



## DOCUMENT D'INFORMATION D'IDC

### Comment Big Data transforme la protection et le stockage des données

---

*Août 2012*

*Écrit par Carla Arend*

*Sponsorisé par CommVault*

---

#### **Introduction : Comment Big Data transforme le stockage**

Big Data constitue l'une des forces majeures dans l'évolution que connaît aujourd'hui l'industrie des technologies de l'information. Les attitudes envers Big Data oscillent entre le sarcasme et l'enthousiasme, mais IDC en est certain : Big Data va transformer la manière dont nous concevons et utilisons les technologies de l'information et, fait bien plus important, Big Data va changer la manière dont sont prises les décisions dans le monde des affaires sur la base de l'exactitude et de la rapidité de mise à disposition des données disponibles pour la prise de décisions. Ce document examine comment l'émergence de cas d'utilisation de Big Data va influencer et transformer les besoins en termes d'infrastructures de stockage.

#### **Les quatre composantes du Big Data**

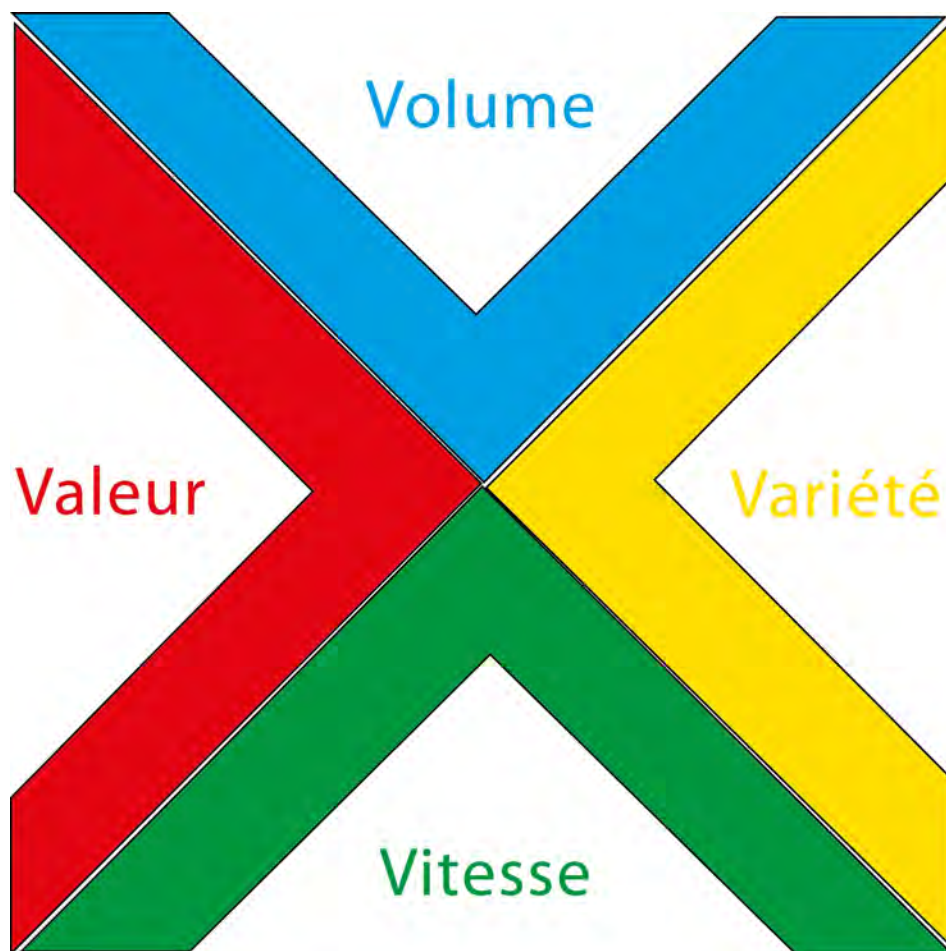
Le concept de Big Data enveloppe une dimension analytique et une dimension de stockage, et les discussions autour de Big Data portent principalement sur la façon dont les entreprises peuvent en tirer un avantage concurrentiel en analysant en temps réel des sources de données existantes et nouvelles. Ces nouvelles exigences en termes d'analyse ont également un impact sur l'architecture du stockage.

IDC définit le concept de Big Data comme suit :

« Les technologies Big Data décrivent une nouvelle génération de technologies et d'architectures conçues pour tirer profit de façon économique de très grands volumes de données de nature diverse, en permettant de réaliser des captures, découvertes et/ou analyses à grande vitesse. »

Figure 1

Les quatre composantes du Big Data



Source : IDC, 2012

Les quatre éléments suivants permettent de décrire Big Data :

- **Volume.** Le défi de gérer des volumes de données toujours plus importants n'a rien de nouveau pour les administrateurs de systèmes de stockage. Cependant, avec Big Data, les entreprises pourraient atteindre les limites de leur architecture actuelle bien plus rapidement.
- **Variété.** Big Data permet aux entreprises d'analyser des données qui ont été générées en dehors de l'entreprise, telles que des données de médias sociaux et des données météorologiques, ainsi que des données générées par des capteurs, des systèmes de point de vente, des étiquettes d'identification par radiofréquence (Radio Frequency Identification Tags, RFID), des caméras de vidéosurveillance, etc. Ces nouveaux types de données posent de nouvelles questions autour de la gouvernance de l'information et viennent s'ajouter à la quantité de données stockées.

- **Vitesse.** Les données se multiplient dans les entreprises à une vitesse croissante : l'analyse Big Data cherche à en tirer parti en temps réel. Par conséquent, la performance est un élément clé de l'infrastructure informatique sous-jacente.
- **Valeur.** L'analyse Big Data a pour objectif de créer un avantage concurrentiel unique pour les entreprises, en leur permettant de mieux comprendre les préférences de leurs clients, de segmenter les clients de façon plus granulaire et de cibler des offres spécifiques au niveau de segments précis. Mais les entreprises du secteur public utilisent également Big Data pour éviter les fraudes et économiser l'argent des contribuables, et offrir de meilleurs services aux citoyens, dans le domaine des soins de santé par exemple. Des cas d'utilisation de Big Data font leur apparition dans tous les secteurs, avec comme points communs l'enthousiasme et la créativité.

Dans l'ensemble, les approches autour de Big Data peuvent être divisées en deux : celles qui visent à optimiser les données et les processus d'analyse actuels grâce à de nouvelles technologies, et celles qui utilisent les nouvelles technologies pour créer de nouvelles opportunités pour leur entreprise et envisager les choses sous un angle nouveau.

### Défis de stockage liés à Big Data

Dans quelle mesure ces quatre paramètres changent-ils les besoins en matière de protection des données ? À quels défis les responsables informatiques font-ils face ?

- **Volume.** L'augmentation des volumes de données est le défi le plus communément admis pour les responsables du stockage. Ils ont fort à faire avec la réduction des fenêtres de sauvegarde, tout en ayant des cycles de sauvegarde encore plus longs en raison des volumes plus importants à traiter. Ils éprouvent également des difficultés à respecter les exigences imposant des processus de restauration plus courts. Big Data accélère ces défis et soulève la question de la réorganisation de l'architecture des processus de sauvegarde ainsi que des questions sur la valeur des données et la question de savoir si toutes les données doivent être traitées de façon égale.
- **Variété.** L'existence de différents types de données, qui ne sont pas toutes générées au sein de l'entreprise, soulève la question de la gouvernance de l'information. Comment protégez-vous les données qui ont été générées sur le Web social ? Comment pouvez-vous appliquer des politiques à des données qui vivent dans le Cloud, sont analysées dans le Cloud, mais constituent la base d'importantes décisions d'affaires ?
- **Vitesse.** La performance est une caractéristique clé de Big Data, et l'un de ses avantages est la réduction du temps de décision. Cela augmente les performances exigées dans l'infrastructure de stockage.
- **Valeur.** L'objectif de l'analyse Big Data est de créer une valeur ajoutée pour l'entreprise. Cela soulève la vieille question de la valeur des données stockées. Différencier les données continue

d'être un défi, et de nombreuses entreprises traitent toutes les données de la même manière, faute d'avoir trouvé une solution alternative efficace. Un autre aspect de valeur est de trouver des données pertinentes et de les rendre accessibles lors du processus de décision, en particulier des informations non structurées.

## **Avantages de Big Data**

Comment le stockage peut-il permettre de créer de la valeur et un avantage concurrentiel à partir des données ? Même si l'avantage concurrentiel naît en grande partie des progrès en matière d'analyse, le stockage joue également un rôle clé dans l'exploitation de Big Data en :

- Offrant une gestion des données reposant sur des politiques. L'architecture et la gouvernance de l'information doivent être revues lorsque les entreprises commencent à adopter Big Data.
- Rendant les données consultables grâce à une indexation intelligente. Trouver des données et les rendre disponibles afin de les utiliser pour la gestion et les prises de décision constitue une autre façon de créer de la valeur ajoutée.
- Garantissant les performances de stockage. La performance est un paramètre clé dans Big Data : la valeur provient d'une analyse en temps réel des données. Une gestion efficace des données garantit des performances de stockage optimales.
- Stockant des données de manière très efficace afin de limiter l'empreinte de stockage (déduplication, instance unique, compression, allocation granulaire de capacité, snapshots). L'efficacité du stockage est un facteur essentiel permettant de générer de la valeur pour l'entreprise. Si les données sont stockées de façon optimale, l'empreinte du stockage peut être réduite à son minimum et l'entreprise peut ainsi libérer des ressources et de l'argent afin de les investir dans l'innovation.
- Fournissant un accès aux données à partir de périphériques mobiles. Les données sont de plus en plus souvent consultées par des travailleurs mobiles et depuis des appareils mobiles intelligents. Ceci est particulièrement vrai pour les données utilisées dans les prises de décision. Les responsables informatiques doivent en tenir compte dans leur organisation.
- Utilisant le stockage en Cloud lorsque cela est approprié. Une partie de Big Data est créée, analysée et stockée en Cloud. Le mouvement de grandes quantités de données sur des réseaux demeure un défi de performance : le stockage en Cloud doit faire partie des solutions de stockage lorsque son utilisation est appropriée.

## **Bonnes pratiques de stockage pour soutenir Big Data**

À partir des cas d'utilisation connus, nous pouvons mettre en évidence certaines bonnes pratiques émergentes autour de la gestion des données pour Big Data :

- Repensez votre architecture de stockage. Certains ensembles de données Big Data nécessitent de multiples copies actives qui sont protégées par réplication au lieu d'être protégées par une sauvegarde traditionnelle. Beaucoup d'entreprises utilisent un mélange de snapshots, de réplication et de sauvegarde pour protéger les ensembles de données Big Data. Prenez un point de départ dans votre infrastructure de stockage actuelle et cherchez à comprendre comment vous pouvez le faire évoluer en tirant parti d'une nouvelle architecture ou de nouvelles technologies.
- Comprenez vos données. Les données ne sont pas toutes aussi importantes les unes que les autres et n'ont pas toutes besoin de la même protection, et ceci est particulièrement le cas avec Big Data. Lorsque l'on examine le processus de Big Data, les données entrantes ont très probablement besoin d'être stockées, mais dans certains cas il ne s'agit que de données transitoires passant dans l'entreprise sans être conservées. Les algorithmes sont généralement la partie la plus précieuse, car ils constituent un facteur de différenciation unique pour les entreprises. Les résultats de l'analyse ne doivent pas nécessairement être stockés, car certains ensembles de données sont plus rapides à créer qu'à restaurer.
- La gouvernance des données est plus complexe. Lorsqu'elles utilisent d'autres types de données dans leurs analyses, les entreprises doivent comprendre les règles de confidentialité associées à ces données. Cela vaut pour les données générées en dehors de l'entreprise, mais aussi pour les données vivant dans le Cloud.

### **Conclusion : Big Data va transformer le stockage. Comment pouvez-vous en tirer parti ?**

Big Data ayant été adopté dans toute l'Europe, les entreprises vont devoir faire évoluer leurs infrastructures de stockage. Cependant, de nombreux défis de stockage résultant de Big Data sont bien connus et bien compris, mais ils le sont simplement à une plus petite échelle. On conseille donc aux entreprises de faire évoluer leur infrastructure de stockage, sans pour autant défaire et remplacer toute leur infrastructure existante. Les fournisseurs de stockage ne cessent d'innover pour relever les nouveaux défis qui émergent et pour certains d'entre eux Big Data est déjà au calendrier. Consultez votre fournisseur de stockage ou votre partenaire de distribution et interrogez-le sur sa vision du marché de Big Data.

IDC vous recommande également de faire appel à des services d'architecture pour comprendre l'impact de l'utilisation de Big Data dans votre entreprise. Les cas d'utilisation de Big Data diffèrent beaucoup selon le secteur et la taille des entreprises, tout comme la valeur des données utilisées pour l'analyse. Comprendre comment créer de la valeur tout au long du processus permet de mettre en place une infrastructure de stockage efficace.

#### AVIS DE DROIT D'AUTEUR

Les opinions, analyses et résultats de recherche présentés dans ce document d'information d'IDC sont directement issus d'études plus détaillées publiées dans IDC Continuous Intelligence Services (Services IDC de renseignements continus). Toute publication d'informations appartenant à IDC dans une publicité, un communiqué de presse ou du matériel promotionnel nécessite l'autorisation écrite préalable d'IDC.

Contactez IDC Go-to-Market Services sur [gms@idc.com](mailto:gms@idc.com) ou la ligne d'information GMS au 508-988-7610 pour solliciter une autorisation de citation ou de publication d'IDC, ou pour plus d'informations sur les documents d'information d'IDC. Consultez le site [www.idc.com](http://www.idc.com) pour en savoir plus sur l'abonnement et les services de conseil d'IDC ou bien [www.idc.com/gms](http://www.idc.com/gms) pour en savoir plus sur IDC Go-to-Market Services.

Copyright 2012 IDC. Toute reproduction est interdite, sauf autorisation.