

LANCER LA TRANSFORMATION  
DE VOTRE ENTREPRISE GRÂCE À  
UN ENTERPRISE DATA CLOUD



ENTERPRISE  
DATA  
CLOUD

RESSOURCES

**Les entreprises axées sur l'information connaissent une croissance de plus de 30 % par an en moyenne**

Insights-Driven Businesses Set The Pace For Global Growth, Forrester Research, Inc. [10/19/18](#)<sup>1</sup> (Les entreprises axées sur l'information imposent le rythme de la croissance mondiale)

**D'ici à 2021, plus de 75 % des entreprises de taille moyenne ou grande auront adopté un environnement multi cloud et/ou une stratégie informatique hybride**

Prédictions 2019 : Increasing Reliance on Cloud Computing Transforms IT and Business Practices, Gartner [12/13/18](#)<sup>2</sup> (Le recours croissant à l'informatique du Cloud transforme l'informatique et les pratiques professionnelles)

**D'ici à l'année 2022, les services cloud publics seront essentiels pour 90 % de l'innovation en entreprise**

Prédictions 2019 : Increasing Reliance on Cloud Computing Transforms IT and Business Practices, Gartner [12/13/18](#)<sup>3</sup> (Le recours croissant à l'informatique du Cloud transforme l'informatique et les pratiques professionnelles)

**Que signifie être axé sur l'information ?**

Les données sont-elles le nouveau pétrole ? La ressource la plus précieuse du monde ? Le carburant qui alimente l'entreprise digitale ? Quelle que soit la métaphore que vous choisirez, la vérité est que les entreprises qui ont le mieux réussi dans le monde ont développé leur activité à partir d'informations précieuses qu'elles ont tirées de leurs données. Les cinq géants de la technologie, que les investisseurs connaissent sous l'acronyme FAANG (Facebook, Amazon, Apple, Netflix, Google) utilisent les données pour mener des stratégies commerciales qui surprennent et enchantent leurs clients au quotidien, et le marché les a largement récompensés. En mars 2019, la capitalisation boursière de ces sociétés se montait à [3100 milliards de dollars](#)<sup>4</sup>.

Être axé sur les données ne se borne pas au seul secteur de la technologie grand public. Les banques s'appuient sur les données pour lutter contre les infractions financières et réduire le taux de churn . L'industrie automobile investit massivement dans les données et l'intelligence artificielle pour privilégier des véhicules autonomes. Les entreprises de fabrication exploitent les données et les analysent pour prévoir et éviter l'arrêt des machines, et permettre ainsi une mise sur le marché rapide de nouveaux produits.

Articuler les données et l'analyse autour de votre stratégie d'entreprise peut paraître évident, mais c'est plus facile à dire qu'à faire. La transformation digitale exige un changement de la culture d'entreprise, ce qui implique de repenser la manière dont les personnes utilisent les données, l'endroit où elles sont stockées et les outils à utiliser pour y accéder. En fait, la difficulté de la tâche est telle que de nombreuses entreprises finissent par faire machine arrière ou renoncent définitivement à leurs stratégies de transformation.

Ce livre blanc étudie certains des défis associés à une stratégie axée sur les données et comment un nouveau type de plate-forme de données (un Enterprise Data Cloud) peut répondre à ces problèmes.

**Pourquoi les sociétés rencontrent des difficultés avec leurs données**

Les cas d'usage d'entreprise les plus importants et conduisant à des changements profonds (ceux qui sont réalisés à partir d'informations uniques ne pouvant être issues que d'une bonne stratégie de données), exigent plusieurs infrastructures d'analyse, outils de data science et algorithmes de machine learning exploitant un ensemble varié de données communes. C'est dans ces charges de travail que les entreprises innovantes sont en mesure de dégager énormément de valeur à partir de leurs données et de dominer leur secteur. C'est plus facile à dire qu'à faire, cependant, car en réalité, les entreprises traditionnelles ne fonctionnent pas ainsi. Au lieu de cela :

- Les traitements d'analyse s'exécutent de manière indépendante dans des silos. Même les entrepôts de données du cloud et les outils de data science les plus récents ne sont pas conçus pour fonctionner ensemble.
- Les données sont partout, dans les data centers, les clouds publics et à la périphérie, et les entreprises ne disposent d'aucun moyen pratique pour exécuter des analyses ou des algorithmes de machine learning sur leurs données.
- Les analyses en silo et les données dispersées en plusieurs endroits rendent quasi impossible une approche cohérente de la confidentialité des données ou de la protection IP. Au mieux, Elles exigent des contrôles onéreux qui limitent la productivité de l'entreprise et accroissent les coûts. .
- Les outils d'analyse et de machine learning basés sur le cloud, qu'ils soient traditionnels ou même modernes, exigent l'utilisation de volumes de stockage et d'algorithmes propriétaires en source fermée, ce qui signifie que les entreprises doivent céder leurs actifs les plus précieux (leurs données) afin d'en extraire des informations ayant du sens.

### Présentation d'un Enterprise Data Cloud

Développer une approche axée sur les données exige la capacité d'appliquer différentes disciplines d'analyse aux données, où qu'elles se trouvent, d'être en mesure de traiter et de lire en continu et en temps réel des données issues de plusieurs points de l'extrémité à la périphérie tout en prévoyant des résultats clés et en appliquant le machine learning au même jeu de données, d'être en mesure de tirer profit de l'infrastructure d'un cloud public pour son agilité, son adaptabilité et de plus en plus, pour la gravité des données et de pouvoir faire le tout sur une plate-forme ouverte sur laquelle la sécurité et la gouvernance des données s'appliquent partout où les données se trouvent et où les analyses sont exécutées. C'est cela qu'on appelle un Enterprise Data Cloud dans le secteur.



Un Enterprise Data Cloud représente une nouvelle architecture de gestion des données, avec des similitudes avec des plates-formes de données modernes telles que Cloudera Enterprise et Hortonworks Data Platform (Cloudera et Hortonworks ont fusionné en janvier 2019). Les caractéristiques clés d'un Enterprise Data Cloud incluent la prise en charge des points suivants :

#### Analyse multifonction :

Les unités d'exploitation doivent être en mesure d'identifier rapidement les données intéressantes, de les assimiler et les transformer, et de déployer des infrastructures d'analyse ayant du sens pour l'entreprise. Offrir une réponse aux problèmes des entreprises du monde réel exige généralement l'application de plusieurs fonctions d'analyse fonctionnant ensemble sur les mêmes données. Par exemple, le développement de véhicules autonomes exige une application de diffusion de données en continu en temps réel en provenance des capteurs du véhicule (la périphérie), l'entreposage de données pour stocker, analyser et rechercher l'automatisation potentielle, puis le développement et le déploiement d'algorithmes d'apprentissage machine en retour vers le véhicule, pour offrir une fonctionnalité de conduite autonome et prendre des décisions intelligentes (IA).

#### Cloud hybride et multi-cloud :

Pour offrir la flexibilité nécessaire au déploiement des cas d'usage dans l'environnement convenant le mieux aux besoins d'une organisation. Un Enterprise Data Cloud doit fonctionner de la même manière sur site et en dehors des locaux et prendre en charge aussi bien les grands clouds publics que les clouds privés. Les données ont un poids. Leur collecte est onéreuse et elles sont difficiles à déplacer, et il est donc préférable de les analyser dans leur environnement naturel habituel. Il faut donc amener les ressources de calcul appropriées jusqu'aux données, qu'elles se trouvent dans un magasin d'objets de

cloud public ou dans le data center d'une organisation.

Dans la plupart des entreprises, les données critiques se trouvent en plusieurs endroits. Par exemple, une société peut choisir de conserver sur site les enregistrements sensibles relatifs à la conformité et, dans le même temps, exécuter des applications de preuves de concept en utilisant des jeux de données rendus anonymes dans le cloud public. Les architectures de cloud hybride et multi-cloud offrent une meilleure maîtrise, en laissant à l'entreprise le choix de l'endroit, du moment et du type de charges de travail d'analyse à exécuter.

Autre aspect important de la séparation du calcul et du stockage : les unités d'exploitation ont désormais le choix entre coût et performance. Par exemple, une analyse critique peut avoir ponctuellement besoin de davantage de ressources de calcul du cloud public pendant un bref moment. Dans ce cas, il faut mettre l'accent sur la rapidité plutôt que sur le coût. D'autres charges de travail analytiques (tableau de bord des programmes, par exemple), peuvent faire l'objet d'une optimisation des coûts en s'exécutant sur site ou dans un cloud privé.

### Sécurité, gouvernance et métadonnées unifiées :

Dans de nombreux environnements, un entrepôt de données ou un outil de machine learning possède ses propres catalogue et stratégie de données propriétaires assurant la sécurité des données et contrôlant l'accès à ces dernières. Pour pouvoir connecter entre eux deux systèmes ou plus liés à un cas d'usage important, l'équipe en charge des données doit extraire les résultats d'un système et réinitialiser le schéma, la sécurité et les stratégies d'accès, pour qu'ils soient conformes à tous les autres systèmes. Cela prend du temps, représente des risques et consomme des ressources qui seraient mieux utilisées ailleurs. Le défi est encore amplifié lorsque les organisations adoptent des infrastructures hybrides.

Une Enterprise Data Cloud offre aux données et aux métadonnées une structure commune qui permet à plusieurs fonctions d'analyse de fonctionner ensemble et d'exploiter des données issues à la fois d'environnements de cloud public et sur site. Cela inclut notamment une sécurité et une gouvernance unifiées appliquées aux données où qu'elles se trouvent pour garantir que l'organisation respecte les contraintes réglementaires. Les sociétés doivent être en mesure d'assurer la sécurité de leurs données et d'appliquer les stratégies aux personnes qui y accèdent, en fonction des circonstances. Elles doivent être en mesure d'assurer un suivi des accès sur la durée, pour être sûres que les personnes agissent dans la légalité et pour répondre aux questions sur la manière et le moment auxquels les données ont été utilisées à des fins d'audit.

Le fait d'avoir accès à plusieurs infrastructures d'analyse, qui collaborent sur différentes infrastructures et systèmes de stockage, exige une gouvernance partagée et des modèles de sécurité cohérents, un moyen unique d'appliquer une stratégie (mise en œuvre dans une infrastructure, indépendamment du système d'analyse) établissant qui peut toucher les données et ce qu'ils peuvent en faire. Les règles appliquées aux données; la sécurité et la gouvernance qui restent liées aux données lors de leur transfert d'une infrastructure à l'autre. Il s'agit de l'unique façon d'être sûr que ces politiques sont uniformément appliquées.

### Plate-forme ouverte :

Les entreprises souhaitent éviter les erreurs du passé et ne pas dépendre d'un unique prestataire informatique. Elles souhaitent assumer la responsabilité de leurs données et des informations qu'elles en tirent. Elles ne veulent pas que leurs décisions analytiques soient limitées à cause de leur environnement d'exploitation. Le terme plate-forme ouverte est synonyme de logiciel open source, d'architectures de calcul et de magasins de données ouverts tels qu'Amazon S3 et Azure Data Lake Storage. Il signifie également intégration et écosystèmes partenaires ouverts. En évitant la dépendance, les entreprises ont le contrôle de leurs données et de leur avenir.

### Cas d'utilisation

Vous trouverez ci-dessous des exemples d'entreprises qui ont appliqué avec succès les principes de l'Enterprise Data Cloud au sein de leur organisation, pour libérer des données précieuses et initier la transformation de leur entreprise.

#### IoT et données connectées

Octo Telematics fournit des solutions télématiques et d'analyse de données pour le secteur de l'assurance auto. La société s'appuie sur les données (11 milliards de points de données générés par cinq millions de voitures connectées parcourant plus de 170 milliards de miles accumulés ensemble), pour changer la façon dont les assureurs évaluent les risques, délivrent des services en cas d'accident et de sinistre, et gèrent leurs relations clients.

La capacité d'utiliser de manière flexible des fournisseurs majeurs du cloud tels qu'AWS, la plate-forme Google Cloud et Microsoft Azure implique que les clients Octo soient en mesure de traiter des projets transitoires mais exigeants en ressources de calcul, et notamment le test de nouveaux algorithmes de tarification ou le développement de modèle de risque sur la base de l'utilisation.

Depuis qu'Octo est passé à une plate-forme de données from the Edge to AI dans une architecture de cloud hybride, l'entreprise a réduit ses délais de mise sur le marché de nouveaux produits de plus de 50 %. La société a également amélioré de manière significative l'expérience client, en améliorant l'efficacité du traitement des sinistres et en permettant une détection plus précise et plus rapide des fraudes.

#### Transformation digitale

Deutsche Telekom est un prestataire de service de télécommunications européen leader offrant des services à plus de 150 millions de clients au niveau mondial. La société a développé sa plate-forme de données moderne sur le cloud afin d'héberger des jeux de données importants, en diffusant des données en continu et en fournissant un environnement capable de prendre en charge l'apprentissage machine, l'analyse rapide, et la fiabilité de niveau d'entreprise. Une des composantes clés du succès réside dans une plate-forme et un écosystème complètement gouvernés dans lesquels les utilisateurs ont la possibilité d'échanger et d'analyser des données et de développer plus facilement, en toute sécurité, des applications multifonctions axées sur les données.

La plate-forme de données moderne Deutsche Telekom génère des résultats tangibles dans l'entreprise :

- **Fraude réduite** : Deutsche Telekom exploite des requêtes de traitement de données haute vitesse et interactives à grande échelle pour améliorer la qualité du réseau et détecter les activités frauduleuses en temps réel. Deutsche Telekom estime que cette visibilité permet de réduire les pertes de revenus provoquées par les activités frauduleuses de 10 à 20 %.
- **Davantage de satisfaction client** : ils ont une meilleure connaissance des besoins et des souhaits du client. Deutsche Telekom a mis en place une vision unique des clients d'entreprise, ce qui lui a permis de lancer des campagnes mieux ciblées, générant des revenus de dizaines de millions d'euros tout en réduisant le nombre de résiliations de clients de cinq à 10 %.
- **Une efficacité opérationnelle améliorée** : L'entreprise est plus rapide avec la plate-forme de données moderne en place, et l'efficacité opérationnelle d'ensemble s'est donc améliorée de 50 %.

### Cas d'utilisation Des soins prédictifs et préventifs

Clearsense permet aux prestataires de soins de santé de générer une valeur mesurable à partir de leurs données. Cette société a développé un écosystème de données de soins de santé sur le cloud sécurisé, qui absorbe rapidement les données émanant d'un ensemble de sources pour offrir une vue en temps réel de n'importe quel environnement de soins de santé. Le produit phare de la société, Inception™, a été spécialement conçu pour les besoins cliniques, financiers et opérationnels des prestataires de soins de santé. Inception identifie efficacement la détérioration de l'état de santé d'un patient, assure le suivi des indicateurs clés de performance pour gérer les coûts et exploiter les données afin d'améliorer l'efficacité opérationnelle.

Les fonctionnalités from the Edge to AI inhérentes à Inception ont permis le développement d'un modèle de surveillance clinique capable de détecter si un patient est susceptible de faire un arrêt cardiaque dans un délai de 4 à 12 heures et d'avertir les médecins en conséquence. La société utilise également des algorithmes de protection pour réduire le nombre de cas d'infection

### Récapitulatif

Ces clients donnent une idée de ce qu'une entreprise axée sur l'information est capable de faire. Les sociétés doivent dans un premier temps réfléchir aux besoins de l'entreprise, puis concevoir des cas d'utilisation axés sur l'information alimentant la stratégie d'entreprise. Une fois que les cas d'utilisation ont été conçus, vous pouvez commencer à réfléchir à vos besoins en matière de données et d'analyse, et enfin, à l'environnement dans lequel votre architecture fonctionnera. Un Enterprise Data Cloud permet aux sociétés d'extraire des informations ayant du sens à partir des données et d'alimenter l'innovation.

Cloudera offre l'ensemble des disciplines d'analyse de données nécessaires pour résoudre les cas d'usage les plus exigeants. Les disciplines de données from the Edge to AI (analyse à la périphérie, analyse de diffusion en continu, ingénierie des données, entreposage des données, analyse opérationnelles, data science et machine learning) fonctionnent toutes ensemble, en toute sécurité, dans les environnements de votre choix (data center, multi-cloud public et cloud hybride).

Nos offres actuelles (Cloudera Enterprise, Hortonworks Data Platform et Cloudera Data Flow) mettent à disposition les fonctionnalités indispensables d'un Enterprise Data Cloud: (cloud hybride et multi-cloud public), l'analyse multifonctions, des services de sécurité et de gouvernance partagés, et des plates-formes open source avec la puissance de calcul et le volume de stockage choisis.

**About Cloudera**

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at [cloudera.com](https://cloudera.com)

**Connect with Cloudera**

About Cloudera:

[cloudera.com/more/about.html](https://cloudera.com/more/about.html)

Read our VISION blog:

[vision.cloudera.com](https://vision.cloudera.com)

and Engineering blog:

[blog.cloudera.com](https://blog.cloudera.com)

Follow us on Twitter:

[twitter.com/cloudera](https://twitter.com/cloudera)

Visit us on Facebook:

[facebook.com/cloudera](https://facebook.com/cloudera)

See us on YouTube:

[youtube.com/user/clouderahadoop](https://youtube.com/user/clouderahadoop)

Join the Cloudera Community:

[community.cloudera.com](https://community.cloudera.com)

Read about our customers' successes:

[cloudera.com/more/customers.html](https://cloudera.com/more/customers.html)

La plate-forme Cloudera Data Platform (CDP) (disponible à partir du second semestre 2019) associera le meilleur des technologies Hortonworks et Cloudera pour proposer le premier Enterprise Data Cloud du secteur. CDP :

- Donne au client la possibilité d'étendre les déploiements actuels de Cloudera et Hortonworks grâce à des services cloud natifs sur AWS et Azure
- Offre un éventail complet d'analyse de gestion de données open source et multifonctions ayant l'agilité, l'élasticité et la simplicité d'utilisation d'une expérience similaire à celle de cloud public
- Offre un plan de commande unique permettant de gérer les infrastructures, les données et les charges de travail d'analyse dans des environnements hybrides et multi-cloud
- Étend les services partagés SDX afin de protéger la confidentialité des données, la conformité à la réglementation et de gérer les menaces de cybersécurité dans l'ensemble des environnements du cloud
- Permet d'être 100 % open source, et de prendre en charge les objectifs d'entreprise visant à éviter la dépendance vis-à-vis d'un seul fournisseur et à accélérer l'innovation dans l'entreprise

Pour plus de détails sur la disponibilité d'Enterprise Data Cloud et de CDP, merci de visiter [cloudera.com](https://cloudera.com)

**Sources**

<sup>1</sup> <https://www.forrester.com/report/InsightsDriven+Businesses+Set+The+Pace+For+Global+Growth/-/E-RES130848>

<sup>2</sup> <https://www.gartner.com/en/documents/3895580>

<sup>3</sup> <https://www.gartner.com/en/documents/3895580>

<sup>4</sup> <https://www.investopedia.com/terms/f/faang-stocks.asp>