Le *cloud*

Introduction

Allongez-vous dans l'herbe et regardez les nuages. On essaie de les identifier.
En vain… ils bougent tout le temps !
Réaliser un dossier sur *le cloud computing* est à la fois fascinant et impossible. Fascinant parce qu'il introduit une approche indispensable sur un sujet vaste comme le ciel. Impossible car le ciel, avec ou sans nuages, ne se résumera jamais en quelques pages.
Au gré d'analyses et de témoignages, de réactions et de retours d'expérience, ce dossier terre à terre prend de l'altitude. Atteindra-t-il des sommets ?

Coordination :
Le comité créatif d'*Archivistes !*


Katell Auguié


Violette Lévy


Hervé Crépet


Charlotte Maday


Charly Jollivet


Céline Souëf

Un ciel et des nuages

Cerner un nuage…

Le mot « *cloud* » est à l'archiviste ce que le mot « archives » est à l'informaticien ; il englobe tout mais ne rend compte de rien.
Depuis de très nombreuses décennies, l'archiviste responsable d'un dépôt d'archives publiques reçoit et transmet des réglementations et des savoirs (savoir-faire, savoir-être). Depuis quelques années et en un laps de temps très court, il se voit confronté à la mondialisation des données et des techniques. Il a l'impression de passer d'un pré-carré à des plaines immenses où l'horizon se dérobe sans cesse. Ces « terres sauvages » portent plusieurs noms : « archives numériques », « métadonnées », « *cloud computing* », pour ne citer que les plus fréquentes.

Le *cloud computing* est un mode d'organisation consistant à donner accès, par un réseau, à une ou plusieurs ressources informatiques physiques et/ou virtuelles, distantes et adaptables aux besoins du client. La force de ce modèle réside donc dans sa souplesse et sa disponibilité : la complexité de la gestion d'une infrastructure, plateforme ou logiciel est masquée et les coûts de location sont présentés comme économiquement plus avantageux.

Deux obstacles importants se présentent à l'archiviste : l'aspect juridique (hébergement, transfert de données, communication) et l'aspect technique (l'informatique est un si vaste domaine…).
L'aspect technique représente d'abord un défi individuel, même s'il trouve des solutions dans le collectif (mutualisation des moyens et des ressources humaines, mutualisation des services d'archives). En France, et pour faire court, les archives et le numérique se sont rencontrés dans les années 1980. Lorsque Microsoft ou Apple pénètrent dans les foyers, ils entrent également dans les collectivités locales et les entreprises. Tout s'accélère à partir des années 2000 : multiplication des messageries (qui entraînent une autre façon de travailler et de communiquer), capacité de stockage des données fortement augmentées, ouverture des réseaux sociaux, etc. Il faut reconnaître que les archivistes nés (professionnellement) dans cette période ont plus d'aptitudes que ceux qui ont pris le train en route. Le métier d'archiviste a subi une forte mutation. Les progrès peuvent être fulgurants et les initiatives de terrains, souvent issues du « privé », finissent par s'appliquer rapidement dans le « public ».
En ce qui concerne les archives publiques, l'aspect juridique colle à la souveraineté nationale (voire européenne). L'archiviste doit avoir connaissance de la législation et de la réglementation, même la plus élémentaire possible. Mais il ne peut agir seul, sans en référer au politique (élus locaux, pouvoir central) et à l'administratif (directions générales des collectivités et des administrations). C'est un travail plus lent, de longue haleine, qui mobilise des acteurs très différents, souvent éloignés géographiquement.
Il est dans ces conditions bien difficile de cerner un nuage. Et pourtant, un nuage, ça se voit !

Le *cloud* nous ramène à l'impérieuse nécessité de bien différencier la gestion, le stockage, la sauvegarde et l'archivage. Nativement, les offres et services *cloud* ne répondent pas à toutes les exigences propres à la gestion des documents d'activité et à l'archivage électronique d'où le travail préalable sur la nature, le besoin de conservation, d'élimination, la traçabilité et la criticité des différentes catégories de données et documents. Si territorialité, sécurité et confidentialité sont souvent citées comme les freins majeurs à l'adoption du *cloud* et concernent également la gestion des documents d'activité et l'archivage, qu'en est-il réellement pour nous, professionnels de l'information ?

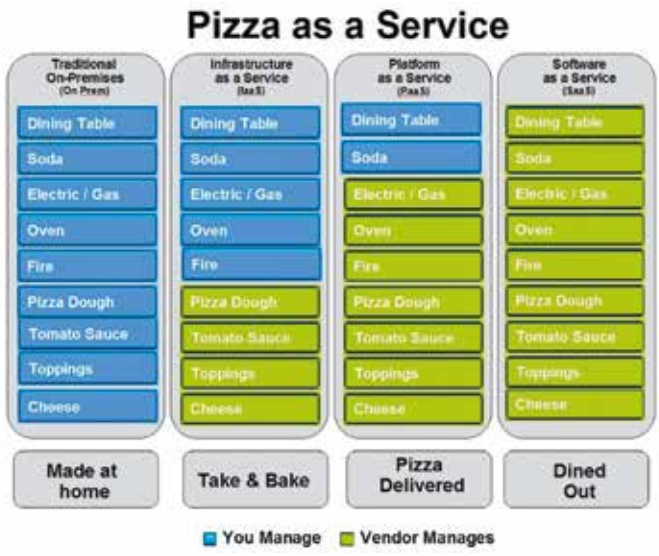
Tant de nuages dans le ciel : un *cloud* ou des *clouds* ?

On peut s'allonger dans l'herbe, laisser l'esprit vagabonder pour voir si tel nuage ressemble à un ours, à une casserole ou à une souris (afin de recoller au sujet). Pour cerner un tant soit peu nos nuages, on ne coupera pas à quelques définitions incontournables[1].
***Cloud* privé (ou dédié)** — Il s'agit du format le plus courant désignant une infrastructure qui est totalement dédiée au particulier. Un serveur privatif gère l'ensemble des données. Il est administré soit en interne soit par un prestataire spécialisé. Il est très proche d'une infrastructure locale. Il a vocation à borner précisément ses limites et à restreindre l'accès à une organisation unique.
***Cloud* public** — À l'inverse, le *cloud* public ne veut pas dire que les données sont accessibles à n'importe qui, mais qu'elles sont hébergées sur une multitude de serveurs eux-mêmes accessibles par un nombre déterminé d'utilisateurs. Ses frontières sont imprécises et il n'y a quasiment aucune restriction pour y accéder. Les *clouds* les plus connus sont d'ailleurs des *clouds* publics (par exemple Microsoft office ou Amazon).
***Cloud* hybride** — Modèle de déploiement mélangeant les *clouds* privés et publics.
***Cloud* communautaire** — *Cloud* public utilisé dans le cadre d'un groupement, réservé à des clients dont les caractéristiques et les besoins sont similaires.
***Cloud* souverain** — Selon une note officielle du ministère de la Santé, le *cloud* souverain correspond à « un modèle de déploiement dans lequel l'hébergement et l'ensemble des traitements effectués sur des données par un service de *cloud* sont physiquement réalisés dans les limites du territoire national par une entité de droit français et en application des lois et normes françaises ». Le client en connaît précisément la localisation.
À ces définitions nuagiques, il faut ajouter les types techniques d'architectures de *cloud*. Trois principales catégories sont communément admises :

1. Définitions données par le service informatique du centre de gestion du Rhône (bases type « Wikipédia »).

- Le SaaS ou *Software as a Service* : un fournisseur de service (éditeur, tiers, etc.) fournit l'usage d'un logiciel à un client pour son activité. Ainsi, la plateforme des marchés publics de l'État est une application SaaS, par exemple.
- Le PaaS ou *Platform as a Service* : un fournisseur de service fournit une plateforme avec de grandes capacités de calcul ou de stockage et diffusion, permettant à ses clients de développer des applications qu'ils peuvent gérer eux-mêmes. Ainsi, par exemple, Amazon Web-services est considéré comme un fournisseur de PaaS.
- Enfin, un modèle assez courant dans les mutualisations d'infrastructures est l'IaaS ou *Infrastructure as a Service*. Seules les infrastructures techniques (serveurs, réseaux...) sont gérées par le fournisseur de service, le reste est à la main du client.

Le ciel prendrait-il tout d'un coup un aspect plus sombre, plus compact : « Orage, haut-désespoir, ô numérique ennemi ? »



Explication des *services on premise*, Iaas, Paas, Saas en utilisant le circuit que peut connaître une pizza.

Source : https://www.episerver.com/learn/resources/blog/fred-bals/pizza-as-a-service – image soumise aux droits d'auteur et utilisée uniquement à fin d'illustration.

Le nuage et l'archiviste

Ces définitions sont indispensables. Spontanément, il ne faut pas, par exemple, associer « *cloud* public » et « archives publiques » ; l'univers professionnel de l'archiviste le rapproche plus d'un « *cloud* privé » lié à l'hébergement de sa collectivité ou de son administration. Par contre, la notion d'archives publiques renvoie automatiquement au *cloud* souverain, français pour l'instant, européen dans un futur proche. Celui-ci vise à garantir la conservation en version originale de toutes données numériques ayant valeur d'archives publiques, donc de Trésor national[2], les plaçant ainsi sous bonne surveillance juridique. Mais « bonne » ne veut pas dire « totale ».

Le Code du patrimoine (article L. 211-2) dispose que « la conservation des archives est organisée dans l'intérêt public, tant pour les besoins et la justification des droits des personnes physiques ou morales, publiques ou privées que pour la documentation historique de la recherche ».
On imagine que cette mission de préservation et de conservation des archives publiques nécessiterait une localisation sur le territoire national pour que l'État puisse garantir aux citoyens des données fiables, intègres et authentiques. Cependant, il ne faudrait pas que cette exigence de localisation soit comprise comme une interdiction de circulation des données. C'est le problème de la cohabitation entre « duplication », « communication » et « réutilisation des données » dans le respect de la réglementation nationale et/ou européenne.

Le cadre « protecteur » existe déjà et rappelle sans détour ce que représentent historiquement et juridiquement les archives publiques (le Code du patrimoine bien sûr, mais aussi la CNIL et la CADA). Sans renier toute la législation mise en place depuis la Révolution française, elles maintiennent comme légitime la conservation des données en version originale sur le territoire français, mais n'interdisent pas l'externalisation de copies à l'étranger (en tant que copies, elles peuvent être supprimées à tout moment).
Il faut également rappeler que le contrôle scientifique et technique implique un droit de regard et de visite sur les lieux d'hébergement d'archives. Il n'y a pas de différence dans ce domaine entre archives papier et archives numériques, ces dernières échappant plus facilement aux différents contrôles de localisation, de conservation et de communication. Les archives publiques en version papier localisées à l'étranger[3] le sont dans un cadre bien précis : cadre juridique défini, regroupement dans des bâtiments identifiés et accessibles, communication soumise à procédure, etc. Les archives numériques sont plus facilement soustraites à ces différents régimes. Rien n'est garanti. D'où le fait de bannir en l'état un hébergement de documents originaux à l'étranger.
L'archiviste n'est cependant pas dupe. Comme monsieur Jourdain faisait de la prose sans s'en rendre compte, l'archiviste peut faire du *cloud* sans le réaliser ou plutôt à son corps défendant ; l'application WeTransfer en est un exemple. Et la question fondamentale est : « Quels types de *clouds* utiliser en fonction des processus métiers ? »

2. Code du patrimoine, article L. 111-5. Note d'information SIAF/DGCL du 5 avril 2016. Depuis la parution de cette note, le SIAF n'a pas enregistré de véritables remontées du réseau au regard des offres qui se sont présentées.

3. On peut distinguer les documents situés sur le territoire français dans un pays étranger (ambassade, consulat) de ceux qui se trouvent dans les grandes écoles françaises (Athènes, Rome…) ou encore de ceux prêtés dans le cadre d'expositions.

Vers un *cloud* souverain européen ?

La bannière bleue aux douze étoiles d'or est-elle sans nuage ? Sur le plan informatique en général et sur celui du *cloud* en particulier, on pourrait dire « oui ». En tous cas, pour l'instant.

C'est en effet du côté du Parlement européen qu'il faut s'orienter pour voir une évolution à court terme : la libre circulation des données à caractère non personnel. Ce projet de règlement « autorisera le stockage et le traitement des données à caractère non personnel dans l'ensemble de l'Union, de manière à stimuler la compétitivité des entreprises européennes et à moderniser les services publics dans un véritable marché unique européen des services de données ».

Deux principes régissent ce cadre : celui de la libre circulation des données transfrontières et celui de la disponibilité de ces données à des fins de contrôle réglementaire, quel que soit le lieu de stockage en Europe. Un code de conduite devrait voir le jour (changement de fournisseur et portabilité des données stockées).

L'objectif premier est de renforcer la sécurité juridique et la confiance. L'objectif plus discret serait de niveler et d'uniformiser le stockage des données pour faciliter la lecture, l'exploitation, la mutation. Ce règlement mobilise les énergies et rappelle d'autres efforts plus anciens : la volonté des monarchies d'Ancien Régime d'unifier des territoires très disparates dans leurs langues, leurs monnaies, leurs unités de mesures et leurs droits coutumiers. L'histoire est un éternel recommencement : unifier pour mieux régner… Un langage commun pour une meilleure collaboration numérique entre les administrations publiques en Europe serait souhaitable. L'exemple de la monnaie unique donne globalement raison à cet argument.

Le « nuage » a aussi une vocation compétitive en termes de stockage ou de distribution des données. Les prix doivent être attractifs pour permettre à l'Europe de « peser » face au géant américain ou à la montée asiatique. Des possibilités de transfert de ressources sont en jeu, avec des garanties techniques et juridiques de stockage et d'exploitation. La notion de culture (française évidemment, européenne éventuellement) doit pouvoir s'exprimer dans ce montage en cours de réalisation. L'union faisant la force, les voix ne seront que plus entendues si elles parlent ensemble et d'un même ton. À son niveau, l'AAF a un rôle non négligeable à jouer.

Uniformiser est une chose. Prendre conscience des écarts juridiques entre les États membres en est une autre. Des États souverains avec un patrimoine culturel et technique différent auront plus intérêt à demander une ouverture des données que ne l'auront des États avec un passé culturel et des structures juridiques plus anciennes. Ce problème est tellement d'actualité.

Hervé Crépet
Archiviste itinérant
Centre de gestion du Rhône et de la Métropole de Lyon

SaintÉ et Google : un projet sans suite mais riche d'enseignements

Cyril Longin, directeur des Archives municipales de Saint-Étienne

En 2015, la Ville de Saint-Étienne et Saint-Étienne Métropole souhaitent harmoniser et mettre à jour leurs systèmes informatiques en passant par un fournisseur privé. Cyril Longin, directeur des Archives municipales, est associé au projet qui prend en compte la mutualisation des messageries, les outils bureautiques et l'hébergement des données. Retour sur une expérience sans suite mais riche d'enseignements.

Pouvez-vous nous préciser le contexte ?

En 2014, le nouveau maire souhaite une mutualisation plus importante des services de la Ville et de la Métropole. Sur le plan informatique, il y a une énorme différence entre les deux systèmes d'exploitation dont la messagerie et les outils bureautiques. La Métropole utilise Microsoft Office tandis que la Ville est sur logiciel libre.

Comment a démarré le projet ?

En 2015, j'ai la chance de bien m'entendre avec les DSI de la Ville et de la Métropole dont les services ne sont pas encore mutualisés. Le DSI de la Ville vient du Grand Lyon où il avait porté des projets *open data* et aussi un peu d'archivage électronique. Il était convaincu de la complémentarité de nos métiers, au-delà de la simple question des archives.

Qui avez-vous contacté ?

Une opportunité s'est présentée via un contact avec Google qui proposait un produit adapté aux collectivités et aux entreprises : Google for Works. Les deux directions me demandent de participer au groupe de travail pour apporter l'expertise de l'archiviste sur la gestion des données et leur sécurité. Sur le plan national, c'est à ce moment que la loi reconnaît les archives comme Trésor national[1].

Comment se passe l'entretien ?

L'équipe Google (cinq personnes) nous rencontre (les deux DSI et moi-même). Elle semble étonnée de me voir dans le groupe de travail. Elle présente des produits de très bonne facture, incluant une gestion des métadonnées intéressante et l'établissement d'un réseau social d'entreprises (RSE). Sont également présentés une sorte de GED et un module dit d'archivage. Sur l'hébergement des données, c'est un peu plus flou : filiale ou association avec une entreprise lyonnaise ?

Il y a d'autres rendez-vous…

Oui. Le deuxième entre plus en profondeur dans l'explication technique des outils. J'interviens moins lors de cette réunion, préparant mon argumentaire pour le troisième rendez-vous, axé sur la sécurité et l'hébergement. Le troisième rendez-vous (juillet 2015) est téléphonique.

1. Loi n° 2015-195 du 20 février 2015 portant diverses dispositions d'adaptation au droit de l'Union européenne dans les domaines de la propriété littéraire et artistique et du patrimoine culturel. Cette loi, véhicule de transposition de trois directives européennes, a modifié la définition des trésors nationaux, dans l'article L. 111-1 du Code du patrimoine, en y inscrivant expressément les archives publiques.

Après une brève présentation de Google, je pose la première question sur la réversibilité et la propriété des données. Google assure qu'elles demeurent propriété de la Ville, mais quelques éléments restent flous. À l'évidence, la notion d'archives publiques leur est étrangère. Suite à la question sur l'hébergement des données, le temps s'accélère puisque, montre en main, le reste de l'entretien va durer… trois minutes. Google annonce que les données sont hébergées en Europe. À la relance de la question « Où sont-elles physiquement hébergées ? », Google répond qu'il y a trois serveurs situés en Irlande, aux Pays-Bas et en Italie. Je rappelle les notions d'archives publiques et leur qualité de Trésor national dans la réglementation française. Google assure respecter la réglementation européenne. Je rappelle qu'en tant que Trésor national, les données ne peuvent être hébergées en dehors du territoire. À ce moment précis, Google clôt la conversation.

Quelle a été votre réaction ?

Ce fut un étonnement que Google ne se soucie pas plus de la notion d'hébergement ou reste flou sur la question. Google avait d'ailleurs démarché une grosse collectivité en France avec qui, nous disait-on, la question n'avait pas été abordée. En appelant cette collectivité, j'ai appris que mes collègues n'étaient pas au courant pour le lieu d'hébergement des données. Google avait en fait une liste de collectivités importantes avec qui il devait travailler et j'ai été, selon eux, le premier à poser cette question sur l'hébergement. Mes DSI n'en revenaient pas.

Et celle des élus ?

Le maire de Saint-Étienne s'est également inquiété de l'hébergement des données. Afin d'appuyer nos interrogations, les DSI m'ont demandé de saisir le SIAF. Une réponse est ainsi adressée au maire de Saint-Étienne et précise le contexte législatif et réglementaire, notamment sur l'impossibilité de conserver des archives publiques en dehors du territoire.

Quels enseignements en tirer ?

Le projet Google a été abandonné, sûrement pour ces raisons, mais aussi pour un coût financier élevé… bien que tout aurait pu être négocié. Quoi qu'il en soit, aussi bien au niveau de la gouvernance stéphanoise que de Google, ce courrier a été déterminant. Cet exemple montre que l'archiviste peut avoir un rôle à jouer.

Cette affaire permet de réfléchir sur la pertinence d'hébergement des données sur le territoire français. Ce n'est peut-être pas fondamental, car ces données ne sont pas mieux protégées sur notre territoire. D'autant que les produits proposés par Google étaient vraiment très intéressants et auraient apporté une révolution sur le plan de l'utilisation, de la gestion et de la transmission des données. Le cœur du problème, c'est ce manque de connaissance que Google a de la législation française. Il leur suffirait de disposer d'un *data center* agréé en France pour y répondre. Peut-être que leur poids suffit à convaincre des collectivités peu regardantes et fières d'être démarchées par la multinationale. Google avait beaucoup insisté sur la tête de gondole que représentait la grosse collectivité française citée plus haut. Peut-être s'agissait-il d'un « cheval de Troie » pour investir le marché français.

Propos recueillis par Hervé Crépet

L'hôtel de ville de Saint-Étienne © Hervé Crépet

Quand le tonnerre gronde… la sécurité des données dans le *cloud*

Les architectures organisées en nuage sont perçues comme aussi bénéfiques que menaçantes pour les données : quels sont les bénéfices et les risques associés ? Qu'en est-il des conséquences sur les données ? L'Agence nationale de sécurité des systèmes d'information nous aide à y voir plus clair.

Qu'est-ce que les infrastructures *cloud* peuvent apporter à la collecte, au classement, à la conservation et à la diffusion des données ?

La collecte, le classement, la conservation et la diffusion de données sont des activités pouvant nécessiter de fortes puissances de calculs et de capacités de stockage. Ces opérations sont souvent complexes et onéreuses à mener sur des infrastructures en propre. L'avantage de l'informatique en nuage, dit *cloud*, est donc de bénéficier à la demande de puissantes ressources de calculs distribuées permettant de traiter des volumes conséquents de données. Ce nouveau paradigme limite le coût d'une infrastructure qui de manière permanente serait fortement coûteuse, tout en permettant une exploitation plus approfondie des données.

Quels sont les gains pour une entité ayant recours aux infrastructures *cloud* ?

Il est courant de dire que l'externalisation dans le *cloud* permet de tirer des gains en termes de coûts, de fiabilité et de flexibilité. Cependant, une analyse des risques est nécessaire en amont d'un tel choix stratégique pour avoir une chance d'en tirer parti.

Sur la notion du coût, une économie ne pourra être réalisée que si les entités adhèrent aux technologies et à la logique du prestataire choisi, quitte à revisiter profondément leur façon d'opérer. Le *cloud* ne représente pas uniquement des ressources de calculs. Il comprend également un éventail d'intergiciels[1] ou de logiciels pré-intégrés permettant aux entités de n'avoir, dans certains cas, qu'à migrer des données (solutions PaaS ou SaaS). Ainsi, en migrant vers le *cloud*, il sera nécessaire d'inclure une phase d'intégration souvent plus longue qu'un projet traditionnel et faisant appel à des compétences avancées en matière d'architecture et d'ingénierie des systèmes.

Concernant la fiabilité, dans le cas de PME n'ayant pas les compétences en interne, une externalisation peut apporter un meilleur niveau de sécurité. Elle apportera une certaine résistance aux attaques informatiques non ciblées et de niveau technique modéré.

Dans le cas d'entités nécessitant une forte puissance de calculs sur un court laps de temps, l'infrastructure *cloud* permet de déployer rapidement des ressources serveurs et de les supprimer lorsqu'elles ne sont plus nécessaires. Cette flexibilité permet également de n'être facturé que sur l'utilisation de ces dernières et d'économiser les coûts de maintien d'une plateforme qui serait surdimensionnée la plupart du temps.

Quels sont les risques identifiés par l'ANSSI ?

Les risques liés à l'externalisation sont multiples.

Sur le plan juridique, il pose la question du droit applicable aux données et applications hébergées, dès lors que l'hébergement est effectué hors du territoire national ou européen, ou que la nationalité du prestataire le soumet à des contraintes juridiques à portée extraterritoriale.

Sur un plan technique, le recours au *cloud* peut offrir au prestataire la maîtrise complète des données et logiciels que lui confient ses clients. La confidentialité ou la disponibilité des données peuvent être ainsi menacées si le prestataire ne les protège pas suffisamment ou y porte atteinte lui-même (par malveillance ou simplement parce qu'il est soumis à des contraintes légales différentes de celles du client).

Sur le plan économique enfin, l'externalisation vers le *cloud* peut conduire à un risque de dépendance technologique accrue envers le prestataire. Ainsi, la commercialisation des logiciels sous la forme de service dans le *cloud* transforme le client, jusqu'alors propriétaire du logiciel, en simple locataire. Il devient ainsi tributaire de la tarification imposée pas son prestataire, souvent sans recours par manque de solutions alternatives ou de réversibilités.

Le *cloud* souverain est-il la meilleure réponse aux enjeux de sécurité autour du *cloud* ? L'ANSSI a-t-elle eu des retours d'expérience d'entités ayant mis en place des infrastructures de type *cloud* souverain ?

La notion de souveraineté dans le contexte du *cloud* est difficile à définir précisément. D'un point de vue économique, il est souhaitable de disposer d'acteurs nationaux ou européens compétitifs dans le domaine du *cloud*. Cependant, sur le strict plan de la sécurité numérique, la nationalité d'un fournisseur de *cloud* n'est qu'un critère d'analyse parmi d'autres. Si le fait que le fournisseur soit français ou européen peut simplifier l'étude du droit applicable aux données, il n'est pas en soi une garantie de robustesse technique de la solution. C'est pourquoi l'approche privilégiée par l'ANSSI est celle du référentiel SecNumCloud[2], qui conjugue les différentes exigences qu'elle juge pertinentes. Celui-ci n'impose rien en termes de nationalité du fournisseur, mais énonce des critères sur la localisation des données client et le droit applicable à celles-ci.

Par ailleurs, une nouvelle tendance se dessine parmi de nombreuses entités avec lesquelles l'ANSSI est en contact : les stratégies hybrides de recours au *cloud*. En plus d'avoir recours au *cloud* public pour des données ou applications peu sensibles, des entités de taille suffisante développent des infrastructures *cloud* internes ou privées pour y héberger leurs données et applications sensibles (par exemple, en matière de confidentialité ou d'importance stratégique). Elles bénéficient ainsi pour partie des apports du *cloud* (mutualisation interne), tout en conservant la maîtrise de leurs outils numériques les plus critiques.

1. Intergiciel ou *middleware* : logiciel tiers qui crée un réseau d'échange d'informations entre différentes applications informatiques. Le réseau est mis en œuvre par l'utilisation d'une même technique d'échange d'informations dans toutes les applications impliquées à l'aide de composants logiciels (définition Wikipédia).

2. SecNumCloud est consultable ici : https://www.ssi.gouv.fr/actualite/secnumcloud-la-nouvelle-reference-pour-les-prestataires-dinformatique-en-nuage-de-confiance/

© Hervé Crépet

Est-il envisagé, au même titre qu'un label CNIL sur la protection des données, la création d'un label ANSSI sur la sécurité des infrastructures *cloud* ?

Au travers du visa de sécurité qui valorise les solutions certifiées et qualifiées, l'ANSSI propose une qualification à destination des prestataires de services *cloud*. Le référentiel SecNumCloud est né de la volonté de promouvoir des offres *cloud* aux entités souhaitant externaliser l'hébergement de leurs données ou applications. Il s'adresse aux prestataires proposant des infrastructures, plateformes ou logiciels sous la forme de services (IaaS, Paas, Saas).
À ce jour, quatre prestataires sont en cours de qualification : CloudWatt par Orange, OoDrive, Outscale et Vendôme Solutions.

Propos recueillis par Charlotte Maday

Cloud computing : y a-t-il un GPS sur les données ou documents ?

Tout d'abord, revenons à la définition juridique de *cloud computing* : « mode de traitement des données d'un client, dont l'exploitation s'effectue par l'internet, sous la forme de services fournis par un prestataire »[1]. Une note vient cependant préciser qu'il s'agit d'une « forme particulière de gérance de l'informatique, dans laquelle l'emplacement et le fonctionnement du nuage ne sont pas portés à la connaissance du client »[2], comme le pointe du doigt notamment la Commission nationale informatique et libertés (CNIL)[3]. Cette note lève le voile sur le talon d'Achille de cette forme de stockage : la localisation[4]. Et pour cause, tout le problème des données ou documents stockés dans le *cloud* est de savoir où ils se trouvent dans l'espace et (par) où ils vont[5].

Pour rappel, à défaut de clause contractuelle le précisant et ce au sein de l'Union européenne, le règlement baptisé « Rome I » mentionne en son article 4 paragraphe 1 point b) que « le contrat de prestation de services est régi par la loi du pays dans lequel le prestataire de services à sa résidence habituelle »[6]. Pour cela, on se fie bien souvent à l'emplacement des *data centers* comme point d'ancrage géographique[7]. Dans l'affaire Microsoft c/ Unites States, les données étaient stockées dans le *cloud* via un *data center* basé en Irlande et donc le mandat américain en vue de la saisie ne pouvait avoir d'effet à l'étranger[8], sans compter que cela pourrait engendrer un transfert non valide (sur simple demande d'un gouvernement) de données à caractère personnel en dehors de l'Union européenne[9]. Dans le bras de fer qui oppose Facebook à la CNIL, Facebook a installé un de ses trois *data centers* européens en Irlande (les deux autres sont implantés respectivement en Suède et au Danemark) et donc accepte de se soumettre au droit européen de la protection des données à caractère personnel[10]. Ceci avait été mis en avant par l'étudiant en droit autrichien Maximilian Schrems contre Facebook dans l'arrêt éponyme de la Cour de justice de l'Union européenne[11], car les utilisateurs européens ont conclu un contrat avec Facebook Ireland incluant un transfert vers les serveurs américains de l'entité-mère[12]. La CNIL suggère d'ailleurs de baliser juridiquement ce point en insérant des clauses contractuelles permettant de connaître la position géographique des serveurs qui stockent les données ou documents visés ou de limiter l'usage à des *data centers* situés dans le périmètre de l'Union européenne, voire également dans des pays reconnus comme offrant une protection adéquate des données à caractère personnel[13]. En effet, on peut craindre un transfert non autorisé de données à caractère personnel en dehors de l'Union européenne ou vers des pays n'offrant pas de protection adéquate, en fonction de la position des serveurs sur lesquels les données ou documents sont dupliqués ou répliqués[14], ou un non-respect des préceptes de sécurité juridique offerte aux archives publiques.

Ainsi, l'externalisation via un support numérique de la conservation des archives publiques aux âges courant et intermédiaire requiert des mesures particulières telles qu'un agrément émis par le Service interministériel des Archives de France (SIAF)[15] pour des prestataires qui ont tous des *data centers* implantés en France, conformément à l'article R. 212-23 du Code du patrimoine français[16]. Dans cette droite lignée, des *clouds* dits « souverains » sont plébiscités et sont les seuls à pouvoir accueillir des archives publiques[17]. Un *cloud* souverain se définit comme un « modèle de déploiement dans lequel l'hébergement et l'ensemble des traitements effectués sur des données par un service de *cloud* sont physiquement réalisés dans les limites du territoire national par une entité de droit français et en application des lois et normes françaises »[18]. Pour l'heure, aucun n'a obtenu l'agrément du SIAF[19].

Moralité : il est vivement conseillé aux entreprises françaises de stocker leurs données ou documents dans des *data centers* basés dans l'Union européenne et aux entités administratives françaises dans des *clouds* souverains et avec/ou dans des *data centers* implantés en France.

Sarah Markiewicz
Doctorante en droit des technologies de l'information
Université d'Aix-Marseille

1. « Vocabulaire de l'informatique et de l'internet », *Journal officiel de la République française,* n° 129, 6 juin 2010, p. 10453.

2. Bénédicte Fauvarque-Cosson, « Le cloud computing, protection des données et nouvelles pratiques contractuelles », *Les Petites affiches*, 2014 ; Anne-Laure Villedieu et Julie Tamba, « Le droit, protecteur des données dans le cloud », *Revues des juristes de Sciences Po*, n° 10, 114, 2015 ; Noé Wagener, « Cloud souverain et archives publiques », *Juris art etc.*, n° 43, 38, 2017.

3. Merav Griguer, « Cloud computing et protection des données personnelles : clôture des débats ? », *Cahiers de droit de l'entreprise*, n° 4, 2012, prat. 20.

4. Anne-Laure Villedieu et Julie Tamba, « Le droit, protecteur des données dans le cloud », *op. cit*.

5. Thomas Lange, « Cloud computing et données personnelles : les clauses à maîtriser », *Dalloz IP/IT*, 459, 2016 ; Romain Perray, « L'externalisation des données des Fintechs : les risques du cloud », *Revue de droit bancaire et financier*, n° 1, 2017, dossier 9.

6. Règlement (CE) n° 593/2008 du Parlement européen et du Conseil sur la loi applicable aux obligations contractuelles (Rome I) du 17 juin 2008.

7. Céline Castets-Renard, « Affaire Microsoft c/ United States : des universitaires européens s'expriment auprès de la Cour suprême des États-Unis », *Dalloz IP/IT*, 80, 2018.

8. *Ibid.*

9. *Ibid.*

10. Anne Debet, « Facebook sommé de se conformer aux règles françaises de la protection des données », *Communication commerce électronique*, n° 6, 2016, comm. 56.

11. « Maximilian Schrems c/ Data Protection Commissioner », Cour de justice de l'Union européenne, aff. C-362/14, 6 octobre 2015, en ligne : http://curia.europa.eu/juris/liste.jsf?num=C-362/14 (consulté le 4 juin 2018).

12. Clara Coudert, « Derrière le cloud, les datacenters : quel avenir fiscal pour ces usines du numérique ? », *Droit fiscal* n° 5, 151, 2018.

13. Merav Griguer, « Cloud computing et protection des données personnelles : clôture des débats ? », *Cahiers de droit de l'entreprise*, n° 4, 2012, prat. 20 ; Thomas Lange, « Cloud computing et données personnelles : les clauses à maîtriser », *Dalloz IP/IT*, 459, 2016.

14. Thomas Lange, « Cloud computing et données personnelles : les clauses à maîtriser », *op. cit.* ; Anne-Laure Villedieu et Julie Tamba, « Le droit, protecteur des données dans le cloud », *op. cit*.

15. Prestataires agréés pour la conservation d'archives publiques courantes et intermédiaires sur support numérique, en ligne : https://francearchives.fr/fr/article/26287437 (consulté le 4 juin 2018).

16. Noé Wagener, « Cloud souverain et archives publiques », *Juris art etc.* n° 43, 38, 2017 ; Code du patrimoine français, article R. 212-23, en ligne : https://www.legifrance.gouv.fr/affichCodeArticle.do?cidTexte=LEGITEXT000006074236&idArticle=LEGIARTI000024240378 (consulté le 4 juin 2018).

17. Valentine Martin, « La République numérique en débat au Parlement : le projet de commissariat à la souveraineté numérique », *Les nouveaux cahiers du Conseil constitutionnel*, n° 57, 107, 2017.

18. Note d'information relative à l'informatique en nuage (*cloud computing*) du ministère de la Culture et de la Communication du 5 avril 2016.

19. *Supra*, note 4.

Le film *Sex tape* (2014), où quand une vidéo envoyée par mégarde sur le *cloud* devient le sujet d'une comédie américaine.
Capture d'écran de la bande annonce du film. Image utilisée uniquement à fin d'illustration. Source : http://www.allocine.fr/video/player_gen_cmedia=19544721&cfilm=205321.html

Comment gère-t-on la question du *cloud* dans les autres pays ?

© Charlotte Maday

La présence de l'AAF à la réunion annuelle du TC46/SC11 de l'ISO en mai à Lisbonne a été l'occasion d'en apprendre davantage sur la manière dont les archivistes et gestionnaires des documents d'activité perçoivent le *cloud*, autour d'un projet de texte technique.

Beaucoup de pays considèrent que le *cloud* est une technologie, ou plus exactement une organisation architecturale technique, qui ne doit donc pas avoir d'incidence sur les grands principes archivistiques : les concepts de provenance, de contextualisation et de classification ne doivent pas être adaptés systématiquement pour chaque type d'architecture ou chaque nouvelle technologie.
Un travail a déjà été mené par différents acteurs[1], et c'est en partie sur cette base que les travaux autour de ce projet de texte technique vont s'appuyer. En effet, la réflexion qui anime la communauté internationale est plutôt d'identifier les particularités des architectures *cloud* pour en cerner les risques et proposer des éléments pouvant les mitiger ou les résoudre.

Un exemple concret : un client bénéficiant d'un outil en SaaS produit par un géant du domaine n'a aujourd'hui pas la force de frappe nécessaire pour demander que ses données, qui lui appartiennent, lui soient restituées dans un format qu'il pourra réexploiter ou que, tout au long du cycle de vie des données dans le logiciel en SaaS, les événements l'affectant soient journalisés et les métadonnées qui sont enrichies ou modifiées soient tracées par le fournisseur de service.
Que peut-on mettre en place en termes de bonnes pratiques, au niveau normatif, pour inviter (fermement) les fournisseurs de service *cloud* à se mettre en conformité de ce point de vue ?
Le travail de ce groupe va donc identifier les risques ou les points d'attention à relever et les propositions de solution en fonction de chaque type d'architecture, que ce soit SaaS, PaaS ou IaaS dans un *cloud* public, privé ou hybride.

Charlotte Maday
Consultante Spark Archives, experte AFNOR/ISO pour l'Association des archivistes francais

1. InterPARES : http://www.recordsinthecloud.org/ ; Nouvelle-Zélande : https://www.ict.govt.nz/guidance-and-resources/using-cloud-services/

Conclusion

La gestion des données numériques est en pleine expansion et en totale mutation. Loin d'être seulement un choix technologique, le *cloud* n'est pas un débat entre informaticiens, mais bien un enjeu de société et un enjeu pour les archivistes. À chacun d'agir et de réagir sur les quelques éléments apportés par ce dossier. L'AAF, avec ses groupes de travail et ses sections, peut jouer un rôle non négligeable sur la mise en place d'un *cloud* responsable et bien pensé. Pour finir, un clin d'œil à tous ceux que le sujet du *cloud* désespère :
« N'oublie pas que chaque nuage, si noir soit-il, a toujours une face ensoleillée, tournée vers le ciel. » [F. W. Weber]

Coordination : le comité créatif d'*Archivistes !*