



Livre blanc technique

Concevoir et construire un pipeline de traitement de données pour vos workflows d'IA

Déployer l'IA, l'apprentissage machine et l'apprentissage profond dans l'ensemble de votre entreprise de la périphérie, au cœur jusqu'au cloud

Santosh Rao, NetApp
Mars 2018 | WP-7264

Synthèse

Dans un contexte où les entreprises sont très intéressées par les technologies d'intelligence artificielle telles que l'apprentissage profond (AP), notamment pour lancer de nouveaux services et mieux exploiter les données qu'elles collectent, il est important que les équipes de data scientists, lorsqu'elles passent du Proof of Concept à la mise en œuvre, élaborent des architectures de données complètes qui éliminent les goulots d'étranglement et facilitent l'itération de modèles d'apprentissage profond.

La conception d'une architecture de données implique de considérer le pipeline de traitement de données dans son intégralité : de l'ingestion des données à l'analytique en périphérie jusqu'à la préparation et l'apprentissage des données dans le data center central jusqu'à l'archivage dans le cloud. Il est essentiel aussi de comprendre les besoins en termes de performances, de datasets et de services de données. Il convient enfin de ne pas négliger les possibilités d'extension et de prise en charge à venir, car le matériel d'apprentissage profond et les approches cloud vont évoluer au fil du temps.

Dans ce livre blanc, nous vous présentons les défis que pose l'infrastructure d'intelligence artificielle. Vous découvrirez également comment NetApp peut vous aider à créer des pipelines de données pour vos workflows d'apprentissage profond actuels tout en assurant la pérennité de l'infrastructure d'IA dans laquelle vous investissez. La planification minutieuse de votre infrastructure permet de fluidifier la circulation des données dans votre pipeline d'AP, d'accélérer le déploiement et d'optimiser votre avantage concurrentiel.

SOMMAIRE

1	Introduction : votre infrastructure est-elle prête pour les workflows d'IA en production ?	3
2	Flux de données dans un pipeline d'apprentissage profond	4
3	Accélérer les E/S dans un pipeline d'apprentissage profond	5
3.1	Éliminer les goulets d'étranglement en périphérie	5
3.2	Éliminer les goulets d'étranglement dans l'infrastructure locale	6
3.3	Éliminer les goulets d'étranglement dans le cloud	8
4	Système de fichiers et architecture de données pour le pipeline d'apprentissage profond	9
4.1	Flux de données dans le cluster d'apprentissage	11
4.2	Autres facteurs de performance	13
5	Technologies NetApp et pipeline d'apprentissage profond	15
6	Pérennisation du pipeline d'apprentissage profond	16
6.1	Planifier l'évolution du matériel dans le cœur	16
Conclusion : prenez le contrôle de votre pipeline de traitement de données et de votre avenir avec l'IA		17

LISTE DES TABLEAUX

Tableau 1) Questions clés et principaux éléments à prendre en compte	10
--	----

LISTE DES FIGURES

Figure 1) Un pipeline de traitement de données conçu pour l'apprentissage profond peut aussi prendre en charge d'autres workflows d'IA et de Big Data	3
Figure 2) Hébergement dans l'infrastructure locale ou dans le cloud, les deux sont possibles pour un pipeline d'apprentissage profond	4
Figure 3) Le tiering des données associé à l'analytique en périphérie permet de classer les données en deux catégories : les données prioritaires destinées au cœur et les données à faible priorité à archiver	6
Figure 4) Pipeline d'apprentissage profond dont le cœur est dans l'infrastructure locale	7
Figure 5) En plaçant vos données à proximité du cloud, vous bénéficiez de la puissance de calcul du cloud tout en améliorant l'accélération des données et en gardant le contrôle	9
Figure 6) Des données non structurées peuvent être agrégées dans le data lake et envoyées vers le cluster d'apprentissage	11
Figure 7) Les données structurées sont lues à l'aide d'E/S aléatoires peu volumineuses et fusionnées dans le cluster d'apprentissage	12
Figure 8) Les technologies NetApp pour Data Fabric	16
Figure 9) Le cœur de votre pipeline IA/AM/AP continuera d'évoluer	17

1 Introduction : votre infrastructure est-elle prête pour les workflows d'IA en production ?

Les technologies d'intelligence artificielle (IA) suscitent un vif intérêt auprès d'entreprises de tous secteurs, qui veulent les exploiter pour lancer de nouveaux services et obtenir de nouvelles informations à partir de leurs données. Mais une fois l'étape du Proof of Concept passée, les équipes de data scientists se heurtent souvent à des difficultés liées à la gestion des données. Par exemple, il peut s'avérer compliqué de déplacer ou de copier des données dans plusieurs référentiels. De même, il peut être difficile de répondre aux attentes en matière de qualité de service, tant du point de vue des performances que de la protection, particulièrement lorsque les datasets sont dynamiques et volumineux.

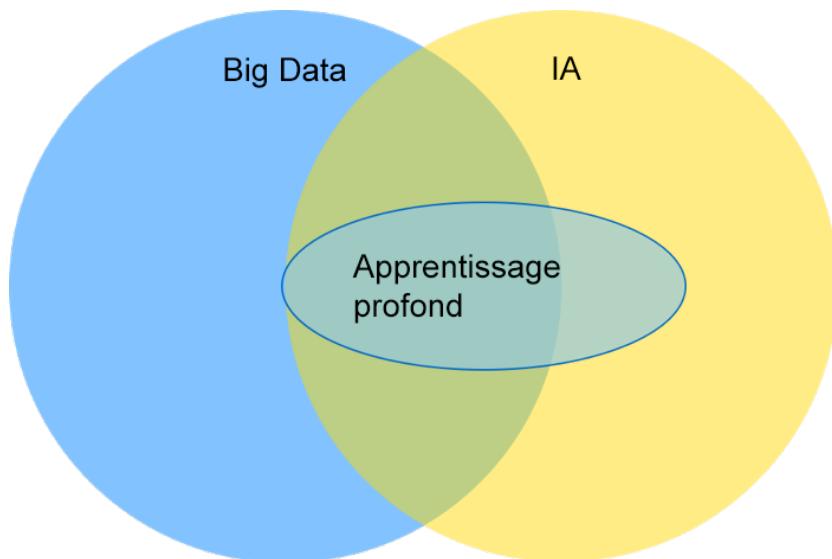
Une partie de la problématique réside dans le fait que le flux de données d'un workflow d'IA n'est pas isolé dans le data center. Alors que les entreprises de tous types adoptent l'Internet des objets et les technologies d'IA, elles doivent faire face aux défis que représentent leurs données, de la périphérie, au cœur jusqu'au cloud.

Par exemple, de nombreux fabricants automobiles ont commencé à collecter des données issues d'un nombre croissant de véhicules (la périphérie). Ces données sont ensuite utilisées afin d'entraîner les algorithmes d'IA nécessaires au fonctionnement autonome des systèmes (le cœur). Les datasets voient leur taille augmenter de manière exponentielle et doivent être stockés afin d'être réutilisés. Par conséquent, leur stockage doit avoir lieu sur une plateforme à la fois évolutive et économique (le cloud). À l'heure actuelle, les entreprises du secteur automobile exploitent les technologies informatiques à leurs limites. Les grandes chaînes de magasins font face à des défis similaires. En effet, elles créent des modèles d'inférence basés sur des données collectées à partir de terminaux de centaines de points de vente à travers le monde.

D'aucuns voudraient vous faire croire que le défi de l'IA en matière de données se résume aux performances. Elles sont certes essentielles au cœur d'un pipeline d'IA, mais s'appuyer sur un pipeline englobant l'intégralité du flux de données, de l'ingestion jusqu'à l'archivage, est nécessaire pour assurer votre succès opérationnel et optimiser les performances, l'efficacité et le coût à chaque étape.

Dans ce livre blanc, nous vous présentons les défis posés par l'infrastructure d'intelligence artificielle. Vous découvrirez également la façon dont NetApp peut vous aider à élaborer un pipeline de traitement de données pour l'apprentissage profond (AP), qui est workflow le plus exigeant en termes de puissance de calcul et d'E/S. Un pipeline conçu pour l'AP peut aussi servir pour d'autres workflows d'IA et de Big Data. (Voir Figure 1.)

Figure 1) Un pipeline de traitement de données conçu pour l'apprentissage profond peut aussi prendre en charge d'autres workflows d'IA et de Big Data.

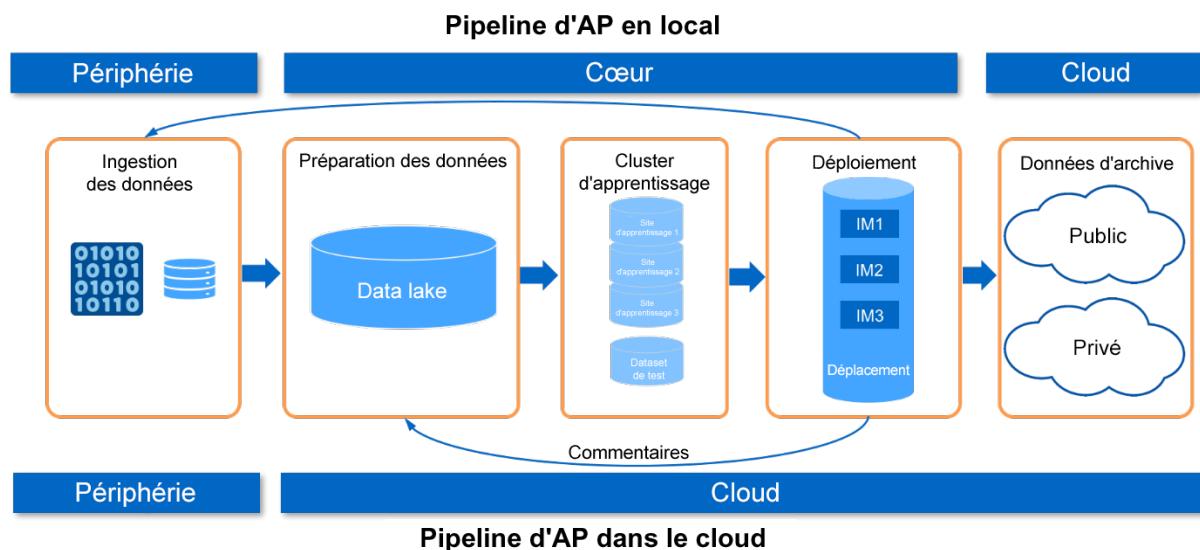


2 Flux de données dans un pipeline d'apprentissage profond

Lors de la conception d'un pipeline de traitement de données pour l'IA ou l'AP, il faut prendre en considération les étapes présentées à la Figure 2 :

1. **Ingestion des données.** L'ingestion a généralement lieu en périphérie : par exemple lors de la collecte de données, affluent depuis des voitures ou des terminaux en points de vente. Selon l'utilisation, une infrastructure informatique peut se révéler nécessaire au niveau du point d'ingestion ou à proximité. Un distributeur peut avoir besoin d'une empreinte réduite dans chaque magasin afin de consolider les données de plusieurs équipements.
2. **Préparation des données.** Un prétraitement des données est nécessaire pour les normaliser avant de procéder à la formation. Ce prétraitement a lieu dans un data lake qui se trouve soit dans le cloud, sous la forme d'un tier S3, soit localement sous la forme d'un magasin de fichiers ou d'objets.
3. **Formation.** Pour la phase de formation critique de l'apprentissage profond, les données sont généralement copiées à intervalles réguliers du data lake vers le cluster d'apprentissage. Souvent, les serveurs utilisés lors de cette phase pour paralléliser les opérations ont recours à des processeurs graphiques (GPU) ou à des puces en silicium personnalisées dont l'appétit en données est gigantesque. La bande passante d'E/S brute joue ici un rôle crucial.
4. **Déploiement.** Le modèle résultant de la formation est testé de manière approfondie puis mis en production. Selon le cas de figure, le modèle peut être redéployé aux opérations en périphérie. Ses résultats en conditions réelles sont surveillés et le retour d'information fourni sous forme de nouvelles données est renvoyé au data lake avec de nouvelles données pour l'itération du processus.
5. **Archivage.** Les données inactives d'itérations passées peuvent être conservées indéfiniment. Nombre d'équipes spécialisées en IA archivent les données inactives dans un système de stockage objet dans un cloud public ou privé.

Figure 2) Hébergement dans l'infrastructure locale ou dans le cloud, les deux sont possibles pour un pipeline d'apprentissage profond.



De nombreuses entreprises ont tenté de construire ce type de pipeline de traitement de données, soit dans le cloud, soit localement ; généralement en utilisant du matériel générique et en adoptant une approche de la gestion des données par la force brute. Le coût du cloud peut très vite devenir prohibitif. Déplacer de gros volumes de données depuis le cloud devient vite coûteux. Une fois que les données y sont déposées, le reste de votre pipeline finira sûrement par s'exécuter dans le cloud. Dans tous les cas, des goulets d'étranglement apparaissent inévitablement lors du passage à la production et de l'augmentation des volumes de données.

Les goulots d'étranglement les plus significatifs surgissent pendant la phase de formation. Le transfert des données vers le cluster de formation d'apprentissage profond requiert une bande passante et un parallélisme d'E/S considérables. Après la phase de formation, les modèles d'inférence produits sont souvent conservés dans un référentiel de style DevOps où ils bénéficient d'un accès à latence ultrafaible.

Pourtant, si le flux de données à travers tout le pipeline n'est pas fluide, à commencer par l'ingestion, alors la productivité de votre pipeline de traitement de données d'IA ne sera jamais optimale. La gestion du pipeline demandera donc de plus en plus de temps à vos collaborateurs.

3 Accélérer les E/S dans un pipeline d'apprentissage profond

Que votre workflow d'IA s'exécute localement ou dans le cloud, les goulots d'étranglement opérationnels peuvent allonger la durée de chaque cycle de formation. Ce temps supplémentaire réduit la productivité de votre pipeline et de vos collaborateurs.

Dans cette section, nous vous proposons de découvrir des solutions aux différents types de goulots d'étranglement :

- Les goulots d'étranglement en périphérie qui ralentissent l'ingestion de données
- Les goulots d'étranglement dans l'infrastructure locale
- Les goulots d'étranglement dans le cloud

Les trois phases au cœur du pipeline (préparation des données, formation et déploiement) créent des exigences d'E/S uniques demandant des réponses spécifiques.

3.1 Éliminer les goulots d'étranglement en périphérie

Le volume de données généré par les terminaux intelligents et la multitude des points d'acquisition peuvent entraîner la saturation des systèmes de calcul, du stockage et des réseaux en périphérie. Ces données peuvent être à l'origine de goulots d'étranglement lors de leur transmission à votre data center ou au cloud.

L'analytique en périphérie permet de traiter et de sélectionner les données au moment de leur ingestion. Cette approche requiert que la périphérie soit dotée d'une infrastructure haute performance et d'un stockage à latence ultrafaible. De nombreux clients de NetApp adoptent une approche hiérarchique avec des infrastructures dites du dernier kilomètre, c'est-à-dire situées au plus près des emplacements de production des données, et des capteurs en périphérique servant de terminaux.

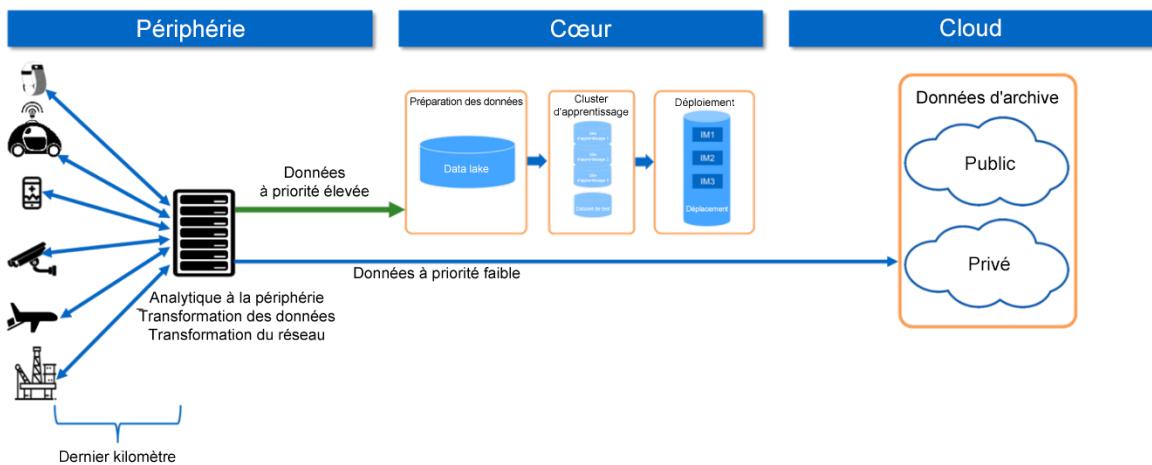
Dans le secteur industriel, des capteurs placés au niveau des équipements alimentent en données les infrastructures déployées dans chaque usine. Les données y sont agrégées et analysées avant de remonter la chaîne après avoir été triées. Cette approche peut également être appropriée pour les véhicules autonomes (chaque terminal peut générer jusqu'à 7 To de données par jour), le retail, et dans de nombreux autres secteurs.

Gestion des données par tier en périphérie

L'analytique en périphérie permet de créer différents tiers de service de données. Grâce à cette approche, certaines données sont priorisées, soit via un simple filtre, soit via une analyse avancée associée à l'IA, puis elles sont transmises de manière efficace au pipeline d'IA/apprentissage machine (AM)/apprentissage profond (AP). Les autres données sont dépriorisées et sont soit supprimées, soit gérées avec un service d'une autre classe.

Selon les exigences, chaque tier de données peut être traité et transformé différemment afin d'atteindre les niveaux de sécurité et d'efficacité de stockage nécessaires. Par exemple, les données à faible priorité sont compressées, dédupliquées, chiffrées et stockées dans un référentiel cloud à des fins de conformité ou en vue d'une éventuelle utilisation ultérieure. (Voir Figure 3.)

Figure 3) Le tiering des données associé à l'analytique en périphérie permet de classer les données en deux catégories : les données prioritaires destinées au cœur et les données à faible priorité à archiver.



La capacité à traiter l'analytique en périphérie est fonction de la puissance de calcul disponible. Actuellement, les fournisseurs de ressources de calcul et de cloud se disputent l'empreinte à la périphérie, déployant pour ce faire diverses stratégies. NVIDIA par exemple amène la puissance des processeurs graphiques (GPU) à la périphérie pour rendre l'IA disponible à des applications telles que les véhicules autonomes. Le point commun de toutes ces solutions du point de vue des données est l'incorporation de DAS génériques dénués d'une gestion intelligente des données. Or le besoin en stockage de données intelligent est ici évident.

Mécanisme efficace de déplacement des données

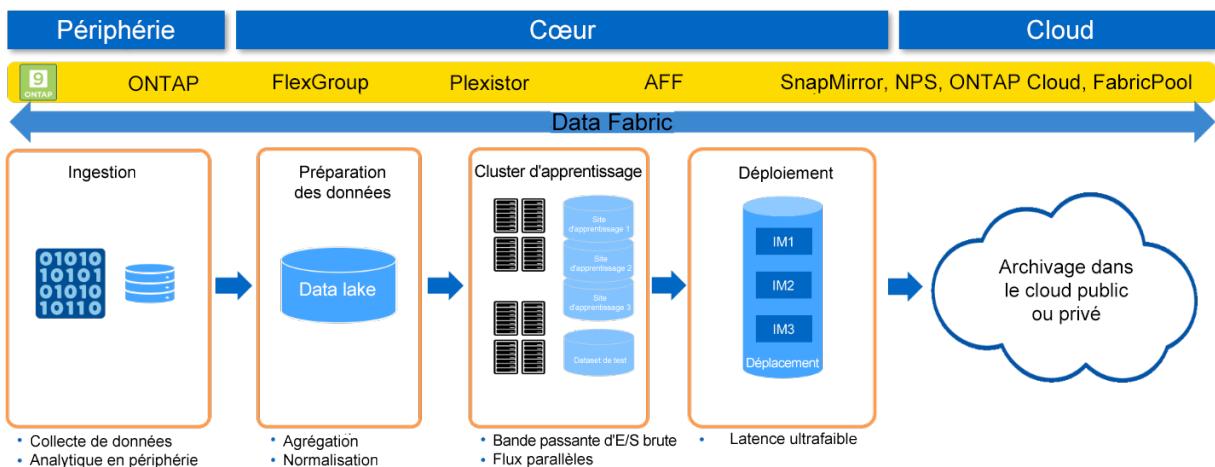
Pour transmettre des données à large bande passante depuis la périphérie, il est essentiel de disposer d'un mécanisme efficace de déplacement des données. Dans les architectures les plus communes actuellement, les données sont déplacées intégralement par le biais d'opérations S3 PUT. Cette approche présente certains inconvénients, car les données sont intégralement déplacées sans être transformées.

Cette méthode brute peut être remplacée par un mécanisme efficace de déplacement des données qui fusionne les données, les transforme afin de réduire leur empreinte, et applique des transformations au réseau de manière à ne déplacer que les blocs modifiés. Un tel mécanisme peut accélérer significativement le déplacement des données et réduire les exigences en bande passante.

3.2 Éliminer les goulets d'étranglement dans l'infrastructure locale

Si le cœur de votre pipeline d'apprentissage profond se situe dans l'infrastructure locale, comme dans la Figure 4, alors vous bénéficiez d'un contrôle direct sur le data lake, le cluster d'apprentissage, et le déploiement du modèle d'inférence.

Figure 4) Pipeline d'apprentissage profond dont le cœur est dans l'infrastructure locale.



Data lake

Les données en provenance de la périphérie sont collectées dans un data lake. Un data lake dont la mise en œuvre n'a pas été correctement réalisée se transforme en goulot d'étranglement lorsque les volumes de données augmentent. Un data lake peut prendre la forme d'une solution Hadoop avec un système de fichiers Hadoop (HDFS), ou être mis en œuvre à l'aide d'un magasin d'objets ou de fichiers. HDFS n'est pas optimisé pour les performances et conserve généralement trois exemplaires des objets de données, ce qui ralentit les performances d'écriture et augmente le coût.

Les magasins d'objets étaient initialement destinés à l'archivage dans le cloud et non à la performance, mais dans de nombreux cas, ils sont devenus de fait les magasins de données des projets de Big Data. Comme vous avez pu le constater précédemment, les magasins d'objets laissent à désirer en matière de performances pour l'apprentissage profond.

Pour ce qui est des magasins de fichiers, les systèmes de fichiers scale-out, tels que Lustre et GPFS, sont conçus pour l'informatique haute performance (HPC) et le traitement par lot. Ils ne réagissent pas de manière optimale face à des charges de travail de fichiers peu volumineux. Les données qui sont envoyées par des appareils périphériques intelligents au data lake se présentent généralement sous la forme d'une multitude de fichiers peu volumineux pour lesquels ces systèmes ne sont pas optimisés. C'est pourquoi les performances en souffrent.

Le système NetApp® AFF, notamment lorsqu'il est combiné à NFS et des volumes ONTAP® Flexgroup, permet de surmonter les limites des autres approches de data lake. Les groupes FlexGroup assurent de hautes performances à la fois pour les charges de travail séquentielles orientées bande passante, et pour les charges de travail de fichiers peu volumineux. Les autres solutions de data lake mentionnées, à savoir HDFS, le stockage objet, Lustre, GPFS et d'autres stockages de fichiers scale-out, peuvent éventuellement faire l'un ou l'autre. Cependant, ils ne sont pas en mesure d'être performants à la fois pour les E/S séquentielles et aléatoires.

Cluster d'apprentissage

Une configuration scale-out avec 32 à 64 serveurs et 4 à 8 GPU par serveur constitue le nec plus ultra pour les clusters de formation d'apprentissage profond. Du point de vue des E/S, il est impératif de maintenir les GPU occupés à 100 %, ce qui signifie qu'il faut fournir à chaque cœur de processeur un flux parallèle en E/S. Ensuite, chaque cœur de processeur a une affinité avec un GPU. Le processeur traite le flux, fusionne les E/S et alimente le GPU en données.

Ce processus provoque des goulets d'étranglement des E/S, car :

- Les données doivent circuler rapidement et efficacement du data lake au cluster.
- Jusqu'à 256 flux parallèles d'E/S (32 à 64 serveurs avec chacun de 4 à 8 GPU) doivent être chargés et indexés pour alimenter les GPU de façon à ce qu'ils reçoivent des données de manière continue.

L'architecture logicielle NetApp ONTAP satisfait aux deux exigences de manière unique. Le data lake peut être conçu grâce à des nœuds Flash hybrides qui peuvent faire circuler les données jusqu'au cluster d'apprentissage à très haute bande passante. Les nœuds de stockage 100 % Flash qui supportent le cluster d'apprentissage fournissent une bande passante allant jusqu'à 18 Gbit/s par paire haute disponibilité à deux contrôleurs et des latences inférieures à 500 microsecondes. C'est ainsi qu'ils fournissent la bande passante nécessaire pour prendre en charge autant de flux d'E/S en parallèle. NetApp propose également une technologie de feuille de route qui vous permet de continuer à accroître les performances d'E/S de votre pipeline IA/AM/AP à mesure que vos besoins augmentent.

Déploiement

Une fois la formation terminée, les modèles d'inférence produits sont déposés dans un référentiel de style DevOps et soumis à des tests d'inférence et de validation d'hypothèses. C'est à cette étape qu'il est important de déployer des systèmes de stockage qui prennent en charge des latences extrêmement faibles.

Grâce à NetApp, une architecture de stockage unique répond à tous les besoins en performances du cœur de votre pipeline d'apprentissage profond. Bien que cette approche présente des avantages immédiats, la solution optimale pour la plupart des clients consiste à exploiter des clusters distincts pour chaque étape du pipeline. Des pipelines de Big Data peuvent déjà être en place et les data lake déjà déployés. Vous pouvez ne vouloir déployer que les éléments nécessaires à l'apprentissage profond, en tant que projet séparé et copier les données à chaque phase. Cependant, comme le volume de données ne cesse de croître, il vous faudra unifier davantage le pipeline. L'AFF rend cette unification possible également.

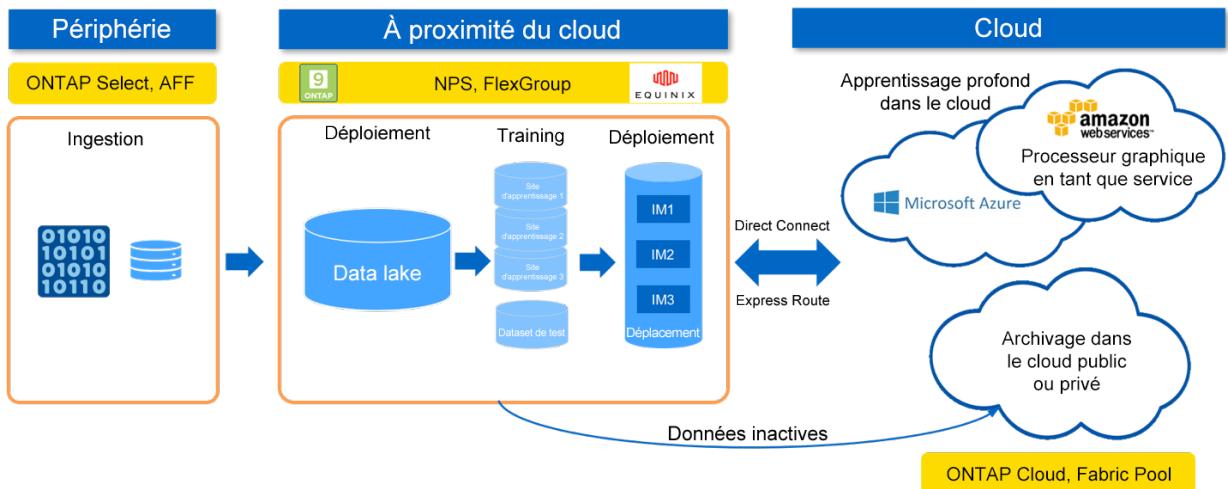
3.3 Éliminer les goulots d'étranglement dans le cloud

Vous pouvez décider de déployer l'apprentissage profond dans le cloud par souci d'agilité et de facilité d'utilisation. Néanmoins, les mêmes goulots d'étranglement potentiels peuvent apparaître lorsque vous exploitez votre pipeline d'apprentissage profond dans le cloud.

- Votre data lake est-il en mesure d'assurer les performances nécessaires pour l'ingestion ? Peut-il faire circuler les données jusqu'au cluster d'apprentissage ?
- Votre fournisseur de cloud est-il en mesure de lui assurer le parallélisme d'E/S nécessaire ?
- Comment assurer les latences ultrafaibles nécessaires aux modèles d'inférence finis ?
- Que se passe-t-il si vous devez vous assurer de la souveraineté des données en cas de données sensibles ?

Le système NetApp Private Storage (NPS) vous permet de stocker vos données à proximité du cloud de manière à pouvoir utiliser les fonctionnalités de calculs du cloud public et d'autres services tout en gardant le contrôle de vos données. (Voir Figure 5.) NPS permet d'utiliser un cloud public tout en bénéficiant de la même architecture et des mêmes performances que celles décrites dans la section précédente. Les problèmes de souveraineté des données sont ainsi écartés et vos données ne sont jamais bloquées dans le cloud.

Figure 5) En plaçant vos données à proximité du cloud, vous bénéficiez de la puissance de calcul du cloud tout en améliorant l'accélération des données et en gardant le contrôle.



Si votre data lake doit impérativement résider dans le cloud, NetApp Data Fabric vous permet de stocker et de gérer de manière fluide vos données NFS depuis le service cloud Azure ou AWS.

4 Système de fichiers et architecture de données pour le pipeline d'apprentissage profond

Dans un pipeline d'IA, il y a différentes caractéristiques d'E/S pour les données affluent de la périphérie, en comparaison avec les données provenant du data lake en direction du cluster d'apprentissage. Dans cette section, vous découvrirez les différents facteurs à prendre en considération pour fluidifier le flux de données dans le pipeline et vers le cluster d'apprentissage.

Imaginez que le pipeline est une piste d'asphalte et le GPU de votre cluster d'apprentissage une voiture ultra puissante. Si vous conduisez sur un circuit d'essai ou sur le périphérique à l'heure de pointe, la différence est considérable. Pour que votre déploiement fournit des résultats optimaux pour l'IA, y compris pour l'AM et l'AP, l'élément le plus important à prendre en compte est probablement le pipeline de traitement de données ; mais il est souvent négligé. L'architecture de données optimale prend en compte les besoins en E/S depuis la périphérie jusqu'au cluster d'apprentissage en passant par le data lake.

Le stockage objet n'est pas conçu pour fournir le niveau de performance exigé par votre pipeline de traitement de données. Les magasins d'objets étaient initialement destinés à l'archivage en cloud et non à la performance, mais dans de nombreux cas, ils sont devenus de fait les magasins de données privilégiés des projets de Big Data. Pour l'apprentissage profond en particulier, les magasins d'objets laissent à désirer en matière de performances.

Le système de fichiers et l'architecture de données que vous choisissez doivent prendre en compte tous les facteurs qui sont importants pour votre environnement d'IA. Le stockage à base fichier demeure le meilleur choix, mais de nombreux facteurs entrent en ligne de compte comme indiqué dans le Tableau 1.

Tableau 1) Questions clés et principaux éléments à prendre en compte

Questions clés	Principaux éléments à prendre en compte
Quels système de fichiers sont à considérer ?	<ul style="list-style-type: none"> • Un système de fichiers scale-out comme Lustre ou GPFS • HDFS, un système de fichiers répandu pour le Big Data • NFS, le système de fichiers partagés le plus largement déployé en vue d'applications techniques depuis les 30 dernières années
Le système de fichiers peut-il prendre en charge et fédérer à la fois des données non structurées et des données structurées de sources variées, sans pour autant sacrifier les performances ?	<ul style="list-style-type: none"> • Données de journaux et de capteurs • Bases de données, y compris RDBMS et NoSQL • E/S aléatoires pour de nombreuses bases de données : analyses de tables, lecture et collecte de documents dans NoSQL, lecture par colonne dans des bases de données orientées colonnes, lectures aléatoires clé-valeur dans des bases de données à clé-valeur • E/S séquentielles pour les bases de données in-memory et moteurs in-memory comme Spark • Journaux d'e-mails • Répertoires locaux • Autres sources
Le système de fichiers assure-t-il les performances requises pour les E/S aléatoires peu volumineuses et les E/S séquentielles ?	<ul style="list-style-type: none"> • Certaines sources de données génèrent des E/S aléatoires, alors que d'autres sont séquentielles • Le système de fichiers doit pouvoir équilibrer les performances pour les deux types d'E/S
Quelles sont les performances et les fonctionnalités des mécanismes de déplacement de données ?	<ul style="list-style-type: none"> • Performances les plus élevées • Déplacement des données le plus efficace
Permet-il d'automatiser le cycle de vie des données ?	<ul style="list-style-type: none"> • Filtrage intelligent pour déterminer quelles données sont destinées au cœur ou aux tiers d'archivage • Performance en temps réel pour les décisions de filtrage
Prend-il en charge les supports de stockage et de mémoire les plus récents pour profiter des avancées de performances et de latence de manière non disruptive ?	<ul style="list-style-type: none"> • Tiers de stockage pouvant satisfaire le rapport prix/performance pour votre magasin de données, incluant les technologies SCM (storage-class memory), NVMe (nonvolatile memory express), Flash, Flash hybride et cloud • Déplacement des données non disruptif à travers les tiers • Conceptions scale-out pour ajouter des performances incrémentielles

4.1 Flux de données dans le cluster d'apprentissage

Outre les éléments à prendre en compte indiqués dans le Tableau 1, certaines variations concernant le flux des données jusqu'au cluster doivent également être prises en considération. Ces facteurs affectent :

- L'endroit où les E/S sont agrégées
- Les exigences d'un namespace unique
- L'évolutivité des métadonnées

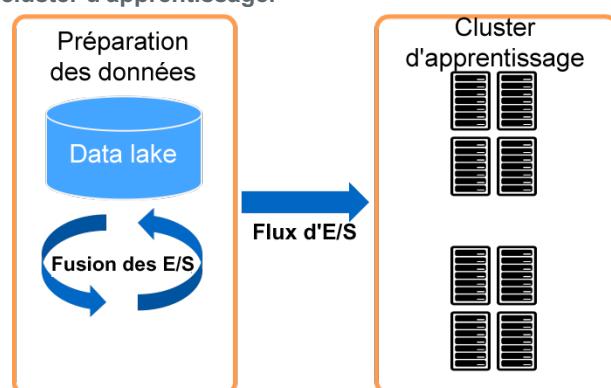
Fusion des E/S

La mise en forme des données dépend de leur source. La fusion des E/S peut avoir lieu à deux endroits :

- Dans le data lake, pendant les phases de mise en forme et de transformation des données ; ce qui produit un flux d'E/S vers le cluster d'apprentissage
- Dans le cluster lui-même ; ce qui produit des E/S aléatoires en provenance du data lake

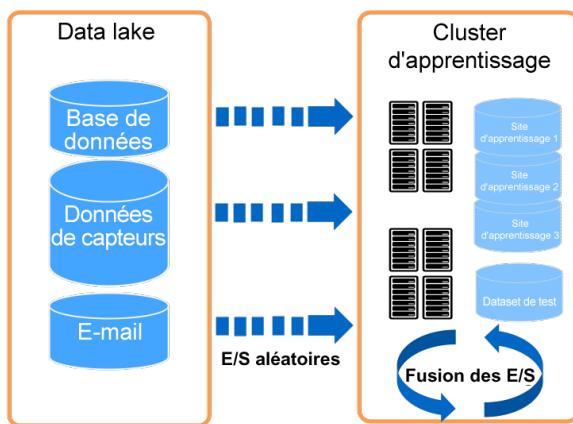
Si vous avez affaire à un data lake non structuré, ce qui correspond presque par définition à un système de fichiers, celui-ci est capable de mettre en forme les données et de les arranger comme un ensemble de flux de fichiers agrégés. Ces derniers peuvent être adaptés au cluster d'apprentissage, permettant ainsi aux données de parvenir directement à son processeur, ainsi que de précharger et d'alimenter les GPU. (Voir Figure 6.)

Figure 6) Des données non structurées peuvent être agrégées dans le data lake et envoyées vers le cluster d'apprentissage.



Contrairement aux sources de données comme les bases de données, les journaux de capteurs, les journaux de fichiers, les e-mails, etc., il peut être impossible d'obtenir des lectures correctement mises en forme permettant un flux direct des données vers le cluster. Dans ce cas, l'accès aux données se fait via des lectures aléatoires et la fusion des E/S a lieu à l'intérieur même du cluster d'apprentissage. (Voir Figure 7.)

Figure 7) Les données structurées sont lues à l'aide d'E/S aléatoires peu volumineuses et fusionnées dans le cluster d'apprentissage.



Selon votre type de sources de données, votre architecture devra éventuellement être capable de fournir au cluster d'apprentissage à la fois des lectures séquentielles de gros volumes et des lectures aléatoires peu volumineuses.

Namespace unique

Les datasets d'IA peuvent atteindre des volumes colossaux, entraînant une prolifération des données considérable. Pour s'adapter à la croissance des volumes, le système de fichiers doit être de type scale-out. Celui-ci doit avoir un seul namespace capable de faire monter en charge les performances de manière linéaire vers un nœud client unique ou vers plusieurs nœuds clients accédant aux mêmes données en parallèle. Il est essentiel de disposer d'une architecture capable d'évoluer au moment où vous ajoutez des ressources de calcul et de la capacité.

Il peut y avoir plusieurs types d'accès client à ce namespace unique, ayant chacun un impact sur les performances. Certains modèles de formation sont considérés comme « asynchrones ». C'est-à-dire que le dataset présente un partitionnement statique dans l'ensemble des nœuds du cluster d'apprentissage avec un accès à nœud unique aux régions du namespace ; ce qui produit un scénario dit de « client unique actif ».

Les autres modèles de formation opèrent de manière synchrone. Ils sont étroitement associés à leur dataset qui est partagé entre tous les nœuds du cluster avec un accès simultané. Ce scénario de « multiclient actif » est le plus exigeant du point de vue des performances.

Il existe d'autres cas où un réseau neuronal multicouche assure la formation des couches du réseau sur plusieurs nœuds. Les nœuds servent de pipeline modèle où le modèle d'inférence progresse de nœud en nœud. Ainsi, l'intégralité du dataset est lue de manière répétée, un nœud à la fois et passant de l'un à l'autre.

La diversité des charges de travail pour lesquelles NFS a été employé est très grande, et vous pourrez le constater lors de vos recherches d'un système de fichiers capable de prendre en charge ces modèles d'utilisation. Il peut aussi bien s'agir d'informatique haute performance, à laquelle NFS était initialement destiné, que de répertoires locaux et de bases de données comme Oracle et SQL qui fonctionnent avec un stockage NAS. Les systèmes SAP et plus récemment la virtualisation et le Big Data sont également concernés. NFS dispose d'une longue expérience sur des charges de travail variées, ce qui lui permet de gérer non seulement les E/S aléatoires, mais aussi les séquentielles. Ces E/S peuvent être générées par plusieurs modèles d'accès au namespace, en particulier lorsqu'elles sont associées aux avantages du stockage 100 % Flash dans un cluster scale-out linéaire.

HDFS étant un système de fichiers relativement récent, l'éventail de charges de travail et de caractéristiques de performances auquel il a déjà été confronté est encore limité. Les fournisseurs de Big Data ont entrepris des réécritures significatives (et propriétaires) afin de répondre aux besoins en performance lors de la transition de MapReduce à Spark. L'IA représente une nouvelle ombre au tableau dans l'histoire de HDFS.

L'utilisation de systèmes de fichiers étudiés pour le Big Data tels que HDFS signifie généralement plus de copies de données et de silos puisque vous devez recopier les données depuis HDFS vers un système de fichiers scale-out haute performance pour l'IA.

Performance des métadonnées

Les modèles d'accès mentionnés plus haut ont aussi un impact sur les performances des métadonnées. Chaque nœud du cluster d'apprentissage peut interroger les métadonnées de manière indépendante. Ainsi, les performances d'accès aux métadonnées doivent monter en charge de manière linéaire pour accompagner la croissance du système de fichiers. L'accès aux métadonnées via des systèmes de fichiers comme Lustre ou GPFS peut donner lieu à des goulots d'étranglement puisqu'ils s'appuient sur des serveurs et un stockage des métadonnées distincts.

4.2 Autres facteurs de performance

Les facteurs que vous devez prendre en considération lors du choix d'un système de fichiers pour votre pipeline d'IA sont nombreux et affectent à la fois les performances et la facilité d'utilisation. Ils incluent :

- La facilité de gestion
- La qualité de service (QoS)
- Les fonctionnalités de clonage
- L'écosystème de solutions de mise en cache côté client
- La capacité à mettre en œuvre l'IA/AP avec un système de fichiers unifié du data lake aux tiers d'IA/AP
- La prise en charge des supports
- La pérennité

Facilité de gestion

Avant de choisir un système de fichiers, il est important de s'interroger sur sa gestion. Est-il en mesure d'évoluer de manière autonome et automatique sans intervention externe ? Combien de temps faut-il allouer à sa gestion et quel est le niveau d'expertise requis ? Est-il facile de trouver des collaborateurs disposant des compétences requises ?

La configuration de systèmes de fichiers scale-out comme Lustre et GPFS peut représenter un défi, de même que leur entretien, leur contrôle et leur gestion. En comparaison, NFS offre une gestion aisée et le marché regorge d'experts en la matière.

Qualité de service (QoS)

La qualité de service peut représenter un élément important de votre architecture de données. Le coût de construction de vos clusters d'apprentissage mutualisés peut se chiffrer en millions de dollars. La qualité de service joue un rôle primordial dans votre capacité à assurer cette colocation, qui permet à des activités multiples de partager les mêmes ressources.

- Le système de fichiers assure-t-il la qualité de service ?
- Est-elle intégrée de bout en bout ?
- Est-il possible de mettre en place des limites et des valeurs maximales d'utilisation sur le stockage, le réseau et la puissance de calcul pour partitionner les niveaux de services pour plusieurs modèles de formation ?

Fonctionnalités de clonage

La colocation a notamment pour but de satisfaire aux exigences de plusieurs fonctions dans l'entreprise. Vous pouvez ainsi avoir plusieurs modèles de formation à des étapes de développement variées pour des cas d'utilisation divers :

- Formation anticipée
- Validation du modèle

- Prédéploiement
- Déploiement en production

La possibilité de cloner les datasets et d'attribuer des paramètres de qualité de services différents selon le clone vous permet de proposer des performances adaptées aux niveaux de services selon le cas d'usage. Le clonage compact est donc un incontournable si l'on possède un cluster mutualisé.

Mise en cache côté client

La mise en cache côté client vous permet d'améliorer encore les performances en fournissant une mémoire tampon de données. Elle permet au flux de données de circuler en continu pendant que les nœuds de cluster accèdent au dataset d'apprentissage. Les avantages peuvent être considérables si le système de fichiers prend en charge un écosystème de solutions de mise en cache côté client, qu'elles soient propriétaires ou open source.

Dans les deux cas, les options existent et sont nombreuses pour un stockage basé sur NFS, alors qu'elles sont encore peu nombreuses pour les systèmes Lustre, GPFS ou HDFS, et quasiment aucune n'est disponible en open source ou très répandue.

IA/AP sur place grâce à un système de fichiers unifié

Il peut arriver que vous souhaitiez utiliser les mêmes données à la fois pour une charge de travail d'analytique Big Data et pour d'autres d'IA, d'AM ou d'AP. Pour l'IA appliquée post-traitement, comme c'est le cas notamment pour la surveillance ou la détection des fraudes, un système de fichiers approprié vous permet d'exécuter les deux charges de travail sans avoir besoin de copier les données. Le dataset réside dans un emplacement unique. Le traitement des calculs est appliqué non seulement pour l'analytique sur place, mais aussi pour l'IA/AM/AP (éventuellement avec une solution de mise en cache côté client, comme indiqué plus haut), sans copier les données dans des systèmes de fichiers dédiés à votre data lake et votre cluster d'apprentissage.

Cependant, si l'exigence qui prime est la performance en temps réel ou s'il s'agit d'un avantage compétitif stratégique, vous ne pourrez sûrement pas vous passer d'une copie des données dédiée au cluster d'apprentissage.

Prise en charge des supports et systèmes de mémoire de pointe

Enfin, choisissez un système de fichiers capable de prendre en charge les dernières avancées en matière de support et de mémoire. Ainsi les performances de votre pipeline de traitement de données peuvent continuer à suivre l'évolution de la technologie. Le système de fichiers est-il optimisé pour la technologie Flash ? Peut-il être étendu de manière fluide afin de prendre en charge de nouvelles technologies ? Les fournisseurs innovent-ils dans les secteurs comme les technologies NVMe over Fabrics (NVMe-oF), NVDIMM et 3D XPoint ?

La technologie Flash atteint aujourd'hui des latences de 500 microsecondes. Le NVMe-oF fait descendre ce chiffre à 200 microsecondes. NVDIMM, 3D XPoint et la mémoire persistante sont en bonne voie pour faire tomber ce chiffre sous le seuil des 100 microsecondes, puis en dessous des 10 microsecondes et finalement les faire diminuer à l'échelle des nanosecondes. Le fournisseur de votre pipeline de traitement de données doit investir durablement pour rester dans la course face à cette évolution des solutions de stockage partagées et sur serveurs.

Architecture de données et système de fichiers pérennes

Dans le domaine de l'IA, les choses évoluent très rapidement, et repartir de zéro tous les six mois à un an peut se révéler peu pratique, voire impossible. C'est pourquoi il est aussi recommandé de s'assurer que vos choix technologiques sont les plus pérennes que possible. Il est important de faire évoluer de manière fluide et non disruptive les différentes couches de technologies, comme le système de fichiers, l'interconnexion, l'emplacement du déploiement, les supports, et le type de mémoire d'une infrastructure choisie. Cette possibilité vous assure un retour sur investissement à long terme et garantit que vous pourrez absorber les évolutions technologiques au fil du temps.

Votre choix de système de fichiers aujourd'hui dépendra sans doute des niveaux de confort de votre équipe, de leurs compétences et de leur expertise préalable. Disposer d'expérience d'un déploiement passé ou d'infrastructures et de déploiements présents constitue un facteur qu'il faudra aussi prendre en compte.

Par exemple, si vous êtes à l'aise et que vous souhaitez déployer en FC ou en InfiniBand, vous pourriez opter pour une architecture SAN et Lustre ou GPFS. Avec le temps, vous pourriez vous rendre compte que la feuille de route 100 GbE ou 400 GbE de NFS répond mieux à vos besoins. Une architecture de données bien planifiée doit pouvoir prendre en charge et pérenniser votre solution. Elle doit vous permettre de changer de système de fichiers facilement, sans remplacer votre infrastructure.

De la même façon, vous pouvez choisir NFS aujourd'hui mais décider par la suite que vous avez besoin d'un système de fichiers basé sur un SAN, des disques NVMe ou NVME-oF, ou bien encore d'une disposition des données basée sur la mémoire. Une architecture pérenne vous permet de faire évoluer vos technologies de magasins de données sans devoir remplacer l'intégralité de l'infrastructure en place.

Les différents points abordés dans cette section constituent les critères de base à prendre en compte pour sélectionner le système de fichiers et l'architecture de données qui correspondent le mieux à vos besoins en termes d'IA/AM/AP. Nous sommes convaincus que NFS associé au stockage NetApp AFF est le meilleur choix possible, d'après notre capacité à satisfaire ces besoins et à évoluer pour prendre en charge les technologies les plus récentes.

5 Technologies NetApp et pipeline d'apprentissage profond

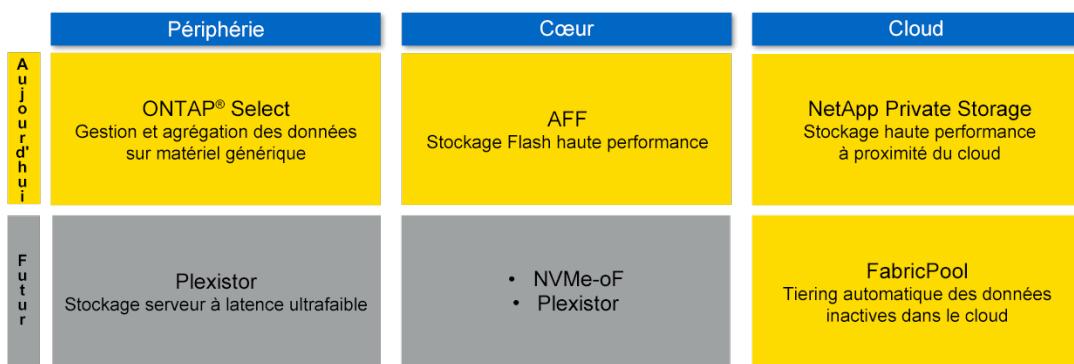
NetApp Data Fabric inclut les technologies de gestion des données pour répondre aux exigences de l'intégralité du pipeline d'apprentissage profond. (Voir Figure 8.) Les fournisseurs de cloud à eux seuls ne prennent pas en charge la périphérie et peuvent donc éprouver des difficultés à fournir des performances d'E/S satisfaisantes. D'autres fournisseurs de stockage tentent de résoudre le problème de la bande passante lors de la phase de formation, mais ne sont pas en mesure de fournir de latences ultrafaibles. Il leur manque par ailleurs la technologie nécessaire pour couvrir l'intégralité du workflow. C'est dans cette situation en particulier que NetApp Data Fabric présente des avantages évidents.

En périphérie, NetApp propose ONTAP Select qui s'exécute sur du matériel générique et assure l'agrégation des données et leur gestion avancée. Notre technologie Plexistor, prochainement disponible, facilitera l'ingestion, en particulier lorsque le taux d'ingestion est extrêmement élevé.

Pour répondre aux besoins en stockage du data lake et du cluster d'apprentissage, le stockage NetApp AFF fournit à la fois de hautes performances et de grandes capacités, mais il réduit également le besoin chronophage d'effectuer des copies de données. NetApp cherche actuellement à étendre les fonctionnalités de l'AFF, en ajoutant les technologies NVMe-oF et Plexistor. Les avantages de NPS sont les mêmes pour les pipelines d'apprentissage profond dans le cloud.

En matière d'archivage de données inactives, FabricPool migre les données automatiquement vers le stockage objet, selon des règles définies.

Figure 8) Les technologies NetApp pour Data Fabric



6 Pérennisation du pipeline d'apprentissage profond

La taille des datasets d'apprentissage profond et les exigences en E/S de votre pipeline vont sans aucun doute continuer de croître à mesure que vous augmentez le nombre de serveurs, et que les processeurs, GPU et puces en silicium dédiées à l'IA continuent de gagner en puissance. La feuille de route de NetApp intègre un certain nombre d'éléments qui vont permettre d'augmenter vos E/S et de rester dans la course. Vous disposez pour cela de nombreux atouts :

- **NVMe-oF.** NetApp fait encore diminuer les latences en intégrant la technologie NVME-oF dans son architecture AFF.
- **Plexistor.** En juin 2017, NetApp a fait l'acquisition de Plexistor, une technologie de stockage côté serveur qui permet de réduire davantage les latences et étend NetApp Data Fabric au serveur. La technologie Plexistor peut être déployée en périphérie, dans le cœur et dans le cloud. Elle accélère l'ingestion des données, l'analytique de la périphérie et la formation.

Par ailleurs, si le cœur de votre pipeline d'IA est hébergé localement, vous pourriez vouloir rechercher des stratégies facilitant l'évolution du matériel de votre cœur. Vous devriez également prévoir la possibilité d'un changement de stratégie à l'avenir avec le passage d'un modèle périphérie/cœur/cloud à un modèle périphérie/cloud.

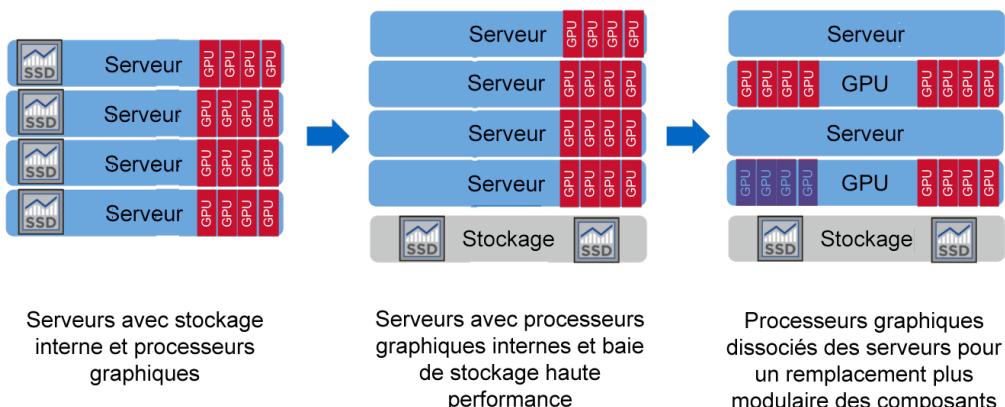
6.1 Planifier l'évolution du matériel dans le cœur

L'un des avantages du cloud est que vous pouvez avoir recours à un service d'apprentissage profond sans pour autant devoir maîtriser les subtilités de la pile matérielle. Ce confort a un prix en revanche : la perte de contrôle. Concernant le cœur du pipeline d'apprentissage profond hébergé localement, l'évolution en continu du matériel requis constitue l'une des tendances clés.

Une bataille fait ainsi rage entre les équipementiers pour déterminer lequel sera le meilleur pour l'apprentissage profond. Tandis que NVIDIA fait pour l'instant figure de leader, de nombreuses technologies émergent. Chaque fournisseur de cloud crée son propre matériel personnalisé. Les TPU de Google en sont un exemple. Beaucoup de startups créent également du matériel personnalisé pour l'IA.

Une tendance à surveiller serait donc la possibilité de séparer l'infrastructure serveur de l'infrastructure GPU, permettant ainsi aux deux d'évoluer de manière indépendante. Une solution qui extrait le matériel GPU du matériel de stockage et des serveurs (voir Figure 9) peut évoluer plus facilement (et à moindre coût) pour tirer parti de nouveaux développements.

Figure 9) Le cœur de votre pipeline IA/AM/AP continuera d'évoluer.



NetApp propose une infrastructure convergée avec plusieurs partenaires serveurs, notamment Cisco ([FlexPod®](#)), Fujitsu ([NFLEX™](#), qui est une infrastructure convergée développée par NetApp et Fujitsu), ainsi que d'autres fournisseurs. Ainsi, vous pouvez aisément exploiter le stockage NetApp combiné avec des plateformes serveur variées lors de la création de votre cluster d'apprentissage profond.

Conclusion : prenez le contrôle de votre pipeline de traitement de données et de votre avenir avec l'IA

Les conseils donnés dans ce livre blanc ont pour but de vous aider à planifier votre pipeline de traitement de données :

- Choisissez le système de fichiers et l'architecture de données les plus adaptés pour satisfaire à vos besoins actuels tout en préparant l'avenir.
- Accélérez votre flux de données dans votre pipeline, qu'il soit hébergé dans une infrastructure locale ou dans le cloud.
- Mettez en place une gestion des données intelligentes en périphérie pour mieux gérer la croissance du volume des données.
- Déplacez les données plus intelligemment et efficacement, de la périphérie jusqu'au cœur ou au cloud.
- Préparez-vous à passer à un modèle allant directement de la périphérie au cloud si une telle configuration peut se révéler nécessaire à l'avenir.
- Créez une architecture matérielle de cœur plus agile, capable d'évoluer rapidement.

Ces actions vous permettent d'éliminer les goulots d'étranglement et d'atteindre un meilleur débit tout en pérennisant vos investissements dans une infrastructure d'IA.

Reportez-vous à la [matrice d'interopérabilité \(IMT, Interoperability Matrix Tool\)](#) sur le site de support NetApp pour vous assurer que les versions de produits et de fonctionnalités mentionnées dans le présent document sont prises en charge par votre environnement. La matrice d'interopérabilité de NetApp définit les composants et les versions de produits qu'il est possible d'utiliser pour créer des configurations prises en charge par NetApp. Les résultats dépendent des installations de chaque client et de leur conformité aux spécifications publiées.

Informations sur le copyright

Copyright © 2018 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUSSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ) QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp n'accepte aucune responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (c)(1)(ii) de la clause Rights in Technical Data and Computer Software de DFARS 252.277-7103 (octobre 1988) et FAR 52-227-19 (juin 1987).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques présentes sur le site <http://www.netapp.com/TM> sont des marques commerciales de NetApp, Inc. Les autres noms de sociétés et de produits peuvent être des marques commerciales de leurs propriétaires respectifs.