CONCOURS
TECHNICIEN TERRITORIAL

EXTERNE

SESSION 2016

EPREUVE DE QUESTIONS

ÉPREUVE D'ADMISSIBILITE :

Réponses à des questions techniques à partir d'un dossier portant sur la spécialité au titre de laquelle le candidat concourt.

Durée : 3 heures
Coefficient : 1

SPECIALITE : INGENIERIE, INFORMATIQUE ET SYSTEMES D'INFORMATION

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 celles figurant le cas échéant sur le sujet ou dans le dossier** ne doit apparaître dans votre copie.
- Seul l'usage d'un stylo à encre soit noire, soit bleue est autorisé (bille non effaçable, plume ou feutre). L'utilisation d'une autre couleur, pour écrire ou pour souligner, sera considérée comme un signe distinctif, de même que l'utilisation d'un surligneur.
- L'utilisation d'une calculatrice de fonctionnement autonome et sans imprimante est autorisée.
- 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 18 pages

Il appartient au candidat de vérifier que le document comprend le nombre de pages indiqué
S'il est incomplet, en avertir le surveillant

- Vous préciserez le numéro de la question et le cas échéant de la sous-question auxquelles vous répondrez.
- Vous répondrez aux questions à l'aide des documents et de vos connaissances.
- Des réponses rédigées sont attendues et peuvent être accompagnées si besoin de tableaux, graphiques, schemas …

QUESTION N°01 (8 points) :

1A. Quels sont les principaux enjeux et défis à relever pour les collectivités locales en termes de Big Data ?

1B. Quels sont les critères de choix entre les différentes solutions logicielles (libres, propriétaires, full-web) ?

QUESTION N°02 (4 points) :

2A. Décrire le principe du BYOD et les enjeux liés à sa pratique au sein d'un Système d'Information.

2B. Quelles mesures préconisez-vous pour assurer la sécurité d'un S.I. auquel sont connectés des terminaux personnels ?

QUESTION N°03 (4 points) :

3A. Qu'est-ce que le « cloud computing » et quels sont les services qu'il offre ?

3B. Quels sont les avantages et inconvénients des solutions « Cloud public » et « Cloud privé » ?

QUESTION N°04 (4 points) :

Quels sont les éléments techniques et juridiques à prendre en compte avant de mettre en ligne des données publiques ?

Liste des documents :

DOCUMENT N°01 « Open data : comment publier les données ? ».
Site Regards sur le numérique / 11/03/2011. (2 pages)

DOCUMENT N°02 « Le BYOD, l'ennemi public numéro un des DSI ? ».
Décision Achats / N° 161 - 25/01/2013. (3 pages)

DOCUMENT N°03 « Guide sur le Cloud Computing et les Datacenters à l'attention des collectivités locales » (Extrait)**.**
Caisse des Dépôts / Juillet 2015. (**5** pages)

DOCUMENT N°04 « BYOD ou COPE : un dilemme pour la D.S.I. ? ».
Site zdnet.fr / 11/10/2013. (1 page)

DOCUMENT N°05 « SIG libre ou commercial ? Présentation et réflexions ».
Géomatique Expert - Mathieu Le Moal / N° 80 - Mai-Juin 2011. (4 pages)

DOCUMENT N°06 « Big Data : enjeu économique ou menace sociale ? ».
Revue l'Hémicycle – Manuel Singeot / N° 460 – 27/02/2013. (1 page)

DOCUMENT N°01

Open data : comment publier les données ?

Site Regards sur le numérique – 11/03/2011.

De plus en plus de collectivités locales envisagent de mettre en ligne leurs données publiques, à l'instar de Rennes et de Paris. Mais une fois la décision prise de se lancer dans le mouvement de l'Open data, d'autres questions subsistent. Nous vous proposons quelques pistes pour réussir au mieux ce lancement.

Tout d'abord, quelles données publier ?

En théorie, la réponse est simple : toutes les données publiques d'une collectivité peuvent être publiées, à l'exception des données nominatives, de celles qui relèvent de la vie privée et de la sécurité. La pratique s'avère beaucoup plus complexe.

D'abord, certaines données sont à la frontière du public et du privé. Se pose ensuite la question de la dangerosité potentielle de certaines informations : les plans des conduites de gaz, par exemple, sont des données publiques… et hautement sensibles à la fois. Une analyse juridique préalable est donc préférable avant toute publication.

Avec quelle licence ?

Le choix de la licence est important pour encourager et protéger en même temps les utilisateurs potentiels. La licence définit ainsi le caractère payant ou gratuit de l'exploitation des données, le degré de liberté accordé aux utilisateurs, les modalités d'accès aux données, ou encore les contraintes pour l'émetteur concernant la mise à jour des données.

Aujourd'hui, plusieurs types de licence coexistent, proposant différents cadres de réutilisation. En complément, l'APIE a élaboré deux modèles de licence type pour les administrations lorsque la réutilisation est soumise à des conditions particulières et/ou au paiement d'une redevance.

Comment publier les données ?

Trois cas de figure sont possibles.

1) **Publier les données sous un format facilement accessible**

   La publication sous la forme de fichiers dits *« plats »*, fichiers Excel par exemple, que l'utilisateur télécharge intégralement. Ce mode de publication convient aux données cartographiques.
   Mais pour les données plus complexes, comme celles qui sont livrées en temps réel (le temps d'attente dans la file d'un musée par exemple), il n'est pas forcément le plus facile à utiliser, notamment par les développeurs qui veulent concevoir des services ou des applications. Ces derniers doivent en effet héberger les données chez eux et les *« reprogrammer »* avant d'envisager une quelconque exploitation.

2) **Publier des API**

   Pour faciliter le travail des développeurs et les encourager à utiliser les données, il est possible de publier des API (interfaces de programmation) qui vont leur permettre de faire directement des requêtes au sein des données, sans avoir besoin de les télécharger intégralement.

Seul bémol : les API sont spécifiques aux données publiées. En d'autres termes, il existe autant d'API que de jeux de données, ce qui ne va pas dans le sens d'une mutualisation des efforts des développeurs qui voudraient utiliser les données de plusieurs sources différentes.

3) **Utiliser un protocole standard**

   Enfin, on peut avoir recours à un protocole basé sur des standards du web, qui permet d'utiliser le même langage et le même format quelles que soient les structures de données. L'avantage de cette solution, c'est qu'elle permet de développer plus rapidement des applications et favorise l'émergence d'un écosystème de développeurs autour des jeux de données.

Parmi ces protocoles, on peut citer SPARQL, développé par l'organisme de standardisation W3C, ou OData (solution développée par Microsoft, l'éditeur de RSLN, NDLR) adopté, par exemple, par la ville d'Edmonton, au Canada.

Où héberger les données ?

Reste enfin la question de l'hébergement des données. L'Open data suppose une plateforme capable d'héberger un nombre croissant de données et d'applications, totalement sécurisée, accessible à tous et en mesure de supporter de forts pics de trafic. Le nuage s'impose de plus en plus ici comme un partenaire naturel.

DOCUMENT N°02

Le BYOD, l'ennemi public numéro un des DSI ?

Décision Achats N°161 – 25/01/2013

A marche forcée depuis trois ans, le phénomène BYOD n'en finit plus de voir ses adeptes se multiplier... au risque de fragiliser la sécurité du SI.
Le danger est réel mais les parades existent, à condition de bien s'y préparer.

Plus d'un salarié sur deux déclare utiliser son terminal personnel (laptop, smartphone, tablette, etc.) dans le cadre d'activités professionnelles, selon l'Institut Forester. Même son de cloche pour l'Ifop qui annonce que deux tiers des personnes équipées d'un terminal dit «intelligent» l'utilisent aussi à titre professionnel. En revanche, parmi cette foule de personnes reliées au SI de l'entreprise, il n'est nullement précisé combien d'entre elles ont été dûment autorisées et identifiées à s'y connecter. Selon une récente étude publiée par Absolute Software et Vanson Bourne, près de 40 % des entreprises françaises n'ont pas mis en place de politique, ni même de règles pour gérer l'utilisation des appareils mobiles sur le lieu de travail. En perdant le contrôle sur le parc de terminaux connectés, la DSI semble désarmée pour limiter les risques potentiels inhérents à ce parc mobile qui circule «en toute impunité». Pour autant, le tableau n'est pas aussi noir qu'il n'y paraît.

Le BYOD étant un sous-ensemble de la mobilité, la DSI aura donc pour responsabilité de déterminer les données «autorisées» à sortir de l'entreprise.

Thierry Karsenti, Check Point

«Il est impossible de savoir, dans un premier temps, si le système du terminal est sain ou s'il s'agit d'une «passoire» qui possède des logiciels espions... qui peuvent se propager dans le SI de l'entreprise.»

Evaluer les risques potentiels

«*Avec la perte de contrôle des DSI sur les terminaux mobiles, des risques supplémentaires sont susceptibles d'apparaître, le SI devenant plus vulnérable*», souligne Jérôme Saiz, rédacteur en chef de Qualys Security Community (acteur de la sécurité informatique). Les risques du BYOD sont liés aux capacités délivrées par les terminaux mobiles, au même titre que ceux qui sont fournis par les RSSI (responsables service des systèmes d'information), à une différence près: la DSI, en n'ayant aucune visibilité sur les terminaux, ne maîtrise pas le moment où le collaborateur vient se connecter au SI de l'entreprise. Par ailleurs, la nature technologique de l'appareil et le niveau de sécurité configuré lui sont inconnus. «*Il est impossible de savoir, dans un premier temps, si le système du terminal est sain ou s'il s'agit dune «passoire» qui possède des logiciels espions, des chevaux de Troie... qui peuvent se propager dans le SI de l'entreprise*», explique Thierry Karsenti, directeur technique pour la zone Europe chez Check Point (spécialiste de la sécurité réseaux). D'autre part, le terminal n'appartenant pas à l'entreprise, un mélange de données à la fois personnelles et professionnelles est possible mais cette «cohabitation» ne peut pas être permise, lorsque les données ont un caractère confidentiel et sensible. Sans parler du droit de regard des managers sur les données professionnelles, qui ne doit pas s'étendre aux données personnelles: «*Il y va de la responsabilité de l'entreprise de bien différencier ce qui relève de la bulle professionnelle de ce qui concerne uniquement la sphère privée*», confirme Thierry Karsenti.

Agostinho Rodrigues, Interdata

«En installant une brique de sécurité logicielle, il est possible de cloisonner les environnements personnel et professionnel et de maîtriser ce dernier. »

Des solutions humaines et technologiques adéquates

Deux types de mesure sont envisageables pour limiter les risques liés à la sécurité du BYOD. Une communication préventive permettra d'avoir d'abord une action sur le comportement des collaborateurs. Les managers (DSI, RSSI) doivent alerter et sensibiliser les utilisateurs sur les risques potentiels de leurs terminaux. «*Les RSSI peuvent également organiser des sessions de formation pour éduquer les collaborateurs à l'usage de leurs terminaux*», précise Jérôme Saiz. La DSI dispose également d'un «arsenal» technologique pour pallier les mauvais comportements involontaires, et parfois volontaires, des utilisateurs afin de limiter ou de prévenir tout dégât potentiel sur le SI. A commencer par la maîtrise du terminal. En effet, des outils de MDM (mobile device management) sont disponibles pour gérer les appareils à distance. On peut ainsi activer le chiffrement des données sur les OS des terminaux ou exiger l'entrée d'un mot de passe (ou d'un pin) pour lancer le terminal. Un second niveau consiste à faire appel à des éditeurs spécialisés ou à se lancer dans du développement spécifique. Il s'agit en l'occurrence de mettre en place une brique de sécurité logicielle qui permet de cloisonner les environnements personnel et professionnel et de mettre ce dernier dans une bulle sécurisée dont l'accès est soumis à une autorisation/authentification. «*Etant confinée, cette partie du terminal peut ainsi être maîtrisée par l'administrateur, tout en réglant les droits d'accès aux logiciels métiers en fonction du profil de l'utilisateur*», explique Agostinho Rodrigues, directeur des nouvelles technologies chez Interdata, intégrateur réseau et sécurité. Il est également possible de lancer à distance un effacement total du contenu de l'appareil, suite à la perte/vol de son mobile. Compte tenu des risques potentiels, la question de la sécurité du BYOD est donc devenue un enjeu majeur. D'autant que les ventes de smartphones deviendront majoritaires en 2013, selon le cabinet iSupply, et que celles de tablettes ont déjà dépassé celles des PC au second semestre 2012.

Trois questions à ... Hervé Doreau, responsable offre sécurité chez Symantec : *« Le BYOD représente un phénomène subi pour les DSI »*

Quels sont les comportements des DSI par rapport au BYOD ?

En théorie, de nombreux DSI et RSSI sont contre l'utilisation de terminaux mobiles apportés par les collaborateurs au coeur de l'entreprise, ou dans le cadre d'un service de mobilité. En réalité, ce sont les mêmes qui, bien que s'y opposant farouchement, ferment finalement les yeux, en raison d'une évidence: ils ne peuvent pas faire grand-chose pour contrer le phénomène. Ils risquent d'ailleurs de déclencher une contre-productivité, si des mesures sont prises pour véritablement interdire le BYOD. Le BYOD représente donc un phénomène subi.

Comment se passe finalement l'intégration des terminaux mobiles ?

L'intégration des terminaux au SI de l'entreprise est finalement assez classique puisqu'ils sont contrôlés via des outils de MDM (mobile device management).

Autrement dit, diverses applications vont permettre de gérer les terminaux, à l'instar de ceux appartenant à l'entreprise. Or, comme l'usage est mixte, à la fois privé et professionnel, cette solution s'avère dépassée car l'utilisateur ne va pas forcément accepter toutes les contraintes liées au contrôle de son terminal (présence d'un mot de passe au démarrage du terminal par exemple).

Que faut-il faire en conséquence pour assurer la sécurité du terminal ?

La priorité est de pouvoir gérer les applications et les données qui font partie de la sphère professionnelle et de ne pas imposer de contraintes lorsque l'utilisateur se sert de son mobile pour un usage privé. Le mot de passe doit être placé non pas lorsque le terminal démarre, mais lorsque l'utilisateur veut accéder à son espace professionnel. Le DSI aura alors un contrôle complet sur ce qui est autorisé ou pas. Il aura la main sur la manière dont sont stockées les données, à quel endroit, et si elles doivent être chiffrées/cryptées, etc. Cette stratégie va au-delà du MDM, il s'agit davantage de MAM (mobile application management) et de MIM (mobile information management).

DOCUMENT N°03

Guide sur le Cloud Computing et les Datacenters à l'attention des collectivités locales

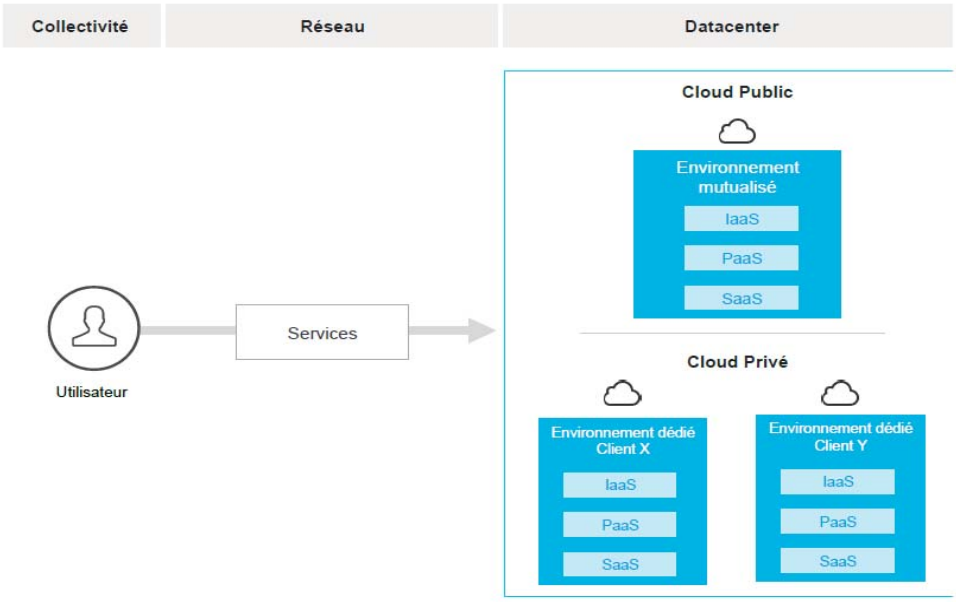
Caisse des Dépôts – Juillet 2015 (Extrait).

1.2. Qu'est-ce que le *Cloud computing* ?

Une définition pour le *Cloud Computing* ?

« Le *Cloud computing* est un modèle d'accès à travers le réseau internet à un ensemble de ressources numériques, pouvant être allouées et libérées à la demande et pour lesquelles le fournisseur du service assure l'ensemble des activités de maintenance, de support et d'exploitation7 »

Ce modèle offre des services de différentes natures, allant des services d'infrastructure (location de capacités de stockage ou de calcul), des services de plateforme (location d'environnements de développement préconfigurés) ou de services d'applications (location d'applications).



Ces services sont opérés par le fournisseur et rendus accessibles au travers du réseau Internet.
Ils sont déployés sur des environnements informatiques partagés et mutualisés. Pour garantir une meilleure qualité de service et assurer la haute disponibilité de leurs services numériques, les fournisseurs de *Cloud computing* hébergent ces environnements dans des infrastructures de type *Datacenter*.

Le succès d'un service *Cloud computing* repose sur trois composantes clés :

- la qualité de la solution logicielle et du support associé ;
- la disponibilité de la solution, liée au niveau de performance du *Datacenter* qui l'héberge et à la qualité de la connexion internet ;
- la sécurité de la solution, reposant en partie sur le niveau de sécurité du *Datacenter*.

Les caractéristiques du *Cloud Computing*

La principale spécificité, et le principal intérêt, du modèle *Cloud computing* est que le fournisseur opère un ensemble de ressources informatiques mis en oeuvre pour fournir le service. Le client final n'a plus la responsabilité des opérations de maintenance et d'exploitation des couches basses, tant sur la partie matérielle que logicielle. Ce modèle se rapproche ainsi des stratégies d'hébergement et d'infogérance mises en oeuvre depuis les années 1990 et plus largement adoptées dans les années 2000.

Le *Cloud computing* se distingue toutefois des solutions « traditionnelles » par les caractéristiques8 suivantes :

- **une large accessibilité via le réseau** : les services sont accessibles en ligne et sur tout typede support (ordinateur de bureau, portable, smartphone, tablette). Tout se passe dans le navigateurinternet ;

- **mesurabilité du service** : l'utilisation du service par le client est supervisée et mesurée afin de pouvoir suivre le niveau de performance et facturer le client en fonction de sa consommation réelle ;
- **une solution multitenant ou multiclient** : une même instance d'un logiciel est partagée par l'ensemble des clients de façon transparente et indépendante. Tous les clients utilisent la même version du logiciel et bénéficient instantanément des dernières mises à jour. Chaque client dispose d'un paramétrage utilisateur qui lui est propre ;
- **une disponibilité à la demande** : le service peut être souscrit rapidement et rendu opérationnel automatiquement avec un minimum d'interaction avec le fournisseur ;
- **l'élasticité immédiate des ressources** : des ressources supplémentaires peuvent être allouées au service pour assurer la continuité du service en cas de pic de charge, ou être bien réallouées à un autre service dans le cas inverse ;
- **la mutualisation des ressources** : les ressources utilisées pour exécuter le service sont mutualisées pour servir à de multiples clients. Les multiples serveurs sollicités, totalement interconnectés, ne forment plus qu'une seule ressource virtuelle puissante et performante.

Cette caractéristique est permise lorsque la solution déployée est multilocative.

Les contraintes du *Cloud Computing*

La mise en oeuvre d'une solution de *Cloud computing* implique cinq principales contraintes qui doivent être prises en compte lors du choix puis du déploiement du service :

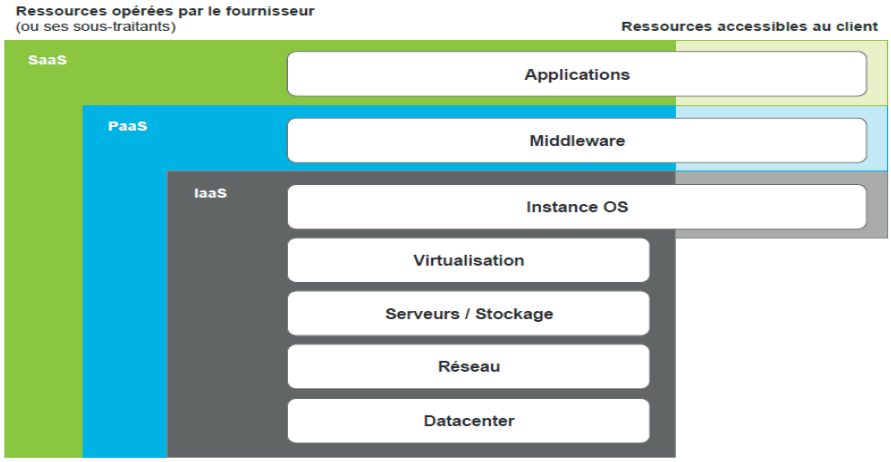
- **la sécurité des accès et des données** : les risques de sécurité sont accrus par la délocalisation des ressources. L'accès au service induit donc des connexions sécurisées et authentifiées pour les utilisateurs et des garanties sur la confidentialité, l'intégrité et la traçabilité des données confiées stockées sur l'infrastructure du fournisseur. L e risque de perte de donnée doit également être anticipé via une procédure de sauvegarde adaptée ;
- **la localisation des données** : le *Cloud* n'ayant pas de frontière, il faut pouvoir disposer d'un engagement sur les lieux de stockage des données, et s'assurer que les règlementations des pays sont en conformité avec les règlementations auxquelles le client est soumis ou bien qu'il souhaite avoir ;
- **le niveau de qualité du service** : le client doit disposer d'un droit de regard sur la qualité des prestations et le niveau de performance du service, avec un engagement contractuel de la part du fournisseur ;
- **le coût à long terme** : si les dépenses d'investissement sont limitées ou nulles, les dépenses d'exploitation doivent être correctement évaluées en disposant de métriques précises et à l'avantage du client pour mesurer la consommation réelle des ressources ;
- **la réversibilité du service** : en cas de rupture de contrat ou de changement de fournisseur, le client doit s'assurer de la récupération et de la destruction de ses données sur l'infrastructure du fournisseur après sa migration.

Selon les modes de déploiement des services de *Cloud computing*, ces éléments peuvent évoluer (voir plus loin).

1.3. *IaaS*, *PaaS* et *SaaS* : des offres de services à valeur ajoutée

Les Services de *Cloud Computing*

Les offres de *Cloud computing* se décomposent en trois familles de services : *IaaS*, *PaaS* et *SaaS*. Pour les distinguer, il faut considérer les différentes couches fonctionnelles qui composent un service numérique.



IaaS – Infrastructure as a Service, des serveurs virtuels disponibles à la demande

Il s'agit de l'offre la plus basique dans le portefeuille *Cloud computing* correspondant à la location de capacités de calcul et de stockage

Dans un service *IaaS*, le fournisseur met à disposition et administre les ressources matérielles virtualisées comprenant :

- la puissance de calcul ;
- les unités de stockage des données ;
- les réseaux ;
- les couches de virtualisation ;

Les systèmes d'exploitation.

- Le client prend en charge la gestion et l'exploitation de toutes les couches supérieures, *middlewares*, bases de données, applications.
- La souscription à une offre *IaaS* permet au client d'externaliser son parc matériel serveur et de s'affranchir des compétences de conception et d'exploitation des infrastructures techniques.

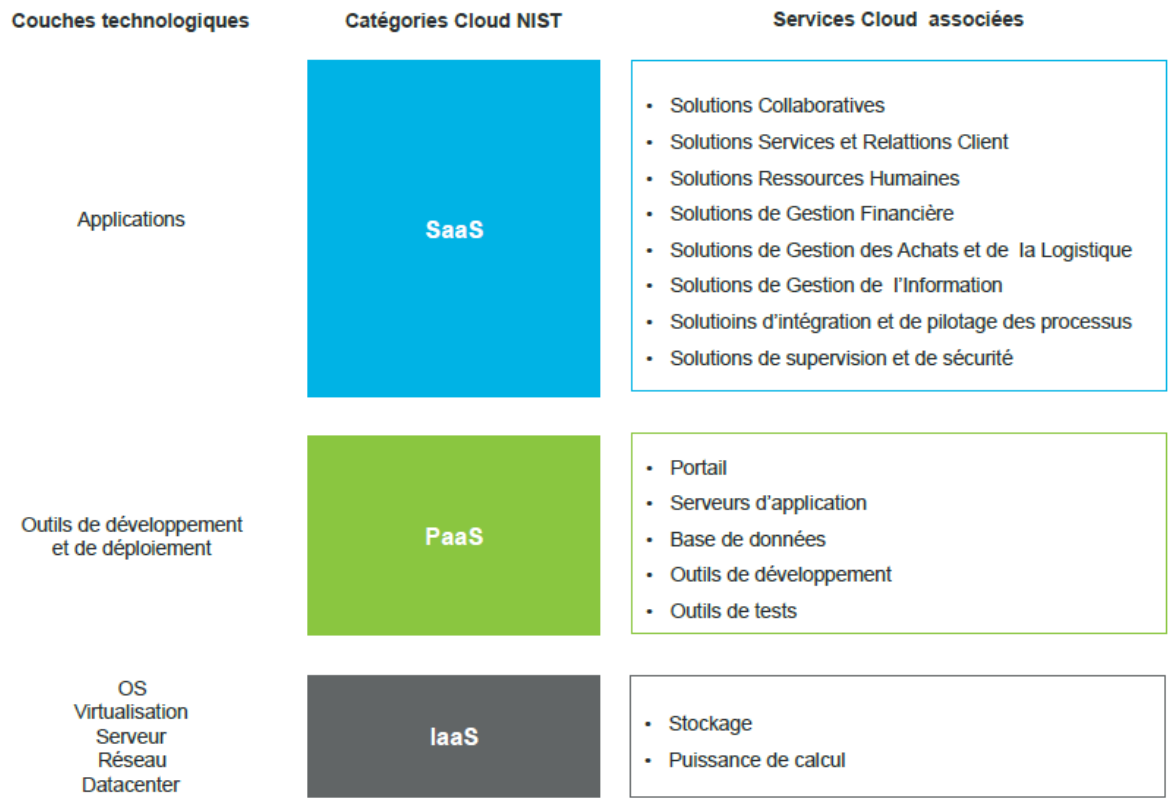
PaaS – Platform as a Service, des plateformes de développement prêtes à l'emploi

- Il s'agit de l'offre intermédiaire dans le portefeuille *Cloud computing*.
- Le fournisseur met à disposition une plateforme middleware opérationnelle, incluant des serveurs d'applications, des bases de données et les outils permettant au client de développer et de déployer ces propres applications.

- Cette configuration est très employée pour disposer de plateformes de développement ou de tests disposant de l'ensemble des outils et middleware nécessaires, en évitant ainsi les taches de construction et de maintenance de ces plateformes non critiques. Elle se destine donc naturellement avant tout aux développeurs.

SaaS – Software as a Service, des services métiers à la demande

Il s'agit d'une offre « tout compris ». Le prestataire met à disposition une application qu'il administre et configure en majeure partie. Le client externalise ainsi ses applications (logiciels métiers ou solutions techniques à destination des DSI), auxquelles il accède à la demande. Il paie à l'usage, selon le nombre d'utilisateurs et/ou le temps d'utilisation du logiciel.



Cloud Computing vs Hébergement et Infogérance

Le *Cloud computing* se différencie des prestations d'hébergement et d'infogérance traditionnelles.
Dans le cadre de la fourniture d'une solution logicielle en mode *Cloud computing*, l'éditeur fournisseur du service s'engage un effet à prendre à sa charge l'ensemble des opérations d'hébergement et d'infogérance associée à ce service.

- **Hébergement** : mise à disposition par le prestataire de ressources (serveurs physiques ou serveurs virtuels) avec une capacité fixée par contrat. Ces ressources sont mises à la disposition exclusive du client. Le fournisseur est responsable de la maintenance des équipements serveurs.

- **Infogérance** : contrat de prestation incluant l'exploitation et de maintenance des ressources numériques d'un client. L'infogérance peut s'appliquer sur plusieurs niveaux :
  - Infrastructure IT : exploitation et maintenance des serveurs ;
  - *Middleware* : exploitation et maintenance des ressources middleware (serveurs d'application bases de données …) ;
  - Application : exploitation et maintenance des applications métiers.

- **Services de *Cloud Computing*** : mise à disposition par le prestataire de ressources à la demande impliquant l'ensemble des activités d'exploitation, de maintenance et de support associé et pouvant couvrir plusieurs niveaux :
  - Infrastructure IT : solutions *IaaS* ;
  - *Middleware* : solutions *PaaS* ;
  - A pplication : solutions *SaaS*.

Dans le cas de l'utilisation de services de *Cloud computing*, les fournisseurs de services doivent également mettre à disposition l'ensemble des outils permettant aux clients de :

- piloter automatiquement l'allocation ou la désallocation de ressources ;
- suivre la consommation de ses ressources (et contrôler ainsi la facturation du service, fourni à la demande) ;
- suivre la qualité du service et le niveau de performance.

Les modes de déploiement du *Cloud*

Selon le degré de mutualisation des ressources, on distingue plusieurs modes de déploiement, garantissant une réservation plus ou moins exclusive au client utilisateur des capacités physiques qui lui sont allouées.
Dans tous ces modes, l'étanchéité entre les ressources allouées à chaque client est toujours garantie.

Cloud public – des capacités partagées à bas prix

Dans le cas du *Cloud* public, le fournisseur propose un environnement informatique avec une mutualisation optimale des ressources : l'environnement est ainsi virtuellement partagé avec un nombre illimité de clients.
Une telle offre privilégie l'élasticité et la flexibilité, et propose, avec la massification des clients des prix plus attractifs.
Cependant, la mutualisation généralisée entraîne une difficulté plus grande pour le fournisseur de garantir un niveau de service spécifique à un client.

Cloud privé ou *Cloud* dédié – des capacités et des niveaux de services exclusifs

Le *Cloud* privé vient répondre aux exigences des clients souhaitant des garanties exclusives sur un contrat de service. Le prestataire met en place des ressources dédiées à un client donné, sur un environnement hébergé à demeure chez le client ou à distance chez le fournisseur.

Ainsi, l'infrastructure réservée au client bénéficie des technologies de virtualisation afin de pouvoir être mutualisée et les ressources disponibles automatiquement allouées en fonction desbesoins des applications internes. Le prestataire s'engage également sur des niveaux de services garantis spécifiques.

Cette architecture est particulièrement adaptée pour un client souhaitant un environnement spécifique tout en bénéficiant d'une allocation à la demande des ressources. Il pourra ainsi contractualiser avec le fournisseur des niveaux de services adaptés à ses besoins, en termes de sécurité ou de performance.

Principales différences entre *Cloud* public et *Cloud* privé

	Cloud public	*Cloud* privé
Type d'infrastructure	Environnement externe multilocatif	Environnement interne ou externe dédié au client
Élasticité / disponibilité en capacité	Illimitée (en théorie)	Limitée
Allocation de ressources	Immédiate	Immédiate dans la limite des capacités disponibles
Localisation des données	Dépend des partenaires du tiers fournissant le service en *Cloud* public	Définie par contrat
Sécurité et niveaux de services	Identiques pour tous les clients	Les niveaux d'exigences peuvent être définis de façon spécifique en fonction des besoins du client
Coût à long terme	Inférieur au coût d'un *Cloud* privé	Élevé en raison des niveaux de services spécifiques

En termes de sécurité, les différences entre *Cloud* public et *Cloud* privé sont uniquement liées aux engagements pris par le fournisseur :

- sur un *Cloud* privé, un client peut exiger des niveaux de services particuliers ;
- sur un *Cloud* public, le niveau de sécurité est le même pour tous les clients. La certification de la solution permet d'évaluer le niveau de sécurité réellement offert.

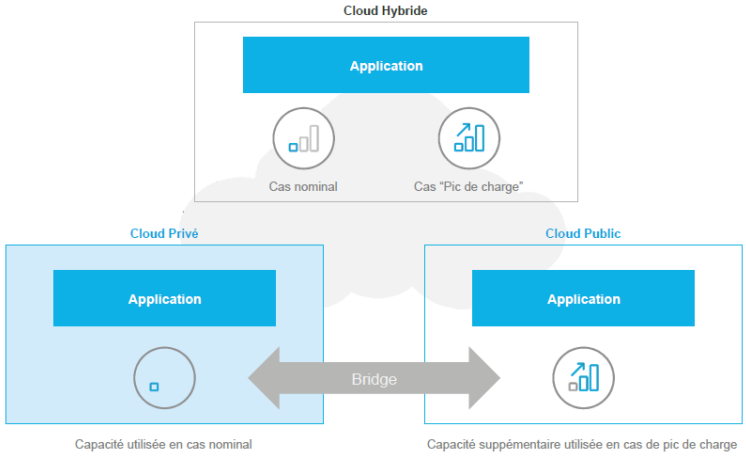
Il est faux d'affirmer qu'un *Cloud* privé est systématiquement plus sécurisé qu'un *Cloud* public.
C'est le niveau de certification qui doit faire la différence.

Cloud communautaire – des capacités mutualisées pour une communauté d'intérêt général identifiée

Le *Cloud* communautaire correspond à un *Cloud* privé partagé par un groupement d'acteurs.
Ce mode de déploiement vise à abaisser la barrière à l'entrée du *Cloud* privé tout en bénéficiant d'un niveau de service spécifique. Il est adapté à la mutualisation au sein d'un écosystème où la cogouvernance est envisageable (regroupement de collectivités ou d'acteurs publics, communautés d'universités, GIE interentreprises).

Cloud hybride – un environnement mixte pour gérer les besoins ponctuels

Le *Cloud* hybride est un mode de déploiement combinant l'utilisation d'un *Cloud* public avec un environnement en *Cloud* privé. Il s'agit dans ce cas de faire cohabiter ces environnements afin de pouvoir absorber plus facilement des pics de charge ponctuels : l'environnement privé est réservé aux systèmes courants, tandis que les capacités du *Cloud* public sont utilisées pour absorber ponctuellement les montées en charge.



Le terme d'Architecture hybride s'applique également pour désigner les environnements hétérogènes mixant des ressources hébergées sur un environnement interne (ou parle de ressource on-premise ou à demeure) et un *Cloud* public ou plusieurs environnements *Cloud* public.

DOCUMENT N°04

BYOD ou COPE : un dilemme pour la D.S.I. ?

Site zdnet.fr - 11/10/2013

La mise en place d'une politique de BYOD est-elle la seule issue pour permettre à la fois souplesse d'utilisation aux utilisateurs et contrôle de l'information par la DSI ?
Le concept de "Corporate owned – personally enabled" en prend le contrepied.

D'un côté, le BYOD propose de laisser les terminaux personnels pénétrer dans le système d'information de l'entreprise. PC portables, smartphones ou tablettes des utilisateurs peuvent se connecter, télécharger des données, bref, interagir avec le SI.

Pourquoi un tel engouement ? D'une part parce que le coût du terminal est supporté par l'employé. Libre à l'employeur de lui fournir une enveloppe prenant en charge tout ou partie du montant du terminal. Cependant, les aspects juridiques et de gestion des ressources humaines de la mise en place des politiques de BYOD impliquent des investissements qu'il est aujourd'hui difficile d'évaluer. Voir une réduction du TCO dans cet aspect est donc illusoire.

Conserver la logique du terminal fourni par l'entreprise

Mais l'arrivée des terminaux personnels dans l'entreprise permet aussi à la DSI de suivre l'évolution rapide des usages et des outils mobiles grand public. Un excellent outil de veille en somme, qui permet ensuite de s'approprier les meilleures pratiques et les meilleurs usages concernant la mobilité.

Face au BYOD, se dresse désormais un autre concept, qui propose lui à la DSI de conserver une prise sur le parc d'appareils mobiles des employés qui se connectent au SI (lire en anglais l'article de Wired sur le sujet).

Baptisée "COPE", pour Corporate owned – personally enabled, il prend le contrepied du BYOD en promettant que l'utilisateur pourra en toute sécurité pour l'entreprise et pour lui-même utiliser un terminal de l'entreprise pour un usage personnel et professionnel. Certains assurent que 70% des grands comptes dans le monde utiliseront cette politique dans les trois années à venir.

Vers un usage hybride, en fonction des profils ?

Côté DSI, le COPE garantit pleinement la maîtrise de la plateforme mobile ; ce qui devait être négocier dans le cas du BYOD. Côté utilisateur, si la DSI assure un rôle de contrôle du patrimoine applicatif et informationnel, de l'entreprise sur le périphérique, liberté peut être laissée ensuite pour installer des applications « autres », certifiées toutefois par la DSI sur la dimension sécurité.

Alors, COPE ou BYOD ? L'idée est de laisser la possibilité aux DSI de s'approprier l'un ou l'autre des modes de fonctionnement, voire pourquoi pas les deux en fonction des publics de l'entreprise, et de leurs exigences ; des solutions hybrides émergent sur le marché. Dans ce cadre, le DSI peut proposer ainsi des magasins d'applications « privés », sécurisés, et dont le périmètre est défini par ses soins.

Géomatique Expert - Mathieu Le Moal / N° 80 - Mai-Juin 2011.

SIG libre ou commercial ? Présentation et réflexions

MATHIEU LE MOAL ■ *AXES Conseil* ■ lemoal@axes.fr, www.axes.fr

En une quinzaine d'années (1995-2010), les systèmes d'information géographique (SIG) sont passés d'architectures simples et majoritairement propriétaires, de type station de travail, à des architectures plus complexes orientées *Web* et Services. Adossée à l'informatique, la géomatique en a simplement suivi les évolutions avec un léger décalage dû à ses spécificités : volume des données, volet graphique…

Un SIG répondant à *l'état de l'art* actuel s'appuie donc sur divers composants logiciels permettant le stockage des données, leur gestion, exploitation et diffusion. Pour la très grande majorité de ces composants d'un SIG, il existe maintenant une offre propriétaire et libre (ou encore *« open source »*). Cependant, la répartition entre des deux gammes pour les composants demeure hétérogène.

Dans certains domaines fonctionnels ou techniques des systèmes d'information, le libre a pris une ascendance visible sur l'offre commerciale (exemples : les CMS, les serveurs *web*). La situation dans le monde SIG est plus ambiguë, les offres commerciales, bien que déjà très consolidées, tirant toujours leur épingle du jeu.

Cet article présente donc un état des lieux de l'offre libre et propriétaire dans le contexte français et s'appuie en partie sur des statistiques internes, issues de l'observatoire des logiciels SIG créé par *Axes Conseil* (http://www.axes.fr/observatoire).

Définition des licences de type libre ou propriétaire

Il convient préalablement de préciser les notions de licence libre et propriétaire.

Une licence de type *« propriétaire »* est un contrat encadrant strictement l'utilisation du produit en limitant voire interdisant ses duplications ou copies par exemple ; elle est le plus souvent payante. Il existe cependant des produits gratuits sous licence propriétaire (*Flash*, par exemple).

Une licence dite *« libre »* est un contrat permettant à toute personne morale ou physique d'utiliser, d'étudier, de copier et de modifier l'objet du contrat (logiciel, données…).

Il existe de très nombreux types et sous-types de licences libres en fonction du domaine, de l'objet et des objectifs : *copyleft* ©, libre diffusion…

Un éditeur peut décider de basculer un produit d'une licence propriétaire vers une licence libre. Il doit alors *« ouvrir »* son produit en donnant par exemple accès aux codes sources dans le cas de logiciels. Ce fut le cas d'*AutoDesk* avec son utilitaire de publication *web MapGuide*.

Pour la suite de l'article, nous distinguerons les solutions diffusées par des éditeurs majoritairement sous licence propriétaire des solutions issues du travail d'une communauté placée sous licence libre.

Les éléments d'un SIG

Très schématiquement, l'architecture d'un SIG moderne s'articule autour de trois types de composants logiciels :

- Un système de gestion de base de données (SGBD) ;
- Une solution serveur permettant la diffusion et l'exploitation des données en mode *web* ;
- Une solution de type bureautique dédiée à la gestion et traitement des données.

À ce triptyque peut être ajoutée une solution mobile de consultation et de mise à jour. Le fonctionnement de ce système nécessite bien évidemment des données et des utilisateurs !

La base de données relationnelle (SGBDR)

L'offre en bases de données spatiales relationnelles est maintenant relativement restreinte. Peuvent être citées les principaux systèmes suivants :

- *Oracle Locator et Spatial (Oracle)* ;
- *PostGreSQL/PostGIS (Refraction/OSGeo)* ;
- *SQLServer* 2008 *(Microsoft)* ;
- *MySQL (Oracle/Sun)*.

Certes, des acteurs historiques des bases de données ont tenté de (*Ingres*, abandon début 2011) ou mis à disposition (IBM, avec DB2) des fonctionnalités géospatiales, mais les implémentations connues sont fort limitées. Des acteurs apparaissent (et disparaissent…) régulièrement, (exemple *Cassandra*, de la fondation *Apache*, *MariaDB* créée par le fondateur de *MySQL*). Ces systèmes sont généralement nativement dédiés au domaine classique de la gestion de données alphanumériques et ajoutent ensuite une fonctionnalité géospatiale à leur moteur.

Le marché de la base de données relationnelle spatiale en France se limite quasiment à *Oracle* (*Locator* ou *Spatial*) et *PostGreSQL/PostGIS*. Les deux autres systèmes (*SQL Server* et *MySQL*) sont théoriquement présents, mais, en pratique, dans des proportions nettement moins importantes.

La solution propriétaire proposée par Esri (*ArcGIS Server/ArcSDE*) tient une place particulière dans ce paysage. Ce produit est en effet indépendant du SGBDR qui le supporte. *ArcGIS Server* peut ainsi être installé sur *Oracle, PostGIS, SQL Server, Informix, DB2*…

Sur les quatre produits les plus présents en France, *PostGreSQL/PostGIS* s'est imposé comme LA solution libre. Les trois autres, dont *MySQL*, acquis à l'époque par *Sun*, elle-même rachetée par *Oracle*, appartiennent à des éditeurs bien établis. Il est d'ailleurs licite de se poser la question de l'avenir de la fonction géospatiale de *MySQL*, compte tenu de sa concurrence potentielle avec celles d'*Oracle*.

L'utilisation d'*Oracle* s'explique le plus souvent par l'histoire. Une entité déjà équipée d'*Oracle* dispose de compétences en interne et d'une expérience qu'elle souhaite valoriser. Si aucun pré requis technique, historique ou contrainte n'est posée, la solution *PostGreSQL/PostGIS* est le plus souvent proposée en priorité comme outil de stockage de données SIG (c'est, par exemple, le cas à l'IGN).

Devant la confirmation de *PostGreSQL/PostGIS* comme SGBDR libre de référence dans le monde SIG, les intégrateurs et éditeurs ont investi du temps et des développements sur son exploitation. Cet effet boule de neige renforce le poids de ce duo !

Base de données « *fichier* »

La solution *SQLite* et son module spatial *SpatiaLite* perce doucement dans le monde des SIG. Complémentaire aux SGBDR *Client/Serveur*, cette base de données relationnelle fichier, implantée dans de nombreux produits (*FireFox, Skype, MacOS X*…), devient un mode de stockage de données SIG intéressant. Son approche relationnelle permet une structuration de données plus fine et plus « *intelligente* », chose impossible dans de simples fichiers de données « *à plat* » de type SHP, TAB... Ces avantages dans le contexte de la mobilité ou d'échange sont évidents.

Esri propose une offre similaire avec la *GeoDataBase File*. Cette base de données permet le stockage conjoint des données et de valeurs ajoutées (relations, domaines de valeurs, etc.). Dans une démarche d'ouverture, l'éditeur américain a publié l'API version ß permettant l'exploitation de la *Géodatabase* fichier en dehors du monde *ArcGIS*.

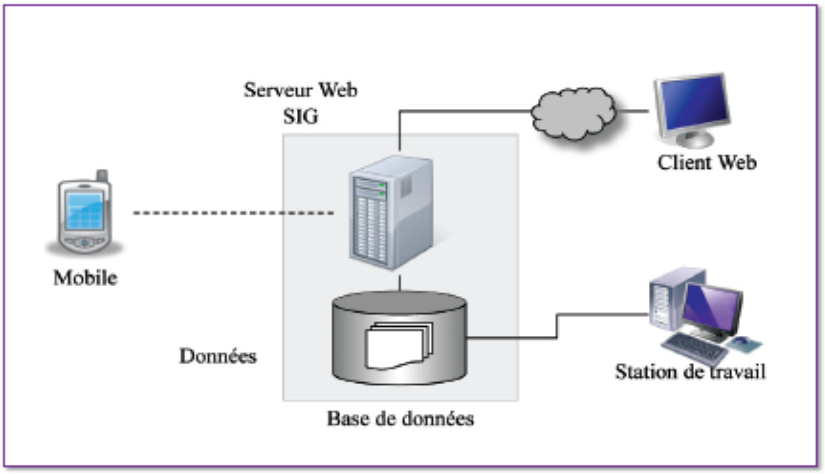
La station de travail (SIG bureautique)

Il convient de différencier trois types de cas :

- Les solutions propriétaires payantes ;
- Les solutions propriétaires gratuites ;
- Les solutions libres.

Solution propriétaire payante

Il s'agit de produits souvent complets proposant une très large plage fonctionnelle. Les solutions SIG génériques les plus répandues sur le marché français, sont, dans l'ordre alphabétique : *ArcGIS*, *Elyx* (*Star-Apic*), *GeoConcept*,*GéoMedia* et *MapInfo*.



Il existe également aux côtés des cinq majeurs précités, de très nombreux produits, souvent thématiques (urbanisme, assainissement, éclairage public, cimetière…), de moindre importance en terme de diffusion.

Solution propriétaire gratuite

Ces produits sont le plus souvent des *viewers* (« *visualisateurs* ») plus ou moins évolués permettant de lire certains formats de données

propriétaires ou libres et des services *web*. Les opérations d'édition sont parfois possibles. Peuvent être cités notamment : *ArcExplorer*, *TatukGIS*, *DivaGIS*.

Solution libre

Les solutions les plus connues sont *GvSIG*, *Quantum GIS* (*QGIS*), *Grass*, *µDIG*.... Si *Grass* est bien implanté dans les milieux scientifiques et universitaires notamment, en raison de son ancienneté (projet démarré en 1982 !), *QGIS* [1] et *GvSIG* sont ponctuellement rencontrés au sein de certaines collectivités ou bureaux d'études.

Il convient de noter que la confirmation du *web* SIG tend à limiter peu à peu les offres bureautiques au profit des offres *web*, notamment sur les applications métier.

Les solutions propriétaires payantes restent souvent incontournables dès qu'il s'agit de traiter des données SIG ou de produire des analyses et cartes de qualité. Les solutions libres sont soit limitées fonctionnellement pour répondre à ces demandes, soit d'un usage complexe, à l'image de *Grass*, décourageant les premières velléités du néophyte ! Les solutions commerciales gardent donc une avance confortable en raison de leur champ fonctionnel ou de l'ergonomie. Cependant, les solutions libres présentent des arguments intéressants pour les petites entités ayant des besoins raisonnables en SIG, mais ne pouvant s'offrir des solutions propriétaires. Au surplus, les solutions libres sont capables d'exploiter pleinement les données stockées dans *PostGIS*. En appui des outils d'administration de cette base (*pgsql*, *pgAdmin*), elles peuvent être utilisées comme des outils de gestion de données complémentaires (chargement, mise à jour de données) plus simples d'usage et ergonomiques.

Le serveur web SIG

L'offre dans le domaine des serveurs *web* SIG a très fortement progressé ces dernières années tant du coté propriétaire que libre. Les diverses couches et composants techniques en jeu (moteur *web* et service, bibliothèque…) rendent difficile l'état des lieux et nécessiteraient quelques articles !

Tous les grands éditeurs de solutions SIG bureautiques du marché français proposent désormais des solutions serveurs associant moteur *web* SIG et son outil de paramétrage. Ces solutions proposent une très haute compatibilité avec leur outil bureautique associé. Se sont également développés des produits *web* SIG libres ou commerciaux indépendants de toute solution bureautique.

Les solutions existantes sont construites sur des composants libres, propriétaires, parfois un mélange des deux. Toutes les combinaisons de composants (serveur d'application, moteur *web*…) sont possibles, allant du tout propriétaire au tout libre. Cependant, pour les strates profondes du système, de nombreuses offres libres et propriétaires s'appuient des composants techniques de serveurs *web* et serveurs d'application issus du libre, tels que *Apache* et *TomCat*. De même, les moteurs *web* SIG du libre (*MapServer*, *GeoServer*) sont utilisés dans de nombreuses offres d'éditeurs.

L'ensemble des ces composants est le plus souvent invisible pour l'utilisateur. Ce dernier jugera son *web* SIG sur des critères de performance, d'ergonomie et de fonctionnalité. L'apport des éditeurs (commerciaux ou libres) porte donc sur leur compétence technique, sur l'assemblage des composants (architecture) et sur les outils de paramétrage, d'administration (frontal administrateur) et d'utilisation du *web* SIG (frontal utilisateur). Et ces derniers sont souvent sous licence propriétaire.

Discussion

La part des logiciels sous licence propriétaire reste majoritaire en France. Il convient cependant de nuancer le constat en fonction des produits. La base de données *PostGreSQL*/*PostGIS* est devenue un composant courant dans le monde des SIG. Cela est dû à ses performances comparables à *Oracle* dans la majorité de ses contextes d'utilisation (volumétrie, utilisateurs) et à une communauté très active soutenant le projet.

Sous l'impulsion du *web* SIG, la plage d'utilisation de l'outil bureautique se réduit. Dans le monde propriétaire, l'offre s'amenuise. Il ne reste que quelques éditeurs proposant des solutions fonctionnellement très complètes (type station de travail) parfois surdimensionnées pour des utilisateurs non aguerris.

Le nombre de produits *web* SIG a explosé, mais la grande majorité des solutions complètes (moteur *web* SIG avec son générateur de code / console d'administration) sont dans le monde propriétaire.

Trois éléments permettent une lecture des tendances décrites ci-dessus :

- La pérennité ;
- L'implication ;
- L'ouverture.

Pérennité

Au-delà des aspects techniques, la pérennité est la question centrale et récurrente pour toutes acquisitions de solutions informatiques. Cela concerne tant les solutions

1. Le nom QGIS signifie en réalité Qt-GIS en raison de l'utilisation du toolkit Qt. Ce dernier étant une marque déposée de la société norvégienne Trolltech, récemment rachetée par Nokia, le créateur du produit a dû se rabattre sur un nom alternatif (NDLR).

commerciales que libres. Dans le cadre d'une solution commerciale, l'éditeur est identifié et un contrat précise les modalités de réalisation entre le client et son fournisseur. Dans le cadre du libre, il n'est possible pas d'identifier un éditeur ou fournisseur unique. Le produit est issu d'une communauté. Même si des distributeurs proposent parfois ce type de solution, ce manque de visibilité perturbe souvent la perception du libre et rend les entités frileuses. Cette réserve est renforcée par la nécessaire implication des utilisateurs.

Implication

Le choix d'une offre commerciale place l'utilisateur en consommateur : le fournisseur l'accompagnera, dans les limites du contrat. Le choix du libre implique l'utilisateur : il devra suivre et agir. L'implication des utilisateurs, dans un domaine technique qui n'est pas nécessairement le leur, est requise et peut être importante (veille technologique).

Cette approche n'est pas conforme au fonctionnement de très nombreuses entités, où il est très souvent demandé d'avoir un fournisseur identifié, permettant d'éventuels recours. Peu de sociétés se sont encore placées sur le secteur du suivi sur site des produits *« libre prêt à l'emploi »*.

L'utilisation de produits dits *« sur étagère »* (*GvSIG* ou *QGIS* par exemple) semble rester relativement marginale et provient d'initiatives individuelles. Une telle démarche impose aux utilisateurs d'assurer eux-mêmes le suivi du produit (changement de version, application des correctifs…).

Ouverture

Disposer de composants libres est une garantie supplémentaire d'interopérabilité et un sentiment d'indépendance vis-à-vis des éditeurs. Paradoxalement, il semble donc que les composants libres soient bien acceptés quand ils s'insèrent dans un système ou une solution d'éditeur. C'est le cas des moteurs *web* SIG libres intégrés dans des générateurs d'applications *web* SIG propriétaires. Dans ce cas, le suivi du moteur *web* SIG est inclus dans la maintenance de la solution l'exploitant. Cette configuration est conforme au mode de fonctionnement de nombreuses entités publiques, notamment.

Une évolution des offres ?

De très nombreuses expériences le montrent, l'approche participative et collaborative du monde de l'*Open Source* a bousculé le monde de l'informatique : les utilisateurs s'habituent à une *« certaine gratuité »* des produits mais également des données. Confrontés à ces nouvelles attentes, les éditeurs de solutions propriétaires s'adaptent et trouvent de nouvelles réponses. Face aux capacités techniques et aux potentiels d'évolution, les éditeurs ont tout naturellement intégré dans leurs solutions des composants libres. Ils doivent alors accepter les conditions d'usage imposées par les licences libres. Leur plus-value monnayable glisse de la simple fourniture d'un produit à un renforcement des volets accompagnement et la personnalisation. ■

Quelques Liens

http://www.opendefinition.org/okd/francais/
www.opensource.org/osd.html

DOCUMENT N°06

Revue l'Hémicycle – Manuel Singeot / N° 460 – 27/02/2013.

Big data : enjeu économique ou menace sociale ?

Big data. Ce concept informatique, engendré par la mise en place d'un monde totalement connecté, est perçu par beaucoup comme un nouvel eldorado en termes de création d'entreprises, d'emplois et de richesses. Mais il fait aussi peser de sérieuses menaces sur la vie privée des citoyens.

L'informatique s'installe partout. Depuis l'ouverture d'Internet au grand public, les ordinateurs ont colonisé la vie privée puis ont commencé à changer de forme. Ordinateurs portables, consoles de jeux, smartphones, tablettes, nos compagnons du quotidien, connectés, nous permettent de lire, mais aussi de créer une masse gigantesque d'informations. Selon Cisco, le trafic mondial de données sur terminaux mobiles sera multiplié par 13 durant les cinq prochaines années, pour atteindre 11,3 trillions d'octets par mois.

L'Internet des objets va encore augmenter cette masse d'informations créées en permanence. Voitures, compteurs électriques, immeubles, produits en magasin vont tous émettre des données sur leur état, leur localisation, leur éventuelle destruction. D'ici à une dizaine d'années, le monde sera littéralement submergé d'informations, en quantité beaucoup trop grande pour être stockées dans les bases de données telles que nous les avons conçues depuis 30 ans. Une nouvelle approche s'impose : le big data.

Réseaux téléphoniques, systèmes de santé ou boursiers, l'humanité contemporaine génère une masse toujours plus grande d'informations.

Un fort potentiel économique et social

Toutes ces données que les systèmes d'informations vont stocker ont une grande utilité : elles permettent de mettre en lumière des comportements de masse. Par exemple, l'Institut d'Asie du Sud à Harvard a dépêché une équipe de spécialistes pour étudier l'un des plus grands pèlerinages au monde, en Inde, le *Kumbh Mela*. Quatre cliniques temporaires, installées dans la zone du pèlerinage, vont pendant plusieurs mois récolter des données sur la santé des quelque 100 millions de personnes participant à ce rassemblement dans la région d'Allahabad. La surveillance en temps réel permettra, selon l'équipe, de mieux comprendre les mécanismes d'épidémies de dysenterie qui se créent lors de ces regroupements temporaires de population, notamment en monitorant des informations sur les origines régionales ou les traitements suivis.

L'analyse de ces énormes masses d'informations va demander des compétences précises et un métier informatique nouveau est en train d'émerger : le data analyste. Les besoins en Europe sont évalués à 160 000 postes nouveaux d'ici à 2014, 500 000 aux États-Unis pour la même période. Selon Fabien Terraillot, chef du bureau du logiciel à la Direction générale de la compétitivité, de l'industrie et des services, le big data est un enjeu non négligeable dans le redressement productif de la France.

Les outils actuels de stockage et de traitement de ces données sont dépassés et de nouveaux logiciels doivent être mis en place, créant un marché, pour les entreprises de solutions logicielles, qui est évalué par plusieurs études à près de 25 milliards d'euros d'ici à 2015. De même, les entreprises dont l'activité génère ces masses de données vont obtenir un gain de compétitivité important pour celles qui sauront correctement les exploiter.

Par exemple, AT&T, opérateur de téléphonie mobile aux États-Unis, a mis en place un système logiciel très avancé qui rend son réseau capable de s'évaluer lui-même afin de détecter les zones de congestion et d'optimiser les flux d'appels. Baptisé SON, ce réseau dépend des données envoyées par l'ensemble des téléphones mobiles des abonnés AT&T ainsi que des relais afin de faire ce diagnostic et gérer au mieux le réseau. AT&T attend des gains substantiels de cette technologie : les essais préalables ont montré que SON pouvait faire baisser de 10 % le nombre d'appels inaboutis, augmentant ainsi la consommation des abonnés, donc le chiffre d'affaires de l'opérateur. Un réseau mobile mieux optimisé réduit aussi les besoins en investissement dans de nouveaux relais.

Les collectivités publiques possèdent, elles aussi, des masses de données inexploitées : suivi de la circulation routière, informations issues des services sociaux ou du système de santé, le potentiel de régulation et d'optimisation des services publics par l'analyse big data des informations fournies en permanence par les administrés est immense. Cela peut peser lourd dans les arbitrages budgétaires publics en période de disette financière. Du coup, certaines grandes métropoles veulent aller plus loin et souhaitent non seulement être en capacité de gérer finement leurs administrations, mais aussi devenir les grands pôles d'implantation des entreprises spécialisées dans le traitement big data. New York vient par exemple d'accorder une subvention de 15 millions de dollars à l'université Columbia, qui met en place l'Institute for Data Science and Engineering. Qui dit laboratoire de recherche dit création d'un tissu de startups.

Menaces sur les libertés individuelles ?

Mais le phénomène big data ne peut pas exister sans la collecte de cette masse de données. Or ces données proviennent des utilisateurs, qui ne sont pas conscients de ce qu'elles révèlent sur eux. Habitudes de consommation, géolocalisation, données comportementales, l'utilité du big data repose en partie sur l'analyse des faits et gestes du citoyen-consommateur. Nous entrons là dans l'univers de Big Brother. Quelles seront les données collectées ? Seront-elles anonymes ou nominatives ? Combien de temps seront-elles stockées et qu'en feront les entreprises qui les collecteront ? Pourra-t-on s'opposer à leur collecte ? Ces questions relèvent de la Commission nationale informatique et liberté. Or cette institution ne compte qu'une centaine d'agents, ce qui est largement insuffisant face aux enjeux. Les élus français ont devant eux un sujet qui va mériter un examen très minutieux. Comment favoriser le développement en France d'un secteur économique prometteur sans mettre en place un contrôle permanent de la société ? Les décisions prises seront lourdes de conséquences.

Manuel Singeot