

Présentation de la solution

Big Data et Cloud : des technologies convergentes

Créer des avantages concurrentiels grâce à l'analyse Big Data basée sur le Cloud

En quoi ce document peut vous être utile

Cet article décrit comment les technologies Cloud et Big Data convergent pour proposer un modèle de distribution performant et rentable pour les tâches d'analyse. Il couvre également les questions suivantes :

- En quoi le Cloud Computing facilite-t-il la mise en place d'un système avancé d'analyse Big Data ?
- Comment un service informatique peut-il assumer le leadership de l'analyse Big Data dans le Cloud, en agissant comme intermédiaire au sein de l'entreprise ?
- Application des modèles AaaS (*Analytics-as-a-Service*) pour l'analyse Big Data sur le Cloud
- Recommandations pratiques pour bien engager votre initiative d'analyse Big Data sur le Cloud

Présentation de la solution Big Data et Cloud : des technologies convergentes

Créer des avantages concurrentiels grâce à l'analyse Big Data basée sur le Cloud



Table des matières

- 3 Le Cloud, moteur de l'analyse Big Data
- 6 Cloud et Big Data : une combinaison gagnante
- 10 Le service informatique comme intermédiaire des services Cloud
- 13 Étapes suivantes pour le service informatique

Le Cloud, moteur de l'analyse Big Data

Partout dans le monde, deux grandes questions préoccupent actuellement les entreprises en matière d'informatique : l'analyse Big Data et le Cloud Computing. L'analyse Big Data peut potentiellement faire émerger de précieux éléments de compréhension, qui permettront de créer un avantage concurrentiel, de susciter de nouvelles innovations, et d'induire la croissance du chiffre d'affaires. En tant que modèle de distribution des services informatiques, le Cloud Computing peut potentiellement améliorer l'agilité métier et la productivité de l'entreprise, renforcer son efficacité et réduire ses coûts.

Ces deux technologies ne cessent d'évoluer. Le stockage des Big Data n'est plus un problème en soi pour les entreprises, qui cherchent aujourd'hui plutôt à déterminer comment elles peuvent conduire des analyses qui apporteront des réponses pertinentes et concrètes à leurs besoins. À l'heure où le Cloud Computing poursuit son évolution, de plus en plus d'entreprises développent des environnements Cloud puissants et agiles, et les prestataires de services Cloud continuent à élargir leur offre.

Dans un tel contexte, on comprend mieux pourquoi les entreprises informatiques s'appuient sur le Cloud Computing en tant que structure d'encadrement de leurs projets Big Data. Les environnements Big Data nécessitent des clusters de serveurs afin de prendre en charge les outils qui traitent les données dans des volumes importants, à une vitesse élevée, et dans des formats variés. Les Clouds sont déjà déployés sur des pools de serveurs, des infrastructures de stockage et des ressources réseau, et peuvent être développés ou réduits en fonction des besoins. Le Cloud Computing offre un moyen économique de prendre en charge les technologies de Big Data et les applications d'analyse avancées capables d'apporter de la valeur ajoutée à l'entreprise.

Cet article couvre les thèmes suivants :

- En quoi le Cloud Computing facilite-t-il la mise en place d'un système avancé d'analyse Big Data ?
- Comment un service informatique peut-il assumer le leadership de l'analyse Big Data sur le Cloud, en agissant comme intermédiaire au sein de l'entreprise ?
- Application des modèles AaaS (*Analytics-as-a-Service* / analyse en tant que service) pour l'analyse Big Data sur le Cloud
- Recommandations pratiques pour bien engager votre initiative d'analyse Big Data sur le Cloud

Tendances en matière de Big Data

Pourquoi le Cloud Computing constitue-t-il un modèle de distribution si rentable pour l'analyse Big Data ? En quoi la convergence des technologies fait-elle de l'analyse Big Data sur le Cloud une option raisonnable ?

Les données sont de plus en plus précieuses. Aujourd'hui, la question n'est plus « quelles données devons-nous stocker ? », mais « que pouvons-nous faire avec ces données ? ». Les entreprises cherchent à libérer toute la valeur potentielle de leurs données afin d'en tirer des avantages concurrentiels. Gartner prédit que le volume des données d'entreprise va s'accroître de 800 % entre 2011 et 2015, avec 80 % de données non structurées (e-mails, documents, vidéos, images, contenus des plateformes de réseaux sociaux, etc.) et 20 % de données structurées (transactions par carte de crédit, coordonnées, etc.).¹

Toutes ces données peuvent potentiellement faire émerger de nouveaux éléments de compréhension, générateurs de compétitivité. Les entreprises doivent pour cela trouver de nouvelles approches pour traiter, gérer et analyser leurs données, qu'il s'agisse de formats structurés comme ceux issus des systèmes de gestion de base de données relationnelle (SGBDR) conventionnels, ou de formats plus variés et non structurés. Le croisement de différents types et sources de données peut par ailleurs permettre de découvrir des relations et tendances jusque-là inexplorées et extrêmement intéressantes.

Qu'est-ce que l'analyse Big Data ?

L'expression Big Data fait référence à de gigantesques ensembles de données de plus grande taille (volume) ; plus diversifiés, comprenant des données structurées, semi-structurées, et non structurées (variété) ; et arrivant plus vite (vitesse) que ce à quoi vous ou votre entreprise avez eu à faire face auparavant. Ce flot de données est généré par les appareils connectés – depuis les ordinateurs et smartphones jusqu'aux capteurs tels que les lecteurs RFID, en passant par les caméras de circulation. De plus, il est hétérogène et inclut des formats variés tels que texte, documents, images, vidéos, et bien d'autres.

La valeur réelle du Big Data réside dans les informations qui résultent de leur analyse – la découverte de principes de comportement, la déduction d'une information, la prise de décision et enfin la proposition d'une réponse adaptée.

L'analyse Big data désigne un ensemble de technologies de pointe conçues pour prendre en charge de gros volumes de données hétérogènes. Elle a recours à des méthodes quantitatives sophistiquées telles que l'apprentissage machine, les réseaux neuronaux, la robotique, les mathématiques computationnelles et l'intelligence artificielle pour explorer les données et faire émerger des relations et tendances cachées.

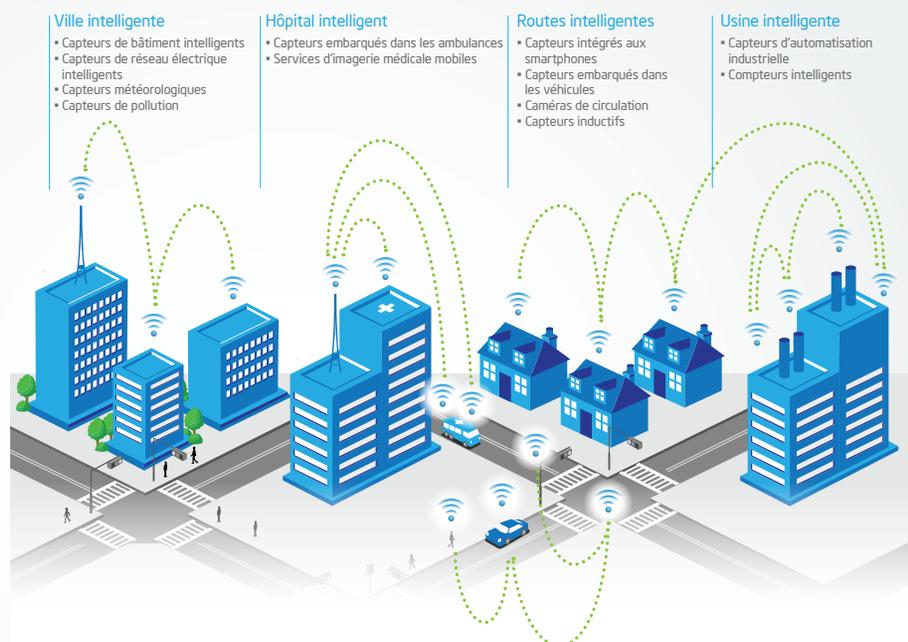
Évolution de l'analyse de données : du traitement par lots au temps réel. Dans son enquête menée en 2012 auprès de 200 responsables informatiques de grandes entreprises, Intel a constaté que, si la répartition entre traitement par lots et en temps réel est aujourd'hui égale, la tendance est toutefois en faveur du temps réel, qui devrait représenter les deux tiers des tâches de gestion des données d'ici 2015.² Le traitement en temps réel ou en quasi-temps réel n'est plus seulement un concept à la mode : sa phase de maturation est d'ores et déjà bien engagée.

Le traitement en temps réel contribue à l'analyse prédictive. L'analyse prédictive permet aux entreprises d'avoir une idée beaucoup plus claire de l'avenir et peut ouvrir d'excellentes opportunités de génération de valeur à partir du Big Data. Le traitement des données en temps réel ouvre la perspective d'une analyse prédictive rapide, précise et flexible, qui s'adaptera rapidement à l'évolution de la conjoncture. Plus vite vous pourrez analyser vos données, plus pertinentes seront vos conclusions et plus grande sera leur valeur prédictive.

Le champ d'application du Big Data continue à s'élargir. L'intérêt précoce qu'a suscité l'analyse Big Data concernait essentiellement les sources de données commerciales et sociales : e-mails, vidéos, tweets, messages Facebook*, avis, comportements en ligne, etc. Le champ d'application de l'analyse Big Data s'élargit aujourd'hui, pour inclure des données issues de systèmes intelligents (*infotainment* embarqué, kiosques, compteurs intelligents, etc.) et de capteurs installés en périphérie des réseaux. Ces systèmes génèrent des Big Data qui comptent parmi les plus volumineuses, les plus rapides et les plus complexes. Le développement et la connectivité omniprésente des capteurs et systèmes intelligents ont ouvert un nouveau champ d'informations, inédit et d'une grande valeur. L'intérêt suscité par l'application de l'analyse Big Data aux données issues des capteurs et systèmes intelligents va croissant, les entreprises cherchant à faire émerger des éléments de compréhension plus riches, cela toujours plus rapidement et de manière toujours plus rentable, afin d'optimiser leur prise de décision et de personnaliser les expériences client.

Exemple d'application concrète du Big Data : la ville intelligente

En plus des données transactionnelles, sociales et géographiques générées par les utilisateurs, les capteurs intégrés aux appareils génèrent, en temps réel, d'autres types de Big Data, comptant parmi ceux qui se développent le plus fortement. Ces données précieuses peuvent être traitées et analysées en déployant une infrastructure informatique dédiée, Cloud ou intégrée, appuyée par des solutions de stockage et de traitement haute performance.



Les technologies Cloud parviennent à maturité

Le Cloud Computing devient aujourd'hui une réalité pour de nombreuses entreprises, avec en particulier une forte poussée du déploiement des Clouds privés. La technologie Cloud atteint son niveau de maturité, les obstacles à son adoption étant progressivement éliminés au fil des améliorations en matière de sécurité et d'intégration des données. De leur côté, les entreprises informatiques évoluent et s'adaptent pour prendre en charge les services Cloud. On assiste donc à un renforcement croissant de la confiance des entreprises envers les modèles de distribution Cloud. Ainsi, dans une enquête récemment menée auprès de 200 responsables informatiques, Intel a constaté que près des deux tiers des entreprises comptent se tourner vers le Cloud comme modèle de distribution de leurs systèmes vitaux au cours des cinq prochaines années.³

Les entreprises stockent de plus en plus de données dans les environnements de Cloud Computing, constituant ce faisant d'immenses et précieuses mines d'informations. Les architectures Cloud constituent en outre pour les utilisateurs métier des ressources évolutives, capables de s'adapter à leurs besoins. L'association entre des serveurs et systèmes de stockage basés sur les derniers processeurs Intel® Xeon® (accompagnés d'unités Intel SSD et des ressources réseau Intel 10 GbE dédiées aux environnements de Cloud Computing), avec des outils de traitement Big Data tels que le logiciel Apache Hadoop distribué par Intel*, fournit la puissance de calcul requise pour analyser de grandes quantités de données de manière efficace et rentable. L'exploitation de Hadoop* dans des environnements virtualisés continue d'évoluer et de gagner en maturité, avec notamment des initiatives telles que le projet open-source Serengeti de VMware*.

Cloud et Big Data : une combinaison gagnante

Une récente enquête conduite par GigaSpaces a révélé que 80 % des responsables informatiques pour qui le traitement Big Data est important envisagent de transférer leurs systèmes analytiques sur un ou plusieurs modèles de distribution Cloud.⁴

Les modèles de distribution Cloud offrent une flexibilité exceptionnelle, qui permet aux responsables informatiques d'évaluer la meilleure approche en fonction des demandes spécifiques des utilisateurs métier. Les organisations qui ont déjà mis en place un environnement Cloud privé interne pourront par exemple ajouter l'analyse Big Data à leurs offres de services internes, utiliser un prestataire de services Cloud, ou mettre en place un Cloud hybride qui permettra de protéger certaines données sensibles au sein d'un Cloud privé, tout en bénéficiant de précieuses sources de données et applications externes issues de Clouds publics.

Le recours à l'infrastructure Cloud pour l'analyse Big Data est justifié pour plusieurs raisons :

Les investissements requis pour l'analyse Big Data peuvent être importants, ce qui nécessitera la mise en place d'une infrastructure efficace et rentable. Les ressources nécessaires pour servir un modèle de traitement distribué en interne sont en général hébergées dans des centres de données de taille moyenne à grande. Les Clouds privés peuvent offrir un moyen plus efficace et plus rentable de mettre en œuvre l'analyse Big Data en interne, tout en complétant les ressources internes par des services de Cloud public. Cette approche hybride permet à l'entreprise de disposer d'un espace de stockage et d'une puissance de calcul à la demande, via des services de Cloud public, pour certaines de ses initiatives d'analyse (projets à court terme, par exemple). Elle offre en outre une marge de capacité et d'évolutivité supplémentaire si nécessaire.

Le Big Data permet d'associer les sources internes et externes. Si les entreprises conservent en général leurs données les plus sensibles en interne, d'énormes volumes de Big Data (qu'elles appartiennent à l'organisation ou qu'elles soient générées par des fournisseurs tiers et des prestataires publics) peuvent être stockés en externe (une partie étant en fait déjà conservée dans un environnement Cloud). Le transfert de certaines sources de données derrière le firewall de l'entreprise peut exiger des ressources considérables. L'analyse des données au niveau de l'emplacement de stockage (que cela soit dans des centres de données Cloud internes ou publics, dans ces systèmes périphériques ou sur des appareils clients) s'avère souvent préférable.

Les services de données sont nécessaires pour extraire la valeur des Big Data. Selon les besoins et le scénario d'utilisation, le déploiement d'un modèle AaaS (*analytics as a service* / analyse en tant que service), appuyé par votre Cloud privé interne, par un Cloud public ou par un modèle hybride, pourra permettre une utilisation optimale de votre budget informatique.

Libérer le potentiel du Big Data dans les environnements Cloud

Les modèles de Cloud Computing peuvent permettre d'exploiter plus rapidement le potentiel des solutions analytiques évolutives. Les environnements Cloud offrent la flexibilité et l'efficacité requises pour accéder aux données, faire émerger des éléments de compréhension et renforcer la création de valeur. L'analyse Big Data sur le Cloud n'est pas une solution à taille unique, cependant.

Les entreprises qui déploient des infrastructures Cloud pour leurs solutions AaaS disposent de plusieurs options. En fonction de leurs impératifs de charge de travail, de coût, de sécurité et d'interopérabilité, les responsables informatiques peuvent choisir d'utiliser leur Cloud privé pour minimiser les risques et assurer un contrôle maximal ; d'utiliser des infrastructures, une plateforme ou des services analytiques de Cloud public pour améliorer l'évolutivité ; ou de déployer un modèle hybride associant des ressources et services de Cloud privés et publics.

Au final, cependant, quel que soit le modèle de distribution Cloud le plus approprié, et quels que soient ses besoins et budgets, toute entreprise peut libérer le potentiel des Big Data à partir d'un environnement de Cloud Computing.

Le cadre d'analyse AaaS

L'architecture AaaS Cloud permet de répondre à tous les besoins des utilisateurs en matière d'analyse : de la livraison et la gestion des données jusqu'à leur consommation. En élaborant une stratégie Big Data Cloud complète, vous pouvez définir un cadre d'analyse et optimiser la valeur totale des données de votre entreprise.

Un cadre d'analyse AaaS inclura les capacités clés suivantes :

- Capture et extraction de données structurées et non structurées issues de sources fiable, en donnant la priorité aux données essentielles, ainsi qu'en identifiant les éléments à conserver et de la durée de leur conservation
- Gestion et contrôle des données conformément à des règles de gouvernance claires s'appliquant à l'ensemble de l'entreprise, ainsi qu'aux exigences spécifiques du secteur d'activité
- Exécution de l'intégration, de l'analyse, de la transformation et de la visualisation des données pour distribuer les bonnes informations, au bon endroit, au bon moment

Types de service Cloud pour AaaS

Les solutions AaaS peuvent être déployées via différents types de services de Cloud Computing. La détermination du mix optimal entre les différents services nécessitera une évaluation des besoins des utilisateurs par rapport aux ressources internes existantes (environnement Cloud privé, par exemple).

Les solutions AaaS Cloud de base incluent notamment : infrastructure en tant que service (IaaS) ; plateforme en tant que service (PaaS) ; et logiciel en tant que service (SaaS).

Infrastructure en tant que service (IaaS)

Déployée sur place ou par l'intermédiaire d'un prestataire Cloud, une solution IaaS permet d'allouer ou d'acheter du temps sur des ressources de serveur partagées, souvent virtualisées, pour répondre aux besoins de traitement et de stockage des tâches d'analyse Big Data. Les systèmes d'exploitation Cloud gèrent des serveurs, un réseau et des ressources de stockage de pointe.

Les solutions IaaS sont au cœur des services de Cloud Computing de nombreuses entreprises. Elles nécessitent toutefois des investissements plus conséquents dans les ressources informatiques, dans le cadre du déploiement des capacités d'analyse Big Data. Votre organisation devra ainsi installer son propre logiciel, par exemple le framework Hadoop, ou une base de données NoSQL comme Apache Cassandra*, MongoDB* ou Couchbase. Votre équipe sera également responsable de la gestion des ressources vous ayant été assignées (cela pourra être facilité par des outils automatisés de gestion et d'orchestration des ressources).

Exemples de solutions IaaS

Les offres ci-dessous sont des exemples de solutions IaaS proposées dans l'écosystème des technologies Cloud.

- Amazon* Web Services
- Citrix* CloudPlatform
- Windows Azure* et Microsoft* System Center
- Logiciel OpenStack*
- Rackspace*
- Savvis*
- Verizon* Terremark*
- Suite VMware vCloud*

Plateforme en tant que service (PaaS)

Une solution PaaS met à la disposition des développeurs les outils et bibliothèques requis pour construire, tester, déployer et exécuter des applications sur l'infrastructure Cloud. Cela permet de réduire la charge de travail de gestion en éliminant la nécessité de configurer et d'adapter les éléments de votre implémentation Hadoop ; cela servira également de plateforme pour le développement d'applications d'analyse avancées.

Logiciel en tant que service (SaaS)

Une solution SaaS peut permettre de mobiliser des applications spécifiques pour les tâches d'analyse Big Data sur le Cloud. Il pourra être nécessaire d'utiliser plusieurs applications SaaS pour couvrir tout l'éventail des scénarios d'utilisation. Un logiciel bien adapté à l'analyse des sentiments ne sera, par exemple, pas nécessairement approprié pour évaluer la gestion des risques ou la performance des actifs. Une solution SaaS peut être proposée comme une application autonome ou en tant que composant d'une solution Cloud plus large. Karmasphere propose par exemple une application *pay-as-you-go* qui analyse les données stockées sur Amazon S3* avec Amazon Elastic MapReduce.

Technologies d'Infrastructure Intel pour l'analyse Big Data sur le Cloud

Les serveurs basés sur les familles de processeurs Intel® Xeon® E5 et E7 sont au cœur de l'infrastructure à la base des environnements Cloud et Big Data, proposant un outil informatique de référence, extrêmement efficace et ultra-performant et doté des avantages suivants :

- Les serveurs de stockage basés sur la famille de processeurs Intel Xeon E5 prennent en charge des fonctionnalités de stockage avancées telles que la compression, le cryptage, la hiérarchisation automatisée des données, la déduplication des données, le codage d'effacement et l'allocation granulaire. Ils se prêtent idéalement au stockage et au traitement de gros volumes de données. Ces tâches de stockage très exigeantes en ressources de calcul garantissent une sécurité accrue, une plus grande efficacité et un coût total de possession plus avantageux, grâce à une empreinte de stockage réduite.
- Les SSD (*Solid-State Drives*) Intel sont des disques durs haute performance à haut débit pour le stockage brut.
- Les adaptateurs réseau convergents Ethernet 10 Gigabit Intel fournissent des connexions à haut débit pour les grands ensembles de données.

Intel fournit également des fonctions de sécurité matérielle renforcée, notamment Intel Advanced Encryption Standard (Intel AES-NI)⁵, qui permet de multiplier par dix la vitesse de cryptage et décryptage des données.^{6,7} La technologie Intel d'exécution fiabilisée (IntelTXT)⁹ permet en outre de garantir la fiabilité des données au niveau du matériel, et de s'assurer que les données sont traitées sur (ou migrées vers) des groupes de serveurs fiables.

Exemples de solutions PaaS

Les offres ci-dessous sont des exemples de solutions PaaS proposées dans l'écosystème des technologies Cloud.

- *Force.com*
- *Google* App Engine*
- *Red Hat* OpenShift**
- *VMware Cloud Foundry*
- *Windows Azure**

Exemples de solution SaaS

Les offres ci-dessous sont des exemples de solutions SaaS proposées dans l'écosystème des technologies Cloud.

- *Amazon* Elastic MapReduce*
- *Cetas* (solutions analytiques VMWare*)*
- *Services Google* BigQuery*
- *Service RackSpace* Hadoop**
- *Windows Azure* HDInsight**

Intel et le framework Apache Hadoop*

Le framework Apache Hadoop* est un logiciel open source qui s'est imposé comme la plateforme dédiée à la gestion du Big Data à grande échelle. Le logiciel Hadoop* peut traiter des ensembles de données volumineux sur des clusters d'ordinateurs utilisant des modèles de programmation relativement simples.

Le logiciel Apache Hadoop* distribué par Intel® inclut Apache Hadoop ainsi que d'autres composants logiciels optimisés par Intel avec des performances matérielles améliorées et des fonctionnalités de sécurité. Conçu pour permettre un large éventail d'analyses de données sur Apache Hadoop, le logiciel Hadoop distribué par Intel est optimisé pour les requêtes Apache Hive*, fournit des connecteurs pour R* afin de permettre le traitement statistique, et permet l'analyse graphique en utilisant Intel Graph Builder pour le logiciel Apache Hadoop, une bibliothèque destinée à traiter de grands ensembles de données sous forme de graphiques afin de faciliter la visualisation des interactions entre données. La distribution Intel inclut Intel Manager pour Apache Hadoop, qui offre une console de gestion simplifiant l'installation, la configuration et le contrôle d'un déploiement* Hadoop.

La distribution Intel est désormais disponible pour évaluation dans le monde entier. Le support technique est actuellement disponible aux États-Unis, en Chine et à Singapour et le sera dans d'autres zones géographiques plus tard dans l'année.

En savoir plus sur la [distribution Intel](#).

Le département informatique comme intermédiaire des services Cloud

Le Cloud Computing et la myriade de services Cloud publics proposés aux entreprises permettent aux grands consommateurs de tâches d'analyse (en particulier les chefs de produit et les responsables marketing) de se procurer directement les services, sans nécessairement passer par le service informatique. S'ils ne sont pas suffisamment informés, toutefois, certains utilisateurs pourront être tentés d'acheter des « analyses instantanées » ; or les approches d'adoption *ad hoc* de services Cloud publics au sein de votre organisation peuvent entraîner des problèmes importants, tels que le choix d'un mauvais fournisseur, la perte de contrôle sur les données sensibles, ou encore un retour sur investissement insatisfaisant, pour n'en citer que quelques-uns. Les responsables informatiques peuvent recommander des services, proposer des perspectives et apporter des compétences spécifiques, qui pourront réduire les risques liés à l'utilisation des Clouds publics et permettre une meilleure utilisation des ressources de Cloud privé existantes.

Le Big Data exige par ailleurs un nouvel ensemble de compétences au sein de l'entreprise, or celles-ci résident le plus souvent au sein du département informatique. Les services informatiques disposent du savoir-faire technologique nécessaire pour que les tâches de Cloud Computing et de gestion des Big Data fonctionnent correctement dans votre organisation, grâce notamment aux administrateurs et développeurs de composants Hadoop tels que la base de données Apache HBase*. Les projets d'analyse Big Data nécessitent des équipes multidisciplinaires, et une collaboration active doit être engagée entre le service informatique et les *data scientists* (ou chercheurs en données, une autre fonction qui a émergé avec le Big Data). Les *data scientists* sont chargés d'appliquer le Big Data à des problèmes métier complexes et d'interpréter les résultats. Ils peuvent aussi bien faire partie des équipes opérationnelles que du service informatique.

En tant qu'intermédiaire des services de Cloud Computing, le département informatique peut travailler avec les utilisateurs métier de façon à mettre en place une solution d'analyse Cloud optimale, prenant en compte tous les facteurs critiques.

Domaine	Questions
Gestion des données institutionnelles	Quels sont les fournisseurs de Cloud évalués ? La prise en charge du stockage, de la gestion et de l'analyse des données par plusieurs fournisseurs et sans supervision ferait courir un risque énorme à l'entreprise.
Propriété des données	Qui possède les données stockées et gérées par votre prestataire ? Votre société conserve-t-elle la propriété des données ?
Sécurité	Quel est le niveau de sécurité assuré ? À quels niveaux de la pile de solutions ? Les prestataires qui proposent une sécurité intégrée en profondeur au sein de l'infrastructure et de la plateforme, ainsi qu'au niveau de l'application, sont en mesure de garantir un niveau de sécurité supérieur.
Conformité	Les questions de conformité sont-elles prises en compte, de manière générale et par rapport aux exigences spécifiques de votre secteur d'activité ? Comment les données sont-elles anonymisées pour protéger la vie privée ?
Intégration des données	Comment les données sont-elles intégrées ? À quel coût ?
Migration des données	Quel est le volume des données à transférer ? À quel coût ? La migration de gros volumes de données vers et depuis le Cloud peut entraîner des coûts prohibitifs.
Streaming des données	Le streaming de données sources est-il constitué d'informations en temps réel ? La gestion des données en temps réel nécessite d'énormes ressources. Les données qui arrivent en un flux interrompu seront sans doute mieux prises en charge en interne.
Évaluation des technologies	Quelles sont les solutions de stockage, de traitement et d'analyse fournies pour le Big Data ? La performance des différents composants est-elle optimisée ?
Compétences requises	Quelles sont les compétences requises pour identifier les sources de données appropriées, appliquer les bons modèles statistiques et analytiques, et interpréter les résultats ? L'offre du prestataire de services inclut-elle l'accès aux technologies et l'assistance à l'analyse ? À quel niveau ? Dans quelle mesure cette offre est-elle complémentaire des compétences internes ?
Retour sur investissement	Quel est le retour sur les investissements réalisés dans la technologie, le modèle de distribution, la sécurité et les méthodes d'intégration des données ?

Avec l'analyse Big Data, le département informatique renforce son rôle

Dans un nombre croissant d'entreprises, les utilisateurs métier exploitent déjà l'informatique en tant que service. Le département informatique peut élargir son rôle en agissant comme intermédiaire des services d'analyse Big Data basés sur le Cloud. Dans un tel cadre, votre rôle consistera à évaluer les besoins des utilisateurs par rapport aux différentes solutions disponibles. Cela consistera notamment à développer une stratégie pour les services privés, publics et hybrides ; instaurer une approche rigoureuse pour la sélection des prestataires de services de Cloud Computing ; et négocier les contrats avec les prestataires potentiels. Sur le plan organisationnel, cela pourra permettre de réduire les risques et d'optimiser l'utilisation des investissements existants dans les technologies de Cloud privé. Les utilisateurs individuels sont gagnants dans la mesure où ils sont assurés de disposer de la solution la mieux adaptée à leurs besoins.

Le département informatique peut démontrer rapidement la valeur de la solution proposée pour l'entreprise, en travaillant en partenariat avec les utilisateurs pour :

- Choisir l'implémentation de Cloud privé ou public la mieux adaptée à leurs besoins en définissant les besoins technologiques, en évaluant le risque, et en précisant les exigences de mise en œuvre compte tenu des règles de gouvernance d'entreprise et des exigences de conformité réglementaire. Certaines charges de travail pourront par exemple être gérées dans un Cloud privé, à un endroit précis.
- Concevoir une solution ou travailler efficacement avec un partenaire technologique pour développer les services requis, selon les besoins.
- Évaluer et valider les services externes dans les domaines suivants : conception, livraison, personnalisation, tarification, confidentialité, intégration, sécurité et assistance.
- Déployer les services issus de sources internes et externes de façon fluide pour les utilisateurs.
- Développer les relations avec les prestataires de services de Cloud Computing retenus.
- Gérer les services existants, notamment les accords de niveau de service (SLA) et le cycle de vie des services.

En tant qu'intermédiaire, le département informatique collabore avec les entités opérationnelles afin d'identifier la meilleure façon d'exploiter les technologies pour créer un avantage concurrentiel. Avec l'analyse Big Data sur le Cloud, l'objectif doit être de fournir la solution la mieux adaptée compte tenu, d'une part des besoins des utilisateurs, et d'autre part des règles de gouvernance en place, des ressources informatiques existantes, des exigences de performance et des objectifs globaux de l'entreprise. Dans la plupart des départements informatiques actuels, la mise en place d'une telle approche consultative nécessitera une réorganisation afin d'éliminer les silos, d'embaucher de nouveaux collaborateurs ou de permettre aux membres existants d'acquérir de nouvelles compétences, et d'encourager un partenariat solide avec les équipes opérationnelles.

Toutes ces mesures porteront nécessairement leurs fruits : les projets d'analyse Big Data nécessitent en effet un degré important de collaboration entre experts informatiques, utilisateurs métier, data scientists et autres intervenants pouvant contribuer à l'élaboration d'un plan et d'algorithmes d'analyse performants, permettant d'obtenir des informations exploitables à partir des données.

Étapes suivantes pour le service informatique

Le service informatique occupe une place unique au sein de l'entreprise. En dépit (ou peut-être à cause) de la croissance exponentielle des données, des technologies émergentes et du rythme effréné des évolutions, vous devez assurer au sein de votre entreprise le leadership en matière d'analyse Big Data. Il convient en premier lieu de s'interroger sur l'évolution possible du rôle de l'informatique en tant qu'intermédiaire pour les services d'analyse Big Data sur le Cloud au sein de votre entreprise. Vous pouvez également :

- **Engager dès aujourd'hui un partenariat avec les responsables des entités opérationnelles** afin de les aider à déterminer dans quelle mesure le Big Data peut être mis à profit pour résoudre les défis auxquels est confrontée l'entreprise et exploiter les opportunités qui se présentent à elle. En tant que partenaire pleinement engagé, vous pouvez ainsi les aider à évaluer les options, faire le bon choix technologique et instaurer les meilleures pratiques.
- **Explorer les options technologiques en matière d'analyse Big Data sur le Cloud** (modèles de distribution privés, publics et hybrides, notamment). Tenez-vous bien informé des tendances, suivez le marché, et assurez-vous d'avoir une idée claire des coûts en jeu.
- **Créer ou mettre à jour la stratégie Big Data** définissant le processus d'implication du département informatique dans les projets d'analyse Big Data. Gardez à l'esprit le fait que votre rôle doit consister à permettre aux utilisateurs métier de travailler rapidement et simplement, sans quoi les entités opérationnelles prendront leurs propres mesures.
- **Réfléchissez à l'organisation du département informatique** afin d'impliquer plus efficacement les utilisateurs métier et renforcer la collaboration et la consultation sur les projets Big Data.

Pour en savoir plus sur le Big Data et le Cloud Computing, n'hésitez pas à consulter les ressources disponibles sur l'Intel IT Center. Rendez-vous sur intel.com/bigdata et intel.com/cloudcomputing.

Notes :

1. Groenfeldt, Tom. «Big Data—Big Money Says It Is a Paradigm Buster.» *Forbes* (6 janvier 2012). forbes.com/sites/tomgroenfeldt/2012/01/06/big-data-big-money-says-it-is-a-paradigm-buster/
2. *Peer Research: Big Data Analytics: Intel's IT Manager Survey on How Organizations Are Using Big Data. Intel® IT Center (août 2012).*
3. *Peer Research: Mission-Critical Workload Migration. Intel (janvier 2013).* intel.com/content/www/us/en/risc-migration/mission-critical-workload-migration-peer-research-report.html
4. Gardner, Dana. « GigaSpaces Survey Shows Need for Tools for Fast Big Data, Strong Interest in Big Data in Cloud. » *ZDNet BriefingsDirect* (10 décembre 2012). zdnet.com/gigaspace-survey-shows-need-for-tools-for-fast-big-data-strong-interest-in-big-data-in-cloud-7000008581/
5. Intel AES-NI nécessite un système informatique doté d'un processeur compatible AES-NI, ainsi que des logiciels tiers qui permettront d'exécuter les instructions selon la séquence requise. AES-NI est disponible sur les processeurs Intel® Xeon®, Intel® Core™ série i5-600 pour PC de bureau, et les processeurs Intel Core séries i7-600 et i5-500 pour PC mobiles. Pour connaître les disponibilités, contactez votre revendeur ou le fabricant de votre système. Pour plus d'informations, lire intel.com/content/www/us/en/architecture-and-technology/advanced-encryption-standard--aes-/data-protection-aes-general-technology.html.
6. Les logiciels et charges de travail employés dans les tests de performance peuvent avoir été optimisés pour les processeurs Intel. Les tests de performance tels que SYSmark* et MobileMark* portent sur des configurations, composants, logiciels, opérations et fonctions spécifiques. Les résultats peuvent varier en fonction de ces facteurs. Pour l'évaluation d'un produit, il convient de consulter d'autres tests et d'autres sources d'informations, notamment pour connaître le comportement de ce produit avec d'autres composants.
7. Source : les tests réalisés sous Oracle* Database Enterprise Edition 11.2.0.2 avec chiffrement TDE (Transparent Data Encryption) AES-256 indiquent un gain de vitesse de 10x en insérant un million de lignes 30 fois dans un tableau vide avec le processeur Intel Xeon X5680 (3,33 GHz, 36 Mo de RAM) qui utilise les routines Intel Integrated Performance Primitives (Intel IPP), par rapport au processeur Intel Xeon X5560 (2,93 GHz, 36 Mo de RAM) sans Intel IPP.
8. Aucun ordinateur ne peut garantir une sécurité à 100 % en toutes circonstances. La technologie d'exécution fiabilisée Intel TXT nécessite un ordinateur doté de la technologie Intel Virtualization, un processeur et un BIOS compatibles TXT, un chipset compatible, des modules de code authentifiés, et un environnement MLE compatible avec Intel TXT. Intel TXT nécessite également que le système contienne un module TPM v1.s. Pour en savoir plus, rendez-vous sur intel.com/go/intelxt.

Partagez ce document avec vos collègues



Ce document est à titre informatif seulement. CE DOCUMENT EST PROPOSÉ « EN L'ÉTAT », SANS GARANTIE QUELLE QU'ELLE SOIT, Y COMPRIS LES GARANTIES CONCERNANT LA QUALITÉ MARCHANDE, L'ABSENCE DE CONTREFAÇON OU L'ADÉQUATION À UN USAGE PARTICULIER OU ENCORE QUI DÉCOULERAIENT D'UNE PROPOSITION OU D'UN DEVIS, D'UNE SPÉCIFICATION OU D'UN CAHIER DES CHARGES OU BIEN D'UN ÉCHANTILLON. Intel décline toute responsabilité, y compris quant à d'éventuelles violations de droits de propriété intellectuelle, quels qu'ils soient, découlant de l'utilisation des présentes informations. Aucune licence, implicite ou explicite, par effet d'estoppel ou autre, sur les droits de propriété intellectuelle n'est accordée par le présent document.

Copyright © 2013 Intel Corporation. Tous droits réservés. Intel, le logo Intel, Intel Core, Intel Sponsors of Tomorrow, le logo Intel Sponsors of Tomorrow, et Xeon sont des marques d'Intel Corporation aux États-Unis et/ou dans d'autres pays.

* Les autres noms et marques peuvent être revendiqués comme propriété de tiers.

Microsoft est une marque déposée de Microsoft Corporation aux États-Unis et/ou dans d'autres pays.

