



# Trois transformations qui révolutionnent les environnements de données modernes

Implications pour les responsables IT

# Sommaire

Un problème qui se complexifie .....	3
1. Pensez en termes de pipelines, plus en termes de compartiments .....	4
2. Déplacement des données en fonction des besoins d'utilisation .....	6
3. Passer du statut de protecteur de données à celui de mentor .....	8
Le problème est plus complexe, mais pas impossible à résoudre .....	10

## Un problème qui se complexifie

Fournir aux entreprises des données fiables afin d'optimiser la prise de décision est une démarche qui n'a pas fondamentalement changé depuis des décennies. Malgré des avancées technologiques massives et de nouvelles stratégies, les services IT qui gèrent aujourd'hui les infrastructures de données ont une mission globalement identique : assurer la mobilité des données dès leur création, ainsi que leur accessibilité et leur clarté auprès des décisionnaires au moment où ils en ont besoin.

Toutefois, si l'objectif est resté le même, les obstacles liés à la création et à la gestion d'une source d'informations exploitables valide au sein d'une entreprise sont beaucoup plus difficiles à surmonter.

Les nouvelles sources de données qui ont généré des volumes de données de sortie sans précédent, souvent avec peu (voire pas) de structure, sont sans doute la principale difficulté qui a émergé au cours de ces dernières années dans les environnements de données modernes. Entre les parcours de navigation, les journaux de serveur, les sources de médias sociaux et les relevés des capteurs et des machines, le déferlement des données provenant de ces canaux est littéralement écrasant. Du point de vue économique et des performances, les entrepôts de données d'entreprise traditionnels n'ont tout simplement pas les moyens de s'adapter à ce flot de données.

C'est pourquoi il a fallu entièrement repenser les stratégies d'analyse et de capture de données, et créer une nouvelle génération de solutions de stockage de données assurant une capture sans schéma, l'évolutivité du matériel et le rapprochement de la capacité de calcul des banques de données, voire leur superposition.

Malgré leur jeune âge comparé aux bases de données relationnelles, ces nouvelles solutions non relationnelles ont beaucoup gagné en popularité au cours de ces dernières années et ont évolué rapidement afin de répondre aux besoins des multinationales les plus complexes. Bien que ces solutions aient essentiellement vu le jour afin de pallier les limitations des infrastructures des entrepôts de données d'entreprise, elles représentent néanmoins un écosystème de données beaucoup plus complexe à gérer pour les services IT.

La disponibilité des données dans les applications cloud s'ajoute à la difficulté qu'ils ont à garantir l'intégrité permanente des environnements de données. De nombreuses entreprises utilisent des applications telles que Google Analytics, Salesforce, Netsuite, Zendesk et d'autres au sein de leur infrastructure. Les données qu'elles génèrent sont essentielles au reporting des entreprises. L'intégration des données provenant de ces solutions cloud et leur facilité d'accès sont devenues des exigences standard pour les services IT.

Comme les entrepôts de données d'entreprise traditionnels ne peuvent plus constituer le seul point de destination des données, la question de savoir quand, où, comment et si les entreprises vont intégrer des données d'applications cloud dans leur environnement de données fait l'objet de discussions continues et animées.

Enfin, à mesure que l'analytique en libre-service s'impose comme la nouvelle norme pour les entreprises de toutes tailles, de plus en plus d'utilisateurs non techniques (sans formation IT ou sur l'exploitation des données) font des découvertes dans leurs données, exécutent leur propre reporting et se chargent même parfois des tâches de préparation ou d'analyse avancée. Les entreprises qui adoptent cette approche constatent souvent une réduction considérable (voir une

élimination complète) des responsabilités des services IT en matière de création d'analyses. Bien que cette mutation soit essentielle à la réussite globale d'une société guidée par les données, cette situation fait naître de nouveaux défis pour les services IT qui doivent permettre un accès plus élargi aux données. Tout cela s'ajoute au besoin d'adapter les technologies aux besoins de l'entreprise, mais aussi aux attentes en matière de sécurité et de gouvernance.

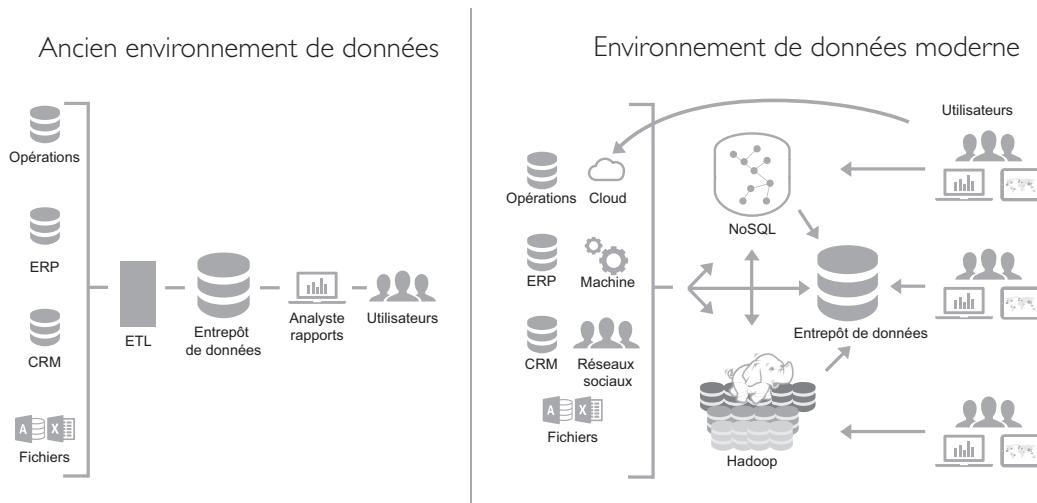
Pour relever ces défis, beaucoup de services IT adoptent très vite de nouvelles technologies et stratégies sans évaluer l'impact potentiel sur la façon dont ils doivent envisager la gestion des données « de leur création à leur utilisation ». Les solutions de Big Data, l'intégration des données cloud et l'analytique en libre-service sont toutes des réponses possibles aux problèmes technologiques principaux. Toutefois, pour permettre leur déploiement efficace au sein d'une entreprise, la démarche des services IT doit changer.

Ce livre blanc vise à souligner les trois principaux mouvements de pensée qui régissent les environnements de données modernes et que les responsables IT doivent comprendre afin de permettre une prise de décisions guidée par les données au sein de leur entreprise.

## 1. Pensez en termes de pipelines, plus en termes de compartiments

Les entrepôts de données d'entreprise ne sont pas morts. Mais ils ne sont plus seuls. Des clusters Hadoop aux bases de données NoSQL, les entrepôts de données relationnels ne sont plus les seuls emplacements officiels où les données peuvent être hébergées.

Autrement dit, le processus de transfert des données n'a plus besoin d'être centralisé autour d'un seul emplacement. D'ailleurs, aujourd'hui, une plate-forme de données moderne s'apparente davantage au tissu nerveux qui connecterait les hémisphères du « cerveau de données » d'une entreprise. Associez à cela les solutions cloud désormais omniprésentes qui permettent à l'infrastructure et aux services de mettre en place des pipelines et des projets ETL en quelques heures, vous obtenez la recette idéale pour assurer une mobilité inédite des données au sein de l'entreprise.



Malheureusement, comme bon nombre des manuels encore référencés en matière de gestion des données reposent sur le concept dépassé de « compartiment de données unique », nombreux sont les services IT qui passent à côté de cette opportunité.

Pour axer son approche non plus sur les compartiments mais sur les pipelines, les parties prenantes doivent comprendre qu'il n'est pas possible de répondre à toutes les questions de données d'une entreprise à partir d'une seule source de données. Un entrepôt de données d'entreprise n'a pas toujours la capacité suffisante pour offrir le niveau de granularité des données requis pour analyser l'équivalent de cinq années d'e-mails tirés directement des journaux de serveur de messagerie. Par ailleurs, lors de l'exécution de requêtes sur des données de vente, un déploiement Hadoop ne parvient pas toujours à fournir des réponses en moins d'une seconde, délai requis pour proposer des analyses en temps réel. Toutefois, dans le monde réel, pour réaliser un projet de données, nombreuses sont les problématiques métier qui exigent à la fois granularité des données et rapidité d'interrogation d'une ou de plusieurs sources, mais à des périodes et selon des séquences différentes.

Face à ce constat, la question pour les services IT n'est plus de savoir dans quel compartiment les données doivent aller, mais quel est le stade de données requis et comment permettre aux utilisateurs de passer facilement d'un stade à l'autre.

C'est là que les pipelines entrent en jeu, car l'emplacement réel des sources de données peut changer à mesure que le projet passe d'un stade de réalisation à un autre. La mentalité orientée pipeline, qui investit le domaine du workflow dans les entreprises, va également bien au-delà de l'univers traditionnel d'intégration des données.

Comment les utilisateurs trouvent-ils des réponses dans leurs données ? La plupart des services IT savent que des référentiels de données individuels non autorisés sont conservés dans des feuilles de calcul aux quatre coins de l'entreprise. Mais il n'est pas toujours facile de réprimer une gestion locale des données cloisonnée et non gouvernée. Les utilisateurs ont souvent l'impression que la seule solution pratique pour remplir leurs objectifs est de contourner l'environnement de données d'entreprise. C'est donc ce qu'ils font.

C'est uniquement lorsque les services IT comprennent le cycle complet que suivent ces utilisateurs pour rechercher des données, les rendre exploitables, les analyser et les présenter, et qu'ils décident de faciliter le processus (tout en prônant la gouvernance) qu'une mentalité orientée pipeline est envisageable. Le fait d'accepter qu'assurer la mobilité des données, qu'il s'agisse des lots de données traditionnels ou des données issues des requêtes ad hoc des utilisateurs, est plus important que de bâtir un bunker fortifié autour des données représente un moment charnière pour les groupes technologiques.

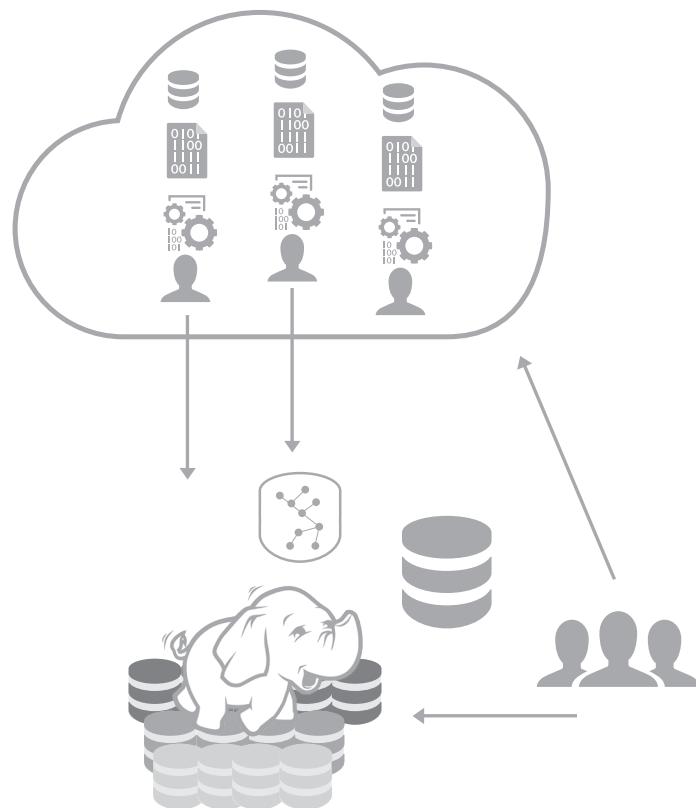


Dans le cadre d'une stratégie internationale, il est de plus en plus important que tous les acteurs puissent accéder à ces informations et les comprendre. Grâce à notre entrepôt de données logique, que l'on combine à Tableau, les utilisateurs peuvent facilement collaborer et accéder à toutes ces informations.

— TIM NALL, DSI, BROWN-FORMAN

## 2. Déplacement des données en fonction des besoins d'utilisation

Les données des applications cloud ne sont pas toujours destinées à arriver directement dans les entrepôts de données d'entreprise. La décision dépend toujours des besoins de l'entreprise. Ou, pour être plus clair, la décision dépend de ce qui est nécessaire pour rendre les données utiles à l'entreprise.



Une fausse idée couramment répandue au sein des services IT, même chez ceux qui ont adopté une approche non centrée sur un compartiment unique de données, est que les données d'applications Web doivent être automatiquement intégrées dans l'entrepôt de données d'entreprise. De prime abord, cela peut paraître logique, étant donné que de nombreuses applications cloud performantes font partie intégrante de l'infrastructure de vente, de marketing et d'assistance. Les données y sont souvent bien structurées, facilement exploitables et accessibles. Le placement de ces ressources dans un entrepôt de données d'entreprise, où le prix par octet est le plus élevé, offre sembla-t-il un retour sur investissement convaincant.

Toutefois, les questions clés que les services IT doivent se poser sont les suivantes :

- Les données cloud sont-elles immédiatement prêtes pour l'analyse ? En d'autres termes : sont-elles exploitables et cela vaut-il la peine de les déplacer ?
- Les données cloud en question n'ont-elles un intérêt réel que si elles sont combinées avec d'autres données ?

En fonction de ces considérations, l'emplacement auquel les services IT choisissent d'envoyer leurs données cloud (dans la mesure où ils choisissent de les envoyer sur site) peut varier d'une entreprise et d'une application à une autre.

Tout service IT responsable d'un déploiement Salesforce.com de taille respectable vous dira immédiatement qu'autoriser des enregistrements d'opportunités non dédupliqués dans l'entrepôt de données d'entreprise conduit tout droit au désastre. Du versement incorrect des commissions aux projections de chiffre d'affaires nettement surestimées, ces problèmes ne représentent que la partie visible de l'iceberg lorsque les données de vente ne sont pas bien nettoyées.

**“** En mettant les données à la disposition des utilisateurs finaux, nous leur donnons la possibilité de créer rapidement les rapports de base dont ils ont besoin. Comme ils sont en première ligne, ce sont les plus à même d'identifier rapidement les améliorations éventuelles à apporter à un produit ou à une application.

— SHARON GRAVES, ÉVANGÉLISTE DES DONNÉES D'ENTREPRISE, GODADDY

De même, si l'exactitude des enregistrements n'est pas garantie, les données de site Web et de marketing des sources cloud comme Google Analytics, Eloqua et Marketo peuvent nuire à la capacité d'une entreprise à suivre la génération de clients potentiels et à calculer le coût d'acquisition des clients.

Ce problème d'exactitude devient encore plus sérieux en cas de fusion des sources. Il se pose souvent pour les entreprises qui doivent brosser un tableau complet de leur entonnoir de conversion, du premier clic vers le site Web à la décision d'achat. Bien que presque tous les services IT sachent combien il est important de disposer de données exactes, que ce soit dans le cloud ou ailleurs, ils oublient souvent de penser à l'endroit où elles arrivent en premier lorsqu'elles sont extraites d'une application Web.

Les services IT dont l'environnement de données est basé sur plusieurs compartiments ou sur des pipelines, envisagent l'intégration des données cloud en fonction de la valeur de ces ressources pour l'entreprise au moment de leur arrivée sur site. Si le niveau de précision d'une source de données cloud est satisfaisant et qu'elle présente une très grande valeur pour l'entreprise à l'arrivée, les services IT sont bien avisés d'envoyer les données là où elles sont le plus rapidement accessibles (entrepôt de données d'entreprise relationnel).

Toutefois, s'il est question de données d'applications Web nécessitant un traitement important et/ou complexe avant de pouvoir être considérées comme sûres pour l'entreprise, les services IT peuvent tirer profit d'environnements à forte capacité de calcul et faible coût par octet, tels que Hadoop. Avec cette seconde approche, il est possible d'optimiser les ressources en termes de nettoyage et de transformation des données, et ce sans ralentir l'entrepôt de données d'entreprise. Les services IT, en collaboration avec les divisions métier, peuvent ensuite décider si les données nettoyées doivent être transférées vers un entrepôt de données d'entreprise et/ou accessibles directement au sein d'un environnement plus vaste.

Toutefois, il n'est pas toujours nécessaire de déplacer les données des applications cloud. Il est de plus en plus courant que les applications Web fournissent des points d'accès faciles à leurs référentiels

backend. Les utilisateurs métier peuvent ainsi se servir des outils de reporting et d'analyse en libre-service pour faire leurs propres recherches sur des données disponibles en direct. Pour les services IT qui souhaitent conserver une couche intermédiaire d'autorisation et de gouvernance dans ces scénarios, certaines des solutions d'analytique en libre-service les plus flexibles proposent également un scénario de connexion proxy. Cette approche répond potentiellement à tous leurs besoins, qu'il s'agisse des accès utilisateur de base ou de logiques applicatives stratégiques. Dans ce contexte, la première question à se poser concernant l'intégration des données d'applications cloud ne concerne pas leur emplacement, mais la nécessité ou non de les déplacer. Est-il vraiment nécessaire de déplacer ces données (nettoyage, valeur ajoutée, etc.) ou peuvent-elles rester là où elles sont et peuvent-elles être intégrées au niveau des utilisateurs métier ?

**“** Nous avons choisi Tableau parce que c'est une solution robuste avec une feuille de route claire, en phase avec notre propre vision à long terme, qui nous aide à atteindre nos objectifs à chaque étape de notre stratégie.

— STEVEN JOHN, DSI, AMERIPRIDE

### 3. Passer du statut de protecteur de données à celui de mentor

L'approche visant à proposer l'analytique en libre-service au sein des entreprises est bénéfique pour les services IT. Et ceux qui sont les premiers à l'adopter mettent toutes les chances de leur côté pour réussir.

Tout comme l'environnement de données moderne a évolué en intégrant d'autres éléments que des banques de données relationnelles, l'environnement analytique s'est enrichi de nouveaux outils particulièrement intéressants pour les utilisateurs métier. Cette avancée permet aux entreprises de s'éloigner du processus traditionnel selon lequel toutes les fonctions analytiques sont réservées à une poignée d'experts hautement qualifiés.

Un partenariat qui fonctionne

#### Rôle informatique

Sécurité

Architecture de données

Évolutivité

Formation

Centre des opérations

Plate-forme Intranet

#### Rôle de l'entreprise

Travail analytique créatif

Promotion de l'acquisition des données

Partage de l'expertise

Demande d'aide en cas de besoin

Promotion

Catalyseur d'action

Habilitation



Exécution

Par conséquent, du point de vue des données et de l'analytique, les services IT visionnaires gagnent le statut de conseiller avisé et précurseur au sein de leur entreprise et redéfinissent leur mode d'interaction avec les équipes métier afin de résoudre les difficultés liées aux données. Ce passage du statut de « protecteur des données » à celui de « mentor » est peut-être la transformation la plus significative que connaissent les environnements de données modernes. Ce facteur est déterminant dans la réussite ou l'échec des entreprises en matière de gestion des données.

La redéfinition de cette relation entre les services IT et les utilisateurs métier implique l'acceptation de la gouvernance des données. Généralement, à moins qu'une autorisation spécifique leur ait été accordée, l'accès des utilisateurs métier aux données était systématiquement refusé. Ce nouveau modèle suppose que tout utilisateur, où qu'il soit dans l'entreprise, puisse accéder aux données tant qu'il respecte les règles de conformité.

L'impact de ce changement d'attitude est colossal. Il s'illustre par la façon dont les utilisateurs métier se mettent à chercher de nouvelles données pour traiter de nouveaux problèmes. Leur motivation est fonction de la facilité avec laquelle ils peuvent accéder à ces nouvelles données. Lorsqu'ils disposent d'une solution d'analytique en libre-service simple d'utilisation et qu'ils savent que leur service IT est en faveur d'un accès étendu aux données, ils peuvent devenir des acteurs du changement. Et ce changement est un signe que la société se rapproche du statut d'entreprise guidée par les données.

Toutefois, il ne suffit pas d'ouvrir l'accès aux données et de déployer de nouveaux logiciels pour résoudre le problème de fond, à savoir s'assurer que les entreprises disposent des données dont elles ont besoin pour améliorer la prise de décision. Les services IT qui ont réussi à transformer leur entreprise en moteur de décisions basées sur les données ont atteint cet objectif en endossant le rôle de mentor en matière de données. Souvent, les membres de l'équipe IT qui jusque là contrôlaient le processus analytique deviennent des conseillers pour le reste de l'entreprise. Ils ont les compétences requises pour former les utilisateurs métier sur des fonctions essentielles telles que l'analyse des données et le reporting.

Cette transmission de connaissances accroît considérablement la puissance de reporting analytique de toute l'entreprise. Elle lui permet également de réaliser un retour sur investissement plus élevé et plus rapide sur les solutions d'analytique en libre-service. Mais surtout, l'utilisation des données pour mieux comprendre l'entreprise dans son ensemble devient partie intégrante de son modèle d'exploitation.

## Le problème est plus complexe, mais pas impossible à résoudre

Entre la gestion de nouvelles solutions de traitement et de stockage des données et l'évaluation de la nécessité d'intégrer des données cloud dans l'écosystème, du moment adéquat pour le faire et de l'emplacement auquel les envoyer, les responsabilités des services IT en matière de gestion des données d'entreprise sont de plus en plus complexes. Si vous ajoutez à cela l'avènement de l'analytique en libre-service, l'impact de toutes ces transformations est tel un raz-de-marée dans la manière dont les entreprises gèrent les données.

Elles ont donc besoin que leur service IT leur montre le chemin. La plupart des incertitudes que connaissent les services IT sont essentiellement dues au fait que leur rôle n'est plus d'intégrer de nouvelles technologies dans les processus métier existants. Le Big Data, les technologies cloud et l'introduction d'un modèle d'analytique en libre-service représentent des changements radicaux dans la façon dont ils doivent envisager les environnements de données modernes.

Pour les entreprises qui ont déjà commencé à adopter cette nouvelle approche des données et de l'analytique pour s'adapter à ces transformations, si les obstacles sont bien réels, les découvertes le sont également. Le changement n'est simple ni pour les utilisateurs métier, ni pour les services IT. Lors de la redéfinition de la relation entre ces deux groupes, des résistances se feront sentir dans chaque camp. Mais il y a également ceux qui accueilleront l'idée du changement avec enthousiasme.

À mesure que les obstacles et les processus rigides qui régissent l'accès aux données d'entreprise évolueront, les services IT trouveront aux quatre coins de l'entreprise des alliés inattendus prêts à défendre cette nouvelle approche. Les responsables IT les plus avisés proposeront à ces nouveaux alliés de collaborer avec leurs experts en matière de données, qui feront office de mentors. Les alliés en apprendront davantage sur l'écosystème des données en évolution constante, tandis que les services IT découvriront les problèmes que les utilisateurs métier essaient de résoudre. Ces premières étapes jettent les fondations d'une culture de l'analyse et de la découverte de données, dans laquelle les décisions seront fondées sur la fiabilité et la précision des environnements de données modernes.

## À propos de Tableau

Tableau vous aide à voir et comprendre vos données, peu importe leur volume, le canal qui les produit ou la base de données dans laquelle elles sont stockées. Vous pouvez rapidement vous connecter à vos données, les fusionner et les visualiser facilement aussi bien sur ordinateur que sur tablette. Créez et publiez des tableaux de bord marketing avec des mises à jour de données automatiques et partagez des informations en temps réel avec vos collègues, équipes, dirigeants, partenaires ou clients. Aucune connaissance en programmation n'est requise.

## Ressources supplémentaires

Télécharger la version d'évaluation gratuite

+ [Tableau Online](#)

+ [Tableau Desktop](#)

## Livres blancs connexes

[Guide d'évaluation des solutions BI et analytiques modernes](#)

[Le libre-service gouverné et adapté aux besoins](#)

[Redéfinir le rôle de l'IT dans la nouvelle approche de la BI](#)

[Tableau pour l'entreprise : présentation pour les services informatiques](#)

