

Gestion durable des données : point sur les enjeux et proposition d'une démarche de pilotage de la performance appuyée sur un balanced scorecard thématique

Frédéric LAURA, Fabien COELHO¹ & Marie-Hélène DELMOND²

¹Mines ParisTech

²Groupe HEC

□ Résumé

L'objectif de cet article est de faire le point sur enjeux, acteurs et facteurs clés de succès de la mise en œuvre d'une gestion des données minimisant l'impact environnemental et de proposer une démarche de contrôle et de pilotage des performances applicable par les entreprises. Cet article s'appuie sur une recherche au cours de laquelle, après une revue bibliographique, nous avons analysé 70 appels d'offres concernant des solutions de stockage et mené 12 entretiens semi-directifs auprès d'experts du green IT et du secteur du stockage, de représentants de DSI et responsables d'exploitation. Les principales contributions de cet article sont d'une part de mettre l'accent sur la nécessité – au-delà de l'optimisation des infrastructures de stockage – de mettre en œuvre une gouvernance durable des données dont nous définissons les grands principes et de proposer, en nous appuyant sur l'approche du balanced scorecard, une démarche de contrôle et de pilotage des performances en ce domaine.

Mots clefs : Green IT, gouvernance des données, balanced scorecard, stockage des données, environnement.

□ Abstract

This paper aims at identifying key issues, contributors and success factors while deploying an ecologically responsible data management scheme. It also defines, using a balanced scorecard approach, a way to integrate environmental performance within the context of corporate strategy objectives.

Our research is based on the study of 70 requests for proposal on storage solutions, followed by 12 semi-directive interviews with green IT storage experts, CIOs and IT Executives.

The main contributions of this paper are: (1) to identify keys issues regarding sustainable data governance, going further than the deployment of technical storage solutions; (2) to define a balanced scorecard designed to drive and check the performance of a green storage policy

Key-words: Green IT, data governance, balanced scorecard, data storage, sustainability.

Introduction

Le thème du « green IT » rencontre aujourd'hui un intérêt profond de la part de la communauté professionnelle pour laquelle il se traduit par un ensemble d'opportunités, de risques et de nouvelles réglementations (Gartner, 2007). Dans ce débat, les enjeux sont à la fois écologiques et économiques et réunissent des acteurs multiples (entreprises, particuliers, fournisseurs de produits et services informatiques, institutions politiques et réglementaires, associations « vertes » spécialistes des TIC¹, associations de DSI comme en France le Cigref). Une étude récente (Gartner, 2010) recense un ensemble de plans d'actions menés par les entreprises et concernant à des degrés divers : efficacité énergétique et le refroidissement des data centers, recyclage et retraitement des déchets, éco-bâtiments, réduction des impressions, stockage des données, gestion des postes de travail et sensibilisation des employés à une utilisation responsable des technologies de l'information. La même étude indique que seulement une petite minorité d'entreprises (14%) disposent d'un système organisé d'évaluation des performances et d'indicateurs robustes leur permettant de piloter ces évolutions.

L'objectif de cette communication, centrée sur la problématique de la gestion durable des données, est :

- de faire le point sur les enjeux écologiques et économiques liés au stockage des données et sur les parties prenantes,
- d'identifier les bonnes pratiques actuelles, très orientées vers l'efficacité énergétique et spatiale et d'analyser les limites – à cet égard, nous montrerons comment à notre avis, ces pratiques ne trouveront leur efficacité que si elles s'accompagnent d'une gestion responsable et durable des données dont nous définissons les principes,
- de proposer, au travers d'un balanced scorecard thématique d'évaluation des pratiques de gouvernance durable des données, un outil de pilotage permettant aux entreprises d'intégrer et de développer les problématiques environnementales dans leur stratégie.

Pour cette étude, nous avons analysé 70 appels d'offres de solutions de stockage émis de mi 2007 à début 2009 et avons réalisé douze entretiens semi directifs auprès d'experts du green IT et du secteur du stockage, de représentants de DSI et responsables d'exploitation².

1. Enjeux et acteurs d'une gestion durable des données

1.1. La prolifération des données numériques et ses conséquences.

Selon une étude IDC³ (2008a), le total des données numériques pouvait être estimé en 2007 à 281 exa-octets (Eo)⁴, soit 281 millions de téra-octets. D'ici 2011, l'univers numérique devrait être égal à 10 fois la taille de celui de 2006, soit 1800 exa-octets pour une croissance annuelle de 60%. Le développement accéléré de l'usage des caméras (ex : vidéosurveillance⁵) et des chaînes de télévisions numériques sont parmi les facteurs les plus contributeurs de cette formidable croissance. Chacun d'entre nous peut se rendre compte simplement de la part de sa contribution « personnelle » au développement de cet univers en testant le programme « *Personal Digital Footprint Calculator* »⁶ développé en 2007 par IDC. Ce programme évalue nos activités numériques quotidiennes et fournit le résultat de nos volumétries de données générées par nos photographies, appels téléphoniques, téléchargements de musique ou vidéo et courriers électroniques⁷. Il convient d'ajouter à cet impact personnel notre « ombre numérique » générée par nos opérations bancaires et téléphoniques, nos relevés de consommations énergétiques, nos données médicales (bilans, scanners), la liste étant ici loin d'être exhaustive...

Les conséquences économiques et écologiques de cette prolifération des données numériques sont fortes. Selon l'étude IDC (2008b), « sur les 5 prochaines années (2008-2012), l'industrie du disque dur fournira 8 fois ce qu'elle avait fourni dans les 11 dernières années » ; « la capacité totale des disques durs livrés atteint 1 642 péta-octets, augmentant de 51,8 % d'une année sur l'autre ». La même étude précise que « avec un coût moyen mondial de 0,07 \$ par kilowatt/heure en 2007, le coût global pour alimenter et climatiser tous les disques durs dépasse le milliard de dollars ». Dès 2007, le Gartner Group évaluait la pollution globale des TIC à 2% des rejets en dioxyde de carbone liés à l'activité humaine⁸, soit autant que celle provoquée par toute l'aviation civile. En particulier, la consommation énergétique des data centers est de 10 à 30 fois supérieure

³ Le cabinet IDC a produit ces dernières années des études bien informées sur le thème de l'évolution des données numériques.

⁴ Progression des quantités informatiques en puissance de 1000 : 1 octet \approx 1 caractère, 1 Ko (kilo-octet) \approx 1000 octets (*un mail simple*), 1 Mo (mega-octet) \approx 10⁶ octets (*le texte d'un petit roman, une photo, un morceau de musique*), 1 Go (giga-octet) \approx 10⁹ octets (*un court métrage en DVD*), 1 To (tera-octets) \approx 10¹² octets (*un gros disque dur*), 1 Po (péta-octet) \approx 10¹⁵ octets, 1 Eo (exa-octet) \approx 10¹⁸ octets, Zo (zetta), Yo (yota)...

⁵ Législation française LOPSI (Loi d'Orientation et de Programmation pour la Sécurité Intérieure)

⁶ Disponible sur www.emc.com/digital_universe

⁷ On peut obtenir raisonnablement un total quotidien de 2 Go

⁸ Avec une hypothèse moyenne de production de l'électricité à partir d'hydrocarbures. Cette proportion varie fortement entre les pays : moins de 20% en France, 80% aux États-Unis, 93% en Inde... cf Agence International de l'énergie, *world energy outlook* 2006, <http://www.iea.org/>

¹ Technologies de l'Information et de la Communication

² Cette étude a été menée dans le cadre d'une thèse professionnelle de maîtrise spécialisée (Laura, 2009).

au m² à celle d'un bureau standard et est souvent mal mesurée par les entreprises qui l'intègrent dans leur facture globale d'énergie (Daim et al., 2009).

1.2. Les parties prenantes.

Les entreprises et les directions des systèmes d'information

La nécessité de la prise en compte des enjeux du développement durable, au départ considérée comme une contrainte par les entreprises, devient aujourd'hui une source d'opportunité et un impératif stratégique pour les entreprises (Raman, 2009). De même, la gestion durable des technologies numériques s'intègre de plus en plus dans le cadre de réflexion des entreprises. Dans les résultats de leur étude, Molla et al. (2009) indiquent ainsi que 65% des directions des systèmes d'information interrogées⁹ se déclarent concernées par la responsabilité de l'IT dans les émissions de gaz à effet de serre. Les mêmes auteurs évaluent en revanche les démarches comme étant à un stade peu avancé de maturité. En France, les associations professionnelles comme le CIGREF, le CRIP (Club des Responsables d'Infrastructure et de Production) ou le CESIT (Comité des Exploitants des Salles Informatiques et Telecom) réunissent leurs compétences et forment des groupes de travail « Green IT ».

Les acteurs politiques et réglementaires

Les programmes, normes et certifications (RoHS, WEEE, ISO 14001, EMAS, Code Of Conduct, label Energy Star¹⁰) ne cessent de se développer. Cependant, les accords mondiaux environnementaux semblent extrêmement complexes à réaliser¹¹. En Europe et en France, les institutions ont commencé depuis peu à encourager les initiatives vertes dans les technologies de l'information (par ex. plan de développement de l'économie numérique française France Numérique 2012). Parallèlement, elles poussent les autres secteurs à se servir des nouvelles technologies pour réduire leur impact et améliorer leur efficacité.

Les fournisseurs des technologies de l'information

Comme toute entreprise, ils sont confrontés aux mêmes problématiques et enjeux « verts ». Les constructeurs, intégrateurs et sociétés de services tentent pour la plupart de capter ces nouveaux marchés et de répondre aux nouvelles attentes afin de se distinguer de la concurrence.

⁹ Etude menée auprès de 143 entreprises américaines, australiennes et néo-zélandaises.

¹⁰ Restriction of the use of certain Hazardous Substances in electrical and electronic equipment; Waste Electronic and Electrical Equipment; Eco Management and Audit Scheme; Code de bonne conduite européen du Data center; label Energy Star de l'agence américaine pour la protection de l'environnement.

¹¹ Cf. les résultats mitigés du sommet de Copenhague organisé par l'ONU en décembre 2009

rence. Ils créent ou rejoignent des associations comme la SNIA (Storage Networking Industry Association¹²) qui a lancé la « *green storage initiative* » pour élaborer les réponses de l'industrie aux demandes réglementaires de plus en plus pressantes.

Les associations ou clubs d'utilisateurs

De récentes associations « vertes » spécialistes des T.I.C. comme *The Green Grid* ou *Climate Savers Computing*¹³, contribuent à sensibiliser les différents acteurs sur les problématiques « vertes » et à faire des recommandations pratiques. Les particuliers voient augmenter fortement leur consommation électrique liée aux technologies numériques et ils cherchent les moyens de la réduire. Ils souhaitent également montrer un comportement plus vertueux et respectueux de l'environnement, et privilégient par exemple l'achat d'équipements électroménagers ou de voitures « économes » en énergie. Mais, à part pour de nouvelles générations de micro-ordinateur, les rayons informatiques ne mentionnent que très rarement les performances énergétiques des composants de stockage et le consommateur peut rarement en faire un critère de choix « vert » (Laura, 2009).

2. L'évolution actuelle des pratiques

L'évolution actuelle montre une prise en compte croissante des problématiques liées à l'optimisation des infrastructures de stockage, qui contraste avec une faible maturité quant à la mise en œuvre d'une gouvernance durable des données.

2.1. Collecte de données.

Les 70 appels d'offres analysés sont indicatifs de deux années de consultations de clients grands comptes publics et privés de tous secteurs, menés au sein de l'entreprise (acteur important de l'offre de prestations de stockage) de l'un des co-auteurs¹⁴.

| Date d'émission | De mi 2007 à début 2009 |
|--|--|
| Répartition par secteur | 45% secteur public 35% grands comptes privés, tous secteurs d'activité 20% grandes PME |
| Montant d'investissement | Varié de 50.000€ à 8M€, la moyenne étant de 300.000€. |
| Données concernées | Dans la majorité des cas de toute nature : structurées, non structurées, multimédias et archives |
| Objectif principal de l'appel d'offres | Étendre ou consolider avec des parties virtualisées les capacités de stockage |
| Volumétrie de données | Varie de quelques téra-octets à plusieurs centaines de téra-octets |

Tableau 1 – Caractéristiques des 70 appels d'offres étudiés

¹² www.snia.org/forums/green

¹³ www.thegreengrid.org; www.climatesaverscomputing.org;

¹⁴ Nous ne ferons ici qu'une synthèse très rapide des résultats. Le lecteur intéressé pourra se reporter à Laura(2009) pour un exposé plus complet.

Sur ces appels d'offres, nous avons mené l'étude des critères les plus demandés et classé ceux-ci quant à leur impact potentiel sur l'environnement. Nous avons ainsi distingué 7 natures d'impact :

- Optimisation ou réduction de la consommation énergétique (E)
- Optimisation ou réduction de l'empreinte spatiale (S)
- Réduction des émissions de chaleur (C)
- Réduction des données ou de leur volumétrie (V)
- Optimisation des performances (P)
- Recherche d'une conformité légale « green » (L)
- Recherche d'une démarche de développement durable (D)

Le tableau 2 récapitule les critères susceptibles d'avoir un impact environnemental les plus demandés dans les appels d'offres analysés. Ceux-ci sont présents seulement dans 25 à 30% des cas au début de la période étudiée mais se généralisent rapidement par la suite (plus de 70% des cas à partir de mi 2008).

| Type de critère | Critère | Nature de l'impact |
|----------------------------------|---|--------------------|
| Composants matériels | Baie de stockage capacitive – multi-technologies disque | E, S, P |
| | Contrôleur multi protocoles ¹⁵ | E, S, P |
| | Disque dur basse consommation | E, S, C |
| | Virtual Tape Library (optimisation sauvegardes) | E, S, P |
| Architectures et fonctionnalités | Consolidation des espaces de stockage | E, S, P |
| | Déduplication (mode bloc ou fichier) | V, E, S |
| | Compression (disque dur ou bande magnétique) | V, E, S |
| | Thin Provisioning | V, E, S |
| | Architecture par « Tiers » ¹⁶ | E, P |
| | Virtualisation | E, S, C, D |

Tableau 2 – Critères à impact environnemental les plus demandés

Cette étude nous a également permis de faire plusieurs constatations :

- Bien que des critères susceptibles d'avoir un impact environnemental apparaissent dès le début de la période étudiée (mi 2007), la préoccupation des entreprises à cette époque est clairement de réduire les coûts de stock

¹⁵ Permettant d'interconnecter différents type de données par un réseau SAN -données structurées – transactionnelles- et NAS -données non structurées, fichiers- avec une seule baie de stockage.

¹⁶ Organisation en espaces de stockage aux performances et consommations énergétiques distinctives.

age en optimisant les infrastructures : la réduction de l'impact environnemental n'est pas une fin en soi, mais apparaît davantage comme un « bénéfice heureux » ou un sous-produit de l'optimisation économique. Ainsi par exemple, si les appels d'offres demandent d'indiquer la consommation énergétique totale, ce n'est pas tant en vue de la réduire que de pouvoir dimensionner techniquement les besoins en alimentation et en refroidissement des nouveaux équipements (dont on cherchera cependant à réduire le coût).

- Cette période est riche en innovations techniques allant dans le sens d'une optimisation économique et écologique du stockage. La diffusion et le rythme de compréhension et d'adoption de ces technologies par les entreprises clientes est remarquable, comme le montre le tableau 2, où les mêmes critères, présents seulement dans un quart des appels d'offres du début de période, se généralisent à 70% un peu plus d'un an après.

- En revanche, des critères liés à la gouvernance des données (réduction de la prolifération, gestion du cycle de vie, définition de la criticité des données et choix de solutions de stockage adaptées à cette criticité, etc.) sont quant à eux généralement très peu présents, la maturité des entreprises se révélant très faible en ce domaine.

2.2.Des avancées dans la maîtrise des architectures de stockage.

Au sein des infrastructures des systèmes d'information, les systèmes de stockages sont, tout autant que les serveurs, soumis à de forts impératifs d'efficacité. La réponse technologique est celle de la multiplicité, de la spécialisation et la redondance des systèmes, pour toujours plus de disponibilité, de performance et de protection – mouvement amplifié par la croissance constante et souvent mal maîtrisée des données.

L'étude des appels d'offres de solution de stockage nous a permis de constater une forte tendance des entreprises à la surévaluation du dimensionnement des solutions de stockage et à la sur-allocation générale de moyens de stockage, de sauvegarde, d'archivage, liées aux pressions des nouvelles réglementations, auxquelles s'ajoutent les précautions prises par les directions d'entreprise. En effet, l'enjeu est de protéger l'activité (plan de secours ou de continuité) déjà suffisamment menacée par la crise économique et la concurrence. S'adaptant à l'évolution actuelle de plus en plus contraignante de l'activité, les ressources de stockage fonctionnant souvent en continu et consomme de l'énergie, même lorsque leurs données ne sont pas accédées : comme les serveurs, le stockage est souvent utilisé à seulement 60 à 70 % de sa capacité. En moyenne, 25 à 30% des données sont sur un espace de stockage inapproprié à leur valeur (coût d'investissement initial du support et consommation énergétique récurrente gaspillée).

Les réponses techniques permettant de limiter l'empreinte écologique du stockage des données sont aujourd'hui disponibles¹⁷, les fournisseurs ayant intégré dans leur offre produit des solutions à base de technologies à basse consommation : utilisation de mémoire flash électronique comme disques (SSD), disques durs capacitifs (S-ATA, réellement plus économes), systèmes permettant de suspendre le fonctionnement des disques durs dont les données ne sont pas accédées (MAID), virtualisation de l'hébergement de machines ou de services, d'optimisation des capacités, ou encore de mutualisation.

L'association *The Green Grid* a proposé un indicateur de mesure de l'efficacité énergétique globale du data center : le PUE¹⁸ (*Power Usage Effectiveness*). De nombreux data centers ont actuellement un PUE de 3 ou plus ; l'optimisation des infrastructures devrait permettre d'atteindre un PUE situé entre 1,6 et 2.

2.3. Mais une gouvernance des données encore immature.

L'optimisation des infrastructures de stockage est certes une avancée intéressante, mais la question aujourd'hui est en fait de savoir si celle-ci pourra compenser d'une part la prolifération incontrôlée des données et d'autre part leur mauvaise gestion due à l'incapacité des acteurs à mettre en œuvre des processus de segmentation et de traitement différencié de ces mêmes données.

Les mécanismes de prolifération

Selon IDC/Bull (2008)¹⁹, la croissance des données gérées par les entreprises devrait encore atteindre plus de 30 % en moyenne en 2008 et en 2009. Les principaux responsables identifiés sont :

- la messagerie : généralisation de l'outil et diffusion de pièce jointe en forte augmentation ;
- la dématérialisation des processus et la mise en œuvre de nouvelles applications souvent critiques pour l'entreprise : applications métiers, applications mobiles, ERP, CRM, décisionnel ;
- l'augmentation des débits réseau, qui favorise les échanges de fichiers multimédias très consommateurs de stockage ;
- les contraintes réglementaires (archivages, conservations plus longues)
- la croissance organique de l'entreprise

À ce flux de nouvelles sources de données s'ajoute un ensemble de contraintes d'exploitation qui ont pour

effet de démultiplier la volumétrie. La SNIA (*Storage Networking Industry Association*) estime ainsi que le volume primaire de données gérées peut être multiplié dans un rapport allant jusqu'à dix fois le volume initial sous l'influence de facteurs comme le choix du type de redondance disque (RAID), la sur-allocation (réserve de volumétrie anticipant les besoins futurs d'une application), les copies instantanées utilisées à des fins de stockage local, la réplication ou miroir vers un site distant dans le cadre d'un plan de secours ou de continuité d'activité, la sauvegarde disque, l'archivage légal, les copies nécessaires aux tests et développements. IDC (2008a) propose l'exemple suivant : un courrier initial de messagerie avec une taille initiale de 0,1 Mo et une pièce jointe de 1 Mo peut ainsi générer une copie locale (+1,1 Mo), une sauvegarde (+1 Mo), un stockage sur serveur redondant (+2,1 Mo), une sauvegarde de l'ensemble sur bande archive (+4,4 Mo), soit au total 9,5 Mo. Envoyé à 4 destinataires, le message peut générer jusqu'à trente fois le volume initial de données.

Une gestion inadéquate des données

Les données gérées dans un data center ont de fait des caractéristiques différentes en termes de criticité, de besoins de disponibilité, de besoins transactionnels et de fréquence nécessaire de sauvegarde. On peut ainsi distinguer :

- les données les plus critiques, haute disponibilité (24h/24 7j/7), traitement transactionnel élevé (bases de données), sauvegarde intensive,
- les données non critiques, issues de bases de données transactionnelles n'exigeant pas de disponibilité 24h/24 ou de sauvegarde intensive,
- les données non critiques et plus rarement accédées, souvent de type non structuré, fichiers, sauvegarde plus espacée,
- les sauvegardes et restaurations sur site (sur bandes ou de disques),
- les sauvegardes hors site, souvent à des fins de reprise d'activité

Une gestion réellement optimisée du stockage ne sera atteinte que si les entreprises parviennent à mettre en regard de ces caractéristiques des choix d'équipement technique différenciés et pratiquent des arbitrages intégrant coût économique, impacts écologiques, gestion des risques.

Les principes d'une gouvernance durable des données

La gouvernance des données (data governance) doit identifier en amont les données numériques, tant les anciens flux que les nouveaux, identifier les responsables de leur production et de leur pertinence, évaluer l'importance de ces données pour l'organisation, mais aussi mettre en place des procédures ou des règles d'usages pour canaliser leur prolifération, gérer leur

¹⁷ Voir Laura (2009) pour une description étendue de ces solutions techniques.

¹⁸ PUE = rapport entre total d'électricité entrant dans la salle informatique et celle consommée par les systèmes informatiques (serveurs, stockage, infrastructures réseaux et télécoms), le reste correspondant aux systèmes de distribution et de refroidissement (alimentations, batteries, onduleurs, climatiseurs...)

¹⁹ Etude réalisée auprès de 100 grands comptes européens

cycle de vie et ajuster voire réduire les ressources de stockage nécessaires. L'objectif n'est pas la protection et la sauvegarde à tout prix (économique et écologique) – assurant un « risque zéro » d'ailleurs souvent illusoire – mais la responsabilisation des acteurs pour orienter et modifier les pratiques.

Khatri (2010) identifie cinq domaines de décisions de la data governance : les principes (rôle et importance de la donnée pour l'entreprise), la qualité (précision nécessaire selon l'usage), les méta-données (pour indexer et éventuellement retrouver la donnée quand on en a besoin), les accès (propriétaire, droits, plan de sauvegarde), et enfin la gestion complète du cycle de vie de la donnée (création, modification, archivage... et surtout destruction finale).

Les processus d'analyse et de décisions de la data governance apparaissent aujourd'hui prioritairement axés sur le contrôle et la qualité des données au travers de la création d'un référentiel pour l'entreprise²⁰, la maîtrise des risques et l'évaluation de la valeur des données (Tallon et Scannel, 2007), la gestion de la sécurité et de la protection de l'information, la conformité aux obligations légales (*data compliance* et *data privacy*)²¹, et la gestion du cycle de vie (Weller 2009).

Nous proposons de compléter ces processus pour intégrer les enjeux « éco-responsables », et contribuer au développement durable de l'entreprise (enjeux économiques, environnementaux et sociaux) :

- Conformité des données : assurer la conformité légale des données mais aussi migrer les données sensibles vers le support adéquat, supprimer définitivement les données qui l'exigent ou qui ne présentent plus de valeur ajoutée ;
- Protection des données personnelles (*data privacy*) : renforcer les règles d'appartenance et d'accès ; respecter les obligations informatique et libertés²², vérifier les durées de conservation, supprimer les données sensibles et périmées ;
- Sécurité des données (*data protection*) : définir une politique globale de protection, de sauvegarde, et de duplication pour le plan de continuité d'activité, qualifier les données éligibles pour chacun des types de protection ;
- Gestion du cycle de vie des données : vérifier à chaque étape du cycle de vie l'alignement des technologies utilisées à la nature des données et aux exigences des activités métiers en tenant compte de l'impact écologique, filtrer les données régulièrement pour éviter l'accumulation inutile, profiter des évolutions et migrations pour canaliser les volumétries

et consolider les données sur une infrastructure de stockage unique pour gagner de l'espace dans le data center et réduire la consommation énergétique globale ;

- Gestion des infrastructures de stockage : recenser la conformité des infrastructures de stockage avec les réglementations comme RoHS, WEEE et développer les procédures de traitement et de recyclage des matériels usagés

Pour être efficace, la gouvernance des données doit être considérée et implémentée au plus haut niveau de l'entreprise, afin de maîtriser en amont et transversalement leur prolifération dans toute l'infrastructure de stockage et réduire les « déchets » que sont les nombreuses données dormantes et persistantes. Certains auteurs préconisent la création d'une direction de la gouvernance des données (Weller, 2009). En son absence, ce rôle est souvent assumé par la DSI, qui peut en particulier :

- *Identifier les propriétaires des données* : des données non utilisées et sans propriétaire deviennent de véritables déchets électroniques, coûteuses car elles consomment de façon récurrente de l'énergie et contribuent à saturer les espaces de stockage et de sauvegarde (voire même d'archivage), développant ainsi constamment le besoin de nouvelles ressources, et sans personne pour prendre la décision de les éliminer.
- *Améliorer la compréhension des « services de stockage »* : afin de rationaliser en amont les demandes de ressources de stockage de ses clients internes, souvent abusives en termes de moyens (volumétrie trop grande, disponibilité surestimée, performance sous utilisée), la DSI peut proposer un catalogue de services de stockage adapté aux besoins métiers, compréhensible par les clients, et pouvant intégrer des niveaux de qualité de service basés par exemple sur ITIL (Chamfrault, 2006).

3. Des outils de pilotage adaptés

Bien que les enjeux environnementaux soient de plus en plus intégrés aux préoccupations des entreprises (Raman, 2009, Barret, 2010) et à celles des Directions des Systèmes d'Information (Molla et al. 2009), ceux-ci se traduisent encore rarement dans les systèmes de gestion et de management, ce qui peut représenter un sérieux handicap à leur réelle prise en compte (Figge et al., 2002).

3.1.L'approche du Balanced Scorecard

L'approche du balanced scorecard (BSC) (Kaplan et Norton, 1996, 2000) a été proposée par un ensemble d'auteurs comme prometteuse pour permettre la prise en compte des enjeux environnementaux et sociaux.

Pour Epstein et Wiener (2001), le balanced scorecard peut aider les organisations à mettre en œuvre une stratégie environnementale et mettre en lumière le lien

²⁰ Master Data Management et Meta Data Management

²¹ Normes et réglementation : Sarbanes Oxley, protocole Bâle II, Loi sur la Sécurité Financière, Loi de sécurité dans l'économie numérique, etc.

²²Loi du 6 janvier 1978 instaurant la CNIL <http://www.cnil.fr/>

entre les objectifs stratégiques environnementaux, les plans d'actions et la performance et les résultats obtenus. Sur le plan opérationnel, ils voient dans le BSC une méthode appropriée pour développer et sélectionner des indicateurs de performance environnementale en lien avec la stratégie de l'entreprise. Johnson (1998) insiste sur l'intérêt du modèle de causalité entre les indicateurs d'action des projets en faveur de l'environnement et les indicateurs financiers et stratégiques de résultat et voit dans le BSC un moyen de sensibiliser les dirigeants au lien entre la performance environnementale et la performance globale de l'entreprise.

Figge et al. (2002) proposent trois modes d'intégration des enjeux environnementaux et sociaux au BSC : une première option consiste à compléter le BSC existant en y ajoutant des objectifs et indicateurs complémentaires liés au développement durable ; la seconde option est de créer une perspective supplémentaire (aux côtés des quatre perspectives classiques – financière, clients, processus et apprentissage) ; la troisième consiste à établir un BSC thématique consacré à la performance durable. Pour Zingales et al. (2002), la création d'une cinquième perspective apparaît contradictoire avec l'un des fondements de la méthode, qui est de mettre en évidence la chaîne de causes à effets qui lie les indicateurs financiers et non financiers. Dias-Sardinha et Reijnders (2005) ont utilisé un BSC thématique pour évaluer la performance environnementale et sociale de 13 entreprises au Portugal et soulignent l'intérêt de cette approche pour mieux comprendre les forces et faiblesses des liens entre les objectifs, les indicateurs, les projets et les résultats atteints.

Le développement d'un BSC thématique, focalisé sur la problématique de la gestion durable des données, nous paraît présenter l'intérêt, outre de permettre d'évaluer la performance et l'atteinte des objectifs, de focaliser l'attention sur les moyens, facteurs clés et plans d'action à mettre en œuvre, notamment au travers des perspectives processus et apprentissage. Speckbacher et al. (2003) définissent trois stades de maturité du BSC, le premier mettant l'accent sur la mesure de la performance (définition d'objectifs et d'indicateurs quantifiables), le second analysant les ressorts de l'élaboration de la performance au travers des chaînes de causalité de la carte stratégique, le troisième, plus abouti, mettant l'accent sur le processus de management et le déploiement de la stratégie via la mise en œuvre d'initiatives (plans d'action). Mendoza et al. (2005) illustrent ainsi dans l'analyse du cas de Sodexo comment le BSC a été utilisé pour structurer de nouvelles orientations stratégiques, opérationnaliser leur mise en œuvre et mobiliser les acteurs concernés.

3.2. Proposition d'un Balanced Scorecard thématique

Nous élaborons ici un BSC thématique visant à accompagner, favoriser et évaluer la performance d'un projet

de mise en œuvre du déploiement d'une politique de stockage durable dans l'organisation. Ce BSC constitue la synthèse de nos propositions pour une démarche de contrôle et de pilotage des performances en matière de green storage, articulée autour des 4 perspectives classiques (financière, clients, processus, apprentissage).

L'étude des appels d'offres de stockage avait mis en évidence des progrès sur le plan des technologies, des architectures et des fonctionnalités, mais également le manque de maturité des entreprises sur la question de la gouvernance des données. Nous avons complété cette étude par une série d'entretiens avec des experts représentants du monde de l'entreprise (clients) et de celui de l'offre (prestataires). Le tableau 3 reprend la typologie des personnes interviewées.

| |
|---|
| <p>Représentants du monde de l'entreprise (clients, 6 entretiens) :</p> <p>DSI grande entreprise (effectif 7000p., secteur services) DSI grande entreprise (effectif 12000p., secteur loisirs) DSI PME (secteur services web) Représentants d'associations de DSI (Cigref, CRIP - Club des Responsables d'Infrastructure de Production, 3 entretiens)</p> |
| <p>Représentants du monde de l'offre (prestataires et fédérations, 6 entretiens) :</p> <p>Responsable stratégie des offres de stockage (SSII), 1 entretien Responsable offre data center (SSII), 1 entretien Consultants spécialisés gestion des données (SSII), 3 entretiens FEDISA (Fédération de l'ILM, du stockage et de l'archivage), 1 entretien</p> |

Tableau 3 – Typologie des personnes interviewées

L'objectif de ces entretiens et leur principal apport a été de confronter l'opinion de ces experts et de dégager les lignes de force des évolutions actuelles dans une vision prospective et qualitative de ce vers quoi pourra évoluer la gestion durable des données et de leur infrastructure.

En ce sens, le Balanced Scorecard que nous proposons en figure 1 n'est en aucun cas le reflet des pratiques actuelles mais constitue plutôt un outil de définition et d'incitation à la mise en œuvre d'une stratégie environnementale comme le proposent Epstein et Wiener (2001).

Cette démarche est convergente avec les travaux du Cigref (Cigref, 2009) qui identifie la dimension écoresponsables des systèmes d'information et préconise la mise en place de bonnes pratiques et d'indicateurs de gouvernance dans ce domaine. Plus globalement, le développement de balanced scorecard thématiques ciblés sur les enjeux environnementaux peut contribuer aux efforts de reporting environnemental soutenus par des organisations comme le GRI (Global Reporting Initiative²³), notamment en explicitant l'approche managériale et en définissant des indicateurs de performance (GRI, 2006).

²³ Site web : www.globalreporting.org

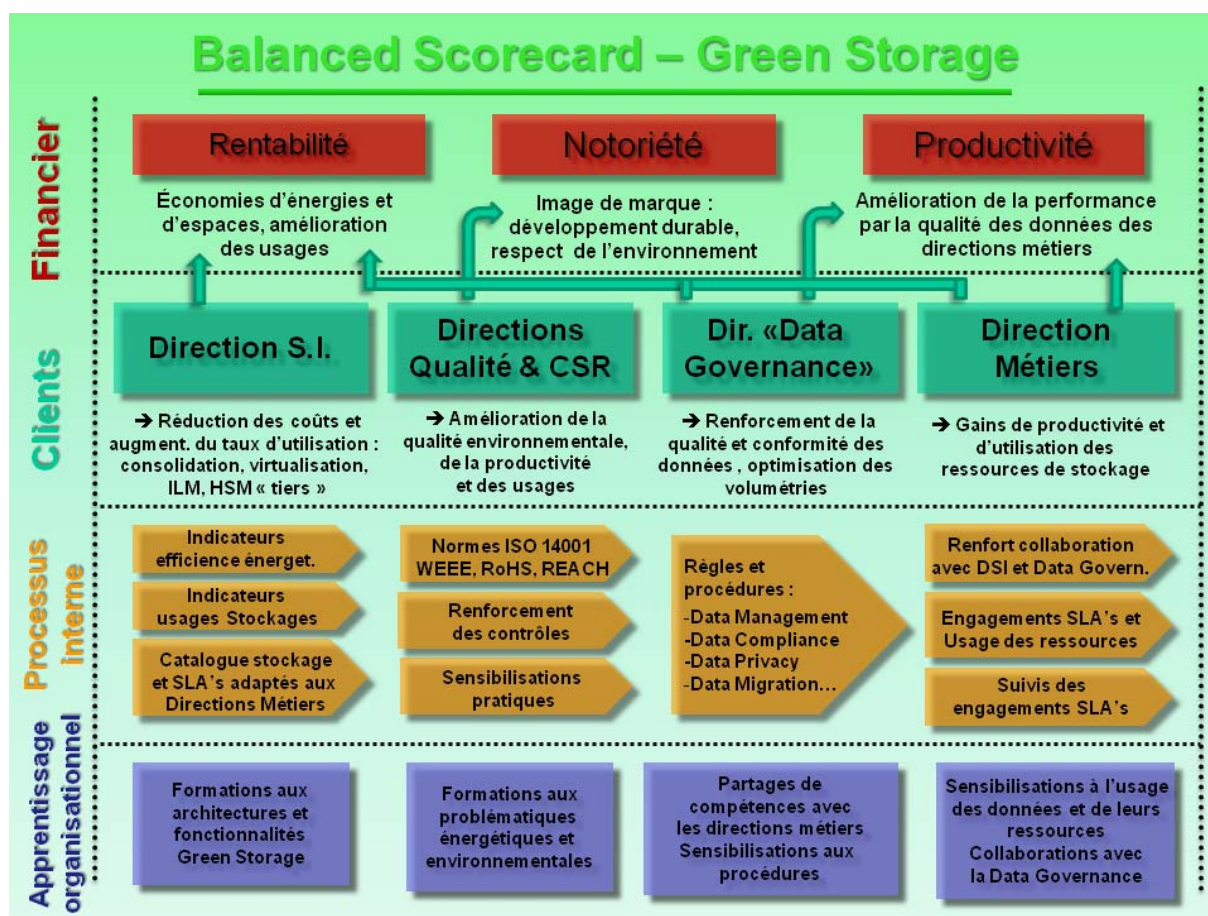


Figure 1 – Proposition d'un BSC thématique pour une gestion durable des données

Perspective financière

Trois objectifs majeurs définis par les entreprises concernant une gouvernance durable des données se dégagent de nos observations :

- Rentabilité (économies d'énergies et gains d'espaces dans le *data center* et amélioration du taux d'utilisation des espaces de données),
- Notoriété/Image (communication sur les efforts faits dans les normes environnementales et le développement durable),
- Productivité (amélioration des performances des directions métiers opérationnelles et fonctionnelles du fait de l'amélioration de la qualité des données)

Les objectifs financiers traduisent les résultats à atteindre (les effets attendus) par les dirigeants et les initiateurs de la stratégie ; ils sont complétés par les attentes des clients du projet de mise en œuvre du déploiement de la politique de stockage durable dans l'organisation.

Perspective clients

Le projet de déploiement de la politique de stockage durable dans l'organisation s'adresse avant tout aux acteurs internes de l'organisation. Nous avons donc

choisi d'analyser la perspective clients en nous centrant sur ces clients internes, à l'instar des BSC thématiques réalisés pour des directions fonctionnelles au service des autres directions de l'entreprise. Différents clients internes peuvent être identifiés, qui montrent la variété des acteurs concernés par la politique de gouvernance des données et de leurs enjeux.

La *Direction des Systèmes d'Information* (DSI) vise la réduction des coûts et la productivité de ses équipements techniques (augmentation du taux d'utilisation)

La *direction de la qualité* (ou, si elle existe, la direction en charge de la RSE) a pour objectifs la sensibilisation de l'ensemble de l'entreprise aux enjeux environnementaux et l'amélioration de la qualité environnementale, de la productivité et des usages. Les résultats attendus peuvent être par exemple l'obtention d'une certification ISO 14001.

Les *directions métiers* (opérationnelles et fonctionnelles) recherchent des gains de productivité et d'utilisation des ressources de stockage, ainsi que l'amélioration de la qualité des données.

Enfin, il apparaît nécessaire de réunir les représentants des différents acteurs au sein d'un *comité de gouvernance des données* (peut-être prélude à une « direction

des données », qui a pour objectif de renforcer la qualité et la conformité des données, mais aussi l'optimisation de la gestion de la ressource.

Perspective processus

La perspective processus permet de définir les actions opérationnelles à mener pour atteindre les résultats définis dans les perspectives financières et clients. Nous avons regroupé ces processus selon les acteurs concernés.

Pour la DSI, il s'agit d'améliorer l'efficacité des infrastructures (indicateurs d'efficience énergétique et de taux d'utilisation du stockage), mais aussi d'élaborer, en partenariat avec les Directions Métiers, un catalogue d'offres et de prestations « stockages » plus compréhensible et adapté aux activités spécifiques des différents métiers.

La Direction Qualité met en place les nouveaux contrôles nécessaires à l'atteinte des objectifs fixés par les normes : elle sensibilise chaque activité de l'entreprise par la diffusion de documents personnalisés et concrets sur les bénéfices apportés par les normes.

Le comité de gouvernance des données précise des procédures et règles strictes dans le management des données, contrôle leur conformité légale, définit les durées de conservation, le respect des données privatives, et stimule la migration de données vers les espaces de stockage adaptés.

Les Directions Métiers, en collaboration avec le comité de gouvernance des données, s'engagent sur des niveaux de service (SLA) ; ces contrats de niveaux de service rationalisent les besoins par des engagements clairs de chaque partie et diffusent une plus grande compréhension des usages permis par les ressources mises à disposition.

Perspective apprentissage

La DSI forme de nouveaux acteurs, « *Data Manager* » et « *Storage Manager* », qui représenteront la DSI et intégreront la Direction *Data Governance* : ces nouveaux responsables suivent des formations spécifiques liées aux activités des Métiers et se forment aussi aux architectures et fonctionnalités « Green Storage ».

La Direction Qualité se forme aux problématiques environnementales, notamment dans les domaines énergétiques et retraitements des déchets : elle forme des « évangélistes » qui sensibilisent chaque direction métier dans leur activité.

La Direction *Data Governance* diffuse à toutes les directions, les règles et procédures constituant la nouvelle politique sur les données par des documents clairs, exposant les procédures de façon pédagogique : elle se forme en échangeant ses compétences avec les directions métiers.

Les Directions Métiers sont formées à un usage plus rationnel des ressources de stockage et des espaces de données : elles sont sensibilisées aux conséquences financières du non-respect des règles et des SLA.

Conclusion

Les données dormantes, redondantes, sans propriétaires, sont autant de « E-déchets » qui envahissent les infrastructures de stockage et donc les serveurs et les data centers. Elles mobilisent des ressources matérielles et consomment de façon récurrente de l'électricité inutilement. Elles polluent aussi considérablement les activités des directions métiers en altérant la qualité globale des données, souvent gérées en silos, difficiles à identifier et à accéder. La gouvernance durable des données peut assumer le traitement, le recyclage ou la suppression des données, en accroître globalement la qualité et la valeur, et limiter les impacts écologiques. Identifier, contrôler, supprimer ces données-déchets est un enjeu capital pour les directions des systèmes d'information.

Dans cet article, nous avons voulu faire le point sur la problématique de la gouvernance des données et montrer que leur gestion durable est d'abord une question de politique de gestion maîtrisée de la prolifération des données, et ensuite seulement une question technique de mise en œuvre de solutions adaptées aux besoins identifiés. L'inversion actuelle de cette proposition a pour effet de mettre en œuvre des solutions techniques pour des volumes et des coûts déraisonnables sans rapport avec les besoins si ceux-ci n'ont pas été évalués, et sans aucune maîtrise si des procédures ou des bonnes pratiques n'ont pas été mise en place pour éviter l'explosion mécanique et souvent inutile des volumes générés par les utilisateurs.

Les entreprises ont besoin d'outils de pilotage de la performance leur permettant de synthétiser ces enjeux et d'orienter leur action. A cet effet, nous avons proposé un Balanced Scorecard thématique dont l'objectif est de mettre en avant ces problématiques de gouvernance, de souligner les enjeux et impacts attendus, d'identifier les projets à mener et de définir un cadre d'évaluation de la performance en ce domaine.

Références

- Anthony R.N. (1965), *Planning and control systems: a framework for analysis*, Harvard Business School Press, Boston, 180 p.
- Baret P., (2010), "Quatre temps pour implémenter une stratégie environnementale", *Management et Avenir*, Issue 29, p242-257.
- Chamfrault T., Durand C. (2006), "ITIL et la gestion des services", Dunod, Paris.
- Chen A.J.W., Boudreau M-C., Watson R.T. (2008), "Information systems and ecological sustainability", *Journal of Systems and Information Technology*, vol.10 n°3, pp. 186-201.

- Cigref, (2009), "Systèmes d'information éco-responsable : l'usage des TIC au service de l'entreprise durable", rapport disponible sur www.cigref.fr
- Daim T., Justice J., Krampits M., Letts M., Subramanian G., Thirumalai M., (2009), "Data center metrics: An energy efficiency model for information technology managers", *Management of Environmental Quality*, vol. 20 n°6, pp. 712-731.
- Delmond, M-H., (2008), "Pilote de la performance et modes de contrôle de la Direction des Systèmes d'Information: analyse des échecs et évolutions", Actes du 14ème Colloque de l'Association Information et Management (AIM), Marrakech, Maroc
- Dias-Sardinha I., Reijnders L (2005), "Evaluating Environmental and Social Performance of Large Portuguese Companies: A Balanced Scorecard Approach", *Business Strategy and the Environment*, 14, pp. 73-91.
- Epstein M.J., Wisner P.S. (2001), "Using a balanced scorecard to implement sustainability", *Environmental Quality Management*, winter, pp.1-10.
- Figge F., Hahn T., Schaltegger S., Marcus W. (2002), "The Sustainability Balanced Scorecard: Linking Sustainability Management to Business Strategy", *Business Strategy and the Environment*, 11, pp. 269-284.
- Gartner (2007), "Green IT: The New Industry Shock Wave", 7 December, n° G00153703.
- Gartner (2009), "Data Centers Focus on Green IT, but Many Neglect Metrics", 29 July, n° G00169058.
- Gartner (2010), "Green IT Initiatives Are Moving beyond Power and Cooling Efficiency", 8 February, n° G00173836.
- GRI (Global Reporting Initiative), 2006, "Lignes directrices pour le reporting développement durable", version 3.0, rapport disponible sur www.globalreporting.org
- Hervani A.A., Helms M.M. (2005), "Performance measurement for green supply chain management", *Benchmarking*, vol. 12 n°4, pp. 330-353.
- Huang C.D., Hu Q. (2007), "Achieving IT-business strategic alignment via enterprise-wide implementation of balanced scorecards", *Information systems management*, vol 24, n°2, janvier, pp. 173-184.
- IDC (2008a), "The Diverse and Exploding Digital Universe", IDC white paper.
- IDC (2008b), "The Real Costs to Power and Cool All the World's External Storage", IDC1807971
- IDC/Bull (2008), "L'optimisation du stockage : enjeux, difficultés et stratégies gagnantes des entreprises européennes", livre blanc disponible sur <http://www.bull.com/fr/storage/wp.html>.
- Johnson S.D., (1998), "Identification and Selection of Environmental Performance Indicators: Application of the Balanced Scorecard Approach", *Corporate Environmental Strategy*, vol.5 n°4, pp.34-41.
- Kaplan R.S., Norton D.P., (1996), *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston.
- Kaplan R.S., Norton D.P., (2000), *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*, Harvard Business School Press, Boston.
- Khatri V., Brown C. (2010), "Designing Data Governance", *Communications of the ACM*, vol.53 n°1, pp.148-152.
- Laura F. (2009), "Green Storage: Enjeux et Facteurs clés de succès", Thèse professionnelle de Mastère Spécialisé MSIT, HEC Mines ParisTech.
- Mendoza C., Delmond M-H., Giraud F., Löning H., De Font Réaulx A. (2005), "Tableaux de Bord et Balanced Scorecards", Groupe Revue Fiduciaire, Paris.
- Molla A., Pittayachawan S., Corbitt B., Deng H. (2009), "An International Comparison of Green IT Diffusion", *International Journal of e-Business Management*, vol.3, n°2, pp.3-23.
- Olson E.G. (2008), "Creating an enterprise-level green strategy", *Journal of Business Strategy*, vol. 29 n° 2, pp. 22-30.
- Raman A.P. (2009), "The New Frontiers", *Harvard Business Review*, July-August, pp.130-137.
- Speckbacher G., Bischoff J., Pfeiffer T. (2003), "A descriptive analysis on the implementation of balanced scorecard in German-speaking countries", *Management Accounting Research*, vol.14, pp. 361-388.
- Tallon P.P., Scannel R. (2007), "Information Life Cycle Management", *Communications of the ACM*, vol. 50, n° 11, pp. 65-69.
- Weller A (2009), "Data governance: Supporting data-centric risk management", *Journal of Securities Operations & Custody*, vol.1 n°3, pp.250-262.
- Zingales F., O'Rourke A., Orsatto R.J. (2002), "Environment and Socio-Related Balanced Scorecard: Exploration of Critical Issues", CMER, Working Paper n°47, Insead.
- Zingales F., Hockerts K., (2003), "Balanced Scorecard and Sustainability: Examples from Literature and Practice", CMER, Working Paper n°30, Insead