

CONCOURS INTERNE DE BIBLIOTHÉCAIRE TERRITORIAL

SESSION 2014

ÉPREUVE DE NOTE DE SYNTHÈSE PORTANT SUR LES SCIENCES EXACTES
ET NATURELLES ET LES TECHNIQUES

À LIRE ATTENTIVEMENT AVANT DE TRAITER LE SUJET

- ↪ Vous ne devez faire apparaître aucun signe distinctif dans votre copie, ni votre nom ou un nom fictif, ni votre numéro de convocation, ni signature ou paraphe.
- ↪ Aucune référence (nom de collectivité, nom de personne,...) autre que celle figurant le cas échéant sur le sujet ou dans le dossier ne doit apparaître dans votre copie.
- ↪ Seul l'usage d'un stylo soit noir soit bleu est autorisé (bille, plume ou feutre). L'utilisation d'une autre couleur, pour écrire ou souligner, sera considérée comme un signe distinctif, de même que l'utilisation d'un surligneur.

**Le non-respect des règles ci-dessus peut entraîner l'annulation de la copie par le jury.
Les feuilles de brouillon ne seront en aucun cas prises en compte.**

Ce sujet comprend 28 pages, y compris celle-ci.

S'il est incomplet, en avertir un surveillant.

CONCOURS INTERNE DE BIBLIOTHÉCAIRE TERRITORIAL

SESSION 2014

ÉPREUVE D'ADMISSIBILITÉ

Durée : 3 heures / Coefficient : 2

Note de synthèse à partir d'un dossier portant sur les sciences exactes et naturelles et les techniques.

Sujet :

Vous êtes bibliothécaire territorial de la ville de X. Votre directeur vous demande de rédiger à son attention, exclusivement à l'aide des documents joints, une note de synthèse sur les centres de traitement de données (datacenters).

SOMMAIRE DU DOSSIER

- Document 1 :** La déferlante des octets, *Fabrice Demarthon, Denis Delbecq, Grégory Fléchet ; CNRS : le journal ; n°269 ; nov-déc. 2012 – 4 pages*
- Document 2 :** Les datacenters ont de l'énergie à revendre, *Julie Nicolas ; Le Moniteur des travaux publics et du bâtiment ; n°5649 ; 2012 – 2 pages*
- Document 3 :** Éléments essentiels devant figurer dans un contrat de prestation de services de Cloud computing, Site www.cnil.fr <consulté le 13/02/2014> – 3 pages
- Document 4 :** Les Datacenter, *Performance des architectures IT : comprendre, résoudre et anticiper / Pascal Grojean, Médéric Morel, Simon-Pierre Nolin - Dunod, 2011 – 6 pages*
- Document 5 :** Telecity Group : Datacenter, Condorcet (extraits de la brochure d'information), Site www.telecitygroupe.fr <consulté le 13/02/2014> – 2 pages
- Document 6 :** Etude Ge-Tic : l'éventualité de la mise en place d'un datacenter dans l'Aube, Site ge-tic.fr mis à jour en septembre 2011 – 1 page
- Document 7 :** Datacenters et création d'emplois : une équation à multiples variables, *Anne-Marie Rouzeré ; Le MagIT : l'informatique pour et par les pros ; 22 juin 2012 – 3 pages*
- Document 8 :** OVH.com offre de la mémoire aux cerveaux de l'université de Lille 1, *Stéphane Hubin, La Voix du Nord, 7 octobre 2013 – 1 page*

Document 9 : Data centers : la donnée écolo, *Morgane Tual, Libération*, 14 avril 2013 – 2 pages

Document 10 : Se chauffer grâce à l'énergie des serveurs informatiques, *Audrey Garric, Le Monde*, 2 juillet 2013 – 1 page

Documents reproduits avec l'autorisation du CFC

Certains documents peuvent comporter des renvois à des notes ou à des documents non fournis car non indispensables à la compréhension

INTERNET,
SCIENCE,
TECHNOLOGIES...



Sur Wikipedia

9 000

nouveaux articles créés

La déferlante des OCTETS



Sur Youtube

104 000

heures de vidéos
mises en ligne



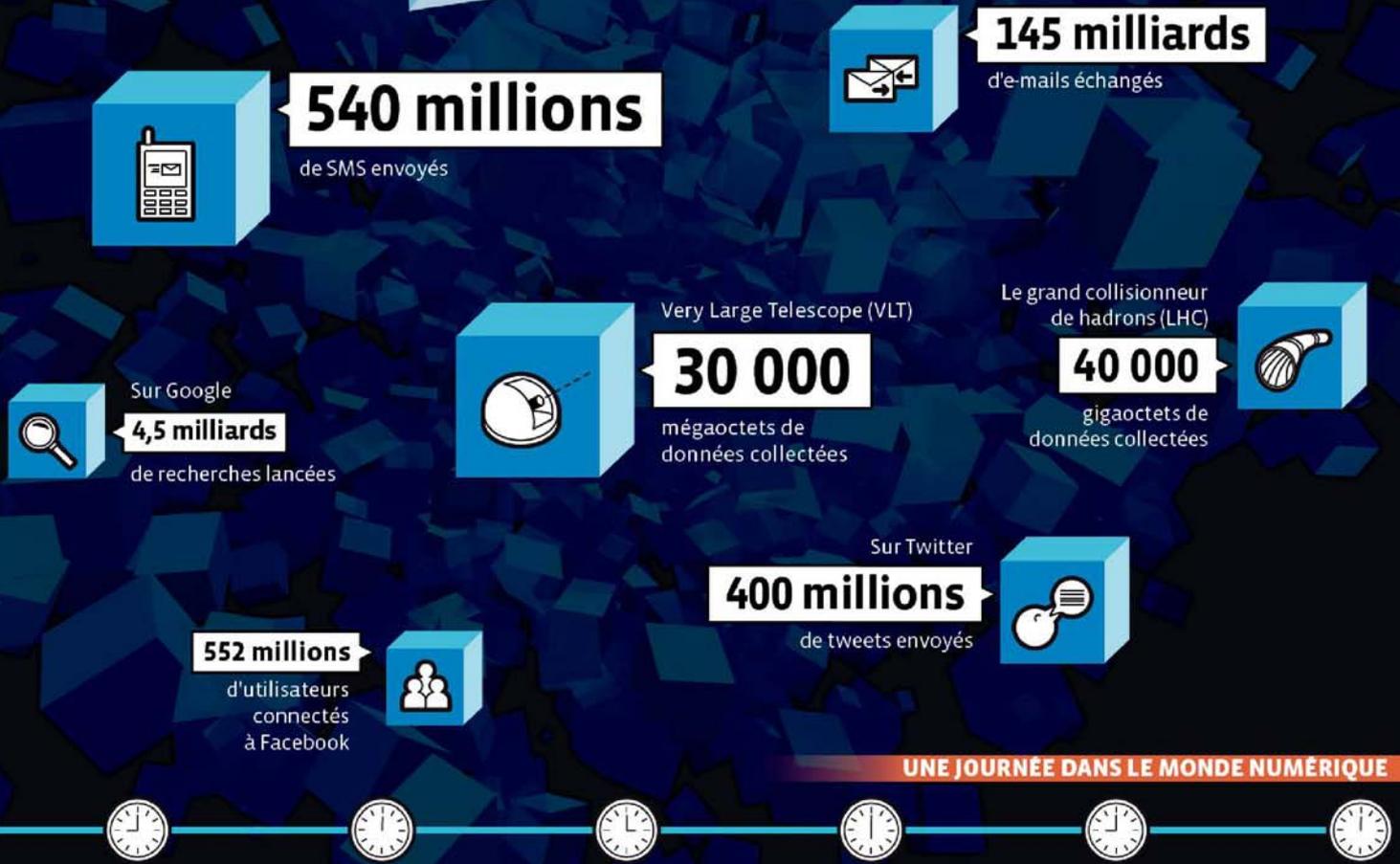
D'Internet aux grands instruments de recherche, le volume de données numériques générées par nos sociétés ne cesse d'augmenter. Un vrai déluge auquel doivent faire face les producteurs de ces données aussi bien que leurs utilisateurs. Comment les classer, les stocker, leur donner un sens? À l'occasion du colloque sur les grandes masses de données organisé par le CNRS le 5 décembre, *CNRS Le journal* se plonge dans l'effervescence du phénomène du « *Big Data* ».

UNE ENQUÊTE DE FABRICE DEMARTHON, DENIS DELBECQ ET GRÉGORI FLÉCHET

Des masses de données à donner le vertige **20** |
Une jungle à défricher **24** |
Le difficile stockage des données **26** |

Des masses

« **Tous les deux jours, nous produisons autant d'informations que nous en avons générées depuis l'aube de la civilisation jusqu'en 2003** », estimait Eric Schmidt, le patron de Google, lors d'une conférence donnée aux États-Unis en 2010. Une déclaration choc sur le grand sujet qui agite les mondes scientifique et économique aujourd'hui : les très grands volumes de données. « *Nous sommes confrontés à un véritable déluge d'informations* », constate Christine Collet, chercheur au Laboratoire d'informatique de Grenoble¹ et professeur à l'Institut polytechnique de Grenoble. Les chiffres parlent d'eux-



SOURCES : THE RADICATI GROUP, YOUTUBE, TWITTER, ITU, CERN, ESO, WIKIPEDIA, FACEBOOK, GOOGLE.

de données à donner le vertige

mêmes. Chaque seconde, plus d'une heure de vidéo est mise en ligne sur Youtube, et plus de 1,5 million d'e-mails sont envoyés². Les scientifiques ne sont pas en reste. En huit ans (2000-2008), Le Sloan Digital Sky Survey, un grand programme d'observation astronomique, a enregistré 140 téraoctets d'images (voir le schéma « L'échelle des données », p. 22). Mais il ne faudra que cinq jours à son successeur, le LSST (« Large Synoptic Survey Telescope »), auquel participent des équipes de l'IN2P3 du CNRS, pour acquérir ce volume ! Le grand collisionneur de hadrons (LHC), lui, amasse chaque année près de 15 pétaoctets de données, l'équivalent de plus de 3 millions

de DVD. Finalement, l'humanité produirait aujourd'hui, par an, un volume de l'ordre du zettaoctet d'informations : presque autant d'octets qu'il existe d'étoiles dans l'Univers³ !

UN ENJEU ÉCONOMIQUE MAJEUR

Avec le succès de l'économie numérique, la généralisation des appareils mobiles, le boom des réseaux sociaux, l'ouverture au public de certaines bases de données (« l'Open Data ») ou encore le développement de grands programmes scientifiques internationaux, le phénomène du « Big Data » va en s'amplifiant. « On assiste actuellement à une véritable démocratisation des données, précise

BIG DATA

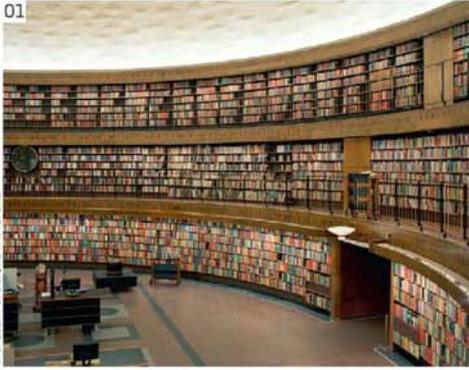
C'est l'appellation couramment utilisée par les spécialistes pour évoquer les grandes masses de données numériques.

MASTODONS : LE CNRS RELÈVE LE DÉFI

Mastodons est l'un des défis lancés par la Mission pour l'interdisciplinarité du CNRS. Ces défis « ont été développés afin de soutenir des sujets de recherche qui ne pourraient pas aboutir sans mettre en commun les connaissances de plusieurs disciplines. Mastodons fait partie de ceux qui impliquent le plus d'instituts du CNRS », explique Mokrane Bouzeghoub, directeur adjoint scientifique de l'Institut des sciences de l'information et de leurs interactions du CNRS (INS2I), qui coordonne Mastodons. Lancé en mai 2012, celui-ci a rencontré un franc succès. En à peine trois semaines, **37 projets** ont été présentés. Dix projets fédérateurs, cinq projets ciblés et un projet d'animation (sur la préservation des données) ont été retenus. Leur financement s'élève à **700 000 euros en 2012**. « L'objectif majeur est de favoriser l'émergence d'une communauté scientifique interdisciplinaire autour de la science des données », précise Mokrane Bouzeghoub. Chaque projet de Mastodons pourrait durer quatre ou cinq ans. **Le premier colloque consacré au défi Mastodons se tiendra le 5 décembre 2012, à Paris**. L'occasion pour les chercheurs de faire le point sur l'avancement de leurs projets.

LA RÉVOLUTION DU BIG DATA DANS LES SCIENCES HUMAINES ET SOCIALES

01



01 L'ère numérique a facilité pour les chercheurs l'accès à l'information, autrefois dispersée dans les bibliothèques.

Christine Collet. Elles sont partout et sont élaborées, commercialisées et consommées comme n'importe quel produit manufacturé. Ces grandes masses de données sont devenues un tel enjeu économique, industriel et scientifique que les gouvernements et les entreprises investissent massivement dans le domaine. Aux États-Unis, le président Barack Obama a dévoilé en mars un plan *Big Data* allouant 200 millions de dollars à la recherche dans ce domaine (« *Big Data Research and Development Initiative* »). De son côté, l'Europe a inscrit la gestion des contenus numériques dans ses priorités pour la fin du 7^e programme-cadre de recherche et de développement technologique. En France,

« Les grandes masses de données ont révolutionné le travail des spécialistes des sciences humaines et sociales. » Bertrand Jouve, mathématicien et directeur adjoint scientifique à l'Institut des sciences humaines et sociales du CNRS (INSHS), voit dans le *Big Data* une belle opportunité pour ses pairs. « Grâce aux bases de données en ligne, nous avons désormais aisément accès à une somme de connaissances qu'il nous fallait trouver dans les bibliothèques souvent dispersées... Et les enquêtes sur Internet ont facilité le travail, des sociologues par exemple. » Mais son enthousiasme ne lui fait pas oublier les difficultés auxquelles sont confrontés les utilisateurs. « Le problème réside le plus souvent dans le traitement des informations brutes,

explique-t-il. Si ces dernières n'ont pas été récoltées par le chercheur lui-même, comment savoir ce qu'elles ont subi avant d'être intégrées à la base de données? » Pour Sihem Amer-Yahia, il s'agit là d'un gros écueil du *Big Data*. « Le traitement des données brutes constitue souvent une boîte noire à laquelle personne n'a vraiment accès. Or on sait déjà que certaines manipulations assez courantes dans ce domaine peuvent supprimer une grosse partie des informations. » L'avènement des grandes masses de données et du tout-numérique soulève d'autres interrogations moins techniques. « Le *Big Data* implique nécessairement des réflexions épistémologiques, note Sandra Laugier, directrice adjointe

scientifique à l'INSHS. Qu'est-ce que cela signifie d'avoir accès à plus de connaissances qu'un esprit humain ne peut en concevoir? Quelles conséquences cette exhaustivité qu'on ne contrôle pas a-t-elle sur notre rapport au savoir? » Se posent également les questions de la propriété des données, des droits d'utilisation, du droit à l'oubli, de l'éthique... Autant de sujets auxquels les chercheurs en sciences humaines et sociales doivent s'atteler, en coopération avec d'autres disciplines, pour le bénéfice des citoyens et, peut-être, contre la mainmise des intérêts privés.

CONTACTS :

Bertrand Jouve
> Bertrand.jouve@cnrs-dir.fr
Sandra Laugier
> sandra.laugier@cnrs-dir.fr

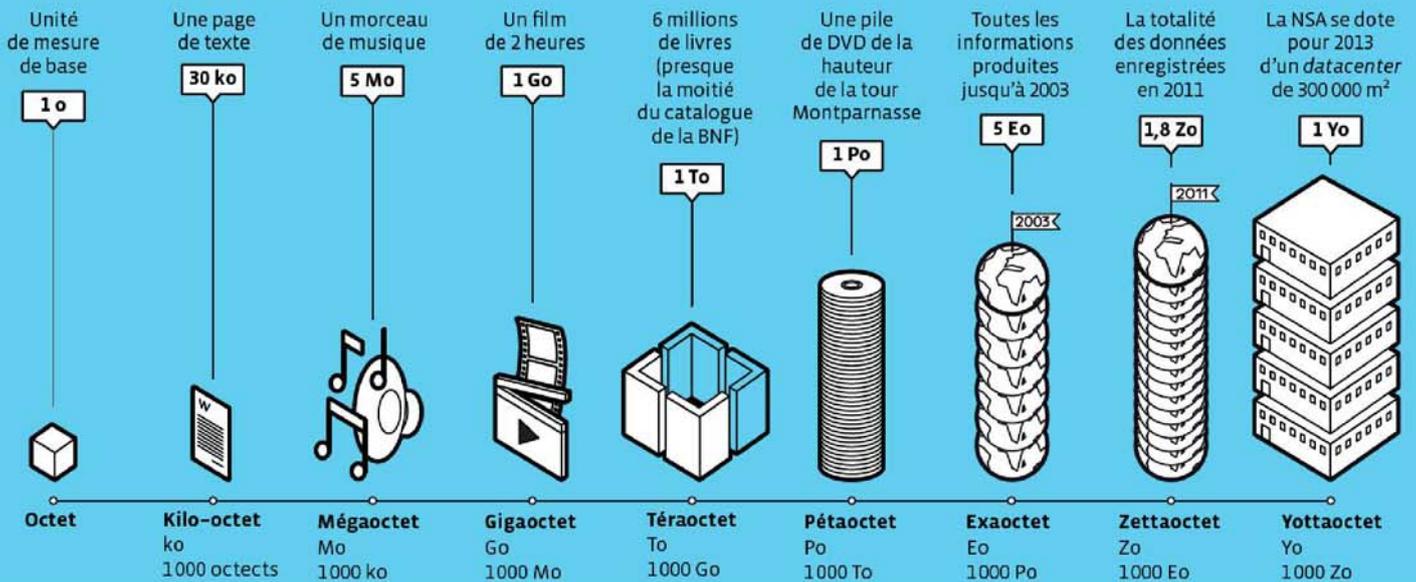
les Investissements d'avenir se sont aussi emparés du sujet. Un programme de 25 millions d'euros est consacré aux technologies d'exploitation des très grands volumes de données.

UN DÉFI POUR LES SCIENTIFIQUES

« Le *Big Data* constitue un défi scientifique considérable qui nécessite des travaux aussi bien en ingénierie que dans les sciences fondamentales », explique Mark

Asch, chargé de mission pour les mathématiques et le calcul numérique à la Direction générale pour la recherche et l'innovation du ministère de l'Enseignement supérieur et de la Recherche. Face à ce constat, la Mission pour l'interdisciplinarité du CNRS a lancé cette année le défi Mastodons (lire encadré p. 21). L'idée est de soutenir des projets interdisciplinaires afin d'identifier où sont les verrous dans la gestion des grandes masses de données

L'ÉCHELLE DES OCTETS



02 Le futur télescope LSST sera installé sur le Cerro Pachon, au Chili (vue d'artiste). 03 Doté de 189 capteurs, il enregistrera des images du ciel d'une taille inégalée. Ici : une simulation d'une photo prise par un seul capteur.



scientifiques. « Comment stocker les données et les pérenniser ? Comment les traiter, les analyser, les visualiser, leur donner du sens ? Comment les protéger, empêcher leur usage abusif et aussi les supprimer ? Autant de questions qui se posent aujourd'hui et dont nous sommes encore loin d'avoir les réponses, détaille Mokrane Bouzeghoub, directeur adjoint scientifique de l'Institut des sciences de l'information et de leurs interactions du CNRS (INS2I), qui coordonne Mastodons. Ces questions concernent autant les aspects algorithmiques que méthodologiques, sans oublier les infrastructures. »

DES QUESTIONS SANS RÉPONSES

Une analyse que partage Farouk Toumani, chercheur au Laboratoire d'informatique, de modélisation et d'optimisation des systèmes⁴ et porteur du projet Petasky dans le cadre de Mastodons. « Prenez le LSST, suggère-t-il. Ce télescope, qui doit être mis en service en 2020, sera capable d'enregistrer des images du ciel de 3 milliards de pixels toutes les 17 secondes. À la fin du programme, les astronomes disposeront ainsi d'une base de données de 140 pétaoctets, avec des centaines de caractéristiques pour chaque objet du ciel. » Or, aujourd'hui, les algorithmes de fouille de données les plus efficaces mettraient des dizaines d'années pour explorer la base et répondre à certaines questions des chercheurs. Les scientifiques engagés dans le programme du LSST savent donc déjà que certaines de leurs interrogations les plus complexes resteront sans réponse. Cependant, une telle base de données constitue le champ

ALGORITHME
Suite d'opérations et d'instructions qui permet de résoudre un problème.

d'application rêvé pour qui veut faire avancer la recherche fondamentale dans le domaine du *Big Data* : c'est précisément le cas des chercheurs du projet Petasky. « Pour contourner les obstacles rencontrés dans la gestion des grandes masses de données, il faudra certes améliorer les technologies de stockage et de calcul, mais aussi inventer de nouvelles manières de manipuler les données », annonce Farouk Toumani. C'est aussi l'avis de Sihem Amer-Yahia, spécialiste des réseaux sociaux au Laboratoire d'informatique de Grenoble. Elle confirme : « La révolution des grandes masses de données, portée notamment par l'explosion du Web social où les citoyens produisent eux-mêmes l'information, bouleverse l'architecture classique des bases de données et de leur traitement. »

UN BOULEVERSEMENT RADICAL

Tout le monde est concerné. Des biologistes aux astrophysiciens, de Facebook au fisc... plus personne n'échappe au phénomène du *Big Data* et aux problèmes

qu'il engendre. Et l'urgence est là : désormais, le volume d'informations produites dans le monde numérique double tous les deux ans et le rythme s'accélère. « Les données sont au cœur de l'économie numérique et de la société du savoir, conclut Christine Collet. Elles représentent une matière première à forte valeur ajoutée, sans laquelle rien ne pourra se passer. » **F. D.**

1. Unité CNRS/Université Grenoble-I/Institut Polytechnique de Grenoble/Université Grenoble-II/Université Grenoble-III
2. Sources : YouTube, Instagram, The Radicati Group
3. Source : IDC
4. Unité CNRS/Université Clermont-Ferrand-II/Université Clermont-Ferrand-I/Ecole des Mines de Saint-Étienne/Institut français de mécanique avancée de Clermont-Ferrand

CONTACTS :

Sihem Amer-Yahia
> sihem.amer-yahia@imag.fr
Mokrane Bouzeghoub
> mokrane.bouzeghoub@cnrs-dir.fr
Christine Collet
> christine.collet@grenoble-inp.fr
Farouk Toumani
> farouk.toumani@isima.fr

Technique et Chantier

Informatique

Les datacenters ont de l'énergie à revendre

Julie Nicolas

Les centres de données informatiques fonctionnent à l'inverse des bâtiments conçus suivant les réglementations thermiques. L'objectif est d'évacuer la chaleur des serveurs informatiques, tout en veillant à la fiabilité de l'ensemble des équipements.

Contemporains de la naissance de l'informatique, les datacenters sont dédiés à l'hébergement de serveurs de données. Leur fonction : garantir une alimentation électrique pérenne et offrir une large connectivité ainsi qu'un haut niveau de sécurité. Pour répondre à ces exigences, le maître mot est la redondance. Tout d'abord pour l'alimentation électrique, assurée en général par au minimum deux lignes de forte puissance. Ce premier niveau est ensuite renforcé par des batteries et des groupes électrogènes, voire par une pile à combustible pour les centres les plus à la pointe. Des onduleurs assurent la conformité du courant électrique aux spécifications des serveurs, en termes de tension, de phase et de fréquence. Connectivité oblige, les serveurs sont également reliés à plusieurs opérateurs télécoms, via des réseaux de fibre optique à très haut débit.

La valeur des données, le manque à gagner en cas d'interruption de service et le coût du matériel justifient les mesures d'ampleur prises pour assurer la sécurité et la pérennité du fonctionnement des datacenters, qui sont conçus pour répondre aux besoins de clients variés. Ce marché très particulier n'a commencé à se structurer qu'à partir de 2008-2009. « Depuis cette période, nous construisons en France des datacenters réellement pensés pour répondre aux attentes de clients différents, qui consomment moins d'énergie tout en prenant en compte l'environnement », indique Stéphane Duproz, directeur général de Telecitygroup France, propriétaire et exploitant de datacenters. La priorité reste toujours d'assurer le fonctionnement des serveurs dans les meilleures conditions, notamment de température et d'hygrométrie. « La difficulté est de prévoir leur évolution, alors que différentes technologies arrivent sur le marché : d'un côté des baies qui concentrent jusqu'à 384 serveurs dans un petit espace, de l'autre des microprocesseurs qui consomment neuf fois moins d'énergie et dégagent donc autant de chaleur en moins pour les mêmes performances », résume Laurent Trescartes, expert de Critical Building, société qui conçoit des datacenters.

Une filière à organiser

Si, dans des pays comme l'Irlande ou l'Islande, certains centres peuvent se passer de climatisation, grâce à des serveurs refroidis par l'air extérieur à 22 °C ou même 24 °C, les évolutions informatiques impactent la puissance frigorifique et les installations nécessaires au refroidissement des bâtiments.

La prise en compte de l'ensemble de ces paramètres fait émerger de nouveaux métiers relatifs à l'exploitation des datacenters. « Difficiles à trouver, ces professionnels ont besoin de formations ad hoc et de filières métier spécifiques », estime Thibaut Siméon, associé de Critical Building.

Autre caractéristique des datacenters : leur conception et leur exploitation vont à l'encontre des préconisations issues des réglementations thermiques. L'enjeu est ici d'évacuer les calories, non de les retenir. A ce titre, l'une des solutions mises en oeuvre à Aubervilliers (ci-contre) consiste à réaliser une isolation mince sur les presque 5 000 m² de toiture, pour profiter du froid hivernal. En été, la membrane d'étanchéité blanche réfléchit les apports solaires et limite les besoins en froid. La climatisation est optimisée avec un soufflage zone par zone et des groupes frigorifiques en toiture installés en fonction des vents dominants.

L'organisation des espaces tient également compte de la sécurité incendie. Si les sprinklers n'ont pas la cote auprès des exploitants, la solution d'extinction par brouillard d'eau, qui diminue la température du foyer jusqu'à son extinction complète, semble préfigurer l'avenir pour les grands datacenters. Le recours aux gaz neutres (air appauvri en oxygène) ou chimiques (qui se combinent avec l'oxygène de l'air) reste néanmoins la solution usuelle pour les sites de taille moyenne, avec toutefois des contraintes de volume de stockage et de coût. Le contrôle d'accès et la vidéosurveillance sont également des éléments clés de la sécurité, au même titre que la discrétion - primordiale pour ces bâtiments qui abritent des données critiques - et l'acoustique. Si certains datacenters optent pour les pièges à sons autour des groupes de froid, la société irlandaise Mercury a choisi de travailler avec un acousticien, qui l'a aidée à diminuer le niveau sonore de ses sites de Croydon, Watford et Woking, dans l'agglomération londonienne. Au final, à chaque étape de la conception d'un datacenter, les exploitants doivent donc combiner plusieurs solutions dont la somme des impacts économisera l'énergie.

Encadré(s) :

Mesurer la fiabilité et limiter les consommations d'énergie - LAURENT TRES CARTES, expert de Critical Building, société conceptrice de datacenters « Simplifier l'exploitation des centres »

Conçu aux États-Unis par l'Uptime Institute, le système « Tier » classe les datacenters en fonction des niveaux de garanties offerts sur leur fiabilité de fonctionnement. Le classement prend en compte la redondance, c'est-à-dire la présence d'équipements capables d'en suppléer d'autres pour réduire l'incidence des pannes et des maintenances. Les datacenters sont classés en fonction de la durée annuelle moyenne d'indisponibilité. Un classement Tier I correspond à une indisponibilité de 28,8 heures, un Tier II à 22 heures, un Tier III à 96 minutes et un Tier IV à 24 minutes. De son côté, l'Union européenne a mis en place en octobre 2008 un code de conduite pour les datacenters axé sur les économies d'**énergie**. Destiné aux opérateurs comme aux exploitants, le code comprend un audit lors du dépôt du projet et des mesures sur place une fois le centre en fonctionnement. Pour les datacenters entrés en fonction avant la publication du code, un audit mesure les progrès réalisés en fonction des objectifs fixés par l'opérateur.

« Aujourd'hui, la conception d'un datacenter performant impose de combiner vertueusement trois facteurs clés :

- l'adaptabilité du bâtiment et de ses infrastructures techniques en raison de l'évolution rapide des technologies informatiques et télécoms. Penser cette adaptabilité nécessite à la fois de connaître et d'anticiper les technologies émergentes sur le marché mais aussi de savoir évaluer leur impact potentiel sur l'alimentation électrique, le refroidissement ou les conditions générales de fonctionnement des serveurs ;
- la fiabilité que l'on est en droit d'exiger pour des installations aussi stratégiques ;
- la prise en compte de l'environnement, grâce à des infrastructures moins énergivores.

Afin de respecter ces exigences, il est nécessaire de maîtriser les process propres à chaque client, qu'il s'agisse de datacenters neutres (qui accueillent des entreprises de toutes tailles avec des contraintes diverses), ou de datacenters privés (où les serveurs sont à l'usage d'une seule société et qui représentent environ 85 % des centres de données en France). Pour ce faire, nos experts misent sur des solutions qui permettent à chaque fois de réduire l'investissement, de simplifier l'exploitation et de consommer moins d'**énergie**, tout en maximisant la démarche environnementale. »

Eléments essentiels devant figurer dans un contrat de prestation de services de Cloud computing

Informations relatives aux traitements

- Respect des principes européens en matière de protection des données personnelles et de la loi Informatique et Libertés (notamment des principes de proportionnalité et de respect des finalités) ;
- Existence d'un système de remontée des plaintes et des failles de sécurité ;
- Moyens de traitement ;
- Destinataires des données ;
- Sous-traitance :
 - Information et obtention du consentement du client en cas d'utilisation de tiers ou de sous-contractants situés ou non à l'étranger pour participer à la réalisation du traitement (Note : *si le prestataire est responsable conjoint du traitement, il devra seulement informer le client et non pas obtenir son consentement*) ;
 - Report dans les contrats de sous-traitance ultérieurs contractés par le prestataire des obligations contractuelles prévues dans le contrat de prestation signé entre le client et le prestataire et organisation de la responsabilité contractuelle des sous-contractants vis-à-vis du prestataire et du client.
- Existence de procédures simples permettant de respecter les droits des personnes concernées vis-à-vis de leurs données (droits d'accès, modification ou suppression, etc.).

Garanties mises en œuvre par le prestataire

- Durée de conservation des données limitée et raisonnable au regard des finalités pour lesquelles les données ont été collectées ;
- Destruction et/ou restitution des données en fin de prestation ou en cas de rupture anticipée du contrat dans un format structuré et couramment utilisé ;
- Devoir de coopération avec les autorités de protection des données compétentes ;
- Lorsque le prestataire est sous-traitant, indication que le client peut procéder à des audits du prestataire afin de s'assurer que ces garanties sont effectivement mises en œuvre.

Localisation et transferts

- Indication claire et exhaustive des pays hébergeant les centres de données du prestataire où les données seront traitées ;
- Assurance d'une protection adéquate à l'étranger (notamment grâce à des Clauses contractuelles types ou à des règles contraignantes d'entreprise « BCR ») ;
- Possibilité de limiter les transferts de données uniquement vers des pays membres de l'Espace Economique Européen ou vers des pays tiers reconnus comme assurant un niveau de protection adéquat par décision de la Commission européenne (*Note : Au contraire des autres éléments, celui-ci est laissé à la négociation des parties. En tout état de cause, un prestataire qui laisse la possibilité à ses clients de limiter les transferts de données vers des pays membres de l'EEE ou vers des pays tiers assurant un niveau de protection adéquat reconnu par la Commission européenne offrira à ses clients des garanties de protection des données renforcées. Toutefois, les clients doivent être conscients que lorsqu'ils choisissent des prestataires localisés dans des pays tiers, les autorités administratives ou judiciaires locales peuvent adresser des requêtes aux prestataires pour accéder aux données*) ;
- Information immédiate du client en cas de requête provenant d'une autorité administrative ou judiciaire étrangère.

Formalités auprès de la CNIL

- Lorsque le prestataire est sous-traitant, obligation de fournir au client toute information utile permettant de procéder à la déclaration du traitement auprès de la CNIL ;
- Lorsque le prestataire est responsable conjoint du traitement, le client et le prestataire doivent déterminer quelle partie sera en charge des formalités pour son compte et pour celui de l'autre partie. Quelle que soit la solution choisie, la partie qui ne déclare pas devra fournir à celle qui effectuera les formalités déclaratives toute information utile permettant de procéder à la déclaration du traitement auprès de la CNIL.

Sécurité et confidentialité

- Indication des obligations incombant au prestataire en matière de sécurité des données et, lorsque celui-ci est sous-traitant, précision qu'il ne peut agir que sur instruction du client ;

- Politique de sécurité et mesures minimales de sécurité :

[Note : le prestataire sous-traitant devra tenir à la disposition du client le détail des mesures mises en place, tandis que le prestataire responsable conjoint du traitement devra seulement garantir que des mesures suffisantes ont été mises en œuvre.]

- Existence d'une politique de sécurité accessible ;
 - Mesures de sécurité et sûreté physique sur le site d'hébergement (protection du site et sécurité des accès, sécurité électrique et système de climatisation, etc.) ;
 - Mesures nécessaires pour assurer la disponibilité, l'intégrité et la confidentialité des données : par exemple, chiffrement des données et procédés garantissant ainsi que le prestataire n'a pas accès aux données qui lui sont confiées (chiffrement côté client, avec un algorithme reconnu et une gestion des clés adéquate, avant tout transfert) et liaison chiffrée avec le serveur de Cloud (connexion de type https ou VPN par exemple), etc. ;
 - Autres mesures de sécurité logique (protection du réseau (pare-feu, antivirus, détection d'intrusion, etc.), gestion des mises à jour, protection du terminal, gestion des habilitations, authentification des personnels, sécurité des développements applicatifs, etc.) ;
- Certifications : preuve de certifications pertinentes par des auditeurs indépendants et qualifiés, par exemple une certification ISO 27001 sur un périmètre incluant intégralement les services fournis, définition rigoureuse d'une politique d'audit du prestataire par le client comprise dans les garanties générales *[Note : au contraire des autres éléments, la certification est laissée à la négociation des parties. En tout état de cause, un prestataire qui dispose d'une certification offrira à ses clients des garanties de protection des données renforcées]* ;
 - Réversibilité/portabilité : garantir la réversibilité ou la portabilité aisée des données dans un format structuré et couramment utilisé, sur demande du client et à tout moment ;
 - Traçabilité : accès aux journaux de traçabilité des actions effectuées sur les données par les personnels du client et par ceux du prestataire et information de toute anomalie détectée par le prestataire ;
 - Continuité de service, sauvegardes et intégrité : système de sauvegarde, redondance des serveurs, etc. ;
 - Engagement de niveaux de services (« *Service Level Agreements* » ou « *SLAs* ») : engagements contraignants pour le prestataire sur le niveau de service, devant notamment prévoir des pénalités pour le prestataire en cas de non-respect des engagements contractuels. Ceci doit être mis en place en particulier pour les clauses relatives à la protection des données (durée de conservation, exercice des droits des personnes concernées, disponibilité du traitement, etc.).

Les Datacenter

Performance des architectures IT : comprendre, résoudre et anticiper / Pascal Grojean, Médéric Morel, Simon-Pierre Nolin ; Dunod, 2011

Objectifs

Les Datacenter (ou centres de traitement de données) sont des installations spécialisées permettant le regroupement des éléments informatiques sensibles comme les applications, les infrastructures (logicielles, systèmes et matérielles), ainsi que les équipements réseaux dans un environnement le plus sécurisé possible.

Ils représentent l'enveloppe du système d'information car c'est en leurs seins que se joue toute la production informatique de l'entreprise. Ils doivent garantir un haut niveau de service afin d'assurer l'intégrité et le fonctionnement des équipements et applications hébergées. Ils doivent également se doter de moyens de pilotage et de suivi nécessaires, que ce soit en termes d'outils de surveillance ou de personnel dédié.

Les problématiques de performance auxquelles ils sont soumis sont principalement liées à la disponibilité. En effet, regrouper la plus grande partie de son système d'information présente le risque de tout perdre si l'on perd son Datacenter et qu'il n'est pas prévu de solution de secours.

Nous nous attacherons, dans ce chapitre, à positionner le Datacenter dans la politique de performance de l'entreprise. Nous aborderons les différents types de Datacenter et nous verrons comment mettre en place un plan de continuité de l'activité en prenant en compte l'hébergement du système d'information.

9.1 DATACENTER ET PERFORMANCE

9.1.1 Contraintes spécifiques

Centre de la production informatique, un Datacenter propose des problématiques particulièrement éloignées des préoccupations classiques des informaticiens. Elles sont la conséquence de la concentration des équipements informatiques en un lieu unique. Ces problématiques se classent en trois catégories :

- **Alimentation électrique** : la plupart des Datacenter ont une consommation électrique en moyenne 10 fois plus élevée qu'un immeuble de bureau. Outre les aspects environnementaux, il s'avère indispensable de permettre au Datacenter d'alimenter tous les équipements informatiques qu'il héberge, au risque de voir ces derniers s'arrêter.
- **Climatisation** : bien que liée à la chaleur engendrée par la consommation électrique, la climatisation joue également un rôle essentiel dans la qualité de l'air qui circule entre les serveurs (humidité et poussière).
- **Capacité réseau** : le regroupement de tous les serveurs de l'entreprise nécessite, de la part du Datacenter, une bande passante réseau à minima supérieure à la consommation maximale de chaque serveur.

Tout cela converge vers un point particulièrement important de l'implication des Datacenter dans la performance du système d'information : la disponibilité.

9.1.2 Gestion de la disponibilité

Dans le cas d'un Datacenter déjà ancien, l'amélioration de la disponibilité peut être traitée sur les deux axes suivants en parallèle :

1. Diminuer le nombre de pannes en investissant sur la résilience des éléments critiques du site, par exemple en mettant en place une voie d'alimentation électrique supplémentaire ou en installant un système de refroidissement plus tolérant aux pannes. Cette démarche aura un retour sur investissement correct si elle est concentrée sur les causes de panne les plus courantes.
2. Mettre en place les éléments et techniques permettant un redémarrage rapide à un niveau de service acceptable (voire en mode dégradé si cela est envisageable) sur un autre site épargné par le sinistre éventuel.

Ces deux approches sont complémentaires : il n'est pas optimal d'investir uniquement sur un seul de ces axes. En effet un site unique, même offrant un très haut niveau de résilience, ne sera jamais à l'abri d'un sinistre de grande ampleur (sinistre régional par exemple).

La mise en œuvre des deux axes de réflexion abordés ci-dessus conduit à des réalisations diverses :

- Mise en place d'un bon niveau de résilience sur chaque site, par exemple via la division du site en zones pouvant se secourir mutuellement afin de parer rapidement à certaines défaillances locales.
- Utilisation de sites « doubles », de type « campus », géographiquement proches l'un de l'autre. Chacun de ces sites fait office de miroir pour l'autre via un réseau à haut débit.
- Installation de sites distants d'au moins 10 km (mais de moins de 50 km), afin de limiter le risque d'être touché par le même sinistre tout en permettant un niveau de synchronisation élevé.
- Installation d'un site de secours réellement éloigné du site principal (au moins de 150 km), qui servira de site de secours en cas de sinistre majeur touchant le site principal ou l'un des sites proches. La distance ayant un impact sur la synchronisation des données, il conviendra de prendre en compte ce point au niveau du PCA.

9.2 LES DIFFÉRENTS TYPES DE DATACENTER

9.2.1 Classification

L'association américaine « Uptime Institute » a établi une classification des Datacenter en quatre groupes, appelée « tiers » en anglais, en fonction du niveau de disponibilité garantie par le type d'installation :

- « **Tiers 1** » : regroupe les Datacenter n'ayant qu'une seule alimentation électrique, un seul système de refroidissement et aucune redondance des équipements qu'ils hébergent. La disponibilité nominale de ce type de centre est de 99,671 %, correspondant à un temps d'arrêt cumulé moyen annuel de 28,8 heures. Il est donc impossible de parler de haute disponibilité dans ce cas.
- « **Tiers 2** » : regroupe les Datacenter ayant une voie unique pour l'alimentation électrique et le refroidissement, mais disposant de moyens de redondance permettant une disponibilité nominale de 99,749 % (soit 22 heures d'arrêt par an). Ayant passé la barre des 99 %, il est envisageable de traiter des applications sensibles. Toutefois le manque de redondance sur les alimentations engendre un risque fort pour l'entreprise car il s'agit là d'un SPOF¹ au niveau du SI dans son intégralité.
- « **Tiers 3** » : regroupe des Datacenter plus complets que les tiers précédents car bénéficiant de plusieurs voies d'alimentation et de refroidissement (dont une seule voie active). Une partie des équipements est redondante, ce qui

1. *Single Point of Failure* : point d'un système informatique dont le reste du système est dépendant et dont une panne entraîne l'arrêt complet du système.

permet de réaliser une bonne partie des opérations de maintenance sans arrêt des machines. Ce type de Datacenter offre une disponibilité de 99,982 % par an, la durée moyenne annuelle d'arrêt étant de 1,6 heure.

- « **Tiers 4** » : regroupe les Datacenter offrant le plus haut niveau de disponibilité (99,995 % soit une durée d'arrêt moyenne annuelle de 0,4 heure). Ils possèdent plusieurs voies actives pour l'alimentation électrique et le refroidissement, ainsi qu'un nombre élevé d'éléments d'infrastructure redondants et tolérants aux pannes.

Cette classification est reconnue par les principaux acteurs du marché des Datacenter et un nombre croissant de Datacenter se font certifier sur cette base. Selon une étude du CRIP¹ (Club des Responsables d'Infrastructure et de Production), réalisée début 2009, la majorité des centres français serait en mesure d'approcher (voir même d'être certifié) Tiers 3 car ils disposent de redondance au niveau :

- des alimentations électriques,
- de la climatisation,
- d'accès opérateur,
- d'accès Telecom.

9.2.2 Impacts

La haute disponibilité n'est pas sans impact sur les coûts. La mise en œuvre des moyens nécessaires pour garantir le niveau de service offert par les différents tiers implique des coûts et contraintes de plus en plus élevés en fonction des tiers :

- Coût de construction, de maintenance et d'équipement du Datacenter.
- Coûts liés à la consommation électrique et aux besoins en refroidissement associés.
- Coût de la surface réservée aux éléments techniques mis en œuvre pour garantir le Tiers retenu (et donc ne pouvant être utilisée pour héberger les éléments du SI).
- Coût induit par un besoin de personnel sur site (croissant avec le niveau du tiers).

Un autre élément important à prendre en compte lors de la création d'un Datacenter est la question des délais de construction et d'équipement : là aussi le délai croît avec le niveau de tiers cible.

1. Association française (Loi 1901) réunissant des responsables d'infrastructures ou de production de grandes entreprises ou d'entités utilisatrices des technologies de l'information.

9.3 LA GESTION DES RISQUES

Les Datacenter hébergeant les éléments techniques nécessaires au bon fonctionnement de l'entreprise, il faut donc les considérer comme faisant partie des ressources critiques. En effet tout dysfonctionnement (voire pire toute destruction) peut avoir un impact négatif considérable sur le SI de l'entreprise. À ce titre ils doivent faire l'objet de politiques de gestion des risques spécifiques, qui traitent aussi bien des risques externes qu'internes aux Datacenter. Typiquement la prise en compte des risques est divisée en trois types (risques naturels, humains et techniques) et selon deux axes : risques intrinsèques et risques extrinsèques.

9.3.1 Risques intrinsèques

Les risques naturels peuvent avoir de trois origines :

- climatique (tornade, foudre, etc.),
- géologique (séisme, éruption volcanique, etc.),
- hydraulique (inondations par exemple).

Ces risques sont liés à la situation géographique au sens large, et influent sur le choix de l'emplacement des Datacenter.

Afin de limiter le risque que deux sites soient touchés par une inondation on évitera de les placer dans la même vallée fluviale.

Les risques d'origine humaine englobent

- les actes malveillants (de type sabotage ou vol),
- les erreurs humaines ou méconnaissances des consignes à appliquer,
- les intrusions en tout genre.

Les risques liés au personnel du Datacenter peuvent être liés à un manque de formation ou encore à l'absence de PCA complet et éprouvé. Par exemple les exercices de reprise sur sinistre peuvent améliorer le niveau de maîtrise des procédures d'urgence en cas de problème grave.

Les risques liés à des personnes tierces peuvent être contenus par la mise en œuvre d'une surveillance et une limitation des accès physiques aux installations.

Les risques techniques :

- panne matérielle,
- mauvais fonctionnement d'un équipement critique,
- usure ou défaillance,
- etc.

Ces risques sont en général bien connus et étudiés dès la conception des installations du Datacenter. Ils sont généralement adressés par la mise en place d'équipements de secours, des solutions de réparation rapide ou de redondance.

9.3.2 Risques extrinsèques

Il faudra aussi tenir compte des risques externes liés à l'environnement géographique du Datacenter (proximité d'une zone dangereuse type Seveso ou encore risques de pollution par exemple) ainsi qu'aux risques de défaillance des fournisseurs critiques (alimentation électrique notamment).

Il faudra également tenir compte de l'environnement politique de la zone où le centre sera implanté. Une forte instabilité de la région ne peut pas être propice à la sécurité de l'endroit et présente donc de forts risques quant à la pérennité des installations.

Au-delà des risques listés ci-dessus, une bonne gestion des risques devra tenir compte des contraintes techniques issues des architectures techniques mises en place au niveau du SI. Ce point conditionne notamment les modalités de mise en place d'un site de secours ainsi que les conditions de basculement.

9.4 DATACENTER ET PCA

9.4.1 Classification des sites de repli

Un des éléments essentiels à prendre en compte lors de l'établissement d'un PCA est la distance entre un Datacenter et le site de reprise (notion de disponibilité). On pourra adopter les dénominations suivantes :

- Datacenter « campus » : le site de reprise est très proche (moins de 10 km en général).
- Datacenter « métropolitain » : le site de reprise se situe à moins de 50 km.
- Datacenter « régional » : le site de reprise se situe à moins de 100 km.
- Datacenter « continental » : le site de reprise est situé à plus de 100 km.

Cette notion est particulièrement importante lors de la prise en compte des sinistres qui pourraient survenir sur le Datacenter. Le tableau 9.1 met en lumière la couverture d'un sinistre en fonction du type de site.

Actuellement on note une nette préférence pour les solutions de reprise locales ou à distance modérée, qui permettent notamment de mettre en place des mécanismes de synchronisation de données entre les différents sites, ce qui allège d'autant la liste des actions à mener en cas de bascule, et donc le temps de rétablissement du service.

DOCUMENT 5extraits de la brochure d'information, Site www.telecitgroupe.fr

Un data center d'exception. Un savoir-faire unique.

Notre engagement de qualité

Nos data centers répondent aux standards internationaux suivants:



ISO 27001

Gestion de la sécurité de l'information.

ISO 9001

Gestion des systèmes de qualité.

ISO 14001

Gestion environnementale.

OHSAS 18001

Gestion de la santé et de la sécurité au travail.

PCI DSS

certified

Gestion de la sécurité des paiements par carte bancaire



EU Code of Conduct for Data Centres – Statut Corporate

Définit les meilleures pratiques en faveur de l'environnement dans l'industrie des data centers.

CERTIFICATS EQUILIBRE+

Certificats EDF Equilibre +

Programme de compensation d'émission de CO₂ et participation à un projet de développement de l'énergie solaire.

Le data center Condorcet vous offre un site à la pointe de l'innovation.

Une flexibilité pour répondre à tous les besoins

Le data center Condorcet a été conçu pour répondre à tout type de besoin. Il peut vous offrir des surfaces d'hébergement à la demande et des hautes densités électriques allant jusqu'à 30kW par baie.

Les plus hauts niveaux de redondance et de résilience

Condorcet a été conçu selon un design de type Tier IV, fournissant les plus hauts niveaux de redondance et de résilience du marché pour vous assurer une continuité d'activité en toute circonstance.

Une réduction des coûts et de l'empreinte carbone

Le design innovant du data center Condorcet permet une réduction significative des émissions de CO₂ qui vous aide à tenir vos engagements en faveur du respect de l'environnement mais également à réduire vos coûts liés à l'hébergement d'infrastructures informatiques.

Un large choix de connectivité

Condorcet donne accès à un large éventail de fournisseurs de services et de points d'interconnexion Internet. Le réseau de fibres optiques TelecitGroup permet également au centre Condorcet d'être connecté aux autres sites français et européens du Groupe.

Une implantation stratégique

Situé aux portes de Paris, le centre de données Condorcet est idéalement localisé au cœur d'un important pôle d'activité regroupant la plus grosse concentration de fibre optique d'Ile de France.

Une amélioration du rendement électrique favorisée par l'utilisation de 3 chaînes électriques indépendantes (verte, bleue, rouge) au lieu de 2 traditionnellement

Conçu pour allier efficacité énergétique et réduction des coûts.

TelecityGroup s'investit pour contrôler l'impact de ses émissions de CO₂ en introduisant de manière régulière des initiatives destinées à maximiser l'efficacité énergétique de ses centres. Condorcet est le data center le plus abouti en matière d'éco-efficacité.

Des choix techniques motivés par le respect de l'environnement

- Utilisation de matériaux éco-compatibles.
- Utilisation d'une pile à combustible (hydrogène) comme source d'alimentation de secours additionnelle des systèmes de sécurité.
- Bannissement des additifs chimiques dans les installations climatiques.
- Toiture blanche pour optimisation de l'albedo selon les recommandations de Steven Chu, prix Nobel de Physique.
- Refroidissement des salles informatiques par free chilling.
- Armoires de climatisation à débit d'air variable et auto-régulé.
- Pompes à débit d'eau variable et auto-régulé.
- Salles techniques refroidies par free cooling.
- Onduleurs à haut rendement.
- Architecture électrique innovante.
- Arboretum utilisant l'air chaud des salles informatiques, hébergeant un programme de recherche scientifique en partenariat avec l'INRA.

Une position éco-responsable innovante

TelecityGroup est engagé dans de nombreuses initiatives visant à réduire ou compenser ses émissions de CO₂.

- L'entreprise est la première à avoir été distinguée par le "Corporate status" dans le cadre de l' "EU Code of Conduct for Data Centres", programme lancé par la commission européenne qu'elle fut également la première à rejoindre.
- TelecityGroup France a été la première de son marché à souscrire aux Certificats Equilibre+ d'EDF attestant que l'équivalent de 21% de l'électricité consommée par ses data centers a été produite à partir de sources d'énergies renouvelables.
- Le Groupe est également membre du consortium mondial "The Green Grid" ayant pour mission d'améliorer l'efficacité énergétique des data centers.
- TelecityGroup a été certifié ISO 14001 pour l'intégralité de ses sites européens.
- Condorcet a été récompensé par le Greenbang Award 2010 du "Data center le plus innovant" pour son architecture éco-efficace.

DOCUMENT 6

Etude Ge-Tic : l'éventualité de la mise en place d'un datacenter dans l'Aube,
Site ge-tic.fr mis à jour en septembre 2011

IV CONCLUSION

Le DATACENTER répond à des besoins de plus en croissants dans l'univers de l'informatique : sauvegarde à distance, externalisation de sa « salle serveur », « virtualisation » de ses applications.

De nombreuses entreprises proposent du libre-service en matière de logiciel de gestion (paie, CRM, ERP...).

On s'est aperçu au cours de cette enquête que de nombreuses entreprises sont potentiellement consommatrices d'un DATACENTER.

En premier lieu, par la sauvegarde à distance qui pourrait devenir un premier levier pour intéresser les chefs d'entreprise de l'Aube.

Le nombre d'applications transférables sur un DATACENTER est conséquent. Pour cela, il faut que les SSII de l'Aube prennent une part importante, et soient une force de proposition. Il existe un tissu important de Web-agency qui pourrait héberger l'ensemble des sites de leurs clients sur un DATACENTER. Des SSII qui proposent des services de sauvegarde à distance pourraient transférer les données sur un DATACENTER aubois. De nombreux services sont potentiellement externalisables.

La multiplication des données, le nombre croissant d'informations à conserver sur support magnétique, la complexité des applications, les ressources matérielles en constante progression sont des raisons qui tendent à nous orienter vers la mise en place d'un DATACENTER dans l'Aube.

En 2010, une étude de Symantec sur les DATACENTER montre une montée en puissance des demandes des entreprises.

Extrait de l'étude :

« Certaines priorités se dégagent en France, reflétant autant d'initiatives, les DSI français ont comme priorité l'archivage et la dé-duplication puisque 78 % d'entre eux les considèrent comme importantes, voire très importantes. 78 % estiment que le cloud computing privé est intéressant pour les problématiques qu'ils rencontrent. La vague verte touche particulièrement les entreprises françaises : les Français sont les plus enclins, avec les Italiens, à intégrer la dimension écologique, en mettant notamment l'accent sur la réduction de consommation énergétique.

Enfin, il s'agit toujours, en 2010, d'optimiser et de réduire les coûts des data centers : 46 % des DSI interrogés considèrent en effet cette réduction comme très importante.

Pour une fois, l'exception culturelle ne s'applique pas totalement puisque les chiffres français varient par ailleurs faiblement des données mondiales de cette troisième édition.

Au cœur des préoccupations des entreprises se trouvent la multiplication des applications, l'augmentation toujours plus rapide du volume des données (qui double tous les 2 ans) et la complexité accrue de leur gestion. 50 % des entreprises considèrent que l'information à manager augmente rapidement, voire très rapidement, générant ainsi des difficultés nouvelles pour atteindre des objectifs d'efficacité dans la gestion des données. Plus d'un tiers des DSI interrogés pensent que la productivité de leurs équipes se heurte à un nombre trop important d'applications et à une complexité croissante des data centers. Il n'est alors pas surprenant que 71 % d'entre eux considèrent avec attention les technologies de réduction de l'information comme la dé-duplication.

La plupart des entreprises ont d'ailleurs au moins 10 initiatives concernant leur centre de traitement de données considérées comme plutôt ou très importantes, et 50 % d'entre elles prévoient des changements "importants" de celui-ci en 2010.

Les problèmes de ressources humaines sont également importants. La moitié des DSI interrogés pensent que ni leurs équipes dédiées à la sécurité, à la gestion des réseaux ou encore à la virtualisation, ni leur budget ne sont assez importants pour faire face aux défis actuels et à venir. 76 % des entreprises ont le même nombre de postes à pourvoir, voire davantage, cette année.

L'étude montre également des résultats surprenants sur les plans de reprise après activité : si 80 % des DSI ont confiance dans leur plan, un tiers d'entre eux ne l'ont pas réévalué pendant les 12 derniers mois et celui-ci ne comporte bien souvent pas d'éléments composants informatiques importants tels que le cloud computing, les bureaux distants ou les serveurs virtuels.

Autre enseignement de l'étude : les entreprises de 2 000 à 9 999 employés anticipent plus fortement les évolutions majeures en matière de traitement de données et sont plus susceptibles d'adopter des technologies de pointe (cloud computing, dé-duplication, réplication, virtualisation du stockage, protection continue des données, etc.).

Les chantiers 2010 en matière de centres de données semblent se dessiner, selon l'étude Symantec. Les DSI interrogés considèrent en effet la sécurité contre les menaces comme primordiale (83 %), 79 % d'entre eux portent leur attention avant tout sur la sauvegarde et la restauration de l'information et 76 % ont classé la protection des données comme une priorité.

« Les entreprises restent confrontées à la même complexité inhérente aux centres de traitement de données. Et cette complexité ne fait qu'augmenter avec l'adoption de nouvelles initiatives », explique Vincent Videlaïne, responsable de la gestion du stockage chez Symantec. « La standardisation autour de solutions inter plates-formes permettant de gérer les nouvelles technologies et d'automatiser les processus réduira immédiatement les coûts et facilitera leur tâche sur le long terme. »

Cette étude nous permet de constater que les besoins des entreprises que nous avons contactées sont des préoccupations nationales. Par la mise en place d'actions conjointes entre les institutions de notre département (CCI, Conseil Général) et des SSII aubois, nous pouvons rendre viable la mise en place d'un DATACENTER dans l'Aube en passant par la sensibilisation des entreprises aubois et devenir ainsi un département en pointe dans l'évolution des TIC.

Datacenters et création d'emplois : une équation à multiples variables

[Anne-Marie Rouzeré](#) Publié: 22 juin 2012

Une friche industrielle, un noeud de réseau à très haut débit à proximité : tels sont les éléments mis en exergue par l'hébergeur OVH dans son choix d'investir 180 millions sur six ans pour installer son neuvième datacenter à deux pas des industries lourdes de Dunkerque. De là à parler d'une implantation d'activité source d'emplois ! C'est le pas que franchit Global Switch en présentant son campus parisien comme un « *exemple de ré-industrialisation réussie* » qui « *développe l'emploi au coeur des villes* ».

Certes, l'annonce d'OVH se double d'une communication autour de son plan de recrutement : 500 embauches d'ici à 2015. Mais à y regarder de plus près, le nouveau site nordiste ne devrait occuper qu'une cinquantaine de personnes (à terme) et une dizaine cette année. L'objectif « *500 embauches sur trois ans* » va de pair avec l'implantation de bureaux à Paris, Strasbourg, Lyon, Brest, Sophia Antipolis et Gand. Histoire de se rapprocher des clients mais aussi des viviers d'ingénieurs et de techniciens supérieurs. « *Il est de plus en plus difficile d'attirer les bons profils d'administrateurs systèmes et de développeurs à Roubaix* », explique Soushila Mookien, DRH du groupe OVH, « *il faut aller les chercher* ». D'où cet objectif de renforcer l'effectif (au total 350 personnes en France) par des équipes de proximité (« *10 à 15 ingénieurs et techniciens par bureau* », précise la DRH), virtualisation des opérations aidant. Entendons bien : ce n'est donc pas le fonctionnement au quotidien des centres de données (au sens, salles informatiques) qui génère cet appel d'air mais plutôt les services associés. Faire tourner les 120 000 serveurs hébergeant actuellement les sites web et applications des clients occupe moins de techniciens (une soixantaine) que répondre à la demande qui implique 200 à 300 installations de serveurs par jour (selon le pdg Henryck Klabo, cité par les Echos). Peu importe où travaillent ces installateurs-administrateurs. De préférence, là où il est plus facile d'attirer de nouvelles recrues. Ikoula, autre hébergeur en croissance à deux chiffres, suit la même logique en implantant ses centres et équipes d'administrateurs à Reims, à proximité de cursus (écoles, IUT, BTS).

Autour des capitales européennes

Tout autre est le positionnement -y compris en matière d'emplois -de Global Switch, Digital Realty Trust, Interxion, et plus généralement des grands data centers mutualisés (hébergement des plateformes informatiques de grandes entreprises, multinationales). Les raisons de la localisation sont semblables (forte capacité énergétique et noeuds de réseau des opérateurs télécom). Ce qui suffit à justifier leur concentration autour des capitales européennes, dont Paris (localisation privilégiée, selon un sondage réalisé en 2011 pour Digital Realty Trust auprès de 205 grandes

entreprises). S'y ajoutent les exigences des clients. « *Ils veulent qu'on leur fournisse un écrin le plus fiable possible pour leurs plateformes, mais à moins de 30 kilomètres de chez eux, capacité des réseaux oblige* », commente Yves Tavernier, directeur général de Global Switch France. La démarche de ce type d'opérateurs ne repose pas non plus sur des équipes importantes en nombre : chez Global Switch, 25 ingénieurs pour le pilotage d'un campus de 48 000 mètres carrés, 50 personnes pour la maintenance 24/7 de l'infrastructure physique (30% d'ingénieurs, 70% techniciens supérieurs). Alors pourquoi parler de ré-industrialisation, voire de contribution à l'animation d'un bassin d'emplois ?

On conçoit que construire et piloter une installation hyper-sécurisée, ultra-redondante, consommant des dizaines de megawatts (« *autant qu'une ville de 50 000 habitants* », selon Fabrice Coquio, président d'Interxion France) et dont 75% des coûts relèvent de ces équipements, repose sur les compétences de *pointures* en génie électrique, en génie climatique, en urbanisation ou optimisation de l'infrastructure mise à disposition des directions informatiques clientes. Mais aussi sur un écosystème d'opération et de maintenance de haute technicité. Pour les six datacenters parisiens d'Interxion, aux côtés de ses 80 salariés (dont 72 cadres) intervient une soixantaine de techniciens en sous-traitance. Le tout dans un dialogue constant entre la partie infrastructure et la partie informatique, dans un climat de culture hybride (facility management, production informatique). Sans oublier les exigences de sécurisation du site, qui plus est en milieu urbain. Ce qui occupe, chez Global Switch Paris, une équipe de seize personnes (prestataire) sous la houlette d'un responsable sécurité (Sun Sundaeswan) recruté chez HP (EDS).

Une sorte de pépinière d'entreprises

« *Un campus comme le nôtre, c'est aussi un jeu de poupées gigognes, avec nos emplois directs, ceux des prestataires in situ qui ajoutent la couche de services spécialistes réseaux, administrateurs des infrastructures informatiques intervenant au quotidien* », relate Yves Tavernier. La location de surface technique en fait de plus une sorte de pépinière d'entreprises qui, selon lui, permet à des petites équipes d'ingénierie informatique d'avoir accès à une clientèle de grands comptes qui ne les auraient peut-être pas repérées autrement. Soit au total, un ensemble d'emplois, dont la présence in situ est d'autant plus recherchée par les opérateurs de datacenters et appréciée des entreprises hébergées que les exigences en réactivité (à tous égards) tendent à s'accroître. « *Les responsables informatiques ont si peu de visibilité qu'ils hésitent longtemps avant d'opter pour un hébergement, et quand ils se décident, il faut aller très vite, aménager en moins de trois mois l'espace dédié* », constate le dirigeant de Global Switch.

Mais en définitive, s'agit-il de création nette d'emplois ? Ou d'un simple transfert, des salles informatiques gérées en interne vers des structures industrialisées avec, au contraire, réduction d'effectif des équipes côté informatique? « *Héberger les données des entreprises est un vrai métier qui exige un vrai savoir-faire* » ; « *Les packages de gestion de l'infrastructure du datacenter ne réduisent pas la taille des équipes d'administration. A effectif constant, l'automatisation autorise plus de complexité* », lit-on dans [le compte-rendu du récent congrès Datacentres 2012](#). Où est évoquée également une évolution de la répartition des tâches contractualisées (entre les équipes de l'hébergeur et des prestataires de services). « *Ce qui intéresse les DSI clientes, ce n'est pas forcément l'externalisation du personnel informatique, mais profiter d'un environnement optimisé, géré de façon plus experte, et surtout beaucoup plus sécurisé* », observe Fabrice Coquio. Autrement dit, une réponse en « ni-ni » : pas forcément plus d'emplois à la clé, ni moins d'ailleurs, ni un transfert systématique. Mais en tout cas, une montée en compétences.

80% des interventions à distance

A voir la fréquentation des congrès et séminaires dédiés (début juin, Data Center Dynamics à Paris, puis Datacentres 2012 à Nice) et la dynamique de clubs (le CESIT, comité des exploitants

de salles informatiques, le CRIP, club des responsables d'infrastructures et de production, dont la convention 2012 vient de réunir à Paris 2000 professionnels), cette montée en compétences concerne aussi les centres de données des entreprises. C'est le troisième cas de figure : les salles blanches des entreprises (infrastructures non mutualisées) et/ou des prestataires d'infogérance, qui représentent près des trois-quarts des datacenters. La question étant alors, à l'heure du cloud et de la virtualisation à tout va, quel sort attend les personnels exploitants (internes ou prestataires) ? Regain d'effectif ou au contraire encore moins de postes ? Des emplois de proximité ou au contraire encore plus délocalisables ?

« *L'équation économique n'est pas la même* », complète Fabrice Coquio. A commencer par les paramètres de la consommation énergétique, et de l'accès aux backbones télécoms dans les meilleures conditions, cruciaux pour les très grands hébergements mutualisés. Alors, que du côté des DSI gérant leurs propres salles ou des infogéreur, peut primer la possibilité de s'implanter là où le mètre carré est moins cher, mais proche néanmoins des viviers de compétences. Réponse nuancée (et chiffrée!) de Luc De Clerck, directeur des services d'Econocom en France : « *on aura toujours besoin d'une équipe technique sur site* ». La règle connue dans le milieu des infogéreur étant de 4 personnes pour 250 mètres carrés de salle blanche. Mais virtualisation et réseaux aidant, « *80% des interventions sur les serveurs ou sur les baies peuvent se faire à distance* ». Autrement dit, une activité confiée aux centres de services, proches ou non des infrastructures. Pour ses activités d'infogéreur/gérant de datacenters, Econocom prévoit d'augmenter les équipes des centres de services (60 personnes à Tours) de 15 à 20%, comme pour l'ensemble de sa division Services (150-200 embauches en 2012).

DOCUMENT 8

La Voix du Nord, Roubaix, Publié le 27/09/2013

LES VISAGES DE L'ACTUALITÉ

OVH.com offre de la mémoire aux cerveaux de l'université Lille1

STÉPHANE HUBIN

Quoi de mieux pour l'avenir de ses produits que de les faire découvrir librement à ses utilisateurs de demain ? C'est ce qu'a imaginé la société roubaisienne OVH.com, leader européen de l'hébergement de données, qui a signé un partenariat innovant avec Lille 1, à Villeneuve-d'Ascq. Le principe : chaque étudiant bénéficie d'un compte de 50 Go de sauvegarde de données, auxquelles il peut avoir accès partout. Une opération gracieuse en échange de laquelle l'ex-start-up de 700 salariés, dont nombre d'anciens élèves de Lille 1, entend bénéficier des remarques de ces utilisateurs pour améliorer ses produits. Avant de partir à la conquête des facs d'Europe.

Déjà très lié au campus villeneuvois où ses cadres interviennent à longueur d'année dans quelques-unes des nombreuses formations en informatique qui y sont dispensées, OVH.com, le troisième hébergeur mondial de données, premier européen, né en 1999 à Roubaix, a renforcé ses liens avec cette forteresse régionale de l'enseignement des sciences et technologies qu'est Lille 1. Depuis une semaine, les 20 000 étudiants, mais aussi les enseignants et les chercheurs de l'université, peuvent ouvrir un compte individuel de 50 Go de sauvegarde de données qui

leur permet d'avoir accès à leurs cours, travaux et autres documents partout où qu'ils soient, sur le campus, dans le train, le métro, ou chez eux, sans rien disposer d'autre que d'un ordinateur portable ou d'un smartphone.

Cette solution de Cloud baptisée hubiC, est mise gracieusement à disposition de Lille 1 par l'entreprise roubaisienne qui compte bien s'y retrouver malgré tout. « Nous prendrons en compte les remontées des utilisateurs pour améliorer l'interface et plus largement nos produits de manière continue. Nous avons beaucoup à apprendre des étudiants de Lille 1 », explique très directement Adelaïde Moulière, chef de produit hubiC, pour qui ce genre de partenariat est appelé à faire tache d'huile.

« Cette première collaboration s'imposait car OVH.com est un partenaire naturel de Lille 1, un lieu à la pointe de la formation des futurs développeurs et administrateurs de sociétés telles que la nôtre et d'où sont issus beaucoup de nos salariés. Mais nous envisageons ensuite de généraliser ce système aux autres universités de la région avant d'aller nouer des relations de ce type un peu partout dans les universités d'Europe », dévoile Adelaïde Moulière.

Philippe Mathieu, vice-président de Lille 1, est lui aussi ravi d'un tel partenariat qui, assure-t-il, coulait de source. « Nous avons identifié depuis des années le besoin d'un espace de sauvegarde pour nos étudiants et nos enseignants. Un besoin devenu plus important avec l'équipement grandissant en informatique nomade. Or, nous connaissons bien et apprécions les gens d'OVH, dont l'offre nous a plu. Elle répondait à nos exigences. »

Comme celle de garantir l'hébergement des données en France, entre autres. « Pour nos chercheurs, la confidentialité des données est un aspect essentiel ». Ou de rendre compatible l'accès à ses comptes depuis n'importe quel périphérique et avec n'importe quel logiciel. Enfin, « il était important de collaborer avec une entreprise régionale qui intervient déjà dans nos formations », souligne Philippe Rollet, président de Lille 1, pas peu fier d'avoir vu le campus prendre le virage de l'université numérique il y a quinze ans, déjà. « Cette orientation continue de porter ses fruits », savourez-t-il. En une semaine, 1 000 étudiants ont déjà ouvert leur compte.

DOCUMENT 9

Data centers: la donnée écolo

Morgane TUAL 14 avril 2013 à 19:06

ENQUÊTE

Ces usines du numérique sont la cause de 2 % des émissions mondiales de CO₂. Les géants du web cherchent pourtant à sevrer leurs serveurs énergivores.

Il y en a des milliers dans le monde. Plus d'une centaine en France. Les data centers sont la face cachée d'Internet ; on y a recours tous les jours. Quand on consulte un site web, quand on envoie un mail, quand on publie une photo en ligne. Toutes ces données sont traitées et stockées dans ces immenses usines du numérique, alignant des armées de serveurs qui n'ont rien de virtuel. Et cela demande de l'énergie. Beaucoup. Pour fonctionner, certains data centers consomment plus qu'une ville de 100 000 habitants. A l'échelle mondiale, ces infrastructures englobent 1,5% de la consommation électrique, l'équivalent de la production de 30 centrales nucléaires.

En 2011, Google possédait à lui seul 900 000 serveurs. Depuis, le million a été largement dépassé. Et tous les géants du Web et de l'informatique suivent le mouvement. Avec l'explosion du *cloud*, qui nous invite à stocker toujours davantage de documents en ligne, le nombre de data centers va encore augmenter. Et leur consommation électrique avec. «*Problème : d'où vient cette électricité ?*» interroge Sébastien Blavier, chargé de campagne Climat-Energie à Greenpeace. «*Aux Etats-Unis, elle provient principalement du charbon et du nucléaire*», déplore-t-il.

Il y a un an, l'ONG a tiré la sonnette d'alarme, rappelant que les data centers étaient responsables de 2% des émissions de CO₂. En ligne de mire : Apple, Amazon et Microsoft, jugés peu regardants sur l'origine de leur électricité. «*Ce sont des entreprises du futur. Pourquoi s'appuient-elles sur des énergies du XIX^e et du XX^e siècle ?*» s'indigne Sébastien Blavier, pour qui «*l'enjeu est énorme*». Face aux critiques de Greenpeace, Apple a promis que tous ses data centers seraient alimentés par de l'énergie 100% renouvelable courant 2013. A commencer par celui de Maiden, en Caroline du Sud, où l'entreprise construit depuis 2011 deux centrales photovoltaïques d'une quarantaine d'hectares.

Eau de mer. Car les opérateurs de data centers n'ont pas tous attendu d'être épinglés par Greenpeace pour s'interroger sur leur consommation énergétique. Et pour cause : celle-ci est leur premier poste de dépenses... «*Nous menons des démarches d'économie d'énergie depuis 2004. Mais pour nous, il s'agit d'abord d'une question de coûts*, explique sans détours Octave Klabka, fondateur d'OVH, leader de l'hébergement internet en Europe. *Quand vous consommez deux fois moins d'énergie que vos concurrents, vos solutions sont significativement*

moins chères et vous faites la différence sur le marché.»

Mais comment y parvenir ? Il faut pour cela limiter l'usage... de la climatisation. Les centres de données contiennent des centaines, voire des milliers de serveurs, qui, comme n'importe quel ordinateur, produisent de la chaleur. Une chaleur généralement évacuée par un système de climatisation extrêmement gourmand en énergie. Dans la plupart des data centers, il faut autant d'électricité pour faire fonctionner les machines que pour les refroidir ! Les entreprises doivent donc innover pour résorber ce gouffre énergétique.

OVH a par exemple inventé son propre système de *water cooling*, ou «refroidissement à eau». «*Nous faisons passer des tuyaux d'eau froide dans chaque serveur*, explique Octave Klabka. *Au début, on se disait : "C'est de l'informatique, on ne peut pas y mettre de l'eau !"*» Mais ce système fonctionne et a considérablement réduit la consommation énergétique de ses onze data centers. Le *water cooling* séduit de plus en plus d'entreprises, parmi lesquelles Google, qui a même envisagé de créer des centres de données en pleine mer. Plus raisonnablement, en 2011, le géant américain en a installé un sur les côtes finlandaises, rafraîchi à l'eau de mer. Une première. Autre innovation, le *free cooling*, qui consiste à refroidir les machines grâce à l'air extérieur. Le fournisseur d'accès Celeste a par exemple construit à Champs-sur-Marne (Seine-et-Marne) un data center tout en hauteur, conçu pour aspirer l'air et le faire circuler efficacement dans le bâtiment. La climatisation ne se met en marche qu'aux périodes les plus chaudes de l'été. Résultat : 35% d'économie d'énergie.

Et quid de la chaleur émise en continu par les data centers ? Dans la plupart des cas, celle-ci est tout simplement évacuée dans l'air : un énorme gâchis énergétique. Puisque cela ne représente pas d'enjeu financier pour les opérateurs, rares sont ceux qui s'en préoccupent. Et pourtant, le potentiel est prometteur. En Seine-et-Marne encore, un data center alimente en chaleur le centre aquatique de Val d'Europe et une pépinière d'entreprises de 1 400 m². «*Ce dispositif nous permet de réduire de 90% les émissions de CO₂ de ces installations*, se réjouit Jean-Paul Balcou, président du Syndicat d'agglomération nouvelle de Val d'Europe. *Et nous avons prévu d'aller beaucoup plus loin, car ce data center a la capacité de chauffer 100 000 m² de bureaux.*»

Energies vertes, réduction de la consommation, recyclage de la chaleur... Ces initiatives suffiront-elles à limiter l'impact environnemental des data centers ? D'ici à 2020, la production de données pourrait être multipliée par 50, ce qui implique «*de plus en plus de big data, de big calculs et donc de gros data centers*», affirme Laurent Lefèvre, de l'Institut national de recherche en informatique et en automatique (Inria). Une perspective qui inquiète les défenseurs de l'environnement, mais aussi les opérateurs du stockage de données : comment gérer une telle masse sans faire exploser la facture énergétique ? *«Actuellement, les data centers les plus performants exécutent 2,5 gigaflops [2,5 milliards d'opérations en virgule flottante par seconde, ndlr] pour 1 watt. Il faudrait qu'en 2020, on atteigne 40 gigaflops pour 1 watt, explique Laurent Lefèvre. Il y a encore beaucoup d'innovations à apporter.»*

L'évolution des composants électroniques devrait aider. Grâce à l'explosion du marché mobile, les fabricants conçoivent du matériel toujours plus puissant, toujours plus petit, pour la même consommation énergétique. De quoi multiplier l'efficacité de chaque serveur sans polluer davantage. Un autre enjeu concerne le «surprovisionnement» : les machines hébergées dans les data centers tournent en permanence à plein régime, même si elles ne sont utilisées qu'à 20% de leur capacité. *«Dans ce cas, 80% de leur alimentation ne sert à rien»,* souligne Laurent Lefèvre. Pas très écolo... Mais difficile de faire autrement : les données doivent être accessibles en permanence et les

serveurs capables de faire face à d'importants pics d'activité. La technique de la virtualisation pourrait toutefois changer la donne : *«Elle permet aux services hébergés de migrer entre différentes machines, en fonction des besoins. Les serveurs non utilisés peuvent alors être éteints, puis rallumés en cas de pic d'activité.»* Mais, dans la pratique, difficile d'allumer et d'éteindre les machines au gré de la demande, au risque de perturber le système de refroidissement. Les ingénieurs ont encore du travail pour résoudre le casse-tête environnemental posé par les data centers.

«Ondulateurs». Pour certains toutefois, ils ne représentent en rien un danger pour la planète, bien au contraire. *«Ce sont des outils majeurs de la réduction de l'impact carbone de l'activité humaine, assure Stéphane Duproz, directeur général de Telecitec France, un important opérateur de data centers. Avec le numérique, plus besoin de se déplacer pour acheter un billet de train par exemple.»* Surtout, l'émergence de gros centres de données permet, selon lui, d'économiser de l'énergie : *«S'ils n'existaient pas, il y aurait une multitude de petits data centers directement installés dans les entreprises. Chacun devrait disposer d'ondulateurs, d'un groupe électrogène... En mutualisant ces équipements, nous réduisons leur impact carbone.»* Cela dit, lui aussi cherche à rendre ses data centers plus écolos, car *«dans cette industrie, tout le monde veut être green !»* Et le faire savoir... La majorité des nouveaux centres s'autoproclament «100% verts», quelle que soit la réalité. Ici aussi, le *greenwashing* a de l'avenir.

DOCUMENT 10

Se chauffer grâce à l'énergie des serveurs informatiques

Le Monde.fr | 01.07.2013 à 16h13 • Mis à jour le 02.07.2013 à 11h16 | Par [Audrey Garric](#)

Chauffer des bâtiments grâce aux ordinateurs. L'idée émerge face au développement rapide des data centers, ces usines du numérique dont les puissants serveurs informatiques émettent tellement de chaleur que leurs opérateurs peinent à s'en débarrasser. A l'échelle européenne, la Commission estimait en 2008 que les centres de données consommaient 56 milliards de kilowatts, dont la moitié pour refroidir les bâtiments.

Lire (édition Abonnés) : [Les centres de données informatiques, gros consommateurs d'énergie](#)

Jusqu'ici, cette énergie était tout simplement évacuée dans l'air au moyen de différents systèmes de climatisation. Mais depuis quelques années, des initiatives de récupération et réutilisation de ces calories sont mises en place pour chauffer des logements, des bureaux, des usines, des piscines ou même des serres végétales.

C'est le cas, à petite échelle, à Amsterdam, aux Pays-Bas, où l'université a accès à de l'eau chaude gratuite provenant du data center de l'opérateur Equinix, qui partage son campus. En France, à Roubaix (Nord), les cinq centres de données d'OVH, le leader français de l'hébergement, chauffent aussi gratuitement les bureaux de l'entreprise, ainsi que quelques firmes voisines. Le data center de l'opérateur britannique Global Switch alimente également une serre tropicale dédiée à la production des fleurs de la ville de Clichy (Hauts-de-Seine). Autre exemple : à Uitikon, en Suisse, une piscine publique est chauffée avec un centre de données d'IBM.

RÉSEAU DE CHAUFFAGE URBAIN

Mais l'expérience la plus large, en Europe, se déroule dans le [parc d'activités de Val d'Europe](#), à Marne-La-Vallée (Seine-et-Marne), tout près du parc Eurodisney, une zone en développement d'une surface prévue de 180 hectares (dont 40 ha aujourd'hui bâtis).

Depuis septembre, la chaleur émise par le data center voisin de la banque Natixis est récupérée pour alimenter un réseau de chauffage urbain. Comment ? L'eau chaude des condenseurs des systèmes de climatisation est récupérée par des échangeurs thermiques, puis distribuée sur le réseau via une centrale de production d'énergie. Au final, ce système fait circuler une eau à 55 °C dans un réseau de 4 km de canalisations, auquel doivent être raccordés différents bâtiments.

Pour l'instant, un centre aquatique et une pépinière d'une quinzaine d'entreprises s'y sont connectés. *"On est aussi en train de raccorder deux hôtels près du parc Eurodisney. D'ici à deux ou trois ans, une centaine de logements locatifs devraient faire de même"*, indique Jean-Philippe Buisson, directeur Ile-de-France de Dalkia, filiale de Veolia Environnement et EDF, à l'origine de l'expérimentation.

5 400 TONNES DE CO₂ ÉCONOMISÉES

A terme, le data center fournira 26 millions de kilowattheures par an, à même d'alimenter en chauffage et en eau chaude sanitaire 600 000 m² de bâtiments. *"Cela permettra d'économiser annuellement le rejet de 5 400 tonnes de CO₂, soit les émissions entraînées par 5 000 voitures chaque année"*, se félicite Jean-Philippe Buisson.

Contrairement aux autres expériences menées à plus petite échelle ailleurs en Europe, ce chauffage n'est pas gratuit. *"Nous avons investi 4 millions d'euros dans le projet, justifie M. Buisson. Ce chauffage revient aux clients 8 centimes le kWh, soit un prix inférieur au chauffage électrique, mais un peu supérieur à une installation au gaz ou au fioul. C'est toutefois une énergie décarbonnée, dont le prix n'augmentera pas plus que l'inflation, contrairement au gaz."*

