestion des données à l'échelle de l'entreprise.

Les infrastructures d'entreprise ont suivi pendant de longues années un ordre hiérarchique assez simple. Les logiciels reposaient sur l'infrastructure, et les données comblaient l'espace. Chaque application était hébergée sur sa propre pile, avec l'environnement de production d'un côté, l'environnement de sauvegarde de l'autre et les données analytiques quelque part en marge. Le stockage faisait son travail en silence, alimentant la pile tout en se tenant à l'écart. **Il n'y avait là rien de raffiné, mais ce modèle a très bien fonctionné pendant de longues années.**

Et lorsque les choses tournaient mal, les équipes redoublaient d'efforts pour résoudre le problème. Elles créaient des copies en abondance pour la restauration, d'autres pour les tests et encore d'autres pour la conformité. En cas de dérapage, le personnel intervenait et trouvait une solution. Il semblait alors raisonnable de gérer quatre ou cinq copies de données identiques. Cette approche était inefficace, certes, mais ne représentait aucun danger. **Mais l'IA a changé la donne.**

L'IA évolue à un rythme sans précédent



L'IA ne souhaite pas disposer d'une seule copie de vos données. Elle en exige une multitude, et qui soient toutes disponibles simultanément : des données d'entraînement, des feature stores, des pipelines d'inférence et des sandboxes répartis dans le cloud public, on-premises et dans les environnements de périphérie, non pas une seule fois mais constamment. Les données qui arrivent sont ajustées, transférées, rapatriées et réutilisées. Aucune latence ne saurait être tolérée, et le provisioning manuel ne peut suivre la cadence.



Sous la pression, les organisations ont donc réagi comme elles l'ont toujours fait. Elles ont multiplié les copies.

Elles ont multiplié les scripts et les snapshots. À tel point que les silos, les partages, les réplicas et les solutions de contournement se sont accumulés. Elles ont perdu en visibilité tout en augmentant leurs coûts et en s'exposant à davantage de risques sans même s'en apercevoir. La gouvernance a commencé à perdre le fil de la réalité, et les piles verticales que nous avions joliment mises en place pour les applications ont commencé à ployer sous le poids de données qu'elles n'étaient pas destinées à gérer à une telle échelle.

L'IA n'est pas à l'origine du problème. Elle a mis en lumière ses limites existantes.

Plus qu'une simple question de stockage

L'ancienne approche traitait les données comme un simple subordonné. Les applications avaient tout pouvoir de décision sur ce qui avait de l'importance, et le stockage se contentait de suivre les instructions. Les données résidaient là où l'application les avait placées, aussi longtemps qu'elle en avait besoin et sous le contrôle en place à ce moment-là. **Ce modèle ne tient plus la route.**

Car les données ont pris le pouvoir : ce sont elles qui donnent le tempo. Charge aux applications (y compris les systèmes d'IA) de les interpréter, et non l'inverse. Face à des données fragmentées, mal gérées ou cloisonnées, l'IA ne fait pas que subir des ralentissements : elle devient paralysée pour ne pas dire tout bonnement inopérante.

Voilà pourquoi le moment est venu de songer à moderniser votre stockage et de supprimer l'idée que le stockage est ce que vous devez gérer en priorité.



Nous ne faisons rien d'autre aujourd'hui que gérer des données ; tout du moins, c'est ce que nous devrions faire.

La gestion des données bouleverse les fondations. Elle implique de renoncer aux boîtes et empilages du passé pour commencer à réfléchir par couches et par politiques. Les données doivent pouvoir se déplacer librement entre les différents environnements. Elles doivent être accessibles sans imprudence, être gouvernées sans tout mettre à l'arrêt et être automatisées par défaut, sans être bloquées par des tickets et des connaissances institutionnelles.

L'importance cruciale d'une plateforme de données

En matière de qualité, de disponibilité et de cohérence des données, l'IA se montre sans pitié. Les modèles sont entraînés sur des données accessibles, et non sur des données idéales. Les données tardives ou incomplètes paralysent les pipelines. Lorsque les performances et les accès varient en fonction de la localisation du modèle, la qualité de l'inférence en fait les frais. Et face à un cadre de gouvernance incohérent, les équipes IA tournent au ralenti ou contournent tout simplement les mécanismes de contrôle.

Sans une plateforme de données sous-jacente, chaque workflow d'IA prend l'allure d'une intégration personnalisée. Les données sont copiées, remodelées et revalidées à chaque étape. Les équipes passent davantage de temps à déplacer les données qu'à tenter d'améliorer les modèles. Et à mesure que l'utilisation de l'IA se développe, on observe un réel décalage entre l'entraînement et l'inférence, entre les différents environnements, et entre les attentes de l'entreprise et les résultats concrets que le système peut légitimement produire.

Une plateforme de données bouleverse cette dynamique en donnant à l'IA une base de travail cohérente. Les données résident dans un ensemble unifié qui s'étend à travers les différents environnements de manière à utiliser une seule et même source plutôt que des copies disjointes pour tous les besoins d'entraînement, de réglage et d'inférence. La gouvernance et les contrôles sont appliqués une seule fois à grande échelle et suivent les données dans tous leurs déplacements, sans avoir à les redéfinir système après système. Et les données sont contextualisées ; elles apportent des éclairages sur la façon de les utiliser, de les protéger et de les conserver, afin de permettre aux modèles d'exploiter des informations fiables, quelle que soit leur évolution. **C'est là toute la différence entre les entreprises qui expérimentent l'IA et celles qui ne font qu'en dépendre.**

Les limites de l'approche traditionnelle

En dépit de ce contexte, bon nombre d'entreprises persistent à utiliser le stockage comme un simple disque dur externe : un stockage cloisonné, manuel et enlisé dans des écosystèmes de fournisseurs traditionnels qui lient les données à des plateformes et des charges de travail spécifiques. Si ce modèle fonctionnait très bien avec les applications d'autrefois, il n'est plus adapté aux exigences actuelles imposées par l'IA.



Avec l'IA, il est impossible d'ignorer cette lacune.

Nous avons besoin d'une seule couche de données qui couvre l'intégralité des environnements : production, analytique, on-premises, cloud public, charges de travail traditionnelles et pipelines d'IA. Une couche qui n'a pas à se soucier de la localisation de l'application, car la gouvernance, la protection et les accès sont traités une seule fois et appliqués de manière globale et cohérente.

Ce n'est pas juste de la théorie. Le cloud public a déjà montré ce qu'il est possible de faire avec le stockage à la demande : des politiques plutôt que des tickets, avec un seul pool de stockage virtuel utilisé pour de nombreuses applications.

Car toute la difficulté réside dans la capacité à étendre ce modèle à l'ensemble des environnements, des équipes et des charges de travail. Et il ne suffit pas pour cela de mettre à niveau des plateformes obsolètes. Ce modèle impose d'adopter une approche radicalement nouvelle, adaptée à l'ère du cloud, qui prend en compte d'emblée la mobilité des données et l'application de politiques.

Il implique de déployer une architecture pensée avant tout pour le logiciel, et non pour le matériel, parallèlement à des systèmes capables d'évoluer en continu, sans avoir à les mettre à niveau tous les deux ou trois ans. Cette approche est fondée sur une couche opérationnelle commune qui s'étend à tous les environnements afin d'obtenir des données un comportement homogène. La prise en charge des protocoles est intégrée nativement, et non ajoutée ultérieurement par des fournisseurs vieillissants après des années d'acquisitions et de projets parallèles.

Une fois ces principes établis, la perspective de faire évoluer un modèle cloud au-delà du cloud public ne relève plus de la science-fiction, mais devient bel et bien concrète.

Si ça vous semble excessif, c'est normal. L'IA ne va pas attendre votre sixième copie d'une suite d'entraînement pour terminer sa réplique. Elle ne va pas se mettre en pause le temps que les équipes s'accordent sur la localisation des données sensibles à un moment précis. Vous avez deux possibilités : suivre le rythme ou continuer à créer des copies. À vous de choisir.

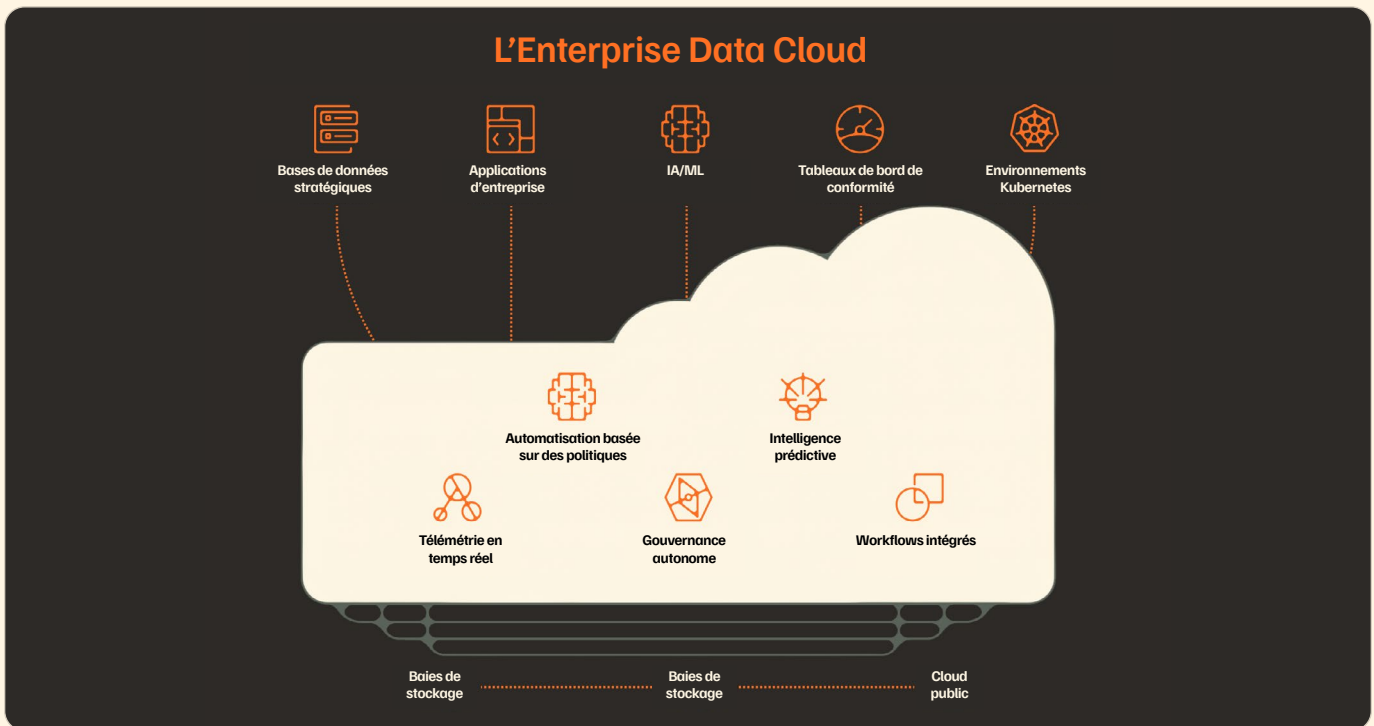
“ Toute l'ère de l'IA s'articule autour de l'accès aux données. Gérer vos données, sans se contenter simplement de les stocker, est la base pour bien se préparer à l'IA. »

Rob Lee, Chief Technology Officer, Everpure, au sujet de l'importance d'une gouvernance des données unifiée dans l'ensemble des environnements.

La bonne approche à adopter

Si vous voulez savoir à quoi ressemble un stockage qui travaille pour les données et non contre elles, l'Enterprise Data Cloud offre une approche concrète. Tout repose sur une idée très simple : les données doivent être traitées comme une ressource gouvernée et partagée, quelle que soit leur localisation ou la manière dont elles sont utilisées.

L'Enterprise Data Cloud offre aux organisations un moyen cohérent d'exploiter leurs données dans l'ensemble de leur environnement, de manière à pouvoir entraîner, déployer et faire évoluer les modèles d'IA sans avoir à entreprendre systématiquement une refonte du socle fondamental. Les politiques remplacent les tickets, et les exceptions laissent place à la cohérence. Les données se déplacent en réponse aux besoins métier, sans compromettre la sécurité, la conformité et les opérations. Les équipes, n'ayant plus à gérer les systèmes un à un, sont en mesure de produire des résultats à grande échelle.



Il ne s'agit pas d'une solution de stockage offrant de nouvelles fonctionnalités. C'est un modèle opérationnel radicalement différent, conçu pour s'adapter en permanence au changement et aux exigences de l'IA.

L'IA n'est pas à l'origine du problème, elle l'a simplement mis en évidence. L'Enterprise Data Cloud est le modèle qui manquait tant aux données.

Envie de vous affranchir de la gestion du stockage et d'exploiter vos données à la manière d'une plateforme ? Le moment est venu d'échanger avec l'équipe Everpure™. Nous sommes là pour aider les entreprises à renoncer aux systèmes de stockage fragmentés au profit d'une approche du stockage et de la gestion de données conçue pour l'IA, pour l'évolutivité et pour le changement constant.

[Découvrez l'avenir du stockage](#)

Visitez notre site Web

+33 9 75 18 86 78

