

# Éviter les pièges et lock-ins des bases de données cloud

Les fonctionnalités essentielles de votre prochaine solution cloud d'analytics

Par  
Steve Sarsfield

*À propos de l'auteur : Steve Sarsfield est un auteur indépendant qui a occupé des fonctions de direction (thought leadership) chez Cambridge Semantics (pour la plateforme AnzoGraph), Talend, Trillium Software et IBM. Il a publié un célèbre blog sur la gouvernance des données, des articles sur Medium.com et un livre intitulé « The Data Governance Imperative ». Il continue de partager ses observations et son avis sur la gouvernance des données et l'analytics.*

## Évaluation de l'exhaustivité de la solution

C'est un fait peu connu : mettre au point un système de base de données complet demande plusieurs années et beaucoup d'heures de travail par personne à une équipe de développement expérimentée. Un SGBD mature et pleinement fonctionnel comprend quant à lui une liste très longue de fonctionnalités difficiles à élaborer, et peu de raccourcis sont envisageables. Lorsque les utilisateurs explorent les premières versions d'une nouvelle base de données SQL, l'exhaustivité qui en fait une solution d'entreprise ne sera atteinte que bien des versions plus tard. La liste qui figure ici à droite présente quelques-unes de ces fonctionnalités de développement à long terme.

Les bases de données cloud ne sont pas différentes. Les bases de données proposées sur le marché étant nombreuses et à différents stades de développement, les services informatiques doivent s'assurer de l'exhaustivité de leur propre système et évaluer le niveau de réponse apportée par la solution aux besoins de leur entreprise.

### Caractéristiques EPIC d'une base de données cloud

- Mise à l'échelle et élasticité
- Sauvegarde
- Gestion des workloads
- Administration et gestion
- Sécurité et chiffrement
- Options de déploiement (SaaS, cloud, conteneurs, sur site)
- Connexions entre écosystèmes
  - Open Source
  - Intégration des données
  - BI et visualisation
  - Formats des fichiers de données
- Optimisation des requêtes

Dans ce livre blanc, je vous présente un modèle pour réaliser cette évaluation. Que vous envisagiez de transférer l'ensemble de votre workload analytique vers un seul et unique cloud public, vers plusieurs clouds ou sur site en utilisant des technologies cloud comme le référentiel d'objets, ou que vous prévoyiez une approche hybride, le moteur analytique que vous choisirez peut avoir une incidence importante sur les coûts et la productivité.

## Un modèle pour choisir les fonctionnalités des bases de données cloud

Dans bon nombre d'entreprises, aucun cadre clair n'a été défini pour la sélection des bases de données. Ce sont souvent les analystes métier qui choisissent les solutions ; et quand leur utilisation s'accroît, elles s'imposent comme plateforme interne. Lorsqu'ils rejoignent une entreprise qui mise déjà sur les outils qu'ils connaissent, les data scientists et les analystes métier apportent leurs connaissances sur les langages, les outils de visualisation et les moteurs d'analyse. Quant aux équipes informatiques et en charge de la gouvernance des données, elles ont tendance à attacher la même importance à des facteurs comme la gouvernance des données, l'efficacité opérationnelle et le contrôle des coûts qu'à la simplicité du déploiement.

Lors de vous évaluez les fonctionnalités et le coût à long terme des plateformes cloud d'analytics, il est fondamental de tenir compte des principaux critères suivants :



## Déploiement

Le déploiement des bases de données cloud varie considérablement, mais les décisions prises lors de la mise en œuvre ont généralement un impact durable sur le long terme. Notez que les options de déploiement que vous choisissez doivent être simples pour tout le monde. Néanmoins, votre plateforme cloud doit répondre aux objectifs de votre entreprise, à savoir limiter les copies inutiles de données, les problèmes de sécurité, la réglementation et la conformité, et les dépassements de coûts liés au cloud computing.

Certains éditeurs font la promotion d'une stratégie « cloud-first » ou « cloud-native » pour leur pile technologique de base de données, ce qui peut généralement être interprété comme une stratégie de développement *100 % cloud*. Avec ce type d'éditeur à vos côtés, vous perdriez la capacité de gérer les workloads à exécuter sur site et les autres options décrites ci-dessous.

Parmi les méthodes de déploiement, on compte :

**Cloud d'éditeur** : de nombreux éditeurs de bases de données SaaS (Software-as-a-Service) se déploient dans un cloud géré par un éditeur, reposant souvent sur un cloud public. Ce sont ces éditeurs qui négocient en votre nom avec les éditeurs de cloud public pour piloter, prendre en charge et sécuriser vos données - et non votre entreprise. Votre entreprise rémunère l'éditeur de la base de données cloud qui ensuite rémunère l'éditeur de cloud. Les données sont désormais sous le contrôle de l'éditeur de la base de données. Si la simplicité de cette approche est un avantage, elle vous lie à l'éditeur, ce qui est un inconvénient.

**Analyse des éditeurs cloud** : une légère variante consiste à utiliser le service SaaS d'un éditeur de cloud public. Des entreprises comme Amazon, Google et Microsoft proposent à la fois l'infrastructure et la base de données. Elles en assurent la gouvernance et le contrôle de la sécurité. Là encore, si la simplicité de la facturation est un avantage, le lock-in avec l'éditeur est encore plus fort. Vous devez vous limiter non pas à quelques déploiements cloud, mais à un seul déploiement cloud.

### À propos du lock-in

Il y a de fortes chances que vos exigences en matière de base de données cloud évoluent au fil du temps. Par conséquent, il est impératif de comprendre ce qu'implique de changer d'éditeur ou de plateforme cloud.

Renseignez-vous auprès de votre éditeur de cloud sur les politiques et procédures de résiliation avant de finaliser votre choix.

**Cloud(s) public(s) contrôlé(s) par l'entreprise** : de nombreuses entreprises qui achètent des bases de données ont mis en place des clouds privés virtuels au sein des clouds publics comme AWS, Azure, Alibaba et Google. Les administrateurs mettent en place des VPN, des contrôles d'accès, des groupes de sécurité, un enregistrement centralisé des accès aux données, et d'autres fonctionnalités de sécurité requises en fonction de leur secteur. Ce phénomène est souvent appelé « Platform-as-a-Service (PaaS) », et certains éditeurs de bases de données optent pour cette stratégie. Grâce au PaaS, le services informatique contrôle davantage la façon dont sont gérées les données et les coûts. Il peut négocier avec les principaux éditeurs de cloud pour obtenir des tarifs plus attractifs, voire profiter de tarifs préférentiels en cas de workloads temporaires. Votre équipe dispose d'un meilleur contrôle sur le type d'instances (vCPU/Mémoire), nécessaire pour compléter le workload analytique.

**Approche hybride sur site et cloud** : conserver certaines données sur site ou exploiter des entrepôts de données publics en ligne à des fins d'analyses se justifie dans certains cas. Pouvoir analyser les données où qu'elles se trouvent se révèle intéressant dans trois cas : si le chargement de l'ensemble de vos données dans le cloud n'est pas pratique ; si il est impossible pour des raisons de conformité légale ou réglementaire ; ou s'il coûte trop cher. La quantité de données à traiter sur site peut rendre impossible leur transfert vers le cloud. La solution pour les administrateurs est de simplifier encore l'approche hybride en donnant au stockage local l'apparence d'un stockage objet sur cloud. ils utilisent pour cela des applications comme Minio ou du matériel spécialisé compatible avec le stockage objet hautes performances (technologies Pure Storage FlashBlade ou HPE Scality RING, par exemple). Ils donnent ainsi aux équipements des centres de données des caractéristiques similaires à celles des entrepôts d'objets cloud, créant deux environnements semblables. Cette approche répond aux prérequis en matière de gouvernance des données lorsque des réglementations ou mandats d'entreprise imposent d'avoir certaines données sur site.

### Pourquoi est-ce si important ?

Étudier ces différentes options de déploiement vous permet de choisir entre la simplicité du SaaS, la flexibilité du PaaS ou la nécessité d'une approche hybride. Les éditeurs de solutions SaaS vous demandent de charger toutes les données dans leur cloud et dans leur format. Mais cela vous enferme dans leur solution et vous empêche de passer facilement à un autre cloud ou de profiter d'une éventuelle solution informatique moins coûteuse. Certains éditeurs n'offrent aucune solution on-premise d'analytics. Si vous en avez quand même besoin, ou si vous devez retirer les workloads sur site à la suite d'évolutions réglementaires, le déploiement SaaS ne sera pas la solution idéale.

Les workloads peuvent nécessiter un environnement à faible latence. Par exemple, si vos big data sont réparties sur des dizaines ou des centaines de nœuds, vous pouvez envisager des nœuds de calcul cloud non standard avec des connexions plus rapides. Les instances Amazon EC2 peuvent varier de « jusqu'à 10 Gbit/s » (et cette vitesse n'est pas garantie) à un débit réseau garanti de 100 Gbit/s. Comme vous pouvez l'imaginer, il peut y avoir des différences considérables en termes de performances analytiques si, pour la même vCPU/mémoire, on opte pour une faible bande passante réseau. Le choix d'un réseau adapté est déterminant pour maintenir les performances. Et certains éditeurs n'offrent pas ce choix.

Enfin, il convient d'examiner attentivement l'émergence *de nouvelles technologies* phares en matière d'infrastructures de déploiement. Si l'une d'elles rend l'analyse encore plus efficace, il faut pouvoir évaluer les coûts pour migrer vos données d'un cloud propriétaire vers celui de la nouvelle solution. Les méthodes de déploiement que vous choisirez auront un impact sur le délai de rentabilité, l'évolutivité, la flexibilité et la sécurité au cours des prochaines années.

## Stockage et workload

Le stockage des données est géré selon différentes manières en fonction des bases de données cloud. Avant de choisir votre solution de stockage et de workload, il faut comprendre les interactions observées entre le data warehouse et le data lake dans de nombreuses grandes entreprises.

Traditionnellement, les data warehouses sont utilisés lorsque la conformité ACID, la prévisibilité et la reproductibilité sont requises. Ils garantissent des métadonnées soigneusement gérées et conformes aux normes établies par l'équipe chargée de la gouvernance des données. Ils sont souvent complétés par des data lakes qui font l'impasse sur certaines des caractéristiques ACID au profit de l'évolutivité et du faible coût. Les data lakes offrent la possibilité d'utiliser un stockage à faible coût, comme Hadoop ou des entrepôts d'objets à l'instar d'Amazon S3, afin d'effectuer des analyses à très grande échelle. Les métadonnées sont gérées de manière plus souple dans un data lake, et les données qui s'y trouvent peuvent être moins structurées que celles du data warehouse.

Tout cela est exact sauf quand votre base de données cloud ne supporte qu'une seule de ces méthodes.

Certaines solutions cloud d'analytics exigent de télécharger toutes les données dans le data warehouse dans un format structuré spécifique et ne proposent aucun data lake, à l'exception parfois d'une pile technologique propre à l'éditeur.

Depuis que les workloads analytiques des entreprises sont uniques, le processus de stockage de l'ensemble des données dans un seul type de base de données peut être erroné. L'équipe de data science a probablement des attentes différentes en matière de performance que les utilisateurs d'un tableau de bord exécutif, par exemple. Les équipes informatiques souhaitent souvent mettre en place une architecture hiérarchisée, c'est-à-dire qu'elles paient plus lorsque le SLA exige une réponse rapide, et moins quand le workload présente un SLA moins critique.

Les données peuvent exister dans de nombreux systèmes de fichiers et dans de nombreux formats. En général, ces formats sont :

- Systèmes de fichiers : Apache Hadoop HDFS, référentiel d'objets cloud ou sur site, Unix NFS
- Formats de fichiers : Texte, CSV, JSON, ORC, Parquet, etc.

### Pourquoi est-ce si important ?

Lorsque vous choisissez votre éditeur de base de données cloud, regardez comment il combine l'échelle et les économies d'un data lake avec la prévisibilité et la reproductibilité d'un data warehouse. Regardez aussi comment votre solution fait la différence entre les workloads. Vérifiez également si elle conserve vos données les plus sollicitées dans des formats optimisés, natifs de la base de données, et celles qui sont rarement consultées dans d'autres formats de type JSON et Parquet ? Ou bien la seule option consiste-t-elle à stocker toutes les données dans le format de la base de données ?

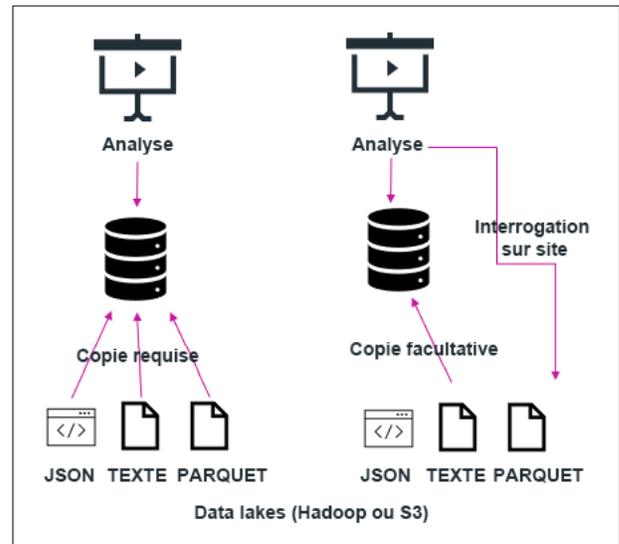


Figure 1 : Votre base de données cloud exige-t-elle que toutes les données soient téléchargées dans les bases de données, ou peut-elle accéder aux données directement dans l'entrepôt d'objets, dans Hadoop ou dans d'autres référentiels de data lakes ?

## Optimisations

Les fonctionnalités des bases de données cloud divergent considérablement lorsque les utilisateurs tentent de résoudre des requêtes à exécution lente. Il est courant que l'optimisation des bases de données cloud soit uniquement basée sur les nœuds. En théorie, si vos requêtes sont trop lentes pour répondre aux besoins de l'entreprise, il suffit d'ajouter des nœuds supplémentaires. La tendance actuelle est de privilégier la disponibilité proposée par les solutions "bon marché" et de gérer les performances en augmentant la puissance de calcul.

Cependant, dans la plupart des entreprises, les workloads analytiques ne sont pas universels. Les performances de votre base de données peuvent être affectées par des rapports trimestriels, une campagne de marketing particulièrement réussie qui génère davantage de données ou encore des requêtes expérimentales ou mal écrites qui impliquent de nombreuses jointures, pour n'en citer que quelques exemples. Au moment de choisir votre base de données cloud, vérifiez les options disponibles pour accélérer les requêtes individuelles. Par exemple :

- **Architecture massivement parallèle** : les administrateurs doivent comprendre comment l'architecture massivement parallèle fonctionne dans votre base de données. Certaines bases de données sont construites sur une base transactionnelle et peuvent être mal dimensionnées pour l'analyse. Et certaines de ces architectures peuvent nécessiter une répartition manuelle et une modification spéciale des requêtes pour tirer parti du cluster.
- **Évolution des nœuds** : face à des contraintes, les administrateurs peuvent-ils faire évoluer les nœuds à volonté pour gérer des workloads exceptionnellement volumineux ? Dans les bases de données modernes, la mise à l'échelle des nœuds est monnaie courante. Votre équipe informatique contrôle-t-elle la taille ou la configuration des nœuds ?
- **Gestion des workloads** : dans un environnement de workloads hautement concurrentiel où des centaines, voire des milliers de requêtes sont exécutées simultanément, votre logiciel vous permet-il d'affecter les ressources de requête (mémoire, unité centrale) à des types de requêtes spécifiques (ou à des groupes d'utilisateurs) ? À titre d'exemple, pouvez-vous vous assurer que le tableau de bord du PDG ne sera pas perturbé par des requêtes erronées ou une science des données ad hoc ? La base de données résout-elle ce problème avec des fonctionnalités d'isolation des workloads qui attribuent aux utilisateurs ou aux workloads une certaine quantité de ressources, empêchant ainsi toute interférence avec d'autres workloads ?
- **Séparation du calcul et du stockage** : une autre technique de partage des données dans un environnement concurrent consiste à séparer le calcul du stockage. Les données sont conservées dans le stockage objet tandis que des groupes de nœuds de calcul sont utilisés pour la simultanéité, la sauvegarde, les tableaux de bord et la science des données. Il se peut que vous en utilisiez une grande quantité à la fin d'un trimestre, et absolument rien pendant les vacances. La séparation du calcul et du stockage est le seul moyen d'obtenir l'élasticité du cloud, nécessaire pour augmenter ou diminuer le calcul en fonction des besoins, à stockage constant.
- **Optimisation des requêtes** : la plupart des bases de données SQL disposent de planificateurs de requêtes qui analysent les requêtes SQL et tentent d'y répondre en sollicitant le moins possible le système. Ceux qui planifient les requêtes vont chercher à limiter les lectures de données et la mémoire nécessaire pour formuler les réponses - mais ils n'y parviennent pas toujours. Vous constaterez qu'accéder aux fonctionnalités d'optimisation des requêtes varie considérablement et que certaines solutions ne permettent pas de le faire. Vous ne pouvez ajouter des nœuds que pour améliorer la vitesse des requêtes. D'autres solutions offrent toutefois des options plus approfondies.

- **Projections/vues matérialisées** : les planificateurs peuvent rechercher des copies des données triées et stockées de manière optimale pour certaines requêtes. Par exemple, supposons qu'un rapport recherche une certaine région pour répondre à une requête. En l'occurrence, le tri préalable de la base de données par région élimine la nécessité de lire l'ensemble de la base de données pour répondre à une requête et accélère donc considérablement l'analyse.

### Pourquoi est-ce si important ?

Toutes les analyses ne sont pas identiques et ne devraient pas être considérées comme telles. Assurez-vous que la base de données que vous choisissez dispose d'options permettant de gérer correctement tous les types de workloads et les attentes en matière de niveau de service. Si seule l'optimisation basée sur les nœuds est proposée, vous risquez de passer à côté de méthodes de maintien des coûts de votre cloud à un niveau bas et d'amélioration des performances des requêtes.

### Profondeur des analyses

Les bases de données cloud n'offrent pas toute la même profondeur d'analyse. Par le passé, les équipes informatiques et un nombre relativement restreint d'analystes métier pouvaient se contenter du langage SQL traditionnel. Mais aujourd'hui, les entreprises étant centrées sur les données, elles doivent répondre à des questions qui dépassent les capacités du langage SQL standard. Il faut donc prendre en compte l'ensemble des groupes susceptibles de questionner les données de votre base de données cloud, ainsi que la profondeur de leurs requêtes.

	Utilisateurs métier	Analystes	Data scientists
Analyses	Indicateurs métier, pipeline	Rapports approfondis, modèles	Perspectives cachées, prévisions, recommandations
Outils	Tableau et autres outils de visualisation	SQL, Tableau et autres outils de visualisation	R, Python, Jupyter Notebook, Apache Zeppelin
Attentes	Réponse en moins de 3 secondes Tableau de bord toujours opérationnel	Intégration d'outils de BI Réponse à la minute Capacité SQL étendue Analyse géospatiale Analyse de séries temporelles Remédiation aux requêtes lentes	Traitement des données Accès aux données brutes Évolutivité Intégration d'outils de ML ou d'algorithmes de ML intégrés

Pour servir ce large éventail d'utilisateurs, vous devez prendre en compte la profondeur des fonctionnalités analytiques offerte dans votre solution cloud de base de données.

Ces types d'analyse comprennent entre autres :

- **Séries temporelles** : fonctionnalités dans lesquelles des fonctions SQL sont intégrées à la base de données pour gérer l'Internet of Things (IoT) et les données de logs enregistrées sur des intervalles de temps déterminés. Prenons pour exemple un cas d'utilisation où un appareil est programmé pour transmettre une mise à jour toutes les 5 minutes. En raison d'une panne de courant, l'appareil cesse d'envoyer des mises à jour pendant plusieurs heures, puis reprend. Votre base de données dispose-t-elle de fonctions de remplissage et d'interpolation pour gérer cette interruption ? Ou devrez-vous traiter cette latence manuellement ? Les fonctionnalités de gestion des séries temporelles varient grandement d'une base de données cloud à l'autre.
- **Données géospatiales** : caractéristiques où les fonctions SQL sont basées sur la latitude/longitude et l'altitude. Pouvez-vous importer des fichiers de forme ? Pouvez-vous utiliser votre base de données cloud pour savoir si un logement se trouve à l'intérieur ou à l'extérieur d'une zone inondable ? Pouvez-vous créer des territoires de vente équitables à partir de vos données ? Pouvez-vous analyser de manière optimale les itinéraires lorsque vous gérez une flotte de camions ? Les caractéristiques géospatiales varient d'une base de données cloud à l'autre. Certaines imposent des coûts de licence supplémentaires pour l'utilisation de données géospatiales.
- **Machine learning** : de nos jours, les analystes ont également besoin d'analyses prédictives. Certaines bases de données cloud permettent d'entraîner, de gérer et de déployer des modèles de machine learning. D'autres encore proposent des licences de machine learning séparées ou facturent par algorithme. Vous aurez besoin de ces fonctionnalités pour effectuer des recherches de motifs, des régressions et des regroupements, entre autres. Examinez attentivement les fonctionnalités de la base de données cloud si vous souhaitez également tirer parti du deep learning.
- **Autres cadres** : souvent, les data scientists apportent leurs propres expériences et outils. Il est donc essentiel de savoir si un accès aux données peut être ouvert tout en prenant en charge des langages type R et Python, ou des interfaces comme Jupyter Notebook. Examinez attentivement le monde de la data science en dehors de SQL et la façon dont il s'aligne avec les fonctionnalités de votre base de données cloud.

**Pourquoi est-ce si important ?**

Préparez-vous à supporter un large éventail de cas d'utilisation de l'analytics et une équipe toujours plus nombreuse à mesure que votre base de données cloud se développe au sein de votre entreprise.

## Gouvernance des données

### Contrôle métier et informatique

L'attrait des bases de données cloud réside dans le fait que n'importe qui, n'importe où et n'importe quand peut créer une base de données et effectuer des analyses pour n'importe qui, n'importe où et à n'importe quel moment. Cependant, la simplicité de ce processus a mis de nombreuses entreprises dans des situations difficiles du fait des réglementations sur les données et des coûts d'analyse des données. Si des utilisateurs métier configurent une instance par rapport à un compte professionnel, ils doivent être suffisamment responsables pour prendre en compte les éléments suivants :

- **Le contrôle d'accès** : traditionnellement, c'est l'équipe informatique qui gère l'accès aux données et en est responsable. L'accès contrôlé aux informations personnelles : par exemple, les numéros d'identification fiscale et les numéros de carte de crédit relèvent du RGPD. Les entreprises doivent veiller à ce qu'un système d'analyse SaaS ne supprime pas le contrôle centralisé des accès aux données par le service informatique.
- **Le chiffrement et la sécurité** : pour atténuer les risques induits par le traitement de données sensibles, un grand nombre de réglementations exigent le chiffrement des données. Vérifiez que votre solution peut chiffrer les données et analyser celles qui sont encryptées. Souvent appelée « chiffrement préservant le format » (FPE, Format-Preserving Encryption), cette fonctionnalité protège les entreprises des sanctions réglementaires, même en cas de violation des données. Le déchiffrement n'est pas nécessaire pour que l'analyse fonctionne. Le chiffrement en mouvement est également essentiel au fonctionnement du cloud. Afin d'éviter les violations de données, si vos données se trouvent dans un cloud public, alors elles doivent être chiffrées même lorsqu'elles se déplacent d'un endroit à l'autre.
- **Le contrôle des coûts** : veillez à ce que votre solution de cloud permette aux utilisateurs de réduire la puissance de calcul lorsqu'elle n'est pas utilisée. On ne compte plus les histoires de dépassement de budget associées aux coûts sur le cloud. Dans un rapport intitulé « The AWS bill heard around the world », un client d'AWS a raconté comment un entrepôt d'objets utilisé pour stocker des weblogs et qui coûtait environ 30 dollars par mois, est soudainement passé à 2 700 dollars le mois. Il a fallu des semaines pour comprendre ce qui s'était passé, l'équipe ayant envisagé des attaques par déni de service et d'autres causes possibles.
- **La gestion des copies de données** : les utilisateurs métier et/ou le service informatique doivent être suffisamment responsables pour limiter les copies multiples des mêmes données afin de réduire les coûts et de garantir leur sécurité. Il se peut que les responsables des budgets consacrés aux services cloud ne se rendent pas compte du gaspillage avant que la facture arrive, ce qui peut provoquer un choc. Le stockage partagé permet à plusieurs équipes d'utiliser les mêmes données sans faire de copie.
- **Les coûts de la mise à l'échelle automatique** : les équipes doivent comprendre comment la solution sera mise à l'échelle automatiquement pour les requêtes longues ou complexes, ou lorsque de nombreux workloads simultanés frappent le système en même temps. Si la base de données génère automatiquement des nœuds supplémentaires, ceux-ci sont généralement facturés par tranches mensuelles. Même si vous n'utilisez ces nœuds supplémentaires que pendant une heure, votre éditeur de solutions peut vous facturer le mois entier.
- **Le contrôle des coûts de sortie** : renseignez-vous sur les coûts de migration des données hors du cloud de votre éditeur SaaS. De nombreux éditeurs facturent une redevance par méga-octet sur les données retirées de leur plateforme et placées sur une autre. Assurez-vous de ne pas placer vos données dans une plateforme qui vous facturera des frais pour les récupérer.

### Pourquoi est-ce si important ?

Les fonctions de gouvernance des données d'une base de données cloud doivent vous aider à respecter les règles sectorielles et publiques relatives à la gestion des données. En suivant les bonnes règles de gestion des données, vous devriez éviter des amendes excessives et de fâcheuses violations de données. Veillez à ne pas faire de copies inutiles de vos données, car les doublons vous exposent à des risques (sécurité, RGPD, autres réglementations). Lorsque votre data warehouse est chiffré, les pirates informatiques peinent davantage à exploiter son contenu, ce qui diminue les risques auxquels votre entreprise est exposée. Enfin, réfléchissez aux choix que vous faites aujourd'hui et à l'impact qu'ils auront lorsque vous devrez transférer vos données d'un cloud vers la nouvelle technologie phare.

## Marche à suivre

Dans ce livre blanc, nous avons abordé un grand nombre de fonctionnalités proposées par les bases de données cloud. Nous avons aussi évoqué des facteurs à prendre en compte pour étudier les solutions disponibles et assurer la réussite de votre entreprise. Votre décision doit tenir compte de ces points importants dont la gestion et le contrôle des coûts, le respect des accords des SLA et les coûts de migration en interne et vers de nouvelles bases de données. Dans votre prise de décision, assurez-vous également qu'un maximum de personnes accèdera à l'ensemble des fonctionnalités analytiques proposées et, par conséquent, que la solution dote l'entreprise d'une certaine flexibilité et d'un réel avantage concurrentiel.