

SÉRIE NOUVELLES ARCHITECTURES

Alerte au tsunami des données :
les entreprises doivent prendre la vague maintenant !



Quels sont les faits qui sous-tendent cette réalité ?

Quelles entreprises sont aujourd'hui en mesure de tirer profit de ce phénomène ?

Mon entreprise peut-elle s'y soustraire ?

A quelle échéance mon SI devra-t-il être prêt ?

De nouveaux moyens techniques sont indispensables pour apporter aux métiers les bénéfices escomptés par la maîtrise des grandes volumétries de données. Cependant, comme pour le web, mon entreprise doit-elle s'y adapter? Ce dossier veut rendre tangible pour tous le phénomène, en décortiquant les faits techniques de ce tsunami qui gagne en puissance à chaque instant.

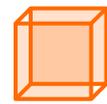
ALERTE AU TSUNAMI DES DONNÉES: LES ENTREPRISES DOIVENT PRENDRE LA VAGUE MAINTENANT !

Une étude CISCO® de mars 2013 révèle que 60% des entreprises confirment que la maîtrise de la donnée globale, produite en interne ou en externe, permet d'améliorer leur compétitivité et de faciliter leurs prises de décision. Si pour ces entreprises la gestion de la donnée est une priorité de 2013, seules 44% d'entre elles ont établi une feuille de route associée. 56% des entreprises n'ont à l'heure actuelle de plan clairement défini pour développer de réelles opportunités du phénomène Big Data.

Les entreprises qui ont engagé des travaux dans le domaine du Big Data n'ont généralement réalisé que des expérimentations ou nommé des responsables du champ d'analyse.

Les raisons exposées dans l'étude CISCO® portent sur des problématiques de sécurité, d'accès aux données, de budget, de manque de temps pour étudier le phénomène, ou encore d'un manque de compétences.

Ce constat matérialise le gap entre la priorisation de la maîtrise des données externes produites en grand volumes et les moyens mis en place par les entreprises. Or **comme avec le web** dans les années 2000, pour décliner une stratégie d'entreprise **il est indispensable** de définir un plan précis **de mise en œuvre de nouveaux moyens techniques** : cela passe avant tout par la compréhension de ce que nous nommons les 4 Dimensions Digitales.



DIMENSION 1 : L'ESPACE / LE VOLUME

Selon une étude IDC sponsorisée par EMC Gartner, les données numériques créées dans le monde seraient passées de 1,2 zettaoctets (Zo, 10^{21}) par an en 2010 à 2,8 Zo en 2012, et s'élèveront à 40 Zo en 2020.



Volume de données générées par jour

La génération de ce volume de données est directement induite par la densité d'appareils en interaction sur internet, ou **Internet of Things / Everything**.

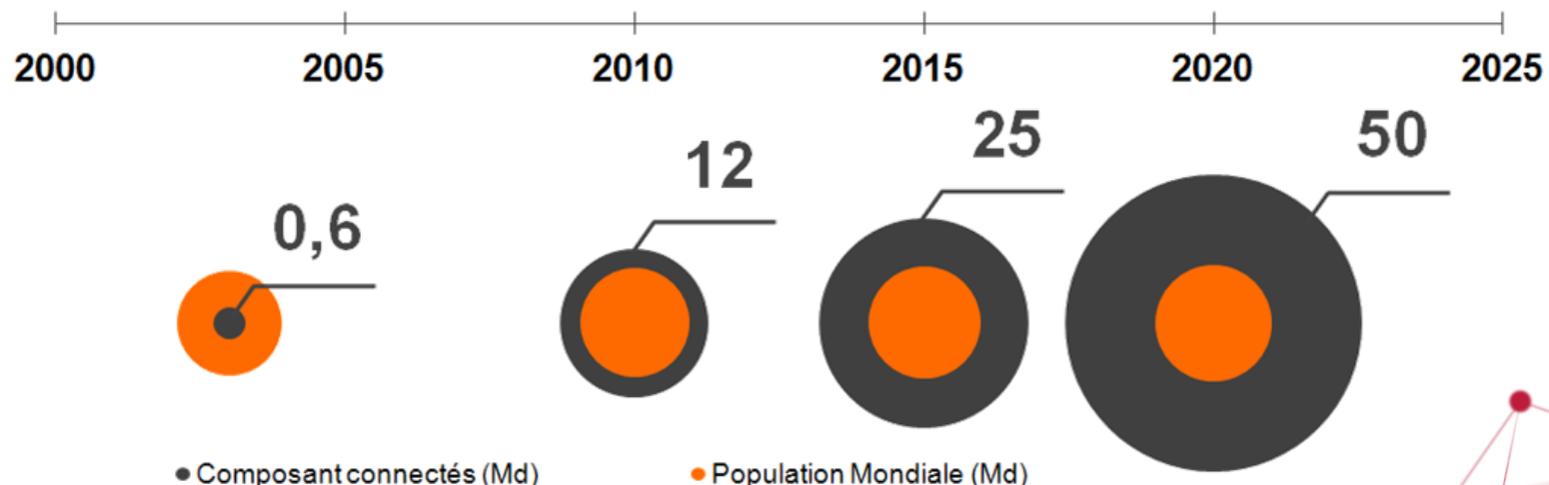
Depuis l'apparition du véritable premier Smart Phone en 2007, l'iPhone d'Apple®, et avec l'avènement de la Google® Glass, de la Galaxy Gear de Samsung®, et des capteurs intelligents (Smart Metering), le nombre

d'objets connectés est passé de 600 millions en 2003 à 12 milliards en 2010, et est estimé à 50 milliards en 2020 !

Au-delà du nombre d'appareils en interaction, il est intéressant de noter que ce sont les grandes puissances émergentes, à savoir les pays d'Amérique du Sud et d'Asie, qui généreront la plus forte croissance du flux d'échange de données sur les 4 prochaines années (resp. 49% et 31% CAGR). En Europe de l'Ouest, la prévision de croissance du flux d'échange de données est de 27% CAGR, et de 22% pour les USA. Il est également intéressant de noter la corrélation de cette projection avec les pays dont les entreprises ont déjà inscrit le Big Data dans leur stratégie : Argentine (89% des entreprises), Chine (86%) et Inde (83%) pour les plus avancées.

On observera donc un déplacement net des volumes de données vers les BRICS, ou les entreprises se préparent à tirer profit du phénomène Big Data.

ÉVOLUTION DU NOMBRE D'APPAREILS CONNECTÉS PAR RAPPORT À LA POPULATION MONDIALE



● Composant connectés (Md)

● Population Mondiale (Md)



LES CHIFFRES CLÉS :

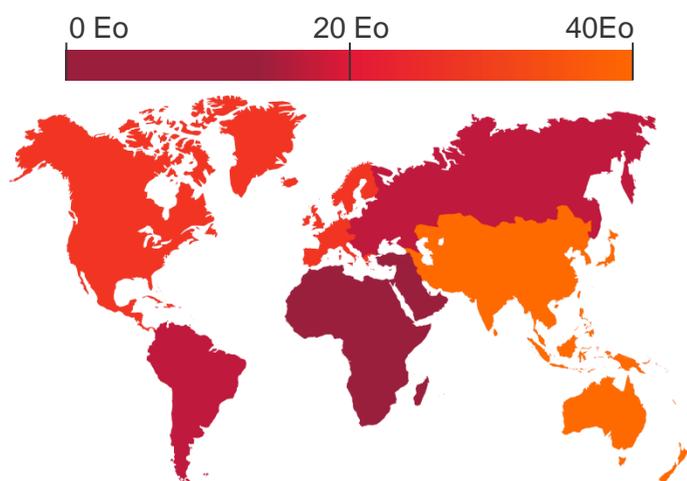
LE BIG DATA EST-IL UNE PRIORITÉ STRATÉGIQUE POUR VOTRE ENTREPRISE ?
(SONDAGE CISCO - 2013)

↑ **2/5** (44%) Oui

→ **2/5** (35%) En cours d'évaluation

↓ **1/5** (21%) Confusion/Incertitude

Etude réalisée auprès de 1800 entreprises dans 18 pays



Projection des flux de données en Exaoctet (10¹⁸ Octet) par mois en 2016

de l'information du cerveau avant 2020.

Il existe enfin une loi pour l'accélération du débit internet, basée sur le développement des points de routage du réseau mondial, la loi de Nielsen, qui prédit une croissance de 50% par an de la vitesse d'échange sur le réseau, et qui est démontrée depuis 1983.

Les technologies de transport, de stockage et de traitement de la donnée sont aptes à accueillir le tsunami attendu, et ce au moins pour les 15 prochaines années.

supporter l'impact du Big Data sur leur réseau, même si l'étude Cisco® révèle que 60% des entreprises ne seront capables de supporter cette croissance programmée sans travaux majeurs (1 DSI sur 2 estime que la charge de leur réseau doublera dans les 2 prochaines années).

DIMENSION 3 : MORPHOLOGIE / VARIÉTÉ

La segmentation des flux d'échange de données est fondamentale, et apporte des informations clés suivant les axes d'analyse. Elle est d'abord 5 fois plus importante chez le consommateur final que dans les entreprises en 2012, et le sera 7,4 fois plus en 2016 pour atteindre 97 Eo/mois pour le premier vs 13 Eo/mois pour le deuxième.

Certes le niveau de raffinement de la donnée est différent, car non structurée chez le consommateur, mais cet axe de segmentation démontre clairement que le

DIMENSION 2 : LE TEMPS / LA VÉLOCITÉ

Le développement du flux de données ne peut s'exécuter sans 3 composants qui suivent une dynamique similaire: le réseau, les moyens de stockage, et les capacités de calculs permettant le traitement de la donnée. Concernant ce dernier, la loi de Moore s'applique (doublement de la puissance de calcul tous les 18 mois) jusqu'en 2020 où un saut technologique majeur se produira (architectures 3D) qui décuplera les capacités des processeurs.

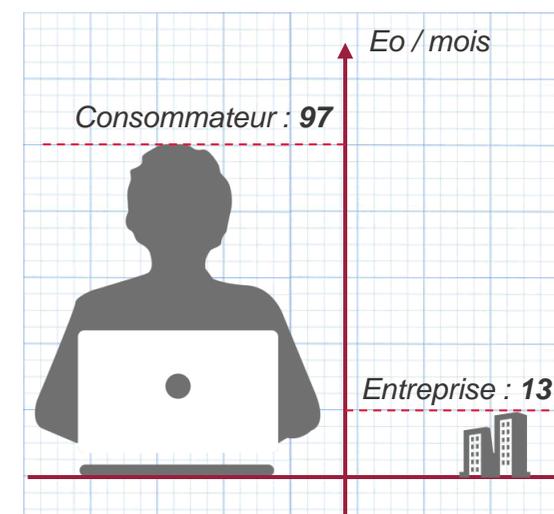
Concernant le stockage, en 2020 un disque de 200To coûtera 150\$, en 2025 les disques Po seront standards et en 2030 les disques Eo seront standards.

De la même façon une barette de mémoire flash dépassera les capacités de stockage

“ 1 DSI sur 2 estime que la charge de leur réseau doublera dans les 2 prochaines années.”

Au-delà de la maturité technologique, la répartition des infrastructures n'est pas linéaire mondialement. Et c'est un atout de l'Europe de l'Ouest et de l'Amérique du Nord que de s'appuyer sur un réseau fortement développé et qui restera au moins 3 fois plus puissant que ceux des pays des BRICS dans les prochaines années. S'agissant des technologies, la 4G est en phase de déploiement global en Europe et en Amérique du Nord, tout comme de la densité de centres de calculs en Cloud.

Les pays membres du G8 sont technologiquement plus à même de

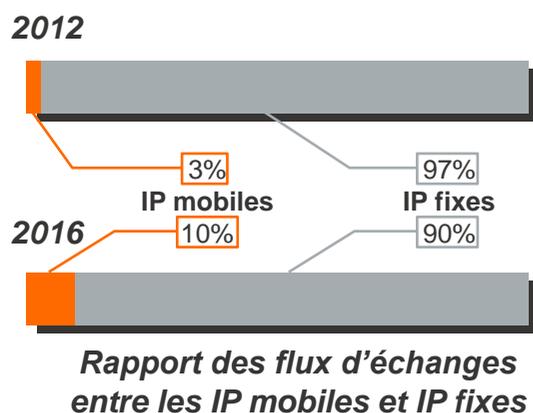


Représentation des segmentations des flux d'échange de données en 2016

point d'extraction de la richesse du Big Data se trouve bien à l'extérieur de l'entreprise.

Si on segmente par mode de connexion, la croissance annuelle des flux d'échange depuis les mobiles est la plus importante (+83% CAGR). En 2012 les flux d'échange de mobiles étaient 33 fois moins importants que les échanges en IP fixe, en 2016, ils ne le seront plus que 9 fois.

L'information sera de plus en plus disséminée et transportée (vs concentrée dans d'énormes entrepôts de données).

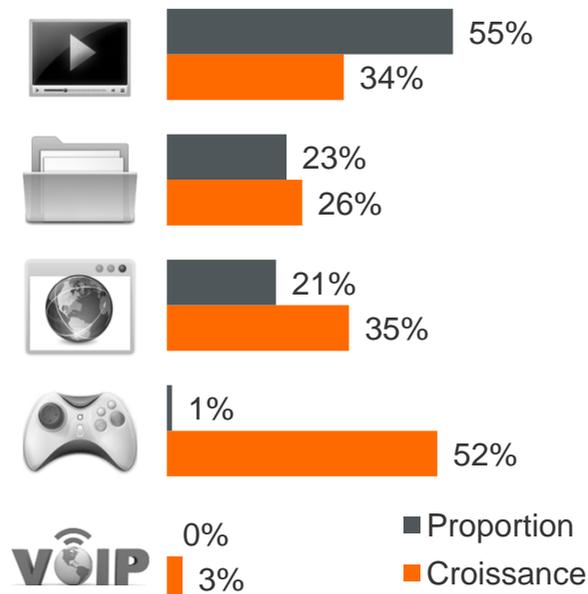


Enfin ici le dernier axe, sans doute le plus commenté, concerne la variété de formes nouvelles de l'information. L'information n'est plus directement exploitable (i.e. structurée), mais intégrée dans des objets complexes qui peuvent porter des informations multiples tels que les tweets, commentaires, blogs, mails, SMS, pour les moins complexes, mais également les vidéos, images ou bandes sonores.

De nouvelles données, générées en temps réel par des capteurs et des senseurs vont petit à petit enrichir ce conglomérat d'informations, souvent à forte valeur ajoutée, mais encore non exploitées.

Parmi les différents formats de données, la vidéo est celui dont la croissance sera la plus importante, globalement 2,6 fois plus élevée que le trafic web/email/données au cours des 4 prochaines années.

Evolution des différents formats de données



Et enfin là les nouveaux capteurs vont générer en temps réel un important volume de données non résilientes encore méconnu : CISCO® a nommé cette capacité d'interaction en temps réel avec ces objets le **Data In Motion**. Bien que les protocoles d'échanges ne soient pas standardisés, certains distributeurs comme Best Buy®

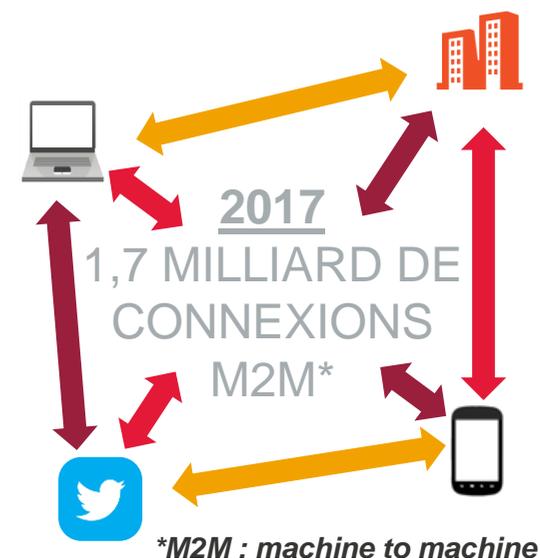
mettent en pratique ce concept.



DIMENSION 4 : L'INTERACTIVITÉ

L'enjeu se situe dans la capacité à corréler efficacement l'information issue de différentes sources. Les applications permettant de tirer les bénéfices du Big Data sont disponibles, comme les outils dits de data crunching qui permettent d'extraire de la donnée pertinente pour l'entreprise à partir de données polymorphes. L'objet de notre série des Nouvelles Architectures est d'aider les entreprises à identifier les moyens techniques devant s'adosser à ces outils, en y intégrant les problématiques de temps réel, de proximité géographique entre le traitement et la donnée, de stockage de l'information et de sécurité.

Cela se traduit par une modélisation de l'élasticité des centres de traitements internes aux entreprises vers des ressources du Cloud. Cela induit également une bonne maîtrise technique des architectures distribuées du Big data (i.e. le MapReduce),



POUR APPROFONDIR LE SUJET :

CISCO® CONNECTED WORLD TECHNOLOGY REPORT (CCWTR) (ICI)

CISCO® VISUAL NETWORKING INDEX: FORECAST AND METHODOLOGY, 2011-2016 (ICI)

CISCO® THE ZETTABYTE ERA (ICI)

FUTURE TIME LINE : 2020 (ICI)





“Une réflexion globale a été menée sur la place du big data au sein du SI de l’entreprise. Cela a permis de corréliser les choix d’infrastructures du SI, avec la trajectoire Big Data donnée aux projets.”

– EDOUARD, CONSULTANT

et des outils de gestion technique associés. **L’intégration de ces technologies en tant que commodités de service SI sont indispensables**, voire un prérequis à l’élaboration d’une stratégie d’entreprise ambitieuse de maîtrise de la donnée.

CONCLUSION

Au travers de cet article, nous avons voulu démontrer que le concept de Big Data est bien une réalité de notre quotidien, tangible, et que les moyens technologiques pour tirer les bénéfices de cette dynamique sont pleinement disponibles et à une échelle industrielle.

Les entreprises phares de la nouvelle économie que sont Google, Yahoo, Amazon, Facebook ou Twitter sont toutes bâties autour du Big Data.

A courte échéance, toutes les entreprises seront tenues d’avoir mis en place l’ensemble des moyens techniques indispensables au traitement de la donnée de masse.

Les entreprises des grandes puissances émergentes ont majoritairement établi leur plan de transformation, et ce malgré un

retard vis à vis de l’Europe ou l’Amérique du nord sur le développement de leur infrastructure. Une accélération de nos entreprises est aujourd’hui indispensable pour entrer définitivement dans cette nouvelle ère, dans tous les secteurs d’activités.

 **MAPS SURVEY**



ALLONS PLUS LOIN ENSEMBLE :

Afin d’enrichir nos informations quant aux **tendances de nos industries**, nous vous encourageons à **participer au survey** en lien avec cet article.

L’ensemble des données collectées nous permettra de construire une analyse étayant localement le point de vue exposé dans le présent article, et **que nous partagerons avec vous.**

ÉCHO-SYSTÈME

Nous accompagnons une entreprise évoluant dans le domaine de l’énergie dans la mise en place d’un système de smart metering à grande échelle.

Sur ce projet d’entreprise, la solution cible doit faire face aux contraintes techniques inhérentes au traitement d’une colossale volumétrie de données : maintien des performances des services et scalabilité des architectures. Ces contraintes colorent la conception de la solution d’une multitude de dispositifs intelligents générant des données et évoluant en permanence. Une telle volumétrie doit faire l’objet d’une réflexion approfondie.

Ces contraintes, amplifiées dans le temps, poussent les équipes d’architecture, au fur et à mesure de la vie de la solution, à remettre en cause certains choix technologiques. Cette étape de re-conception est itérative avec validation de chacun des choix majeurs par la réalisation de prototypes. Ils permettent de démontrer la pertinence de ces nouveaux choix et de challenger les différents produits existants.

L’accompagnement de mon client passe aussi par sa sensibilisation aux solutions Big Data qui émergent : les solutions de cache distribué, par exemple, sont maintenant des solutions éprouvées qu’il convient de prendre en compte en tant qu’architecte. Certaines de ces solutions ont ainsi été identifiées comme des pistes d’amélioration du système.

DATAIN MOTION

INTERNET OF THINGS

HADOOP

3V

CLOUD COMPUTING

ELASTIC DATACENTER / INFRASTRUCTURE

MAP REDUCE

INTERNET OF EVERYTHING

PROCHAIN NUMÉRO

Série Nouvelles Architectures

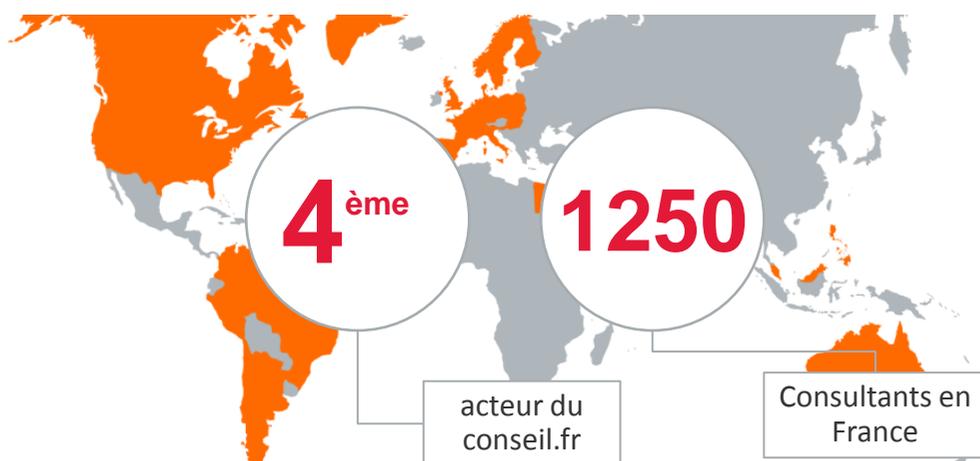
Maturité des offres de

Cloud Computing



LARGE SCALE ARCHITECTURES

La problématique Big Data doit être adressée dès les phases de conception de l'architecture pour pouvoir en maîtriser le positionnement et l'intégration dans le SI. C'est pourquoi dans le cadre de la méthodologie d'architecture Archim@de développée par CGI Business Consulting sur la base des grands frameworks internationaux, nous avons inséré un volet Big Data permettant d'évaluer très tôt la maturité, le besoin et les contraintes de mise en place d'une architecture technique en adéquation avec son contexte. Cette phase d'étude d'opportunité est le moment idéal pour faire émerger un besoin non identifié et lié aux environnements de grandes volumétries de données. Ces éléments sont structurants pour le design de l'architecture technique, garantir la performance de la solution mise en oeuvre, et son exploitabilité.



Notre équipe en chiffres

65
27

Consultants
certifiés TOGAF

25

solutions maîtrisées

15

années d'expérience

12000

J projets en 2012

CGI | **Business Consulting**
La force de l'engagement

L'AVIS DU MÉTIER

“ *L'information est le nouveau pétrole* ”

Aujourd'hui, les entreprises qui créent le plus de valeurs tirent leurs succès des ressources immatérielles. La chaîne de valeur de l'information ouvre différents leviers de maximisation. Or 90% des données accessibles sont non structurées ou semi-structurées. La mise en place de cette chaîne de valeur permet d'exploiter tout le potentiel de cette information brute: **créer**, détenir et protéger **les gisements**, **extraire**, transporter et mettre à disposition l'information brute, **raffiner** par croisement ou mettre en qualité l'information, pour **vendre le produit fini** et ses offres connexes à forte valeur ajoutée.

Pour exécuter ce cycle, de nombreux métiers doivent être combinés. Data Scientists, Consultants en organisation et Experts solutions doivent coopérer pour élaborer une stratégie Big data pertinente localement pour un marché.

Forts de retours d'expérience concrets et de partenariats innovants, nos équipes pluridisciplinaires aident les entreprises à définir leur stratégie Big Data : de la définition de nouveaux services, à l'expérimentation concrète de solution au travers de journées d'assessments, d'expérimentations, de définitions d'opportunités ou de plan de gouvernance.

NOUS CONTACTER : MAPS@CGI.COM

© CGI Business Consulting 2013
www.cgi.fr/metier/conseil