

**CLOUDERS**



**CLOUDERS**





# CLOUDERS

Pierre Cottanceau sous le tutorat de  
Nicolas Nova

2017–2019 Media Design,  
Master of Arts HES-SO  
in Media Design



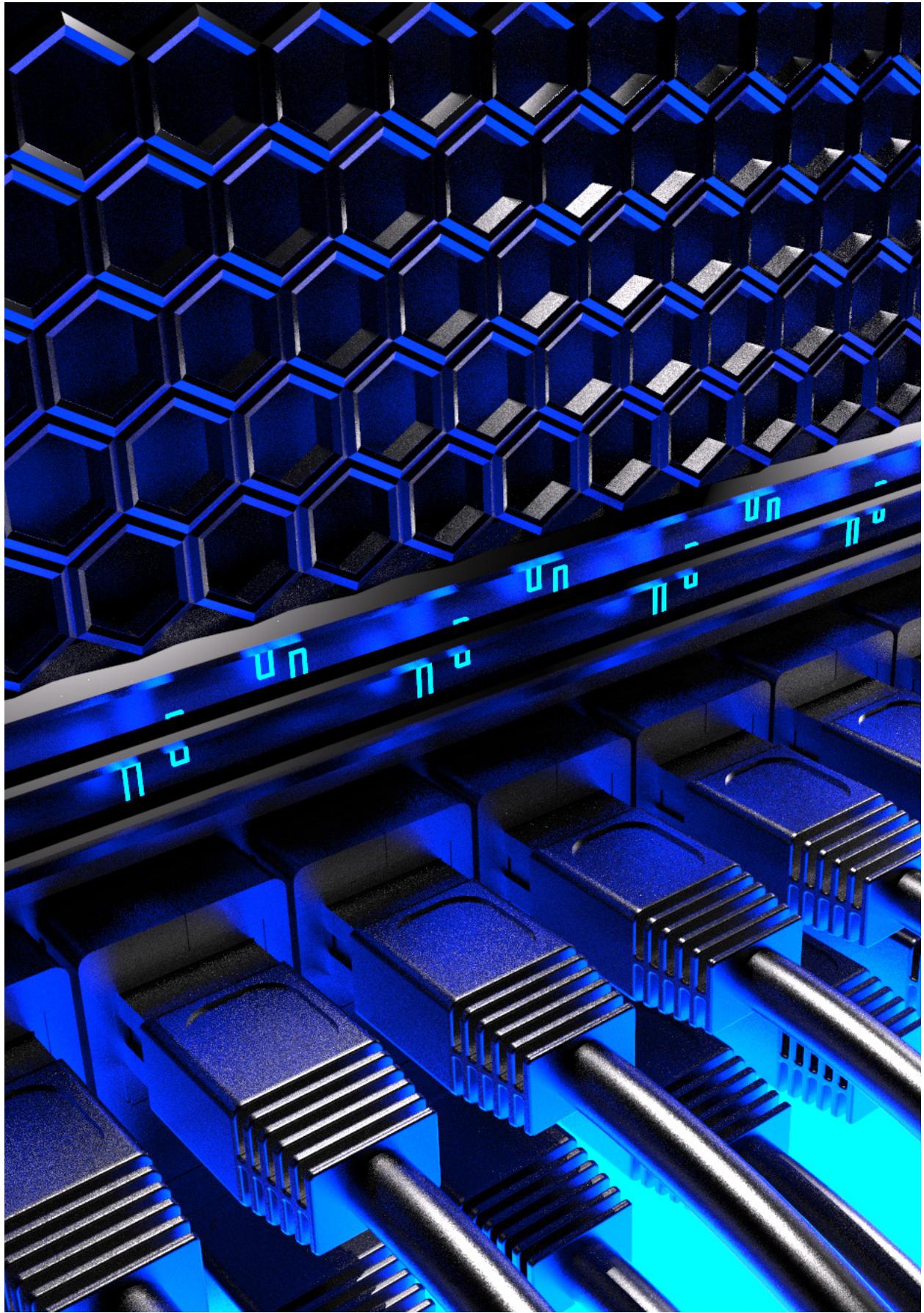
## Sommaire

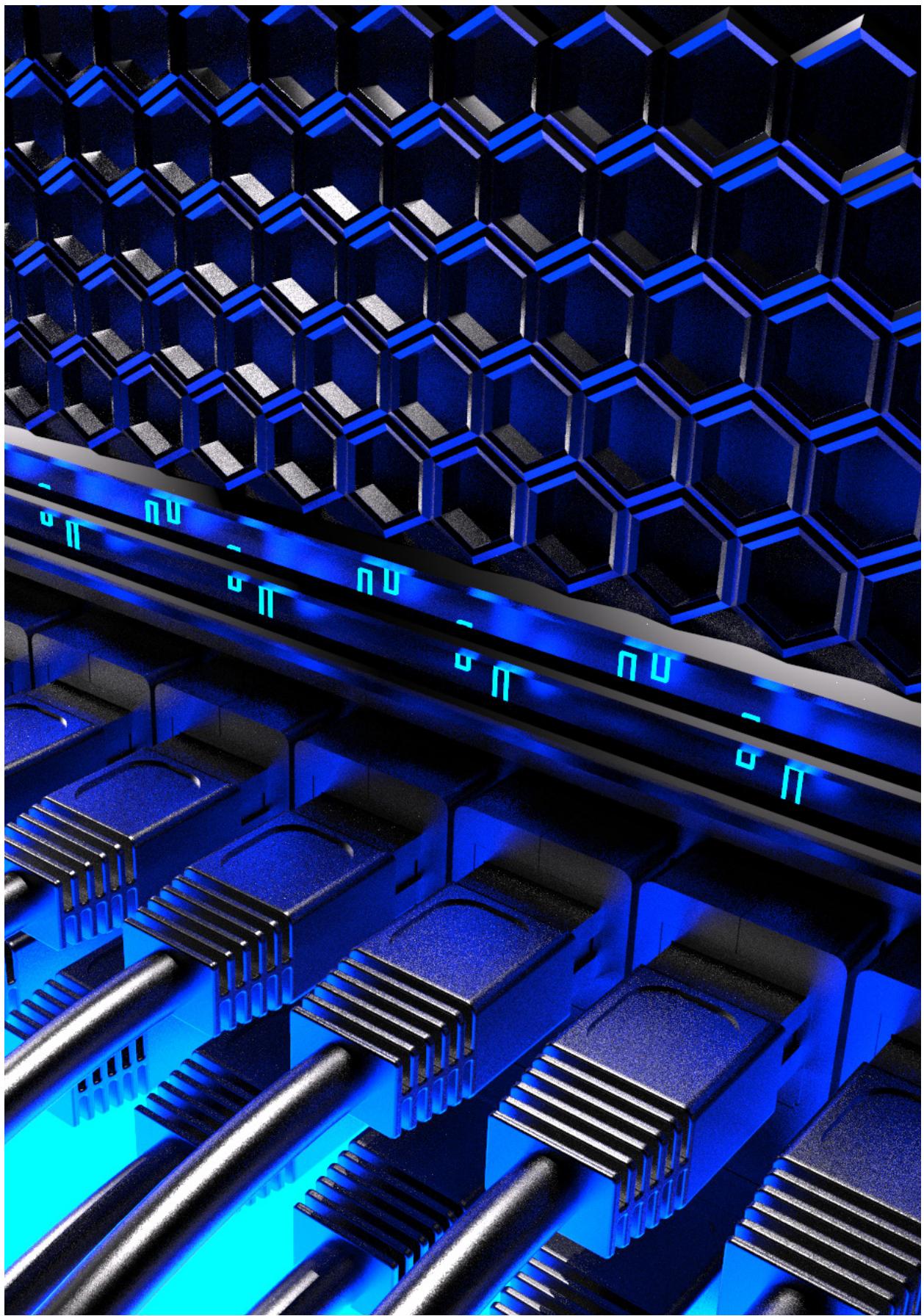
### **Partie 1 : Utilisateur et Utilisation :**

- Etat de l'art des cloud services.
- Utilisateurs et personnes ciblées.
- Construction de l'entretien, définitions des termes.
- Analyse des résultats, habitudes.
- Classifications des données.
- différents types de stockage.

### **Partie 2 : Technologie et comportements**

- Analyse des technologies existantes et en développement.
- Comparaison des technologies et justification d'usage.
  - Réduction des besoins énergétique.
  - Comportement et réduction d'énergie.
- Conclusion





Le cloud storage ou les cloud services tels que Google drive, Dropbox, Github ou bien Soundcloud restent pour beaucoup un mystère et suscitent un manque de compréhension et de connaissance quant à leur fonctionnement. Ainsi la technologie cloud est utilisé pour des télécommunications ou du stockage d'information, les usagers ne connaissent pas clairement les besoins énergétiques et matériels que ces data center requièrent. Etant donné que le contenu Web augmente de façon exponentielle en fonction de l'augmentation du nombre d'usagers, des problèmes vont se poser à très court terme.

Nicolas Nova  
dans «Clouds of  
practices»(2018).

Ces data center nécessitent d'abord une occupation terrestre conséquente, par exemple en Virginie la proximité avec les « backbones » (grand cables de fibre optique faisant transiter les informations d'internet d'un continent à un autre) et les faibles coûts de l'électricité ont attiré les géants d'internet, à tel point que la région de Dulles à quelques kilomètres de Washington est maintenant surnommée la *Data Center Valley*. Google, Amazon, Microsoft, Facebook y possèdent des data center . D'après le site spécialisé « *Data center knowledge* » la surface occupée par les data center seuls et non par les terrains, avoisine le million de mètres carrés avec un tiers de plus en développement livré à la fin de l'année. A lui seul Amazon annonçait la construction d'un centre de 46 000 mètres carrés en 2017 d'après le site spécialisé technical.ly (2018).

À cela nous pouvons aussi ajouter des besoins énergétiques et matériels tout aussi conséquents. Les besoins matériels sont constitués de racks, de serveurs, de disques durs, de climatiseurs et de générateurs électriques (en cas de coupures, les informations stockées doivent rester accessibles) qui nécessitent une alimentation électrique permanente. Ces demandes de ressources matérielles et énergétiques ne font qu'augmenter avec le temps car elles sont corrélées à l'augmentation des usagers. A l'heure actuelle, les disque durs peinent à s'adapter à l'évolution rapide que connaît le monde du stockage des datas center . En effet en 2014 nous dépassions tout juste le milliard d'utilisateurs, aujourd'hui

internet compte 4 milliards d'utilisateurs selon *World Internet Users Statistics* et le nombre de données stockées en ligne augmente exponentiellement en fonction de celui-ci.

Chaque minute 400 heures de vidéo sont chargées sur Youtube, ce rythme exige un nouveau volume d'un péta-octet/jour (1 péta-octet est égal à dix millions de Giga octet). Cette demande est multipliée par 10 tout les 5 ans d'après Google. La consommation électrique de ces infrastructures est immense. En effet, les climatiseurs et les systèmes de refroidissement peuvent consommer jusqu'à 50 pour cent de la demande énergétique d'un data center conformément aux chiffres de Greenpeace. Ces systèmes de refroidissement sont nécessaires au bon fonctionnement des data center , les disques durs accumulés engendrent de la chaleur liée à leurs mécanismes. Si la chaleur devient trop élevée (supérieure à 35 degrés) les disques peuvent s'endommager engendrant la perte des données dans le même temps.

*Disks for Data Center  
White paper for  
FAST (2016).*

*USA Clicking Clean  
report (2017).*

En France par exemple, on trouve 147 data center consommant entre 7 et 10 pour cent de la production annuelle d'électricité soit à peu près 52,9 GW/h selon EDF. En clair, si tous les datas center cumulés représentaient un pays, il serait le cinquième demandeur d'énergie à l'échelle du monde.

En parallèle de cela l'utilisateur n'interagit quasiment pas avec ces data center et donc ses propres données. Chaque fois qu'il sauvegarde une donnée sur l'un de ces services cloud (Dropbox, Google Drive, Github, Soundcloud) l'information quelle qu'elle soit, sera multipliée en très grand nombre afin d'éviter sa perte par crash d'un des disques détenant l'information. Le degré d'importance de l'information n'est, en aucun cas, pris en compte qu'il s'agisse de l'utilisateur ou autres personnes pouvant s'intéresser à cette ou ces données. Cette importance est liée de près au fait qu'elle soit consultée ou non par l'utilisateur lui-même, ou par d'autres qui pourraient y trouver un intérêt. En effet les informations consultées ou non consultées nécessitent la même demande en énergie et en entretien afin de maintenir l'information

pérenne. Il existe aujourd’hui plusieurs technologies de sauvegarde ayant des caractéristiques différentes et complémentaires, mêlant économie d’énergie, pérennité, temps d’accès à l’information rapide et simplicité d’usage.

Face à ces enjeux, une relation entre l’utilisateur, ses données, le data center et une technologie de sauvegarde actuellement démocratiser (disque dur ou disque ssd) pourrait être mise en place afin de réduire l’impact écologique de celui-ci (data center). Nous pourrions ainsi formuler la problématique suivante :

Comment et en quoi les différents supports de sauvegarde numérique (disques dur et disque SSD pour les plus démocratiser) pourraient ils se compléter et sensibiliser les utilisateurs des services cloud, afin de réduire leurs consommations d’énergie matérielle et immatérielle ?

En effet, chaque données est différente. Nous ne consultons pas autant nos photos personnelles, qu’un morceau de musique ou qu’un fichier de travail. Ainsi nous pouvons nous rendre compte qu’un grand nombre de données sont stockées dans ces centres avec des intentions différentes.

En m’appuyant sur une série d’entretiens menés avec des personnes utilisant de façon fréquente ces services (GoogleDrive, Soundcloud, Dropbox, Github), tels que des designers, des architectes, des musiciens, des développeurs ou bien d’autres travailleurs indépendants liés de près à ces technologies de cloud service. J’ai pu identifier les comportements et les habitudes de ces utilisateurs sur ces services de cloud computing. L’entretien était constitué de trois parties. La première partie concernait l’utilisation du cloud chez les personnes interrogées, la seconde sur leur compréhension de ces services, la troisième sur les améliorations et les différences à nuancer des fichiers. La durée de l’entretien variait entre une heure et une demi heure selon les participations de chacun. Afin de rencontrer des personnes liées de près à ces services je me suis rendu dans un espace de co-working à Lyon s’appelant « Soffa » où toutes ces professions se retrouve dans ce même lieu.

*Thèmes discutées  
en annexe. (guide  
d’entretien)*

Ce qui en fait un endroit singulier de par toutes ces professions vivant dans un même lieu les amenants à collaborer de façon plus spontanée. Pour élaborer ces entretiens je me suis servi du travail de recherche que Nicolas Nova à fait dans son projet « *clouds of practices* » (2018) permettant de garder un rapport très ouvert au cours des discussions.

Les utilisateurs que j'ai souhaité approcher pour m'entretenir avec eux sont pour la plupart développeurs, modélisateurs, architectes, designers, musiciens ou réalisateurs. Ces domaines sont quotidiennement confrontés à ces plateformes, soit pour confronter leur travail à leurs pairs ou bien pour le diffuser sur internet, ou encore pour archiver des documents terminés.

À l'issue de ces entretiens j'ai pu avoir des informations significatives quant aux comportements de ces personnes vis à vis de leurs données et de ce que cela impliquait dans leurs gestes de sauvegarde. Grâce à la première partie de ce questionnaire j'ai pu identifier les utilisations faites des services de cloud computing. La totalité des personnes que j'ai approchées m'ont confirmé avoir une utilisation quotidienne de celui-ci, que cela soit professionnel ou personnel. Que ce soit pour écouter de la musique ou bien envoyer des mails chacune de ces activités concerne un des data center de l'une de ces sociétés (Github, Soundcloud, Google drive, Dropbox).

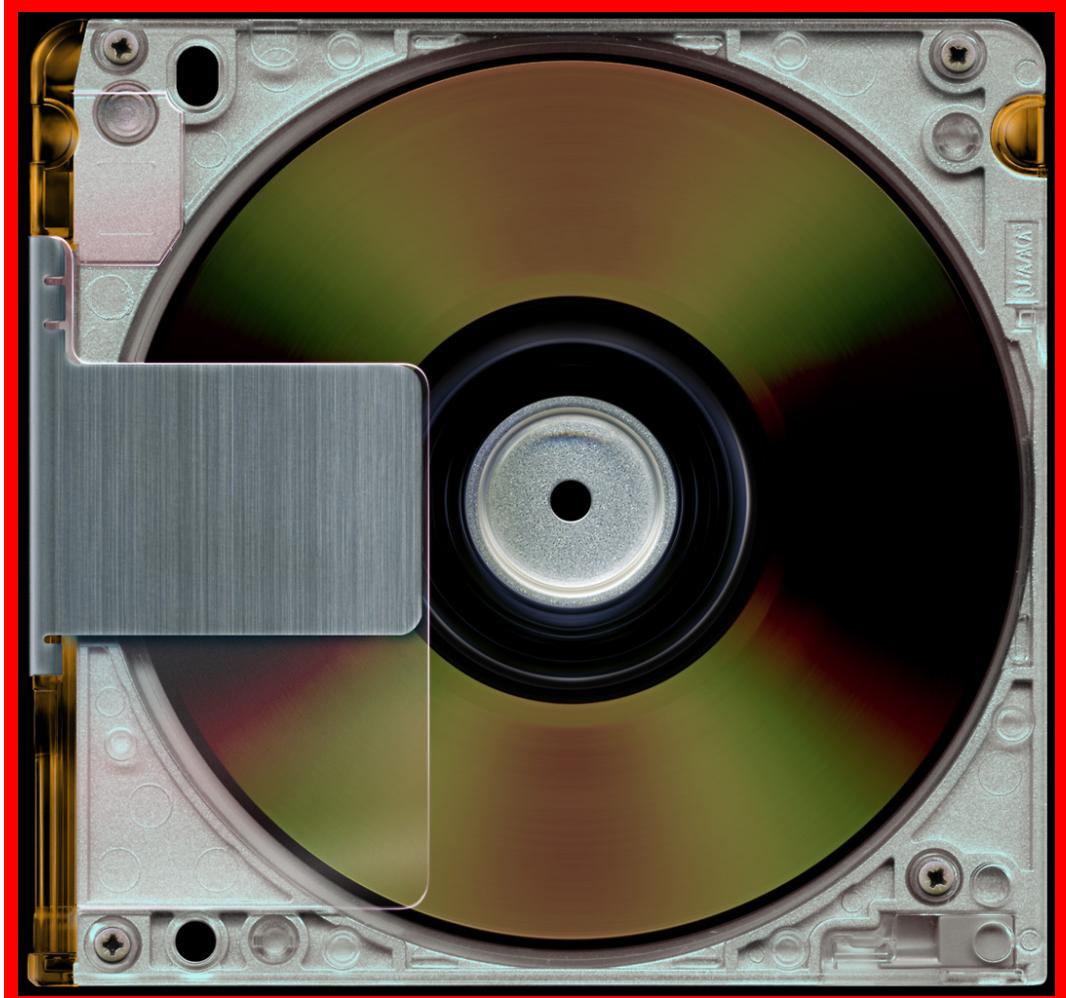
J'aborde ensuite les différentes technologies de sauvegarde afin de déterminer les caractéristiques de chacune, leurs complémentarité ainsi que leurs différences. En effet, en fonction des comportements que nous avons identifiés, nous avons pu déterminer une technologie en fonction de l'intention de l'utilisateur. Cette partie est donc constituée en premier temps d'une étude des différents supports existant et également de ceux étant en développement ou en prototypage. Cela permet de prendre connaissances de certaines caractéristiques comme leurs consommation électrique, durée de vie ou leur temp d'accès à l'information. Ainsi nous avons pu conjecturer des technologies en fonction des différents comportements étudié et de leurs caractéristiques.

le questionnaire se trouve en annexe.

Car comme dit plus haut, nous ne consultons pas un morceau de musique ou bien une photo ou encore un fichier de travail de la même façon.

Du fait qu'une donnée puisse ne pas être consulté par l'utilisateur est un facteur lié de près à sa consommation en énergie matérielle et immatériel (électricité nécessaire pour alimenter les disques durs des data center, la climatisations des infrastructures). Chaque médium de sauvegarde que ce soit les disques durs ou les disques flash (Solid State Drive) ont des caractéristiques différentes. Leurs consommation électrique par exemple peut passer du simple au triple en fonction du médium de sauvegarde. Cette partie se consacre donc à faire une corrélation entre l'intention de l'utilisateur et le médium de sauvegarde lui même afin de réguler au mieux la consommation d'énergie de ces data center. (aborder dans la partie 2)



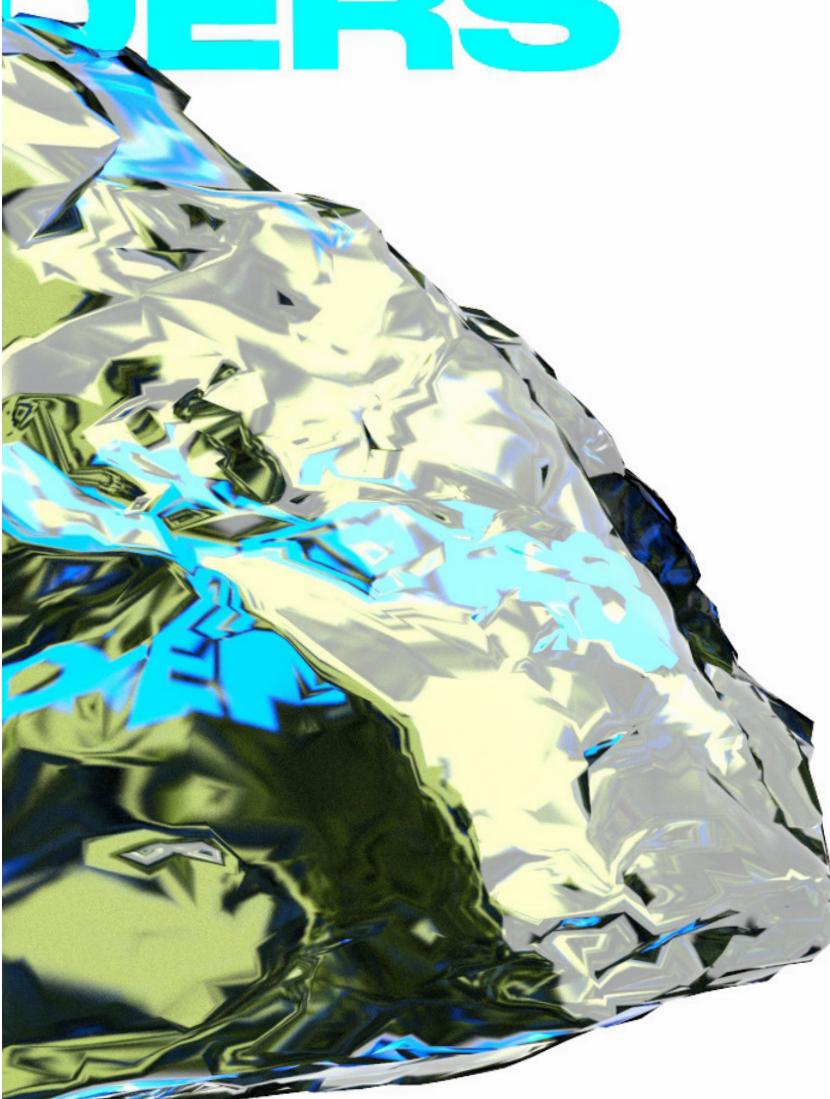


**Partie 1:**

CLOUD

CLOUD  
CLOUD

CLOUD



# DERS

# DERS

**Utilisateurs et utilisation.**

Si le terme *cloud* est couramment utilisé par les utilisateurs de façon général, nous avons remarqué que la plupart d'entre eux ont une notion peu claire de ce que cela représentait. Le domaine du *cloud computing* était difficile à appréhender et plusieurs définitions nous ont été données se rapprochant de ce qu'il désigne réellement, alors qu'il s'agit d'infrastructure fondamentale pour la vie quotidienne numérique. Il semble que ce domaine ne soit pas aussi significatif que ce qu'il englobe. Pour la majorité des utilisateurs le *cloud* se présente à eux de façon graphique (une notification sur un téléphone ou ordinateur les alertant d'un manque d'espace disponible pour sauvegarder un fichier sur leurs machines, les invitant à utiliser des services cloud pour le *stocker*). La plupart du temps ces usagers ont accès à leurs données stockées sur ce *cloud* et interagissent avec lui sous forme de petites icônes et pictogrammes. Cependant pour presque tous nos appareils nous ne pouvons pas faire de réelles différences entre ce qui est sur le *cloud* et ce qui ne l'est pas ou ce qui est réellement dans notre appareil. Cela pose un problème lorsque l'utilisateur n'arrive plus à avoir accès à ses données (problème de réseaux par exemple). Un sentiment de frustration se fait ressentir car celui-ci ne fait pas clairement la différence entre ce qui se trouve bien dans son téléphone et ce qui ne l'est pas.

Ainsi, pour mieux explorer les différents types d'usage de ces services j'ai élaboré un guide d'entretien divisé en trois parties permettant de mieux cibler les attentes et les usages des parties prenantes. Ce questionnaire comprend donc une partie sur l'utilisation, la compréhension et sur les améliorations et différences que les utilisateurs souhaiteraient apporter.

Les motivations et les raisons les poussant à utiliser ces services sont aussi très diverses. Je cite en dessous les différentes réponses que j'ai obtenue :

Je ne peux pas me permettre de perdre un de mes projets. Parfois je suis obligé de revenir sur l'un d'eux si un client me pose une question. Aussi lorsque celui-ci est terminé je préfère le déposer sur ces plateformes et ainsi libérer de l'espace sur mon ordinateur.

*Entretien ayant eu lieu  
à Lyon*

Nicolas (dévellopeur), 29 ans, à Lyon.

Ce comportement résume assez bien le fait que l'utilisateur souhaite pouvoir libérer de l'espace sur son ordinateur pour s'adonner à de nouveaux projets. Il dépose ses travaux sur ces plateformes mais ne les consultera pas forcément de manière régulière, cependant il doit pouvoir tout de même y garder un accès. Mais d'autres comportements peuvent aussi apparaître comme notamment le fait de vouloir *stocker* un fichier uniquement sur le cloud car il ne paraît pas utile d'utiliser de la mémoire sur son appareil pour celui-ci. Parfois lorsque nous ne possédons plus assez de place sur nos appareils nous préférions délocalisé certains fichiers ailleurs que sur notre propre machine. Car les fichiers stockés localement sur la machine de l'utilisateur peuvent avoir plus d'importance que d'autres d'où le souhait de *stocker* des informations moins importantes sur des plateformes de cloud computing. C'est ce que dit Simon :

Parfois quand je travaille avec des clients et que le boulot en lui-même ne me plaît pas, Je ne prends pas la peine d'amener les sources d'inspiration du client sur ma propre machine. Je les consulte directement sur ces plateformes.

Simon (graphiste), 27 ans, Lyon.

Cette particularité montre aussi le fait que parfois certaines informations sont très peu importantes pour l'utilisateur même et les *Oublieras* dans un futur proche. Il existe aussi différentes façons de travailler en groupe via ces plateformes. En effet il se peut que des domaines et des métiers doivent collaborer sur un projet et se conjuguer pour celui-ci. Simon par exemple devait travailler sur un projet de pochette de disque pour un artiste. La pochette devait être en accord avec les désirs de l'artiste et du photographe ayant pris la photo. « Ce qui était intéressant dans ce travail c'était que nous pouvions voir l'évolution du travail de chacun en fonction des retours que nous faisions. » Ce comportement est également intéressant dans le sens où ce travail devait être présent sur les ordinateurs de chaque personne concernée. L'information se retrouve ici multipliée par le nombre de personnes travaillant sur le même projet.

Un quatrième comportement apparaît également. Lorsqu'une personne doit pouvoir être capable de travailler sur un même projet mais sur différents appareils. C'est le cas de Julien, architecte, qui parfois travaille de chez lui sur son ordinateur personnel lorsqu'il est en retard sur une tâche.

Mon ordinateur au bureau est plus puissant et plus adapté pour mes tâches professionnelles. Cependant il m'arrive pendant mes jours de congé de devoir travailler sur un de mes projets et il faut impérativement que je sois en mesure de reprendre la ou je me suis arrêté. Julien (architecte), 32 ans, Lyon.

Cette attitude démontre bien l'importance de pouvoir accéder à ses données n'importe où et de pouvoir reprendre un travail en cours avec n'importe quelle machine. Il existe aussi le fait de vouloir dupliquer des informations de façon numérique. Par exemple à Lyon il existe un espace de co-working où plusieurs professions et plusieurs petites entreprises travaillent dans un même lieu. Lorsqu'une nouvelle personne arrive dans ce lieu pour acquérir un espace de travail il doit remplir un formulaire lui permettant de s'inscrire dans celui-ci. Une fois le formulaire d'adhésion rempli il est numérisé et stocké sur une de ces plateformes (en l'occurrence GoogleDrive). Ces informations ne sont ensuite jamais consultées à part lors d'un départ ou de l'arrivée d'une personne.

Enfin, il peut exister un comportement de sauvegarde visant à être diffusé à grande échelle pour un travail accompli. Que ce soit pour de la visibilité ou des conseils d'autres utilisateurs, une information peut aussi devoir se multiplier en très grand nombre pour être accessible par beaucoup et ainsi augmenter son audience. Edouard étant musicien âgé de 29 ans souhaite, lui, diffuser un maximum ses productions et stocke ainsi la quasi totalité de son travail sur cette plateforme qui est SoundCloud. Nicolas également est confronté à ce genre de stockage lorsque lui sauvegarde des travaux sur Github pour avoir des conseils de réalisation ou d'architecture dans ses programmes de la part de ses pairs.

Sur la base de ces échanges, nous avons défini quatre types de comportements face aux données archivées.

Ces comportements peuvent se résumer en une série de catégories d'intentions de l'utilisateur. Il y aurait donc accroître, *partager*, *Oublier*, *stocker*.

*Keller, P. (2018). Cloud of Cards. ECAL.*

– **Accroître**: Cela vaut pour les données que nous souhaitons *partager* avec le plus grand nombre. Ces données se doivent d'être multipliées dans les data center concernés pour que l'information soit accessible rapidement. Par exemple, lorsque Édouard souhaite que son morceau soit accessible rapidement celui-ci doit être multiplié dans plusieurs zones géographiques différentes en fonction de là où se trouve la personne souhaitant écouter ce morceau.

– **Partager**: Pour les données que nous considérons comme communes à plusieurs usagers. Un projet de travail en commun avec un groupe par exemple.

– **Oublier**: Pour les données que nous mettons sur un de ces services et que nous ne consulterons plus une fois que la tâche à accomplir sera terminée. Un travail en groupe par exemple pourra devenir par la suite un ensemble de données considérées comme à *Oublier*. Il en va de même pour des photos, des mails (le pire étant les newsletters) qui ne seront plus jamais consultés.

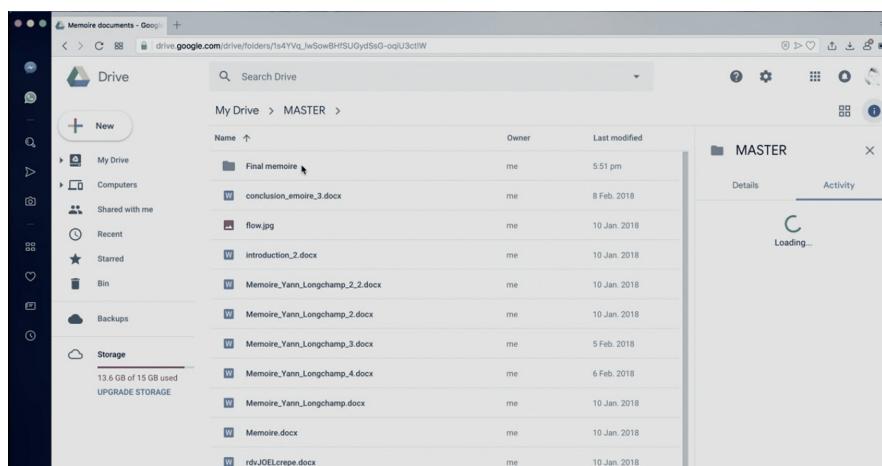
– **Stocker**: Comme l'archivage de fichier papier que l'on souhaite copier numériquement mais que l'on ne consultera pas. Ce genre de données peut être par exemple un fichier de comptabilité d'une entreprise qui souhaite s'assurer de la pérennité des informations importantes concernant l'entreprise en question.

Une fois que nous avons défini les différents types de comportement en fonction de l'intention que l'utilisateur souhaite faire de sa donnée, nous avons analysé à quel type de service ces données appartenaient. En effet, chaque donnée est très différente, une photo ne correspond en rien à une discussion en ligne par exemple.

Grâce aux entretiens nous avons pu définir trois domaines de types de donnés :

### Les sauvegardes de fichiers :

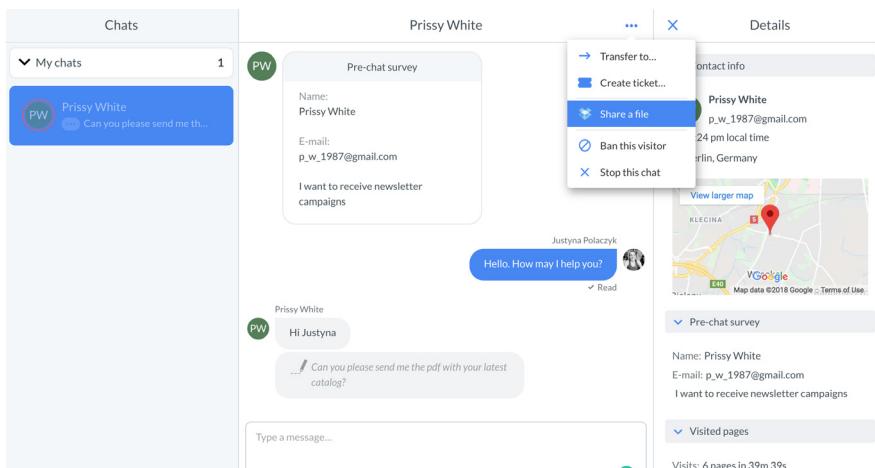
La plupart de ces fichiers sont sauvegardés en cas de perte des données en local afin de s'assurer une copie sur le *cloud*. On trouve aussi les synchronisations de fichiers ou plusieurs utilisateurs travaillent sur un même projet et souhaitent *partager* leurs avancées ou encore l'archivage (*fig1*) de fichiers initialement sous format papier doublés en fichier numérique. Il peut également avoir le cas de certains fichiers qui ne sont sauvegardés que sur le *cloud* et plus de manière locale, comme par exemple un ancien fichier de travail que nous souhaitons archiver de manière sûre tout en libérant de l'espace sur les machines concernées.



(fig1)

## Les communications :

Les communications englobent plusieurs types de fichiers également. Il y a tout d'abord la collaboration de plusieurs utilisateurs sur un projet. Le partage de photos, de plans ou d'images ainsi que les discussions (*fig2*) en ligne pouvant les accompagner. À l'issue d'un projet ces communications peuvent devenir des sauvegardes de fichiers par défaut. En effet ces échanges seront enregistrés mais plus consultés. Ils resteront accessibles malgré le fait qu'ils ne servent plus.



(*fig2*)

## Les services :

Les services sont peut-être les données les moins importantes car elles concernent des ressources informatiques (*fig3*) pour les développeurs ou les modélisateurs lorsque ceux là veulent acquérir, par exemple, une bibliothèque permettant de pouvoir ajouter une commande ou une caractéristique à leur projet ou code. Ces librairies sont des ressources très importantes dans le gain de temps pour accomplir un projet. Même en étant un des services les moins gros en terme de taille de sauvegarde, il fait tout de même partie des plus consultés un grand nombre de fois par les différents utilisateurs

An analyzer library for C# and VB that uses Roslyn to produce refactorings, code analysis, and other niceies. <http://code-cracker.github.io/>

Branch: master	New pull request	New file	Find file	HTTPS	<a href="https://github.com/code-c">https://github.com/code-c</a>	Download ZIP
<a href="#">giggio</a> Merge pull request #711 from adraut/Bug#700CC0052TriggeredWithVariable... <a href="#">...</a>						Latest commit <a href="#">721ddc4</a> 3 days ago
<a href="#">nuget</a>	Update nuget to 3.3 and psake to 4.5					13 days ago
<a href="#">src</a>	Merge pull request #711 from adraut/Bug#700CC0052TriggeredWithVariable... <a href="#">...</a>					3 days ago
<a href="#">test</a>	Merge pull request #711 from adraut/Bug#700CC0052TriggeredWithVariable... <a href="#">...</a>					3 days ago
<a href="#">gitattributes</a>	remove eol=crlf from gitattributes as now appveyor supports init config					a year ago
<a href="#">.gitignore</a>	Update nuget to 3.3 and psake to 4.5					13 days ago

(*fig3*)

Nous pouvons à ce stade dire que les données sont très différentes et que les milieux professionnels et privés sont liés quant à l'utilisation des services cloud. On constate que l'utilisation de ces services est quotidienne pour la plupart des gens possédant des appareils permettant d'avoir accès à ces services. Mais que ces gestes et comportements décrits ci-dessus par les utilisateurs traduit la valeur et la certaine nature de leurs données.

Une fois les comportements et les données analysés, nous les avons présentés aux utilisateurs ayant participé à l'entretien. Il leur est apparu comme logique le fait que des données puissent paraître plus importantes que d'autres quant à

leur sollicitation par les usagers. Ce qui est notable dans cette différenciation, c'est que toutes les informations sont stockés sur les mêmes supports de sauvegarde numérique, les disques durs. Les disques durs consomment énormément d'énergie, de l'ordre de 6 à 12 Watt par heure, sont responsable de la chaleur émise des data center et nécessite par le même coup un apport de froid permanent pour leur bon fonctionnement.

À cela, si nous ajoutons le nombre de données qui ne sont que très rarement consultées voire jamais et la consommation qu'elles nécessitent pour être accessibles et disponibles nous pouvons dire qu'elles représentent un énorme gaspillage d'énergie matérielle (espace pris dans les disques durs) et immatérielle (électricité nécessaire pour garder l'information accessible).

Alors qu'aujourd'hui les centres de stockage arrivent à un stade où énormément de données leurs sont envoyées. Que ces supports (disques durs) après soixante ans de bon et loyaux services commencent à arriver à leur limite physique. Cependant, il existe aujourd'hui des objets qui envoient énormément d'informations à ces data center. Ces objets se font appeler IOT (Internet of things) et nous en possédons tous smartphone, tablette, gps, etc. Les informations générées par les IOT sont liées à nos activités sur ces appareils (consultation internet, recherches diverses) et sont affiliées à nos profils personnels dans des buts commerciaux (publicités ciblées essentiellement).

Cependant, la plupart de ces informations sont redondantes ou non essentielles et peuvent être facilement supprimées. La prolifération de ces objets est une source majeure de croissance du volume des données. Connectés en permanence à des infrastructures de réseau et collectant des données, les appareils IOT devraient dépasser les 20 milliards d'unités d'ici 2020. De plus 80 pour cent de ces données sont considérées comme étant non-structurées. Ce terme signifie que les données se distinguent par l'absence de tout format spécifique. Ces données non-structurées peuvent elles aussi contenir de nombreuses informations commerciales précieuses, mais cela peut ressembler à la recherche

d'une aiguille dans une botte de foin. Environ 10 pour cent seulement des données non structurées méritent d'être sauvegardées pour analyse. Les utilisateurs ne savent pas forcément que ces données sont généralement créées par leurs propres recherches sur ces appareils et ils ne peuvent pas interagir avec elles, ou bien décider de ce qui est important de garder pour eux ou pas. En effet, ce sont des ingénieurs informaticien et des experts qui analysent ces informations et qui décident de garder ce qui leurs semble intéressant (d'après le site spécialisé *au coeur des infrastructures*).

Nous pouvons dire que dans ce cas que l'utilisateur n'est pas forcément conscient de ce qu'il émet comme données. Elles ne sont pas présentées de la même façon que lorsqu'il souhaite enregistrer ses propres documents. Elles émanent de ses appareils sans que celui-ci le sache et représente une énorme source de données stockées sur les mêmes supports que ses propres données, les disques durs.

Nous pouvons nous rendre compte d'avantage que ces médiums que sont les disques durs arrivent au bout de leur capacité tant physique que historique. De plus, si nous combinons les données faites par les utilisateurs de façon consciente (en rapport avec les descriptions analysées plus haut) et de façon inconsciente (générées par les appareils connectés des utilisateurs) nous arriverons finalement à collecter énormément d'informations sur nos sociétés via ces appareils et d'ici 2020 nous pourrions cumuler un total de dizaines de zettaoctets (1 zettaoctet équivaut à 1 milliard de gigaoctet) par an au cours des cinq prochaines années. (d'après le site spécialisé *Cisco*)

Entre temps la technologie a évolué permettant de créer de nouveaux moyens de sauvegarde plus rapides, plus fiables et consommant nettement moins d'énergie. Alors que d'autres, toutes aussi innovantes sont en cours de développement ou de prototypage. Nous pourrions conjecturer un dialogue s'instaurant entre l'utilisateur, ses données, le médium d'archivage et les data center eux mêmes. Ce dialogue s'instaurerait entre l'utilisateur et les données émises consciemment par lui-même, en

fonction de la catégorie de la donnée et de l'intention de l'utilisateur (intention de l'utilisateur et catégorie de la donnée définis en première partie).

Chaque donnée pourrait être stockée sur un médium favorisant le fait qu'elle puisse être consulté ou non. Les informations les moins consultées pourraient ne pas être mises en valeur quant au médium choisi, c'est-à-dire les disques durs. Ces informations pourraient être mises en veille ou supprimer par l'utilisateur afin de réduire l'impact écologique de ces data center .

Donc parmi les fichiers les moins consultés nous aurions les fichiers considérés comme à *Oublier*, à *stocker* et une partie des fichiers à *partager* qui ne seraient plus actifs dû au fait que la tâche qui devait être accomplie est terminée. Quant aux informations les plus consultées par les utilisateurs nous pourrions les *stocker* sur des médiums consommant moins d'énergie, ayant un temps d'accès à l'information beaucoup plus rapide que les moyens utilisés aujourd'hui.

Les informations les plus consultées pourraient correspondre aux catégories : Accroître et *partager* (lorsque les usagers sont encore actifs sur ces projets qui nécessitent de *partager* les informations, ou bien la diffusion de celui-ci sur les plateformes liées à la nature du projet).

Chaque catégorie définie plus haut (*Accroître, Partager, Oublier, Stocker*) représente une intention de la part de l'utilisateur. Chacune de ces intentions traduit une volonté de la part de celui-ci quant à la consultation ou non de la donnée.

Par exemple la catégorie *Accroître*, nécessite une large consultation de la part des utilisateurs et doivent être disponibles dans plusieurs disques se trouvant au plus proche de celui-ci pour que l'accès à la donnée soit le plus rapide possible. Dans ce cas de figure l'information justifie sa consommation électrique par le fait d'être consulté. Cependant comme dit plus haut il existe des médiums de sauvegarde qui nécessite moins d'énergie et étant plus rapide que ceux

utilisés actuellement. Car en reprenant l'exemple d'Édouard (musicien) ses morceaux de musiques étant multiplié dans plusieurs data center nécessite donc une consommation d'énergie importante due au médium de sauvegarde lui même (disque dur). Il serait donc possible de réduire la consommation énergétique des data center en utilisant un médium nécessitant moins d'énergie et étant plus rapide afin d'avoir moins de disques contenant l'information mais la distribuant tout aussi rapidement.

Pour le cas de la catégorie *Oublier*, ces données ne sont plus consulté car l'utilisateur n'en a plus aucune utilité mais continue de la *stocker* sur un de ces services cloud, elle peut être également ignorer de la part des utilisateurs (newsletter, données produites par les IOT). Cette donnée ne justifie donc pas sa consommation électrique dans les data center. Cette catégorie de fichier pourrait donc ne pas être alimenter, mis en veille, ou supprimer afin de réduire l'impact énergétique des data center.

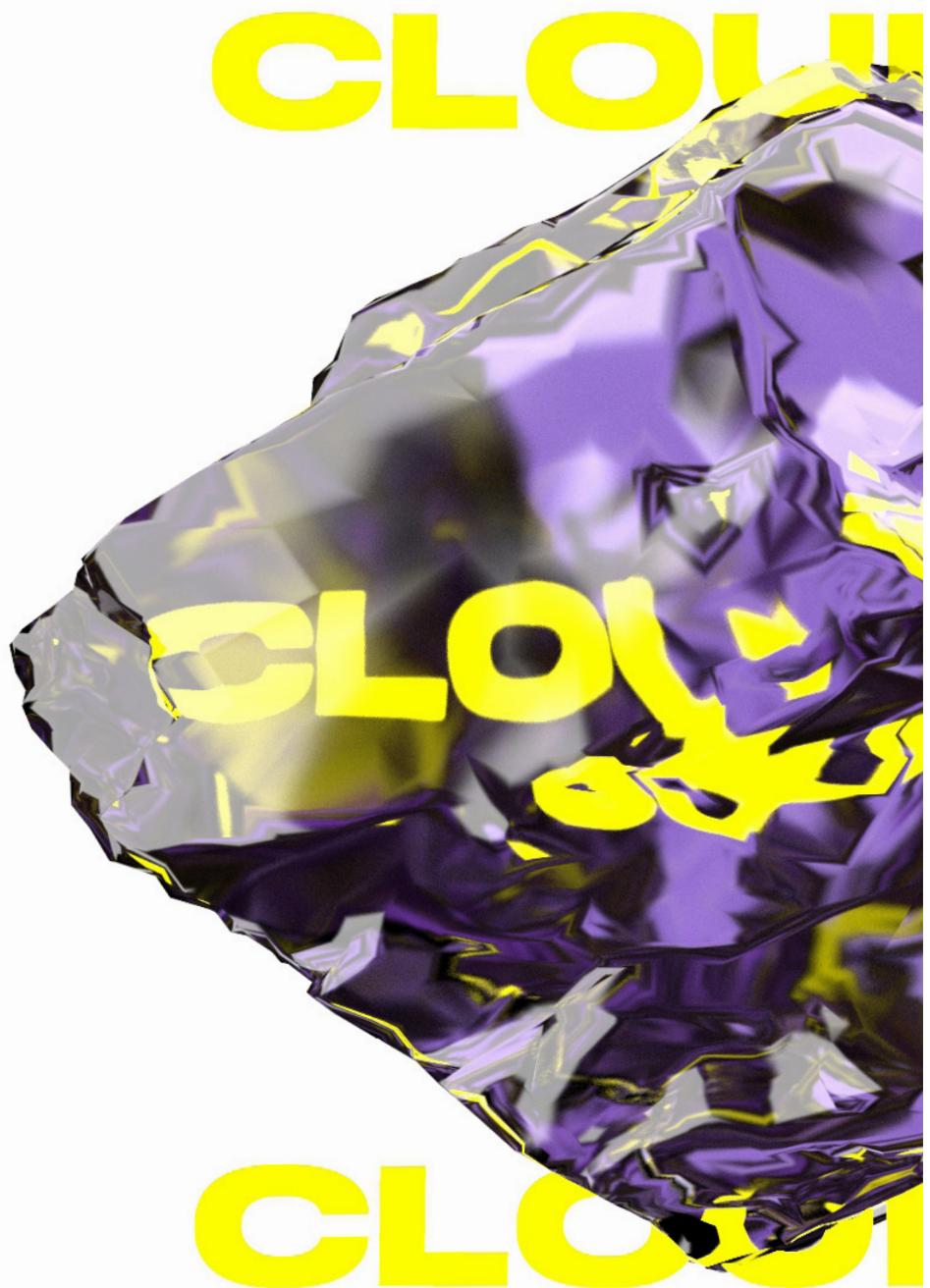
La catégorie *Partager* nécessite également une consultation plus ou moins importante selon le nombre de personnes avec lesquels l'information ou le projet est partagé. Cependant une fois que ces projets sont terminés il ne sont plus consulté. Cette catégorie de données justifie donc une consommation électrique lorsque elles sont consultés par les utilisateurs. Cependant une fois le projet terminé, les informations restent stockés mais ne sont plus consultées de la part des utilisateurs et ne justifie donc plus leur consommation électrique permanente.

Enfin la catégorie *Stocker* ne se fait quasiment pas consulter mais doit être archivé de manière pérenne. Ce type de catégorie ne nécessite pas une alimentation permanente du fait de leur faible consultation de la part des utilisateurs. Cependant ces données restent alimentées de façon permanente sur des médiums consommant beaucoup d'énergie que sont les disques durs (leurs consommation est en moyenne de 6 W/h). Ce type de données pourrait nécessiter une alimentations partiel ou une mise en

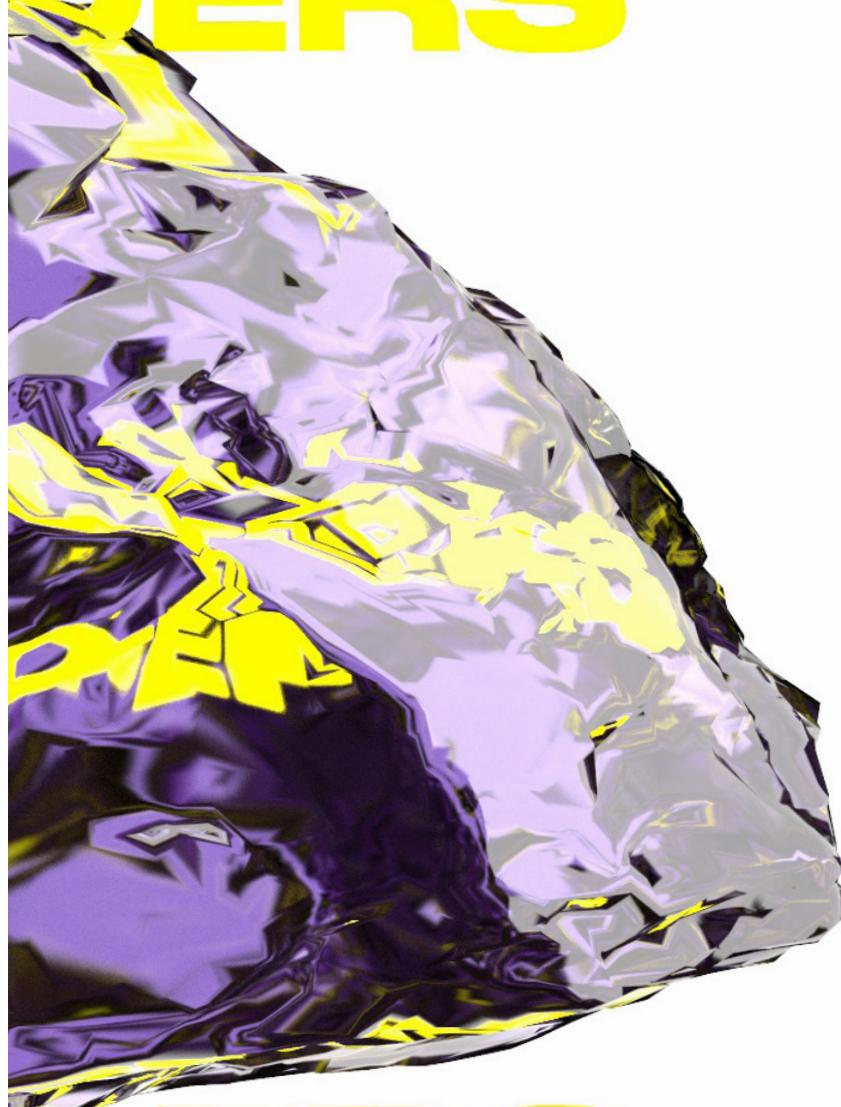
veille afin de réduire la consommation électrique.

À partir de ces enjeux, quelles peuvent être les technologies permettant de mettre en valeur ces informations consultées ou pas par l'utilisateur ? Quelles sont leurs avantages ? Pouvons nous combiner plusieurs technologies afin d'influencer les consommations faites par ces data center ? Quelles solutions sont apportées par ces médiums d'archivage ?

**Partie II:**



# REERS



# DERS

**Technologies et comportements.**

Nous disposons aujourd’hui de différentes technologies, certaines sont largement démocratisées et abordables, d’autres nettement plus onéreuses, mais nettement plus stables et pérennes pour nos informations. Tout d’abord, la première technologie de sauvegarde connue et utilisée aujourd’hui de la part des utilisateurs est sans conteste le disque dur. (fig4) L’histoire de cet objet est aussi intéressante que sa large démocratisation. En effet, cette technologie est marquée par une histoire d’innovation et d’amélioration pour qu’elle coïncide au mieux à une utilisation quotidienne de la part des utilisateurs. Depuis sa création en 1956, la taille du disque dur a été réduite de 57 000 fois, son stockage augmenté d’un million de fois et son coût diminué de 2 000 fois. En d’autres termes, le coût par gigaoctet a été réduit de 2 milliards de fois en environ 60 ans. Toutes ces innovations ont conduit à la démocratisation de cette technologie et à l’appropriation de celle-ci. C’est également la technologie la plus rentable financièrement pour les constructeurs de data center car son rapport Go/\$ est le plus faible par rapport aux nouveaux médiums cités plus bas. Cependant, cette technologie arrive aujourd’hui à ses limites physiques et historiques dues à la croissance exponentielle des utilisateurs et des contenus web. En effet, cette technologie consomme énormément d’énergie (entre 6 et 12 W/h à plein régime). De plus, elle produit énormément de chaleur due à son mécanisme de fonctionnement. Lorsque ces disques tournent à plein régime, ils engendrent énormément de chaleur (environ 55 degrés Celsius) et nécessitent en plus de leurs propres consommations électriques une seconde source d’électricité nécessaire pour les refroidir. Greenpeace estime que la moitié de la consommation électrique d’un data center est due aux apports d’air froid.

*d’après le site spécialisé  
colocationamerica*



(fig4)



Une seconde technologie est également largement plus démocratisée depuis une dizaine d'années et a permis à de nombreux utilisateurs d'augmenter la puissance de stockage et d'exploitation de leurs appareils. Cette technologie se nomme disque SSD(*fig5*) (solid state drive) et nombreux sont ses avantages. En effet, cette technologie consomme trois fois moins d'énergie que les disques durs. Seulement 2 W/h sont nécessaires au bon fonctionnement de celui-ci, le temps d'accès à l'information est également divisé par 10 pour atteindre 0.1 milliseconde, son format est plus compact (2,5 pouces contre 3,5 pour les disques durs) et ils n'engendrent pas de chaleur, car il n'y a pas de système mécanique comme pour le disque dur. Néanmoins, leur principal inconvénient reste leur prix, car ils sont en moyenne 5 fois plus chers que les disques durs. Les utilisateurs se sont malgré cela très rapidement approprié cette technologie. Lors des entretiens, j'ai pu rencontrer des développeurs de jeux vidéo et de projets interactifs faisant partie d'un studio nommé Treenode. Ils m'ont rapporté plusieurs faits :

La première fois que j'ai allumé mon ordinateur qui avait un disque SSD j'ai été stupéfait. J'ai adoré cette vitesse, ce silence, et aussi le facteur « whoa » de cette technologie qui semblait meilleure à tous les points de vue par rapport aux disques durs classiques. J'étais prêt à accepter pleinement la promesse des disques SSD. Mon bureau utilise un SSD pour le démarrage, les applications et les fichiers de travail. Mon ordinateur portable a un seul disque SSD de 512 Go. Lorsque mon ordinateur portable est à mon bureau, il est connecté à un disque dur externe de sauvegarde de 1,5 To. Les disques durs ont toujours une place dans mon environnement informatique personnel comme dans le vôtre. Cependant, rien ne reste longtemps identique, en particulier dans le monde informatique. Certaines technologies de stockage feront leurs apparitions, avec peut-être un facteur « whoa » encore plus impressionnant.

Cet exemple montre bien que l'utilisateur stocke ses

*d'après le site consacré  
aux médias de  
sauvegarde numérique  
storage review*

informations de premier plan (celles dont il a le plus besoin) sur le disque SSD pour tous ses avantages. L'information est nettement plus rapide à atteindre et les temps de traitement des tâches sont plus petits grâce à cette technologie, la capacité de sortie et d'entrée des informations est également supérieure (Input/Output) et sa faible consommation électrique (2 W/h) sont ses principaux atouts. Les catégories de fichiers pouvant s'apparenter à être mis en valeur via ce médium correspondent à *Accroître* et *Partager* (lorsque ces travaux sont en cours de réalisation et non terminés). Ainsi les utilisateurs pourraient avoir accès de la même façon à leurs données les plus consultées tout en réduisant l'empreinte écologique des data center. En revanche comme il le cite « Les disques durs ont toujours une place dans mon environnement informatique » dû au fait que cette technologie reste relativement peu chère sur le rapport Go/Cela souligne le fait que l'utilisateur dans cette situation met les informations dont il n'a plus besoin sur ce médium. Elles peuvent bien sûr être reconsultées, mais plus de manière aussi immédiate qu'avec le disque SSD. Ce geste permet d'économiser beaucoup d'énergie, car si les informations les plus consultées se trouvent sur des disques SSD alors leur consultation n'engendrera pas une consommation électrique comparable à celle utilisée par les disques durs. Un disque dur consomme entre 12 W/h en plein régime et engendre beaucoup de chaleur, engendant à son tour une consommation d'énergie supplémentaire nécessaire à l'apport de froid pour le bon fonctionnement des disques durs. Ainsi, en désignant le disque SSD pour les catégories *Accroître* et *Partager* de l'énergie serait économiser grâce aux caractéristiques de ce médium.

*storage review*

Enfin, les données des catégories *Stocker* et *Oublier* sont redistribuées aux disques durs. Ce qui permet d'engendrer moins d'énergie de la part de ces disques, grâce à une typologie de données qui nécessitent moins de consultation. Par contre, ces médiums d'archivage n'ont pas fait leurs preuves sur le long terme.

Nous savons que les disques durs ont une durée de vie d'une vingtaine d'années tout au plus, que les disques SSD ont une capacité de 100 000

réécritures, qu'à partir de ce chiffre le disque se détériore et n'est plus stable. Nous vivons une époque où l'humanité n'a jamais autant produit d'informations. Pourtant une caractéristique assez étonnante se fait remarquer. Les supports de stockage n'ont jamais été aussi fragiles qu'aujourd'hui. Nous savons qu'une inscription sur la pierre dure 10 000 ans, sur un parchemin 1 000 ans, sur un vinyle 50 ans. L'informatique dont nous dépendons est confrontée à une grande fragilité. En observant cette évolution, des technologies ont déjà disparu (disquette, VHS, CD (en voie)). On se rend compte qu'aucune d'entre elles ne s'est imposée sur le long terme. Cela implique un problème de transmission de nos données aux futurs utilisateurs. Cependant, il existe aujourd'hui plusieurs technologies pouvant avoir des aspects complémentaires entre elles. Ces dernières sont élaborées par des laboratoires dans différents endroits du monde afin d'assurer la pérennité de nos informations, mais aussi dans leur technique d'utilisation simple afin que les utilisateurs puissent toujours avoir les outils nécessaires à leur lecture à l'inverse des technologies conçues jusqu'à présent.

Une technologie en cours de développement à l'université de Kyoto au Japon reprend les caractéristiques de la gravure et de la roche. Ce sont les disques à quartz. L'information est gravée dans la masse du disque la protégeant ainsi de toutes marques pouvant altérer la lecture de l'information en surface. Pour lire l'information contenue dans le disque, il suffit d'un microscope classique. Le fait de garder un outil de lecture simple comme le microscope assure la simplicité d'utilisation pour les futurs utilisateurs et sa pérennité. Son rapport taille/espace de stockage est également notable, l'équivalent d'une pièce de 2 francs peut contenir 360 To (contre un maximum de 1To pour les disques SSD). Ce médium assure une pérennité quasi éternelle, car celui-ci supporte des températures supérieures à 1 000 degrés sans altérer les informations contenues à l'intérieur de celui-ci. Son utilisation pourrait se faire dans le cadre de maintenir des informations importantes de façon pérenne. Ce genre de médium pourrait être utilisé pour des archivages d'informations importantes sur le très long terme. Par exemple pour des

travaux déterminants comme la position de déchets nucléaires pour que les futures générations puissent être capables de localiser facilement ces lieux. Cela vaut également pour des travaux scientifiques, universitaires, professionnels qui font partie de notre patrimoine humain. L'avantage de ce médium est qu'il est possible de ne pas alimenter son outil, le microscope lorsque l'information n'est pas consultée. Si ce médium venait à se démocratiser de la même façon que le disque dur ou bien le disque SSD alors les informations appartenant à la catégorie *Stocker* pourrait ne plus être alimenter et ferait baisser la consommation énergétique des data center de façon conséquente puisque ces données ne sont presque pas consultées ou que de façon ponctuelle. Ce qui serait intéressant serait donc de placer chacune des catégories de donnée sur un médium correspondant à sa fréquence de consultation. Ainsi nous pourrions concentrer les données qui sont les plus consultées (*Accroître* et *Partager*) sur des médiums tels que les disques SSD.

Ce médium consomme trois fois moins d'énergie électrique (2 W/h) pour sa propre alimentation et ne nécessite pas d'apport d'air froid donc d'énergie supplémentaire comme pour les disques durs. En ajoutant à cela un temps d'accès à l'information extrêmement rapide (0,01 ms contre 0,13 pour un disque dur, d'après le journal spécialisé *Colocation America*). Elles pourront être distribuées presque dix fois plus rapidement aux utilisateurs en nécessitant seulement 2 W/h. Quant aux catégories *Oublier* et *Stocker* elles pourront être sauvegardées sur les disques durs. Ces derniers produiraient moins de chaleur et ne gaspilleraient pas d'énergie supplémentaire pour leur refroidissement. Cette relation entre les catégories de données (*Accroître*, *Oublier*, *Stocker*, *Partager*) définies par la consultation et l'intention de l'utilisateur serait une alternative « ponctuelle » au système actuel. Les disques durs sont une technologie vieille de soixante ans et leur émulation ne peut être infinie. Aujourd'hui, le choix optimal pour les data center serait la technologie des disques SSD pour les avantages qu'ils apportent. Pour autant, cette technologie a ses limites. Les disques SSD stockeraient des catégories de données qui ne

Arte

« Nos ordinateurs ont ils la mémoire courte »  
(2016)

seraient pas ou peu consulter (*Oublier, Stocker*) ce qui consommerait tout de même de l'énergie.

Un second médium est lui aussi en développement. La traduction de données numériques en ADN. Cette traduction doit donc prendre en compte le fait que les données sont constituées d'un langage binaire (0 et 1) et doivent être traduites en langage ADN contenant quatre unités de langage (ATCG). L'énorme avantage de cette technologie est que l'outil permettant sa lecture existe depuis la découverte de l'ADN et restera actuel. Aussi, la taille de stockage serait considérablement réduite. Le stockage d'internet dans sa globalité pourrait tenir dans deux mètres cubes plutôt que les millions de mètres carrés utilisés actuellement. Cette technologie réduirait considérablement l'espace pris au sol des data center. Toutefois, cette technologie sera extrêmement onéreuse.

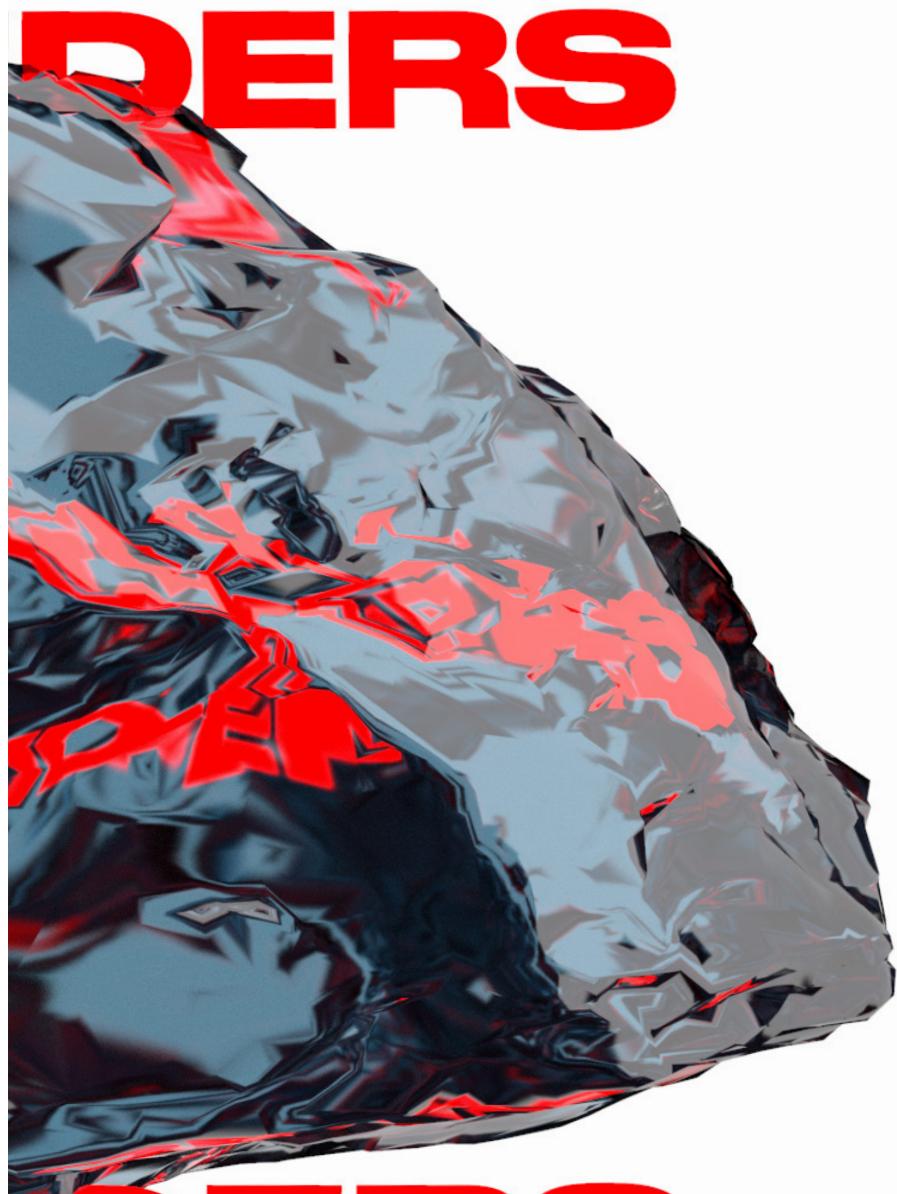
Ces technologies sont intéressantes, mais aucun usage n'a été fait de la part d'utilisateur. Ce qui est notable avec ces deux dernières technologies c'est la volonté de rendre ces données éternelles. Aujourd'hui ce genre de médium ne conviendrait pas pour des usages quotidiens, mais seulement à vouloir faire perdurer certaines de nos données.

Si les disques à quartz ou les supports de sauvegarde ADN venaient à se démocratiser de façon similaire aux disques durs ou aux disques SSD, ils pourraient contenir ces informations moins consultées et réduire l'empreinte écologique des data center .

CLOW



CLOW



# DERS

# DERS

**Conclusion.**

39

Ce mémoire a pour ambition de réduire l'empreinte écologique des data center en se demandant si les différents supports de sauvegarde numérique (disques durs et disque SSD principalement) peuvent se compléter, être utilisés différemment, afin de réduire la consommation d'énergie matérielle et immatérielle des data center. L'approche principale de cette réduction passe par la sensibilisation des utilisateurs sur les services cloud.

Au cours de ce travail, j'ai pu constater que les utilisateurs avaient des intentions différentes lorsqu'ils utilisaient ces services cloud. Chacune de ces intentions traduit la volonté de la part de l'utilisateur de vouloir *partager*, *stocker*, *Accroître* ou *Oublier* une donnée. En d'autres termes, l'utilisateur en fonction de l'intention qu'il a, décide de la condition de la donnée. Sera-t-elle consultée ou pas ?

Dans un premier temps, il a fallu définir les usages des utilisateurs de ces services cloud (Github, Soundcloud, Google drive, Dropbox). Pour cela, une série d'entretiens de personnes ayant recours à ces *cloud services* de façon fréquente a permis de définir les différentes utilisations des usagers. A l'issue de ces entretiens, quatre catégories de classification de fichiers relatant différentes intentions de la part des utilisateurs ont pu être définis : *Partager*, *Oublier*, *Stockier*, *Accroître*.

Il est possible de constater que les utilisateurs ont des intentions différentes lorsqu'ils utilisent ces services cloud. Chacune de ces intentions traduit une volonté de la part de l'usager. La manière dont il souhaite *stocker* et consulter ses données va influencer son choix de service et donc la condition et le futur de la donnée. Les catégories définies sont liées à la consultation des fichiers et à leur consommation énergétique (comme découvert dans la première partie de ce travail, les catégories *Accroître* et *Partager* sont plus consultées que *Oublier* et *Stockier*).

Dans un deuxième temps, un état des lieux des technologies de sauvegarde numérique (disque dur, disque SSD et disque à quartz) a été réalisé afin

de mettre en évidence leurs caractéristiques. Ces technologies varient quant à leur consommation électrique, durée de vie, entretien, temps d'accès aux informations.

Ainsi en répartissant les données appartenant aux catégories les plus consultées, sur les médiums consommant le moins d'énergie et ayant un temps d'accès à l'information plus rapide, les utilisateurs pourraient influencer l'empreinte écologique des data center. En appliquant cela, les disques durs seraient moins consultés, donc ne tourneraient pas à plein régime et nécessiteraient moins d'énergie électrique.

À partir des classifications de la première partie et des analyses technologiques de la deuxième, une relation entre les différents médiums et les intentions d'utilisations a pu être établi afin de proposer une meilleure régulation de la consommation énergétique des data center. L'idée est d'organiser les données en fonction de leurs catégories et de leurs consultations. Les informations les plus consultées se trouveraient sur des disques SSD, pour tous les avantages qu'ils apportent, et les informations moins consultées sur des disques durs classiques. Toutefois un autre problème apparut à ce moment de la recherche car l'espace terrestre nécessaire à ces data center augmenterait continuellement en fonction du nombre d'utilisateurs et de la quantité de contenu Web.

Les technologies réseaux pourraient permettre de réduire l'espace terrestre qu'émettent les data center. En effet, ils occupent un immense espace au sol (dans la région de Dulles aux États-Unis, c'est près d'un million de mètres carrés occupés). Au début des années 2000 une technologie réseau nommée *peer to peer* se démocratise largement grâce aux logiciels comme Napster, Limewire ou encore e-mule. Cette technologie permettait aux utilisateurs de *partager* et télécharger des données entre eux sans qu'elles ne soient stockées dans un data center. Son fonctionnement est différent du modèle que l'on connaît actuellement, c'est à dire du modèle client/serveur. Dans un système de réseau *peer to peer* tel que Napster, chaque utilisateur,

grâce à son ordinateur remplit le rôle de serveur et utilisateur en même temps. Cela signifie que chaque utilisateur de ces programmes attribuait une partie de la mémoire de leurs disques durs pour héberger les ressources de ces logiciels. Ainsi les ressources de ces logiciels étaient partagées entre tous les utilisateurs sans qu'elles ne soient hébergées par des data center. Cela signifie que lorsqu'un utilisateur faisait une recherche pour télécharger un morceau de musique sur Napster par exemple, il consultait d'autres ordinateurs autour de lui qui possédaient le morceau recherché pour ensuite le télécharger.

Les contenus restaient présents sur le système tant qu'ils restaient consultés ou demandés par les utilisateurs. À partir d'un certain temps si le contenu n'était plus demandé le système l'effaçait localement donc de la machine de l'utilisateur. Une caractéristique des réseaux *peer to peer* est que la qualité des données disponibles augmente à mesure que le nombre d'utilisateurs augmente. La valeur du réseau augmente donc avec sa popularité et ne nécessite pas d'infrastructure particulière puisqu'elle se définit par les utilisateurs.

Il serait concevable d'imaginer une relation alliant les réseaux peer to peer à celle entre les intentions des utilisateurs et les data center pour réduire leurs consommations énergétiques matérielle et immatérielle. Cela permettrait de sensibiliser les utilisateurs quant à la consommation des data center tout en réduisant aussi l'espace terrestre occupé par ceux-ci. Cette approche permettrait d'être plus proche des utilisateurs et de leurs habitudes afin qu'ils puissent avoir un impact écologique et énergétique positif en faisant partie du réseau.



## **Articles en lignes:**

### **IICloud(s) – Inhabiting and Interfacing the Cloud(s)**

[http://www.iiclouds.org/author/nicolas\\_nova](http://www.iiclouds.org/author/nicolas_nova)

consulté le 20 octobre 2018

publié le 12 février 2018

### **Report : Virginia is the ‘world’s largest market’for data centers.**

<https://technical.ly/dc/2018/02/12/report-virginia-leading-world-data-centers/>

consulté le 14 octobre 2018

publié le 12 février 2018

### **Datacenter : il faut changer la forme des disques durs affirme Google**

<https://www.zdnet.fr/actualites/datacenter-il-faut-changer-la-forme-des-disques-durs-affirme-google-39833294.htm>

consulté le 28 octobre 2018

### **Internet World Stat**

<https://www.internetworkstats.com/stats.htm>

consulté le 15 septembre 2018

publié le 12 février 2018

### **Les data center , grands pollueurs de l’ère numérique ?**

<https://www.energystream-wavestone.com/2017/02/data-center-grands-pollueurs-numerique/>

consulté le 28 octobre 2018

publié le 12 février 2018

### **Greenpeace**

<https://www.greenpeace.org/international/publication/6826/clicking-clean-2017/>

consulté le 11 septembre 2018

publié le 12 février 2018

### **The future of storage: HDD or SSD**

<https://www.colocationamerica.com/blog/the-future-of-data-center-storage>

consulté le 28 octobre 2018

publié le 12 février 2018

**Ericsson prévoit 20 milliard d'objets connectés  
dont 1,8 milliard par réseau cellulaire en 2020**  
<http://www.vipress.net/ericsson-prevoit-20-milliards-dobjets-connectes-a-liot-dont-18-milliard-reseau-cellulaire-2023>  
consulté le 11 novembre 2018  
publié le 29 novembre 2017

**IOT : Quel impact sur nos infrastructures ?**  
<http://www.aucoeurdesinfras.fr/iot-impacts-infrastructure/5214>  
consulté le 28 octobre 2018  
publié le 1<sup>er</sup> mai 2017

**VNI Global Fixed and Mobile Internet Traffic Forecasts**  
<https://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html>  
consulté le 23 Août 2018  
publié le 1<sup>er</sup> septembre 2017

**SSD vs HDD**  
[https://www.storagereview.com/ssd\\_vs\\_hdd](https://www.storagereview.com/ssd_vs_hdd)  
consulté le 15 Août 2018  
publié le 26 mars 2018

**Reducing the Storage Footprint & Power Use in Your Data Center**  
<https://www.datacenterknowledge.com/archives/2013/03/11/reducing-the-storage-footprint-and-power-consumption-in-your-data-center>  
consulté le 22/09/18  
publié le 11/03/2017

**Livres :**

**What Do You See in the Cloud? Understanding the Cloud-Based User Experience through Practices**

lien :

[https://link.springer.com/content/pdf/10.1007%2F978-3-642-40480-1\\_47.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-642-40480-1_47.pdf)

**Disks for Data Centers White paper for FAST**  
(2016) Eric Brewer, Lawrence Ying, Lawrence Greenfield, Robert Cypher, and Theodore Ts'o  
Google, Inc. February 23, 2016

lien :

<https://static.googleusercontent.com/media/research.google.com/fr//pubs/archive/44830.pdf>

**Clicking Clean: Who is winning the race to build a green internet**  
(2017) Greenpeace

**What Do You See in the Cloud? Understanding the Cloud-Based User Experience through Practices**(2013) John C. Tang<sup>1</sup>

, Jed R. Brubaker<sup>2,3</sup>, and Catherine C. Marshall<sup>3</sup>  
<sup>1</sup> Microsoft Research, One Microsoft Way,  
Redmond, WA 98008 <sup>2</sup> Department of  
Informatics, University of California Irvine,  
Irvine, CA 92697 <sup>3</sup> Microsoft Research Silicon  
Valley, 1065 La Avenida, Mountain View, CA  
94043  
p680-687

## **Reportages:**

### **Trevor Paglen's Deep Web Dive | Behind the Scenes**

(2016)

<https://www.youtube.com/watch?v=h7guR5ei3oY>

Durée: 14 min

Langue: Anglais

### **Les gardiens de la mémoire**

(2014) Directeur: Quentin Domart

Producteur: Gedeon Programmes

Durée: 54 min

Langue: Français

Pays: France

### **Nos ordinateurs ont-ils la mémoire courte ?**

(2017) Arte

Durée: 52 min

Langue: Français

Pays: France

### **Downloaded**

(2013) Alex Winter

Durée: 80 min

Langue: Anglais

Pays: États-Unis

## **Descriptif d'entretien:**

**Voici le descriptif du guide d'entretiens:**

**1//Utilisation du cloud, centre de stockage**

**1.1 Qui êtes vous ? Comment utilisez vous les services**

**«cloud » ? Est-ce de façon professionnelle, personnelle ?**

**1.2// Quelles raisons vous amènent à utiliser ces services? Quels type de services précisément?[stockage, partage de document, collaboration sur un projet]**

**1.3// A quelle fréquence utilisez vous ces services ? Quels sont les avantages et les inconvénients selon vous ?**

**1.4// Cette utilisation est elle courante dans votre milieu ?**

**1.5// Pensez vous à d'autres pratiques : points de vue différents à adopter de ceux utilisés aujourd'hui? Pourquoi?**

**1.6// Quelles sont vos frustrations ? Le dernier problème rencontré?**

## **2// Compréhension du cloud**

**2.1// Qu'est ce que le « cloud computing » évoque pour vous ? (fonctionnement, organisation)**

**2.2// Vous souciez vous de son fonctionnement ? De ce que cela implique en termes de ressources matérielles et énergétiques?**

**2.3// Lors du choix de ces services quels sont les paramètres que vous prenez en compte ? Prix,**

sécurité, réputation, possibilité d'inter agir et de comprendre l'organisation de celui-ci?

### **3 // Amélioration et différence**

**3.1// Si vous deviez classer des données importantes jusqu'aux données les moins importantes comment les classeriez vous ?**

**3.2// Faites vous la différence entre stocker des photos personnelles et archiver des dossiers servant à vous et à d'autres ? Pourquoi et en quoi cette différence ou non fait sens selon vous ?**

**3.3// Est il concevable pour vous d'avoir un temps d'accès à l'information plus lent si l'information est moins consultée et d'augmenter les temps d'accès aux informations les plus consultées par vous même et par d'autres utilisateurs ?**

## **Remerciements:**

L'auteur remercie chaleureusement Remy Opalinski lors de la réalisation de ce mémoire ainsi que Romain Talou. Un grand merci à Nicolas Nova pour sa reactivité, sa patience, sa précision dans ses remarques.

Merci également à tout le corps enseignant et le corps administratif du master Media Design de la HEAD-Genève ainsi que les assistants du pôle prototypage et bois.

Imprimé au pôle impression de la HEAD-Genève sur du papier lessebo 300g/M<sup>2</sup> et du lessebo 100g/M<sup>2</sup>.

Le texte est composé en GTSuperText et Druke Wide.

Je tiens à remercier également ma maman pour ses corrections.



**CLOUDERS**



**CLOUDERS**