

UNIVERSITE LIBERTE

FACULTE DES SCIENCES INFORMATIQUES



Vérité et liberté

ETUDE COMPARATIVE D'UN SYSTEME CLOUD COMPUTING DANS LES PRODUIT GMAIL, YOUTUBE ET FACEBOOK

Par KONDO FUNDJI Paul

Dirigé par Professeur Blaise FYAMA

Co-directeur : As. Yannick ZAGABE

ANNEE ACADEMIQUE 2016-2017

UNIVERSITE LIBERTE

FACULTE DES SCIENCES INFORMATIQUES



Vérité et liberté

ETUDE COMPARATIVE D'UN SYSTEME CLOUD COMPUTING DANS LES PRODUIT GMAIL, YOUTUBE ET FACEBOOK

Par KONDO FUNDJI Paul

Travail présenté et défendu en vue de
l'obtention du grade d'ingénierie en
Sciences Informatiques

Option : Réseaux et Télécommunication

SEPTEMBRE 2017

DEDICACE

*A tout chercheur passionné de ce sujet ayant le souci de bien vouloir l'approfondir,
l'améliorer ou encore poursuivre les recherches.*

KONDO FUNDJI Paul

IN MEMORJUM

A mes regretté grand père, grand-mère, tante paternelle et oncle : LEMBA PAUL, OKELE RACHEL, ESAKALONGA LOUIS, LOKUTSHU JEAN PAUL, MABOKE ANDRE, LOSELE OSUNGU PASCAL, LOSELE PASCAL PASSY, TORGA LOKOHO, OKITAFUMBA KOYANANA RAPHAEL pour la vie et l'amour que vous avez semés en moi.

A ceux-là qui m'ont donné le gout des études et le sens de la vie.

KONDO FUNDJI Paul

AVANT-PROPOS

L'expression de notre profonde reconnaissance s'adresse à l'Université Liberté de nous avoir reçus dans son installation et formés en qualité d'ingénieur en sciences informatiques.

A notre Directeur, le professeur Docteur BLAISE FYAMA, ainsi qu'au codirecteur, assistant YANNICK ZAGABE, qui en dépit de leurs multiples occupations ont bien voulu accepter la direction de ce travail nous leurs sont reconnaissant, ainsi qu'à tous les professeurs, chef des travaux et assistants qui ont défilé devant nous, envie de nous encadrés durant tout notre parcours à L'Université Liberté nous vous disons merci.

A mes parents EKANANU LEMBA Henriette, DANGA LEMBA Vicky et ANTHO LOWALO sans qui ce modeste travail n'aura pas pu voir le jour.

A mes sœurs et frères, OKELE DIVINE, LEMBA PENASHE Paul, OKITA FUMBA Dode, OMEDIKONGA Nouveau, TAMBASHIE OKITAFUMBA Cezar, LOKUTSHU ESAKALONGO Jean-Paul, DJANDE OKITAFUMBA Bénédicte, TSHIATSHIAWO Claudine.

A mes cousins et cousines, OMADJELA Papy, DJOKO Sophie, DIANA Passy, NONGOTAMBA Junior, FUNDA NAOMIE, FRANCYNE OHAKE, PRISCA OHAKE, BELINDA OHAKE, DHO OHAKE, MARCELINE OHAKE, TERRE OHAKE, SAMI OHAKE, ainsi qu'à tout la famille restreinte.

Sans oublier WITAKENGE JEANNOT et sa famille, TSHINYAMA DECA et sa famille, ELOMBE JULIE et sa famille, KAVULU MATTHIEU et sa famille, MUZINGU GEULORD et sa famille, pour tous leurs soutient matériel et moral.

A mes amis et connaissances, NGOMA DOROTHE, MUKOMA DECO, MBULUNGU CHADRACK, WITAKENGE MARTIN, KABUAYI MPOYI, BUKASA BUCKSON, NGALAMULUME PATTY, KADIMA JACQUE, KALENGA JUSTINE, KAJIBA NATHALIE, MUKENDI TOY, NTUMBA SALOMON, JULLE THAWED, NKOSI GRACE, KUMANDA GRACE,

KILONGOZI HERMESS, TRIBELDO TRIBUNAL, BOPE JCHRIS, MWEPA ARSENE, MULOPO ARSENE, WAKAKA NICKQSIN, KADIATA VICTOR, MUKOMA ERICK, NKUFI SMITH, KAJIBA NATHALIE, HIDRISS MWABILU, MADIA.

Nous n'allons pas terminer sans pour autant penser à tous nos compagnons de lutte qui ont accepté de faire route ensemble avec tant de courage et de labeur notamment, KABAMBA DANIEL, KUPULA DIEUDONNE, KABUYI ISAAC, KYAMBEZI MATTHIEU, MOSENGO PATIENT, MPOYI MIKE, KON VERONICA, KALUBISA JOHN, NTUMBA ELIE, KAPANGU FIDEL, KIMBEMBE CHADRACK, MONGA MARDOCHE, KABUNDA JOSEPH, MWILAMBWE JEAN, CILOMBWE ABRINCE, MOTATIANE JUNOIR, KANDANDA EMILIE ainsi qu'à toute la promotion de grade II réseau et télécommunication.

Que tout le monde trouve ici l'expression de nos sentiments d'indéfectible attachement. A Dieu le père tout puissant pour le souffle de la vie.

KONDO FUNDJI Paul

LISTE DES FIGURES

Figure 1 concept du cloud computing -----	7
Figure 2 services du cloud computing -----	11
Figure 3 cloud privé-----	12
Figure 4 cloud public -----	12
Figure 5 : Cloud community -----	13
Figure 6 cloud hybrid-----	14
Figure 7 Centre des données -----	17
Figure 8 architecture 3 Tiers -----	51
Figure 9 chemin de courrier électronique-----	52
Figure 10 cheminement d'un courrier électronique-----	53
Figure 11 : Système Gmail-----	54
Figure 12 connexion Gmail -----	55
Figure 13 se connecter sur Facebook -----	56
Figure 14 se connecter sur YouTube -----	57

LISTE DES TABLEAUX

Tableau 1 : comparaison San et Nas -----	23
Tableau 2 connexion au compte Gmail -----	54
Tableau 3 : Différentes protocoles-----	58
Tableau 4 : Comparative du point de vue convergence et divergence-----	61
Tableau 5 : Comparative du point de vue capacité de stockage-----	61

LISTE DES SIGLES ET ABREVIATIONS

1. ASP : Application Service Provider
2. API: Application Programming Interface
3. AOL: Amirica Online
4. AONT: All Or Nothing Transform
5. BGP: Border Gateway Protocol
6. CAAS: Communication As Service
7. CRM: Customer Relationship Management
8. CIFS: Common Internet File System
9. CNBC: Consumer News and Business Channel
10. CPU : Central Processeur Unit
11. DAAS : Database As Service
12. DAS: Direct Attached Storage
13. DCCO: Distributed Control Centralized Override
14. DNS: Domain Name System
15. EC2: Elastic Compute Cloud
16. ERP : Enterprise Resource Planning
17. EXIF : Exchangeable Image File Format
18. EO : Exa Octets
19. FAH : Fournisseur d'Applications Hébergées
20. FSC : Fournisseur Services Cloud
21. FTP : File Transfer Protocol
22. FBML: Facebook Markup Language
23. GO: Giga octet
24. GPS: Global Positioning System
25. GWS: Google Web Server
26. HTTPS: Hyper Text Transfer Protocol Secure
27. IBM: International Business Machines
28. IGP: Interior Gateway Protocol
29. IAAS: Infrastructure As a Service
30. IT: Information Technology
31. IP: Internet Protocol

-
32. IMAP: Internet Message Access Protocol
 33. IETF: Internet Engineering Task Force
 34. MPLS: Multi-Protocol Label Switching
 35. MDM: Mobile Device Management
 36. NAS: Network Attached Storage
 37. NAAS: Network As Service
 38. NFS: Network File System
 39. OSPF: Open Shortest Path First
 40. PAAS: Platform As a Service
 41. POP: Post Office Protocol
 42. PC: Personal Computer
 43. PFS: Perfect Forward Secrecy
 44. PHP: Hypertext Preprocessor
 45. PNG: Portable Network Graphics
 46. QOS: Quality Of Service
 47. QUIC: Quick Udp Internet Connection
 48. RAID: Raid Array of Independent Disks
 49. RDP: Remote Desktop Protocol
 50. RAM: Random Access Memory
 51. RTMP: Real Time Messaging Protocol
 52. SAAS: Software As a Service
 53. SAN: Storage Area Network
 54. SSH: Secure Shell
 55. SLA: Service Level Agreement
 56. SSL: Secure Sockets Layer
 57. SMTP: Simple Mail Transfer Protocol
 58. SMS: Short Message Service
 59. SOAP: Simple Object Access Protocol
 60. SEO: Search Engine Optimization
 61. UDP: User Datagram Protocol
 62. TCP: Transmission Control Protocol
 63. TIP: Telecom Ifra Project
 64. TLS: Transport Layer Security
 65. TO: Tera Octet

- 66. VG: Volum Group
- 67. VPN: Virtual Private Network
- 68. WEB: World Wide Web
- 69. XCP: Xen Cloud Platform
- 70. XML: Extensible Markup Language
- 71. XMPP: Extensible Messaging and Presence Protocol

INTRODUCTION GENERALE

Les jours, les mois, les années et le temps passent si vite, les sciences informatiques évoluées dans une compétition extraordinaire, dont chaque scientifique ou chercheur veut faire le mieux possible dans l'évolution de la technologie des sciences informatiques au profit de nos besoins pour une bonne amélioration et efficacité. Vers 1969 l'un des principaux scientifiques du projet ARPANET appelle Leonard Kleinrock déclara : '' Qu'à partir de maintenant, les réseaux informatiques sont encore à leurs balbutiements, mais à mesure qu'ils grandissent et deviennent sophistiqués, nous allons voir probablement la propagation de services informatiques qui devient comme les services d'électricité et de téléphone, il servira toutes les habitations et tous les bureaux à travers le pays ''. La vision de Leonard Kleinrock est de fournir les services de l'informatique comme un service public, qui est aujourd'hui une réalité grâce au cloud computing.

Depuis l'apparition de l'homme sur la terre, il est en perpétuelle recherche du développement. Ce facteur lui a permis de se démarquer des autres êtres vivants et de les apprivoiser. Nous constatons que c'est plus aux cours de ces derniers siècles que l'homme a vraiment connu un développement qui dépasse tout entendement, il a pu effectuer des voyages interplanétaires, il arrive à inventer la bombe atomique, l'automobile, découvert de l'électricité, etc. Toutes ces réalisations d'un être humain n'auraient pas pu être possibles sans l'essor de l'électronique qui a déclenché celui de l'informatique.

Le cloud computing est une méthode qui propose des différents types de services sur Internet, comme une application logicielle livrée sur Internet (par exemple Google Drive). D'où ça peut être une plate-forme de développement comme (Google App Engine) sur lesquelles les utilisateurs peuvent créer leurs applications. Il peut être un système d'exploitation (par exemple Amazon EC2) ou les utilisateurs peuvent créer leur propre plate-forme et logicielle. Tous ces services sont accessibles via la toile. Le cloud computing permet de résoudre les problématiques liées à la rareté des ressources informatiques comme le réseau, CPU, RAM etc. Surtout ces dernières années les demandes ont augmenté de la sorte très élevé.

Parlant du terme informatique qui a vu le jour en 1962 par Philippe Dreyfus, l'homme ne s'est pas mis seulement de tout ce que les systèmes de la technologie existante lui offrent en matière de matériels physiques de stockages. Nous racontons la teinte des logiciels étant donné que l'homme veut même la partie matérielle demeure telle qu'elle eut

et que l'on se servait un plus du logiciel pour optimiser en matière de support de stockage. L'informatique définie comme étant le traitement automatique de l'information est aujourd'hui un domaine indispensable, elle a trouvé sa place dans le monde et est parvenue à le conquérir. Jadis, l'informatique ne se prétend pas être transformatrice du système existant en un nouveau système, mais elle vient juste automatiser les flux d'informations qui existent déjà dans l'ancien système, chose qui apporte un gain en temps et en coûts. Etant dans un monde compétitif, le temps est un facteur très important qu'il faut utiliser à bon escient pour un développement intégral.

Le cloud computing grâce à son model infrastructure as service permet aux personnes qui désire exploiter l'entreprise en ligne de louer de l'espace virtuel machine et le fournir le service à leurs besoins, il fournit les ressources informatiques virtuelles. IaaS apporte l'avantage de "pay-per-use", ce modèle pay-per-use est facilement compris par comparaison d'utilisation d'un courant électrique dont payer la quantité de consommation du courant utilisé. De même IaaS permet aux clients de payer les ressources utilisées, cela permet aux clients de n'est plus investir dans les capitaux dans l'infrastructure informatique et le permet de concentrer sur leur QoS à bien fournir à ses clients au lieu de penser à l'approvisionnement et le sous-dimensionnement des ressources informatiques.

L'informatique comme étant la nouvelle et parmi tant d'autres sciences, elle est devenue la grande science à l'heure actuelle incontournable et permet de traiter des solutions possibles pour l'homme à la société.

Le cloud avec son modèle IaaS appelle cloud stakeholders c.à.d. acteurs et intervenants de cloud sont-ce les fournisseurs et utilisateurs du cloud, ils sont responsables de la mise en place de l'infrastructure informatique et ces utilisateurs utilisent les infrastructures fournies pour effectuer leurs tâches.

De nos temps, les grandes firmes développent des technologies et des centralisations des données avec un transfert énorme de flux d'information qui sollicite des matériels pouvant offrir un bon débit. Nous avons plutôt besoin des bons fournisseurs en matière de qualité des bons équipements et en matière de logiciels. L'administrateur système doit assurer un entretien continu de toutes ces fournitures que ça soit en software ou hardware. Voici donc comment qu'intervient d'autre problème que celui de la complexité de plates-formes des systèmes d'informations. Les meilleures ou les avantages des uns aux autres peuvent nous profiter de plusieurs manières tout en sachant que la concurrence est très utile. Elle nous

conduit à développer une certaine spécialité qui fera en sorte qu'une entreprise devienne très utile et intéressante. Sans ignorer, nous savons que l'informatique étant une science de traitement automatique de données qui s'avère être trop bénéfique et important dans tous les domaines qu'ils soient scientifiques, professionnels, privés et/ou publics.

La plate-forme de gestion du cloud est un logiciel qui convertit l'infrastructure informatique physique existant à l'infrastructure de cloud, le nombre de la plate-forme commerciales de gestion cloud augment beaucoup plus de la popularité de Iaas-cloud, tels que vCloud, CloudForms, etc. des nombreuses contributions ont été faites par la communauté open source associé à Iaas-cloud, certaines plate-forme open source sont : XCP, Nimbus, Eucalyptus, Openstack et OpenNebula. Pour résoudre le problème lié au choix de plate-forme de gestion du cloud, un Framework doit être développée. Le Framework regroupe les fonctionnalités offertes par la plate-forme de gestion de différents types de cloud que ce soit en open source ou commerciale afin de faciliter la comparaison. Ce Framework donne trois avantages à savoir, il permet aux consommateurs cloud de faire le choix de la plate-forme adapté au besoin opérationnels, deuxièmement, les éditeurs de ces logiciels de gestion c.à.d. les entreprises qui vendent leurs plate-forme, par exemple OpenStack, Eucalyptus etc. Qui peut inclure des fonctionnalités supplémentaires à leurs logiciels. Enfin les développeurs de logiciels de gestion cloud c.à.d celui qui commence la création d'un nouveau logiciel de gestion cloud peuvent commencer à construire leurs logiciels en mettent les fonctionnalités répertoires dans le Framework. Openstack et Eucalyptus sont l'un des plates-formes populaires du cloud computing, en particulier le nombre de consommateurs du cloud qui télécharger derniers sont plusieurs milliers par mois.

En faisant l'étude d'observation dans des grandes entreprises du monde aujourd'hui nous, nous rendons compte qu'elles réalisent des travaux des plusieurs besoins dans peu de support et de logiciel de la nouvelle technique que l'informatique apporte dans le Cloud Computing, ainsi ce genre de travail ne s'effectue pas aux hasard ni à la main, mais plutôt par des nouvelles technologies de la science de l'informatique via ces supports et ces logiciels. Par contre la science du traitement automatique de l'information, dans ses prouesses, nous apporte un environnement qui a des technologies, des services et des applications qui s'occupaient de gérer et superviser les fonctionnements sur le champ, donc avec des bonnes parties de ces exigences (matérielle + logicielle) et nous propose juste des fournitures comme : la distribution du courant électrique et le service Internet.

C'est à juste titre que nous avons intitulé notre sujet de recherche : « *Etude comparative d'un système Cloud Computing dans les produits Gmail, YouTube et Facebook* ». Le choix de ce thème de recherche est porté sur une évaluation des connaissances acquises tout au long du second cycle de notre formation universitaire en sciences informatiques, en proposant aux entreprises et aux particuliers comment est-il bon d'embrasser les systèmes des services technologiques Cloud computing dans leurs environnements pour être à la page de la modernisation à la mondialisation qui pourront les aider dans leurs besoins à venir ; mais aussi à faire un choix du point de vue stockage des données et partage des informations dans le Cloud sur Gmail, YouTube et Facebook.

Ce sujet nous a permis d'affiner les connaissances sur les fonctionnements et les réalités du déroulement des processus de contrôle des technologies, et d'approfondir nos connaissances acquises. La plupart des entreprises n'utilisent pas les technologies Cloud Computing comme solution, nous chercherons plutôt à faire en sorte que ça soit basé sur un minimum de matériel, une bonne partie qui voudront conserver leurs libertés en informatique puisse jouir de ces technologies présentes.

Ce travail est un couronnement du long parcours universitaire de formation en tant qu'ingénieur en réseaux et télécommunications à **l'Université Liberté** en vue d'obtention du grade académique, et clôture ainsi le long chemin de souffrance, peine, espoir et bonheur d'études commencé pendant un certain moment d'une vie académique pour l'étudiant ou étudiante. Nous voudrions par ce travail de permettre aux futurs chercheurs intéressés par cette question des circonstances digne pour leurs investigations et permettra d'ajouter un plus aux sciences informatiques en vue de l'évolution de la technologie de cette discipline, il nous permet aussi de nous conformer aux exigences d'une formation universitaire en **République Démocratique du Congo**.

Les trois produits Gmail, YouTube et Facebook sont des services les plus populaires. En particulier, le nombre des utilisateurs s'ajoute du jour au jour par mois. Certains des travaux existants se sont focalisés sur la performance de ces plates-formes (OpenStack et Eucalyptus) ou encore la solution déployée de messagerie Exchange basée sur le stockage de cloud, dont l'impact sur la solution cloud au sein de l'entreprise. Cependant, il y a peu d'attention sur la comparaison de ces services.

Du fait que notre étude soit basée sur les fonctionnements de la manière d'accéder aux différentes plateformes à savoir Gmail, YouTube, Facebook et sur la façon dont sont stocké des données.

Les technologies d'internet Gmail, YouTube, Facebook utiliseraient une architecture 3 tiers c'est-à-dire client/serveur. Dans le cloud computing les données de Gmail et YouTube reposeraient sur Google Drive et pour le cas de Facebook ses applications tourneraient sur sa propre infrastructure cloud. Ainsi par cette magie de stockage de la technologie Cloud computing, les données seraient stockées par les fournisseurs qui offrent en plus une gestion sécurisée et confidentielle de nos données.

Ce sujet d'étude, nous pousse au terme de l'évolution de prendre la méthode d'analyse de fonctionnement de différentes technologies. Nous utiliserons la méthode UML pour représenter le fonctionnement logique des différentes technologies.

Ce mémoire porte sur un grand objectif. Qui est celui de montrer les différents fonctionnements sur les technologies Gmail, YouTube et Facebook. A cet exercice, nous pensons pouvoir mettre à la disposition des autres chercheurs un document qui leur servirait d'enrichir leurs connaissances en rapport avec les technologies Internet.

Le reste de ce mémoire est organisé comme suit :

- Chapitre 1 : Concepts et terminologie du cloud computing,
- Chapitre 2 : Les technologies Internet dans la messagerie électronique et les réseaux sociaux,
- Chapitre 3 : Etude fonctionnelle et comparative sur Gmail, YouTube et Facebook.

A la fin nous résumerons notre mémoire avec une conclusion générale.

CHAPITRE I : CONCEPTS ET TERMINOLOGIE DU CLOUD COMPUTING

INTRODUCTION

La technologie sur l'internet s'est développée plus rapidement, au cours de cette dernière année, il y'a eu une nouvelle destination et ayant pour but d'améliorer les services dans le domaine de la technologie de l'information et de la communication. Il s'agit du « *cloud computing* ». Ce dernier est un concept qui regorge plusieurs technologies dans sa diversité qui sert à délivrer différents services. Il peut être schématisé par un ensemble indéterminé des ressources informatiques interconnectées et une coopération dans un réseau bien défini.

L'accès aux services se fait sur demande via une application standard facilement disponible à savoir « *un navigateur Web* ».

Ainsi, dans ce chapitre, nous allons présenter les notions et les concepts fondamentaux du Cloud Computing.

1.1. HISTORIQUE

La faible performance des processeurs : capacités de stockage et d'autres ressources employées dans les systèmes classiques sont limitées et coûteuses ; d'où ils ne permettent pas l'accès à des services et des ressources de façon rapide. Avec le développement actuel de la technologie, cela a ouvert la voie à l'émergence du concept de cloud computing, qui permet donc l'utilisation des serveurs plus efficaces et moins coûteux, augmentant la performance des connexions Internet lorsqu'ils deviennent plus rapides à la demande des utilisateurs. La notion de ce service a été proposée pour la première fois en 1961 lors de la conférence au Massachusetts Institute of Technology par John McCarthy. Il est également l'inventeur du langage Lisp dans lequel un groupe d'ordinateurs se relie pour former un ordinateur virtuel unique permettant ainsi le calcul de haute performance.

Salesforce.com fut le premier hébergeur de cloud en 1999, suivi en 2002 par Amazon qui proposa un ensemble d'hébergements d'applications et de stockages. Amazon va développer ses services en 2005 (Amazon Web Service) et en 2006 (Elastic Computer Cloud ou EC2). En 2007, Google, IBM et des universités lancèrent un projet de recherche sur le cloud qui permet de lui faire gagner en popularité et en consistance. C'est vers les années 2009 que surgit la réelle explosion du cloud avec l'arrivée sur le marché de sociétés comme Google (Google App Engine), Microsoft (Microsoft Azure), IBM (IBM Smart Business Service), Sun (Sun Cloud) et Canonical Ltd (Ubuntu Enterprise Cloud).

1.2. DEFINITION

Le terme cloud désigne des ressources informatiques qui sont mises en commun, puis redistribuées en fonction des besoins des utilisateurs. National Institute of Standards and Technology en français « *Institut national des normes de la technologie* », il définit le cloud computing comme « *l'ensemble des disciplines, pratiques, technologies et modèles commerciaux utilisés pour livrer un service sur demande et le réseau des capacités informatiques (logiciels, plateformes et matériels)* ».

Les définitions de cloud computing choisies dans d'autres littératures :

- **Le cloud computing** : est un concept qui consiste à déporter sur des serveurs distants des traitements informatiques traditionnellement localisés sur le poste client de l'utilisateur ;

- **Génération NT** : « *le cloud est un concept d'organisation informatique qui place Internet au cœur de l'activité des entreprises, il permet d'utiliser des ressources matérielles distantes pour créer des services accessibles en ligne* » ;

- Le cloud est un modèle qui offre aux utilisateurs du réseau un accès à la demande, à un ensemble de ressources informatiques partagées et configurables, aussi avec le cloud n'importe qui peut entrer et à partir de n'importe quel endroit et à tout moment ;

- C'est un concept abstrait, son but est de pousser les entreprises à externaliser les ressources numériques qu'elles stockent. Ces ressources offrant des capacités de stockage et de calcul, des logiciels de gestion de messagerie, et d'autres services ont mises à disposition par des sociétés tierces et accessibles, grâce à un système d'identification via un et une connexion à Internet.

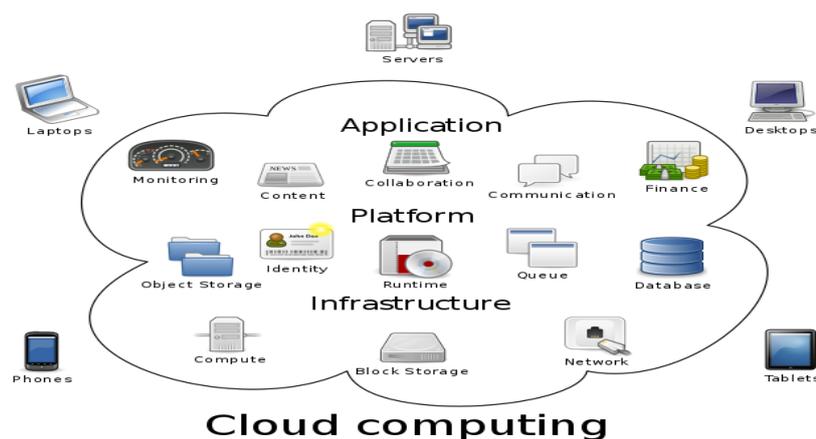


Figure 1 concept du cloud computing

1.3. CONCEPTS DE BASE DU CLOUD COMPUTING

Il y a plusieurs services et modèles qui fonctionnent en coulisse et qui rendent le cloud possible et accessible aux utilisateurs finaux. Ainsi les modèles des travaux pour le cloud computing sont :

- Modèles de services ; infrastructure as service (IaaS), plateforme as service (PaaS) et software as service (SaaS).
- Modèles de déploiement ; cloud public, cloud privé, cloud communautaire et cloud hybride.

1.3.1. MODELES DES SERVICES DU CLOUD COMPUTING

Il existe trois types des services principaux :

A. *Infrastructure As A Service (IaaS)*

En français ‘ l’infrastructure en tant que service’, c’est la mise à disposition par l’internet de machines virtuelles aux ressources « facilement » modifiables à la hausse et hautement disponible. L’entreprise loue ainsi des capacités de traitement, de stockage et autres ressources qu’elle peut structurer et gérer de façon autonome cote logiciel des système d’exploitation.

Pour des calculs complexes ou ponctuel (simulations budgétaires, calculs prédictifs...) ou encore du stockage (notamment dans le cadre de sauvegardes), les administrateurs de ressources informatiques trouvent dans le cloud des environnements aux capacités quasi illimitées et dont la mise en œuvre est quasi instantanée.

- Avantages :
 - Offre une grande flexibilité ;
 - Eviter d’achat et la gestion de matériel ;
 - La possibilité d’installer tous types de logiciel ;
 - Contrôle total des systèmes (administration à distance par SSH ou RDP).
- Inconvénients :
 - L’application du modèle du cloud computing pour les infrastructures (ressources informatiques, serveurs, capacité de traitement, stockage...) implique des changements d’organisation.
 - La migration d’infrastructures existantes vers des Infrastructures as a Service (IaaS) peut poser des problèmes techniques délicats de compatibilité : la compatibilité des services de stockage, la sécurité des machines virtuelles, la sécurité des données, etc.

Parmi les prestataires d'IaaS, nous citons : Amazon avec EC2 qui est le roi de ce genre d'infrastructures, Gogrid et Terremark.

B. Platform As A Service (PaaS)

En français ‘‘ plate-forme en tant que service’’, facturée à la consommation, est un environnement qui permet à l'entreprise de déployer ses propres applications en dehors de sa salle informatique. L'entreprise loué un environnement middleware à l'infrastructure masquée.

L'utilisateur gérer, mais ne contrôle pas l'infrastructure cloud (réseaux, serveurs, systèmes d'exploitation, stockage). Il a ainsi le contrôle sur les applications déployées et la possibilité de configurer l'environnement d'hébergement applicatif.

- Avantages :
 - Le déploiement est automatisé ;
 - Facilités à gérer des développements des applications ;
 - Ne nécessite pas une infrastructure.
- Inconvénients :
 - L'utilisateur dépend fortement de l'infrastructure et du logiciel du fournisseur. Par conséquent, il peut se retrouver lié à un vendeur particulier sans possibilité d'en changer. Par ailleurs, les éventuelles pannes ou les changements de roadmap de développement d'un fournisseur peuvent compromettre les projets des utilisateurs. Si un fournisseur cesse de prendre en charge un langage de programmation particulier, les utilisateurs seront eux aussi forcés de changer de langage, ou de faire appel à un autre fournisseur. Dans les deux cas, il s'agit d'un regrettable contretemps.

Il y a plusieurs fournisseurs des services PaaS, les principaux sont : Microsoft avec Azure, Google avec Google App Engine et Orange Business Service.

C. Software as a service (SaaS)

En français ‘‘ l'application en tant que service’’, est accessible à toutes les entreprises et, il facturé au nombre d'utilisateurs. L'entreprise loué les applications du fournisseur de services. Plus besoin d'acheter un logiciel. Ces applications sont accessibles via différentes interfaces, navigateurs web, clients légers.

De nombreux logiciels sont disponibles dans le cloud en mode SaaS.

- ❖ Messagerie : l'e-mail est certainement l'application la plus utilisée en mode cloud.
 - ❖ CRM : la gestion de la relation client est l'autre application phare du cloud.
 - ❖ ERP : certains logiciels de gestion intégrée sont en mode SaaS.
 - ❖ Collaboratif : les outils de collaboration (partage de documents, réseaux sociaux...) se fonctionnent bien au mode SaaS.
- Avantages :
 - Pas d'installation ;
 - Pas de mise à jour ;
 - Plus de License ;
 - Paiement à l'usage ;
 - Facilité de faire le test de nouveaux logiciels.

Les prestataires de solutions SaaS les plus connus sont : Google avec Google Drive et Microsoft avec Office 365.

- Inconvénients suivants :
 - L'entreprise perd son indépendance informatique. En effet, le prestataire qui lui fournit le service possède les applications et l'hébergement des données de l'entreprise ;
 - Un autre inconvénient majeur est que le changement de prestataire est souvent difficile car il met en cause le fonctionnement de l'entreprise et nécessite le transfert de la base de données. Le retour à un système interne est quant à lui long et coûteux ;
 - Enfin contrairement à ce qu'on peut penser, la multiplication des applications sur les postes associés à la durée du contrat (2 à 3 ans en général) peut être coûteuse, le fonctionnement en mode SaaS peut alors devenir nettement plus coûteux que l'acquisition de l'application.

Nous résumons les trois services principaux du cloud computing avec la figure suivante :

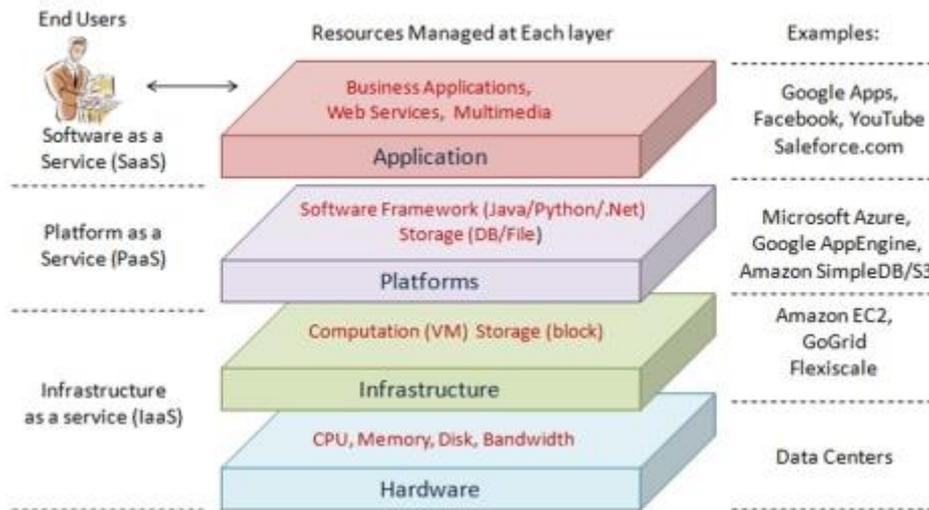


Figure 2 services du cloud computing

Aussi bien, il existe d'autres formes de service du cloud computing à savoir :

- Database as service (DaaS) en français "Base de données en tant que service" offre des fonctions traditionnelles de bases de données, typiquement définition des données, le stockage et la récupération sur une base d'abonnement sur web ;
- Communication as service (CaaS) en français "communication en tant que service", est un service de communication audio et vidéo, services collaboratifs, communication unifiées, messagerie électronique, messagerie instantanée partage de donnée (web conférence) ;
- Network as service (NaaS) en français "réseau comme un service" fournir à l'utilisateur la possibilité d'utiliser le réseau, les services de connectivité de transport et inter-cloud services de connectivité réseau. Parmi des modèles de NaaS, le modèle VPN "Réseau Privé Virtuel", il permet à un ordinateur hôte d'envoyer et recevoir des données à travers des réseaux partagés ou publics, comme s'il s'agissait d'un réseau privé avec toutes les fonctionnalités et politiques de ce réseau.

1.3.2. MODELES DE DEPLOIEMENT DU CLOUD COMPUTING

1.3.2.1. CLOUD PRIVE

En anglais "private cloud" est un ensemble des services et des ressources disponibles à un seul client. Il est divisé en deux cloud, cloud privé Interne qui est hébergé par l'entreprise elle-même, ou bien avec sa branches, parfois partagé ou mutualisé en mode privatif avec les filiales ; cloud privé Externe qui est donc hébergé chez un tiers, il est entièrement dédié à l'entreprise et accessible via des réseaux sécurisés de types VPN.

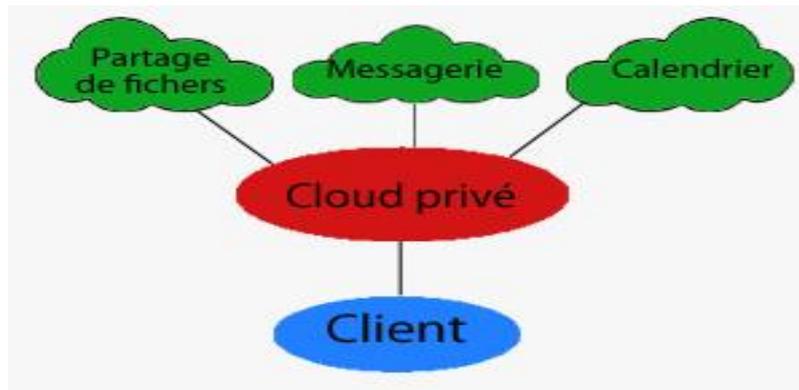


Figure 3 cloud privé

1.3.2.2. CLOUD PUBLIC

En anglais “ public cloud” est un ensemble des services et des ressources accessibles par Internet et géré par un prestataire externe, ces ressources et ces services sont partagés entre plusieurs clients qui utilisent sur demande et à tout moment sans savoir où elles existent, ainsi ces services peuvent être gratuits ou payants. Dans cette dernière possibilité, il existe des contrats Service Level Agreement (SLA) entre les clients et les fournisseurs, SLA est un document qui définit bien la qualité de service requis entre les deux. Exemple de cloud public : Windows Azure Service Platform, Sun Cloud, IBM’s Bleu Cloud, Google App Engine, Amazon Elastic Compute Coud et Amazon S3.



Figure 4 cloud public

1.3.2.3. CLOUD COMMUNAUTAIRE

En anglais ‘‘ Cloud Community’’ est un ensemble des infrastructures qui permet à plusieurs clients ou organisations de partager les différentes ressources qui ont des besoins communs, ces ressources sont généralement spéciales à des organisations, ce type de cloud peut être géré par ces organisation elles-mêmes ou par des fournisseurs externes. Ainsi il permet à un groupe d’utilisateurs de créer leur propre Cloud avec des caractéristiques de Cloud privé tel que la sécurité, ressources dédiées et un cout réduit. Le cloud communautaire peut héberger une application métier très spécialisée, mais commune à plusieurs entités qui décident de fédérer et la gérer, ce cloud est donc plutôt dédié à une communauté professionnelle spécifique incluant partenaires, sous-traitants...pour travailler de manière collaborative sur un même projet ou cloud gouvernemental dédié aux institutions étatiques.



Figure 5 : Cloud community

1.3.2.4. CLOUD HYBRIDE

En anglais ‘‘ Hybrid Cloud’’ est une composition de deux ou trois types de cloud (Privé, Communautaire et Public). Ces infrastructures sont liées entre elles par la même technologie qui autorise la portabilité des applications et des données, exemple l’utilisation des applications dans un cloud public mais ces derniers nécessitent des données stockées sur un cloud privé.

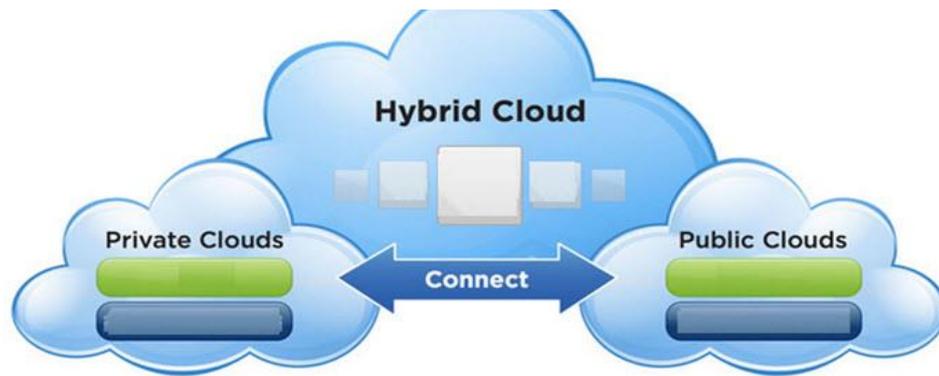


Figure 6 cloud hybrid

1.4. CARACTERISTIQUE DU CLOUD COMPUTING

- **Elasticité** : il a des capacités illimitées qui peuvent être augmentées ou réduites selon l'usage. Par exemple lorsque la personne effectue ses achats en grande surface, elle choisit non seulement la qualité des produits qu'elle souhaite mais aussi la quantité. Cette dernière peut donc varier d'une semaine à l'autre en fonction de besoins. Il est possible de faire l'analogie de cloud non seulement sur le choix des ressources à consommer par le client, mais aussi en déterminant la quantité. Ce qui justifie donc la variation de celle-ci peut à tout moment.
- **Self-service** : la capacité à fournir une ressource informatique automatiquement sans recourir à l'interaction humaine comme fournisseur. Dans un supermarché, les clients opèrent leurs choix sur les produits qu'ils souhaitent sans l'aide du commercial. Ce principe se retrouve également dans le cloud dont les utilisateurs ont accès à un catalogue de service choisi et qu'il souhaite consommer.
- **Pay-as-you-use** : les clients cloud recevront des tarifications les plus précises pour les ressources qu'ils utilisent. Par exemple la tarification d'Amazon pour les micro machines virtuelle 0.02\$ par heure et 3.10\$ par heure pour des grosses machines virtuelles.
- **Mesure de la qualité de services** : évaluer et garantir un niveau de performance et de disponibilité adaptées aux besoins spécifiques des clients.
- **Accès réseau universel** : l'accès aux ressources est très rapide à travers un réseau, par des protocoles standards de façon très élastique.
- **Mise en commun de ressources** : Datacenter fournissant les ressources (machines, stockage etc...) pour les différents clients en mode partagé.
- **Multi-tenancy** : la capacité de fournir un service simultanément à plusieurs clients, cela permet d'augmenter l'utilisation des ressources IT déployée.

- **Disponibilité** : la haute disponibilité de la plate-forme est obtenue par redondance et par la capacité de se remettre rapidement en cas des problèmes.
- **Green IT** : est un concept qui décrit l'ensemble des nouvelles technologies à faible impact environnemental. Le cloud computing, basé sur la virtualisation, qu'il réduise les risques des machines physiques sur un environnement et sa consommation énergétique.

1.5. ELEMENTS CONSTITUTIFS DU CLOUD COMPUTING

1.5.1. INFRASTRUCTURE

L'infrastructure du cloud est un assemblage de serveurs, d'espace de stockage et de composants réseau organisés de façon à permettre une croissance incrémentale supérieure à celle que l'on obtient avec les infrastructures classiques. Ces composants doivent être sélectionnés pour leur capacité à répondre aux exigences d'extensibilité, d'efficacité, de robustesse et de la sécurité. Les serveurs d'entreprise classiques ne disposent pas des capacités réseau, de la fiabilité ni des autres qualités possibles pour satisfaire efficacement et manière sécurisée les accords de niveau de service SLA (Service Level Agreement). Par ailleurs, les serveurs d'un cloud affichent des coûts de fonctionnement moins élevés et ils peuvent être plus fiables s'ils ne sont pas tous équipés de disques internes.

1.5.2. VIRTUALISATION

La virtualisation est la principale technologie dans le cloud, est un procédé pour partitionner une ressource physique en plusieurs ressources virtuelles, exemple un serveur, espace de stockage ou un réseau lors de la création des machines virtuelles. Elle permet d'intégrer les différents serveurs de façons plus flexible pour faciliter l'utilisation. Elle a pour but de faire la transparence d'utilisation et l'efficacité d'exploitation des ressources, d'assurer le fonctionnement des différents services et la séparation entre de multiples locataires (utilisateurs) impliqués dans un matériel physique.

1.5.3. INTERFACE DE SERVICE

L'interface de service placée entre le fournisseur et le client est un élément de différenciation du cloud. Elle représente un contrat qui fait respecter la proposition de valeur décrite par des SLA et des conditions tarifaires. Si le cloud semble nouveau, c'est principalement en raison de cette interface. Elle représente la valeur d'un fournisseur et sert de base à la concurrence. Par l'ajout d'interfaces de libre-service, nous obtenons d'autres optimisations. Les clients du cloud sont en mesure d'engranger des ressources de manière automatisée sans que le service informatique soit un obstacle. L'espace de stockage et les ressources sont présentés au travers d'une interface graphique que l'utilisateur peut

manipuler de manière à obtenir et à instancier une infrastructure informatique virtuelle. Un navigateur web ou une carte bancaire à titre d'exemple.

1.5.4. DATACENTER (CENTRE DE DONNEES)

Un centre de traitement de données en anglais "datacenter" est un site physique sur lequel sont regroupés des équipements constituant le système d'information de l'entreprise (baies de stockage, mainframes, serveurs, équipements réseaux et de télécommunications, etc.). Il peut être interne ou externe à l'entreprise, exploité ou non avec le soutien de prestataires. Il comprend en général un contrôle sur l'environnement (climatisation, système de prévention de contre l'incendie, etc.), une alimentation d'urgence et redondante, ainsi qu'une sécurité physique élevée. Cette infrastructure peut être propre à une entreprise et utilisée pour elle-même ou à des fins commerciales. Ainsi, des particuliers ou des entreprises peuvent venir y stocker leurs données suivant des modalités bien définies.

- Un Datacenter est constitué de trois composants élémentaires :
 - Infrastructure : est constituée de l'espace et des équipements nécessaires aux supports des opérations du Datacenter. Il comprend les transformateurs électriques, alimentations sans interruption, générateurs, armoires de climatisation, système des distributions électrique, etc.
 - Equipements informatique : ensemble d'équipement de base de la puissance du traitement de calcul d'information. Ils sont constitués des racks, serveurs de stockage, câble, outils de gestion de système et équipements réseaux.
 - Opération : est constituée des personnels d'exploitation qui pilotent, entretiennent et réparent les systèmes informatiques et non-informatiques lorsque cela est nécessaire.
- Caractéristique d'un Datacenter :
 - Disponibilité électrique : fournir de l'électricité en quantité suffisante, de haute qualité ainsi que l'électricité de secours.
 - Système de refroidissement : il doit répondre à trois exigences principales dont refroidissement des équipements, de baies ainsi que l'implantation des baies dans la salle.
 - Equipement informatiques : il est là pour traiter, stocker et router des informations sur des serveurs d'applications, serveurs de stockages et éléments réseaux. Brève explication sur les serveurs d'applications, et celle les machines sur lesquelles les logiciels s'exécutent.
- Les serveurs de stockage, repartie en deux selon la technique utilisée :

- Stockage local : l'information est stockée sur le serveur qui l'utilise, le DAS (direct attached storage).
- Stockage dédié : on dédie à certaines machines une fonction de stockage, constituant ainsi une « baie de stockage ».
- Fonctionnement d'un Datacenter répond à 3 grandes missions :
 - Stocker de données
 - Traiter de données
 - Sauvegarder et protéger de données



Figure 7 Centre des données

1.6. DIFFERENCE ENTRE VIRTUALISATION ET CLOUD COMPUTING

- Virtualisation permet aux entreprises d'utiliser une seule pièce matérielle physique pour effectuer le travail d'un grand nombre de machines. Plusieurs instances d'un système d'exploitation s'exécutant sur un seul dispositif matériel, sont beaucoup plus économiques qu'une pièce de matériel pour chaque tâche de serveur.
- Cloud computing est l'accès par Internet aux applications d'entreprise fonctionnant dans un environnement non-local. Le cloud computing peut certainement tirer des avantages de la virtualisation.

1.7. DIFFERENCE ENTRE GRID ET CLOUD COMPUTING

Le grid ressemble au cloud computing dans le grand nombre de serveurs, communication réseau très importante, stockage généralement distribué, etc. ces deux techniques sont assez différentes.

- Dans le cas du grid, il s'agit surtout de mettre à disposition de quelques équipements des ressources informatiques très puissantes pour des périodes de temps données. Par exemple le projet de Grid 5000 en France, qu'il est utilisé pour faire des calculs parallèles sur de très gros volumes de données ;

- Dans le cas du cloud computing on note souvent une notion d'élasticité, de disponibilité et de virtualisation de ressource, on ne sait trop sur quels serveurs fonctionne les applications ainsi on paye en fonction de l'utilisation.

1.8. AVANTAGE DU CLOUD COMPUTING

- Possibilité de partager ses contenus avec son réseau personnel ;
- Possibilité de partager les ressources ;
- Redondance de stockage de données en cas d'un disque dur tombe en panne, les données restent en sécurité et disponibles automatiquement sur un autre disque ;
- Possibilité de spécifier des ressources ;
- Manipuler et configurer l'application en ligne à tout moment ;
- Aucune défaillance c.à.d. qu'il n'y a ni de problèmes de matériel, ni de logiciel de la technologie ;
- Sécurité dans cloud ou chaque utilisateur peut être isolé dans son propre environnement ;
- Pas besoin d'installer un logiciel spécifique pour avoir accès ou manipuler les applications cloud ;
- Plus flexible ;
- Offre le développement en ligne et des outils de déploiement, de l'environnement d'exécution de la programmation par la plate-forme comme modèle de service ;
- Augmente la collaboration ;
- Offre un self-service sur demande, les ressources peuvent être utilisées sans interaction avec le fournisseur de services de cloud ;
- Mobilité, l'utilisateur peut à tout moment et à partir de n'importe quel appareil se connecter à ses applications (Any where-any time) ;
- Les ressources sont disponibles et accessible sur le réseau indépendamment de la plate-forme utilisé par les clients ;
- Contrôle de couts ;
- Il est très rentable car il fonctionne de façon plus efficace avec une plus grande utilisation, il faut juste une connexion Internet ;
- Offre un équilibrage de charge qui le rend plus fiable ;
- Accéder sur Internet à des applications comme les services utilitaires.

1.9. INCONVENIENTS DU CLOUD COMPUTING

- L'accès au cloud passe nécessairement par une connexion internet, par conséquent si votre connexion est interrompue momentanément ou durablement, vous n'avez plus accès à vos documents ou service ;
- Sécurité, la plateforme cloud, si elle est externe (non installée sur le réseau interne ou avec une ouverture extérieure) doit être suffisamment sécurisée pour éviter le risque d'intrusion, de vol des données par piratage.
- Sécurité et vie privée, la plus grande préoccupation étant donné que la gestion des données et gestion des infrastructures en nuage sont fournies par des tiers, il y a toujours un risque de transfert de l'information sensible à ces fournisseurs, bien que les vendeurs cloud s'assurent de bien protéger les comptes de leurs clients avec des mots de passe crypté et surs, tout signe brèche de sécurité se traduirait par le perte des données sensibles des clients et l'atteinte à la réussite des entreprises ;
- Piégé, il est très difficile pour les clients de changer d'un fournisseur de services cloud(FSC) à l'autre. Il en résulté une dépendance avec un FSC en particulier pour le restant de vos activités ;
- Panne d'isolation, ce risque concerne la défaillance du mécanisme d'isolement qui sépare le stockage, la mémoire, le routage entre les différents locataires ;
- Interface de gestion compromise, en cas de fournisseur de cloud public, les interfaces de gestion de la clientèle sont accessibles par Internet ;
- Suppression des données précaire ou incomplète, il est possible que les données marquées pour la suppression ne puissent pas être effacées, cela risque d'arriver à cause de l'une des raisons suivantes :
 - Les exemplaires supplémentaires de données sont stockés mais ne sont pas disponibles ;
 - Un disque endommagé peut également stocker les données provenant d'autres clients.

1.10. SECURITE DANS LE CLOUD COMPUTING

La sécurité et la conformité sont souvent citées comme le rétenteur principal à l'utilisation des services cloud, qu'elles sont les principales préoccupations des informaticiens pour développer cette technologie. La sécurité permet de garantir la confidentialité, l'intégrité, l'authenticité et la disponibilité des informations.

1.10.1. CONFIDENTIALITE

La confidentialité assure que les données d'un client ne soient accessibles que par les entités autorisées. Les différentes solutions de cloud computing comportent des mécanismes de confidentialité comme la gestion des identités et des accès, l'isolation ou le cryptage. La majorité des échanges internes ou externes au cloud sont encapsulés en SSL (Secure Sockets Layer) et authentifiés avec un certificat rédigé et signé par client. Tant que ce certificat assure le contrôle de clé privée de sa cliente, ce mécanisme permet d'augmenter le degré d'assurance et la confiance parce qu'elle est spéciale à ces clients autorisés.

1.10.2. INTEGRALITE

Les utilisateurs de cloud peuvent stocker plusieurs données, et généralement doivent être protégés contre le vol et les modifications non autorisées. Dans le nuage, il existe plusieurs systèmes qui fournissent des mécanismes pour protéger l'intégrité des informations. Par exemple pour le service de stockage de Windows azure, intégrité est définie par les applications utilisant le modèle de contrôle d'accès. Chaque compte de stockage a deux clés qui sont utilisées pour contrôler l'accès à toutes les données dans ce compte de stockage.

1.10.3. DISPONIBILITE

L'un des principaux avantages fournis par des plates-formes de cloud computing est la disponibilité robuste basée sur la redondance réalisée avec des technologies de virtualisation. Par exemple Windows azure offre de nombreux niveaux de redondance fournissant une disponibilité maximale des données et des applications, ou les données sont répliquées sur trois nœuds distincts pour minimiser les problèmes des pannes matérielles.

1.11. TECHNOLOGIE DE STOCKAGE EN RESEAU

1.11.1. STOCKAGE SAN

Le SAN (Storage Area Network) est une technologie de stockage en réseau. C'est un réseau physique en fibre optique, il connecte l'ensemble des unités de stockages et des serveurs. Dans ce réseau les données stockées sont routées et structurées via des commutateurs.

Cette technologie est basée sur Fibre Channel, qui autorise le transfert de données entre périphériques sans surcharger les serveurs. Cette manière de stockage permet :

- Des partages des ressources de stockage entre plusieurs des utilisateurs,
- L'accès à des informations très rapidement,
- La résolution de problèmes de connectivité entre plusieurs serveurs et unités de stockage.

Le SAN est caractérisé par deux propriétés principales suivantes :

- Utilise un réseau Fibre Channel, où le Fibre Channel est un protocole utilisé par des mainframes,
- Permet l'installation distante de sous-systèmes de disques durs et de bibliothèques de bandes.

Parmi les topologies de SAN :

- 1) **Point à Point** : est la plus simple, est un lien croisé entre les deux ports.
- 2) **Boucle Arbitrée (FC-AL)** : est constitué au maximum de 126 ports reliés entre eux, tous ces ports se partagent la largeur de bande et utilisent un concentrateur FC.

L'architecture de SAN se compose par des équipements principaux :

- **Baie de disques** : est un ensemble des disques pilotés par un ou des contrôleurs qui seront regroupés en VG (Volum Group) grâce à un raid.
- **Raid** : est signifié (Raid Array of Independant Disks), est une technologie dans le

SAN, permettant de constituer une unité de stockage à partir de plusieurs disques regroupés de différents niveaux de raid.

- **Switch** ou bien commutateur.

a. Avantage du SAN

- La consolidation des informations au sein d'un réseau de stockage centralisé,
- La simplification des sauvegardes et de la restauration.
- Relie de l'ensemble des ressources de stockage,
- Performances,
- Fourni grand volume de capacités de stockage,

b. Inconvénients du SAN

- Coûteuse : par exemple le prix d'entrée environ de 283 000 € pour un réseau SAN de taille moyenne,
- Manque de standardisation génère certains problèmes d'interopérabilité,
- La gestion des composants est plus difficile en fonction de l'augmentation en taille et en complexité du réseau SAN.
- La complexité du réseau SAN et son expansion augmentent la difficulté de gestion des composants de ce réseau.

1.11.2. STOCKAGE NAS

Le NAS (Network Attached Storage) : est un serveur qui fournit une interface au niveau du fichier pour un stockage pouvant être partagé sur plusieurs systèmes, ce serveur est attaché directement au réseau IP, donc il est accessible pour tous les équipements de ce réseau, généralement le serveur NAS est plus lent que le réseau SAN ou le stockage DAS.

Il utilise un protocole de transport de fichier comme NFS, CIFS et FTP. Il peut accepter de consolider des données malgré la différence des systèmes d'exploitation.

Le NAS se distingue par les caractéristiques suivantes :

- La connexion directe au réseau et permet l'accès des clients aux données sans arriver le serveur d'applications, ...
- Il est basé sur le partage de fichiers dans des réseaux hétérogènes par exemple WindowsNT, Unix, Windows 2000,
- Facile à déployer et à gérer.

a. *Avantage du NAS*

- Facile à mis en œuvre,
- Partage les fichiers facilement,
- Partage multi-environnement lié aux différentes implémentations du protocole que l'on utilise.

b. *Inconvénients du NAS*

- Demande des ressources comme CPU,
- N'est pas valide, en cas des applications qu'ils nécessitent des grosses performances des disques durs.

1.11.3. STOCKAGE DE DAS

- Le DAS (**D**irect **A**ttached Storage) : où chaque disque ou ensemble de disques est physiquement lié à une seule unité de commande, serveur de fichiers ou mainframe.
- Ces méthodes consistent à connecter, directement ou via un bus d'E/S, une ressource de stockage unique à un hôte dédié, tel qu'un serveur ou une station de travail.

Le tableau suivant résume une comparaison entre les deux types de stockages SAN et NAS :

SAN	NAS
Orienté paquets SCSI	Orienté fichier
Basé sur le protocole Fibre Channel	Basé sur le protocole Ethernet
Le stockage est isolé et protégé de l'accès client général	Conçu spécifiquement pour un accès client général
Support des applications serveur avec haut niveau de performances SCSI	Support des applications client dans un environnement NFS/CIFS hétérogène
Le déploiement est souvent complexe	Peut être installé rapidement et facilement

Tableau 1 : comparaison San et Nas

1.12. ARCHITECTURE DU CLOUD COMPUTING

Le cloud est classé en fonction de son emplacement par son déploiement dans son environnement qui le distingue, peu importe le modèle du cloud utilisé par une organisation, ils restent similaires au niveau des éléments qui composent l'architecture, les architectures cloud contiennent une série d'éléments communs qui est :

- Couche de virtualisation : la virtualisation des serveurs et des systèmes de stockage joue un grand rôle dans l'architecture de cloud en offrant l'un des principaux avantages du cloud l'agilité, cette couche de virtualisation permet aux fournisseurs d'autoriser ou de bloquer l'accès aux serveurs cloud, en fonction des besoins des utilisateurs du service ;
- Le stockage évolutif horizontal : l'évolutivité est une autre marque de distinction de l'architecture de cloud computing, en matière de stockage, elle résulte également de technologies exploitant des ensembles importants de composants matériels standard qui peuvent être facilement développés sur demande à moindres frais au fur et à mesure de l'évolution des besoins en matière d'infrastructure et de stockage ;
- Mécanismes de prise en charge d'utilisateurs multiples : un service cloud doit permettre la séparation physique ou virtuelle des données stockées pour chaque

utilisateur et être capable de retracer le suivi de l'utilisation du service par utilisateur. Il est intéressant de noter que la prise en charge d'utilisateurs multiples est une caractéristique intégrale, même dans le cloud privé, dans ce contexte, les divers services ou groupes de travail d'une entreprise représentent les différents utilisateurs ;

- API Web : est l'ensemble d'API Web (utilisant des méthodes standard telles que les appels http RESTful, XML et SOAP) au travers des quelles services cloud peuvent être invoqués. Ces API permettent la mise à disposition des services par le biais d'un navigateur Web standard ou d'une autre application cliente http.

1.13. SERVEUR DU CLOUD COMPUTING

Le serveur cloud fonctionne de la sorte qu'un serveur physique, à la différence toutefois que les fonctions qu'ils fournissent peuvent être différentes. Il y'a deux principales fonctionne d'hébergement à savoir, l'hébergement partagé qui est l'option la plus économique dans la mesure où les serveurs sont partagés entre les clients est hébergé sur le même serveur que les sites web appartenant à d'autres clients. Il présente de nombreux désavantages, y compris le fait que l'installation manque de flexibilité et n'est pas en mesure de gérer de grandes quantités de trafic. L'hébergement dédié est une forme d'hébergement beaucoup plus avancée, ou le client fait l'acquisition de serveur physiques complets cela signifie que le serveur est complètement à lui sans avoir partagé avec d'autres clients. Le serveur dédié permet d'avoir un contrôle total sur son hébergement, leur désavantage consiste à ce que les capacités requises doivent être prévues avec suffisamment de ressources et de puissance de traitement pour gérer les niveaux de trafic attendus.

L'hébergement cloud permet d'adapter les ressources à la hausse ou à la baisse selon les besoins, ce qui en fait un système bien plus élevé, il est possible d'augmenter automatiquement leur capacité pour répondre à la demande, sans devoir payer pour cette capacité de manière permanente. Un serveur dédié peut fonctionner sur un hyperviseur, dont le rôle est de contrôler la capacité et de l'affecter là où elle est nécessaire suivant l'utilisation des systèmes d'exploitation. Le cloud serveur est disponible pour chaque client, cela permet de dédier les ressources informatiques à un client déterminé si cela est nécessaire.

- Avantages clé dans son serveur à savoir :
 - Flexibilité et extensibilité, il est possible d'accéder aux ressources supplémentaires selon les besoins et quand cela est nécessaire ;

- Efficacité des coûts, étant donné que le serveur cloud est accessible quand les clients en ont besoin, ceux-ci ne paient que ce qu'ils utilisent à une période déterminée ;
- Facilité l'installation, le serveur cloud est facile à installer ;
- Fiabilité, vu le nombre de serveurs disponibles, si l'un d'entre eux rencontre un problème, les ressources sont déplacées vers les autres afin que les clients ne soient pas affectés.

1.14. BIGDATA

Il est appelé grosses données ou méga données (recommandé), parfois appelées données massives, qui désignent des ensembles de données devenus si volumineux qu'ils dépassent l'intuition et les capacités humaines d'analyse et même celles des outils informatiques classiques de gestion de base de données ou de l'information.

Big Data, inventées par les géants du web, ces solutions sont dessinées pour offrir un accès en temps réel à des bases de données géantes. Face à l'explosion du volume d'informations, le Big Data vise à proposer une alternative aux solutions traditionnelles de bases de données et d'analyse (serveur SQL, plateforme de Business Intelligence...). Confrontés très tôt à des problématiques de très gros volumes, les géants du web, au premier rang desquels Yahoo, Google et Facebook ont été les premiers à déployer ce type de technologies.

Le Big Data a quatre dimensions :

- **Volume** : les entreprises sont submergées de volumes de données croissants de tous types, qui se comptent en téraoctets, voire en pétaoctets. Transformer les 12 téraoctets de Tweets créés quotidiennement en analyse poussée des opinions sur un produit. Convertir les 350 milliards de relevés annuels de compteurs afin de mieux prédire la consommation d'énergie.
- **Vélocité** : parfois, 2 minutes c'est trop pour les processus chrono sensibles tels que la détection de fraudes, le Big Data doit être utilisé au fil de l'eau, à mesure que les données sont collectées par votre entreprise afin d'en tirer le maximum de valeur.
- **Variété** : le Big Data se présente sous la forme de données structurées ou non structurées (texte, données de capteurs, son, vidéo, données sur le parcours, fichiers journaux, etc.). De nouvelles connaissances sont issues de l'analyse collective de ces données. Utiliser les centaines de flux vidéo des caméras de surveillance pour

contrôler les points d'intérêt. Tirer parti de la croissance de 80 % du volume de données image, vidéo et documentaires pour améliorer la satisfaction client.

- Véraçité : 1 décideur sur 3 ne fait pas confiance aux données sur lesquelles il se base pour prendre ses décisions. Comment pouvez-vous vous appuyer sur l'information si vous n'avez pas confiance en elle. Etablir la confiance dans les Big Data représente un défi d'autant plus important que la variété et le nombre de sources augmentent.

Big Data va bien au-delà de la seule notion de volume : il constitue une opportunité d'obtenir des connaissances sur des types de données et de contenus nouveaux. Aujourd'hui, la plateforme Big Data d'IBM a recours aux technologies les plus pointues et à des solutions d'analyse brevetées, afin d'ouvrir la porte à de nouvelles possibilités.

Les principales technologies de Big Data, elles sont nombreuses pour optimiser les temps de traitement sur des bases de données géantes, plusieurs solutions peuvent entrer en jeu :

- Des bases de données NoSQL (comme MongoDB, Cassandra ou Redis) qui implémentent des systèmes de stockage considérés comme plus performants que le traditionnel SQL pour l'analyse de données en masse (orienté clé/valeur, document, colonne ou graphe).
- Des infrastructures de serveurs pour distribuer les traitements sur des dizaines, centaines, voire milliers de nœuds. C'est ce qu'on appelle le traitement massivement parallèle. Le framework Hadoop est sans doute le plus connu d'entre eux. Il combine le système de fichiers distribué HDFS, la base NoSQL HBase et l'algorithme MapReduce. D'autres technologies visant à tendre vers des traitements plus "temps réel" ont émergé dans la foulée (c'est le cas d'Apache Spark).
- Le stockage des données en mémoire (Memtables) permet d'accélérer les temps de traitement des requêtes.

CONCLUSION PARTIELLE

Le Cloud Computing a certainement contribué à réduire les problèmes, qui ont tourmentés les entreprises comme : espace de stockage, réalisation des différentes applications et la disponibilité de données.

Donc le Cloud est un modèle de distribution et de partage de ressources à grande échelle, qu'il offre aux utilisateurs un environnement plus simple et plus efficace pour travailler, parmi les intérêts du Cloud : faciliter le travail, augmenter la collaboration, réduire le coût de matériel et éviter le déplacement de clients pour obtenir une donnée. On peut exploiter les bénéfices du Cloud à partir des différentes solutions.

CHAPITRE II : LES TECHNOLOGIES SUR INTERNET DANS LA MESSAGERIE ELECTRONIQUE ET RESEAUXSOCIAUX

INTRODUCTION

Dans le chapitre précédant, nous avons présenté les concepts de base du cloud, leurs technologies et leurs modèles des services, de déploiement et les éléments constituant le cloud computing ainsi que ses avantages, ses inconvénients et la sécurité. Nous avons parlé sur la différence entre cloud et d'autres technologies et aussi parlé sur quelques technologies de réseaux, sur l'architecture du cloud, son serveur et la technologie big data.

Quant à ce chapitre, nous expliquerons Google avec ces produits Gmail, YouTube et produits Facebook tout en donnant un état de l'art sur ces produits ci-haute

SECTION 1 : SERVICE DE MESSAGERIE GOOGLE MAIL

2.1.1. PRESENTATION GOOGLE

Google est une entreprise américaine offrant des services sur les technologies de la science informatique, fondé le 4 septembre 1998 en Californie par Larry Page et Serguei Brin. Ce sont eux les créateurs du moteur de recherche Google, il s'est fait connaître à travers la grande force de son moteur de recherche, qui à concurrencé d'autres entreprise. Elle a opéré beaucoup des changements sur les développements et détient aujourd'hui des nombreux logiciels et sites web parmi lesquels : YouTube, le système d'exploitation Android, ainsi que d'autres services tels que Google Earth, Google Maps ou Google Play. Il s'est donné comme mission d'organiser l'information à l'échelle mondiale et de la rendre universellement accessible et utile.

2.1.2. LES ALGORITHMES DE GOOGLE

Le SEO est l'art de comprendre la manière dont travaillent les moteurs de recherche comme Google, pour faire que son site sorte dans le haut de la première page de réponse à une recherche de mots clés. Ce classement est le fruit d'une organisation effectuée par le moteur de recherche sur la base de calculs gérés par l'un des secrets les mieux gardés de la planète : l'algorithme Google (<http://mysocialselling.com/algorithme-google/>, 2016).

2.1. Les principes des algorithmes Google

Google repose sur quatre grands algorithmes :

- Le PageRank : C'est l'élément le plus connu du classement de l'algorithme Google développé par Larry Page and Sergey Brin, qui ont fondé Google. Le PageRank est un indice correspondant aux nombres de liens sur une page,

gage de qualité pour celle-ci. Il est toujours utilisé actuellement, il continue à bénéficier d'évolutions, mais il est largement dépassé, et fait maintenant partie d'un système plus important.

- Les signaux : L'algorithme Google utilise maintenant 200 critères qu'il appelle « *signaux* ». Ces derniers dépendent à la fois du contenu de la page, et de son évolution, des requêtes et du comportement des visiteurs.
- Les classificateurs sont utilisés sur les requêtes : leur but est de restituer le contexte de recherche, le cadre dans lequel elle se place. Par exemple, veut-on rechercher un produit pour l'acheter ou se renseigner sur quelque chose ?
- Les modèles : l'algorithme Google fait appel à plusieurs modèles. A savoir :
 - Les modèles de langues : la capacité à comprendre les phrases, les synonymes, les accents, les fautes d'orthographe, etc.
 - Les modèles de requêtes : il n'y a pas seulement la langue, mais comment elle est utilisée de nos jours.
 - Les modèles de temps : certaines pages répondent mieux quand elles existent depuis 30 minutes, et d'autres quand elles ont subi l'épreuve du temps.
 - Les modèles personnalisés : toutes les personnes ne recherchent pas les mêmes choses, ne le cherchent pas avec les mêmes mots etc.

L'algorithme Google bénéficie d'une puissance de calcul unique au monde. Le calcul de l'algorithme Google dispose de centaines de milliers d'ordinateurs pour indexer les milliards de pages de tous les sites Web au monde, et est capable de mettre à jour l'index tout entier en quelques jours ! Il est important de savoir que les centres de données stockent une copie de toutes les pages du Web pour pouvoir y accéder plus rapidement.

Par ailleurs l'algorithme Google bénéficie d'une recherche d'amélioration constante : Des groupes travaillent à l'amélioration de l'algorithme, d'autres opèrent sur l'évaluation des résultats. L'algorithme Google a bénéficié de plusieurs milliers de modifications.

Les algorithmes de Google en SEO les plus connus de Google sont Panda et Pingouin. Apparus en 2011 et 2012, ils font respecter leurs règles et attention aux faux pas. Enfin, « *hummingbird* » ou Colibri en français apparus en 2013. Cet algorithme va plus loin que les précédents car il ne raisonne pas en fonction des mots-clés tapés lors de la recherche (<https://www.journalducms.com/2017/03/08/seo-algorithmes-google-panda-pingouin-colibri-15229/>, 2017).

- Google Panda : sanctionne les sites de mauvaises qualités au contenu de faible valeur ajoutée. Par exemple, les sites publiant du contenu pour générer des revenus publicitaires ou bien des sites dupliquant du contenu déjà existant ainsi que les comparateurs de prix. La dernière version officielle de Panda date du juillet 2015. C'est un algorithme majeur de Google.
- Google Penguin : s'attaque aux sites « *suroptimisés pour le SEO* » dans le but de booster leur référencement. Cette technique consiste à multiplier des liens sortants (netlinking) non justifiés, mais de multiplier les liens entrants issus d'autres sites (backlinking), enfin de décliner un contenu identique sur différentes pages d'un site. Penguin a été mis à jour le 23 septembre 2016. Il est enfin intégré dans l'algorithme de Google avec des mises à jour « *en temps réel* ». C'est un autre algorithme majeur de Google.
- Google Colibri : « *Hummingbird* » ou Colibri, mis en ligne en septembre 2013, est né pour améliorer la pertinence des résultats par une meilleure compréhension de l'intention de l'internaute. À première vue, colibri s'intéresse aux requêtes composées de plusieurs mots clés. Il serait le plus intelligent des algorithmes car il essaie de deviner la définition qui se cache derrière chaque requête et la traduit en une interrogation courte.

2.1.3. GOOGLE DANS LE CLOUD

2.1.3.1. Stockage Gmail et YouTube dans le Cloud

Le cloud Gmail et YouTube c'est fait par Google avec son produit Google Drive, Google Drive propose à chaque nouvel abonné 15 Go à partager entre le Drive mais aussi Gmail et Google+ Photo. Chez Google un seul compte suffit pour avoir accès à l'intégralité des services de la marque (Google+, YouTube, Gmail, Analytics...), ce qui explique le partage du cloud. Si 15 Go semble être un stockage assez conséquent, Google Drive offre un autre avantage : les photos inférieures à 2048×2048 et les vidéos de moins de 15 minutes ne sont pas prises en compte dans l'espace de stockage du cloud.

Google Drive propose lui 15 Go mais ne compte pas la quasi-totalité de vos photos, vos courtes vidéos ou certains de vos documents, Si 15 Go avec une réserve presque illimitée ne vous suffisent pas, vous pouvez passer à la caisse. Le cloud by Google est d'ailleurs plutôt économique avec différentes formules, même si pour notre comparatif, je prendrai les 100 Go comme mètre-étalon. Ainsi Google Drive propose les 100 Go à 1,99€ par mois. Un tarif tout à fait raisonnable si vous avez besoin d'un stockage considérable. On mettra quand même un bémol à l'offre de Google : si vous souhaitez plus de 100 Go, il faudra passer

directement à un To qui se négocie à 9,99€/mois. Pas d'offre intermédiaire dans le cloud de Google.

Les Services offerts par Google restent la compatibilité totale entre tous ces services (Gmail, Google+...) mais aussi sur toutes les plateformes (compte Google sur PC, Smartphone et Tablette Android et bientôt sur Android Gear ou Google Glass). Google Drive s'intègre parfaitement dans cette suite : les photos et les vidéos des Smartphone Android sont stockées sur le cloud, tandis que des sauvegardes régulières se font également dans cet outil. La synergie créée par l'association avec Chrome ou Gmail fait de Google Drive le cloud indispensable des utilisateurs de produits Google.

2.1.3.2. Fonctionnement de Google Drive

Google drive fonctionne de façon de vous permettent de charger 30 formats des fichiers pris en charge, musique, vidéo, images, feuilles de calcul, document texte etc. pour y accéder de n'importe où. Vous pouvez organiser en dossiers et choisir lesquels vous voulez partager et avec qui. A l'intérieur du cloud vous naviguez facilement en utilisant différents volets de navigation, MonDrive regroupe tous vos éléments, Activité renfermée les fichiers triés par date de modification ou encore Partages avec moi ou vous trouvez des fichiers partagés avec vous par vos amis, etc. vous pouvez faire de recherche d'un document en tapant un mot clé et il existe encore une petite astuce de recherche,

Google Drive est capable de reconnaître les lettres dans une image (si vous avez photographié une boulangerie et que vous tapez boulangerie dans la barre de recherche, il vous trouvera la photo) et un autre point fort de ce service vous pouvez y accéder même hors connexion. Pour cela vous pouvez aller sur drive.google.com et sélectionnez configuration de Google Documents hors connexion, en cliquant sur Autoriser les documents hors connexion et clique sur le bouton installer, à partir de là vous pouvez ouvrir les feuilles de calcul en lecture seule et les autres documents en lecture et écriture. Google vous offre 5 Go d'espace de stockage gratuit, pour les restes il faudra sortir votre porte-monnaie (www.androidpit.fr/, s.d.).

2.1.3.3. Avantages de la technologie Google Drive

- La gratuité : les logiciels de bureautique contenus dans le service Google Drive sont tous totalement gratuits pour peu qu'on ne dépasse pas les 15 gigas alloués lors de la création du compte (ce qui représente entre 50 000 et 100 000 documents de traitement de texte).
- L'importation des fichiers : la suite Google Drive permet d'importer des fichiers du type traitement de texte (Word), Tableur (Excel), ou présentation (Power Point). Le

nombre important de formats pris en compte rend Google Drive souple d'utilisation. Le logiciel intègrera tout seul les évolutions des autres formats évitant aux utilisateurs les mises à jour régulières. Le seul bémol réside dans une possible perte de qualité en cas d'import d'un document qui comprend une mise en forme complexe.

- Pas de mise à jour pour nouvelle version nécessaire : lorsque vous achetez un pack Office ou téléchargez un produit Open Office vous devez vous assurer d'avoir un système d'exploitation compatible aux dernières versions. Ce n'est pas le cas avec le service Google Drive, quel que soit le système d'exploitation (Windows, GNU Linux et Mac OS X) ou le navigateur (Internet Explorer, Safari, Firefox, Opéra, chrome, ...) que vous utilisez vous aurez toujours accès à vos données.

Remarque : vous pouvez même avoir accès en consultation et en modifications à vos données depuis votre tablette ou votre Smartphone.

- Le partage des données : le service Google Drive permet de partager en ligne les documents avec d'autres personnes présentes sur le net soit en mode consultation (il n'autorise le lecteur à modifier le document) soit en mode modification qui autorise le ou les lecteurs à compléter ou corriger le document en ligne. C'est un atout important pour les équipes qui travaillent sur un document en commun car il évite les envois de documents avec des numéros de version Un seul partage de fichier et l'ensemble des collaborateurs « *partagés* » peuvent avoir accès à la dernière version en temps réel.
- La collaboration sur un fichier partagé en temps réel : dans la continuité du partage des données, le service Google Drive permet à plusieurs utilisateurs d'être connectés simultanément sur un document commun et de modifier les données en même temps. Un signalement visuel de couleur permet à chacun de visualiser les modifications que fait chacun des utilisateurs. Cette fonction est intéressante pour de nombreuses entreprises car elle permet à plusieurs personnes de travailler en même temps sur un fichier commun ou centralisé à partir de lieux géographique différents. La seule contrainte étant d'être connectée à Internet.
- La création de formulaire : le service Google Drive propose un service de création de formulaire en ligne qui permet de réaliser des questionnaires ou sondages d'opinion à un groupe de contacts (collaborateurs, clients, prospects, population test, ...).

-
- La disponibilité et la souplesse de Google Drive : les documents étant hébergés en mode SAAS, ils sont consultables et modifiables à volonté, gratuitement et de n'importe quel PC, Tablette ou Smartphone connecté à Internet.
 - Le mode « Hors Connexion » : permet aux utilisateurs de créer un nouveau document Google Drive sans avoir besoin d'être connecté. Le système se synchronisera dès que la connexion sera effective ce qui permettra l'enregistrement des nouveaux éléments dans les serveurs Google. Il ne permet cependant pas de modifier un document existant, puisque n'étant pas connecté il n'a pas accès aux serveurs de stockage.
 - La lecture d'un grand nombre de différents types de fichiers : les services Google Drive permet la lecture de plus de 40 différents types de fichiers :
 - Image & vidéo (Photoshop, ...)
 - Bureautique : traitement de texte, tableur, présentation (Word Excel, Power Point, open office...)
 - Pdf
 - Fichier compressés (Zip, Rare, ...)
 - La connexion à votre messagerie Gmail : étant connecté à votre messagerie Gmail, le service Google Drive vous permet d'ouvrir vos documents directement à partir de votre messagerie. Plus besoin de copier ou d'importer vos fichiers avant de les ouvrir.
 - L'enregistrement automatique : le service Google Drive enregistre vos données automatiquement et régulièrement afin de ne pas perdre les modifications en cas de problème informatique. Cet atout est aussi disponible sur les autres logiciels ce qui n'enlève pas son intérêt.
 - Les exportations de données : Google Drive permet d'exporter les données dans les formats suivants :
 - Logiciels Office (Word, Excel, Power Point, ...)
 - Les Pdf
 - Les .rtf (texte enrichi)
 - Les pages web
 - Les textes bruts

2.1.3.4. Inconvénients de la technologie Google Drive

- Les données sont stockées sur un serveur Google. Si elles sont sécurisées, il est préférable de ne pas mettre de documentation confidentielle (innovation, R&D, ...).
- Le mode hors connexion a ses limites. Il est nécessaire de s'organiser lorsqu'on souhaite travailler dans un endroit où la connexion n'est pas bonne (train, avion, campagne...)
- On est tributaire des modifications de Google. Même si les changements sont progressifs et peu nombreux, les utilisateurs sont obligés de s'adapter aux mises à jour Google. Le point positif est que les utilisateurs s'adaptent au fur et à mesure des petites évolutions. Ce qui évite de devoir se former tous les 5 à 10 ans sur les nouveaux fonctionnements des logiciels de bureautique.
- Les logiciels ont une moindre puissance par rapport aux logiciels habituels. Elle est quand même largement suffisante pour 80% des utilisateurs qui n'emploient jamais l'ensemble des fonctionnalités que propose Microsoft ou Open Office.
- L'ergonomie des logiciels est différente et demande un peu d'adaptation. Une fois que l'habitude est prise certains préfèrent l'ergonomie de Google Drive aux autres logiciels ...

2.1.3.5. Sécurité dans Google Drive

- Les fichiers importés dans Google Drive sont stockés dans des centres de données sécurisés.
- En cas de perte ou de vol de votre ordinateur, de votre téléphone ou de votre tablette, vous pouvez toujours accéder à vos fichiers avec un autre appareil.
- Vos fichiers restent privés, sauf si vous décidez de les partager.

2.1.4. ECHANGE DE DONNEES GOOGLE

La firme américaine Google est utilisée par de nombreux internautes, pour son moteur de recherche, mais aussi pour ses nombreux services gratuits (Gmail, Drive, YouTube, Google Maps...). Puisque Google exploite vos données sans que vous n'en ayez toujours conscience (www.ouest-france.fr/, s.d.).

Ce sont plus de 200 services proposés gratuitement par l'entreprise. Pour la plupart d'entre eux, la seule contrepartie demandée est l'ouverture d'un compte Gmail, le service de messagerie en ligne maison. L'adresse e-mail et le mot de passe associé deviennent alors vos sésames pour vous identifier et entrer dans la sphère Google, depuis n'importe quel terminal à travers le monde, en échange de vos données personnelles.

Toute cette gratuité a cependant une face cachée : l'exploitation commerciale de nos données personnelles. En effet, elles représentent une manne financière des plus importantes. En acceptant les « conditions générales d'utilisation », que nous ne lisons quasiment jamais, nous donnons le droit à Google de tracer et d'utiliser tout ce que nous faisons sur Internet : les sites visités, les achats effectués, les lieux dans lesquels nous nous rendons, les films regardés, les livres lus, la musique écoutée...

L'ensemble de ces données est alors analysé par les puissants ordinateurs de la firme, dans le but créer une sorte de carte d'identité très précise de chaque utilisateur. Ces profils, compilant de très nombreuses données, se revendent à prix d'or aux marques désireuses de cibler au mieux leur publicité. C'est ce que l'on appelle le « Big Data ».

Pour profiter gratuitement des services de Google, comme ceux de nombreux autres acteurs des nouvelles technologies, nous devons donc rogner sur notre vie privée, en abandonnant la confidentialité de nos données personnelles. Il existe une formule qui résume parfaitement cette pratique : « si c'est gratuit, c'est que le produit c'est vous ! »

2.1.5. PROTECTE DE DONNEES GOOGLE

Ce site utilise Google Analytics, un service d'analyse de site Internet fourni par Google Inc. (Google). Google Analytics utilise des cookies, qui sont des fichiers texte placés sur votre ordinateur, pour aider le site Internet à analyser l'utilisation du site par ses utilisateurs. Les données générées par les cookies concernant votre utilisation du site sont en règle générale transmises et stockées par Google sur des serveurs situés aux Etats-Unis. Si vous avez choisi de rendre anonyme votre adresse IP sur ce site, votre adresse IP sera toutefois transmise par Google, mais tronquée dans les pays membres de l'Union européenne ou dans les autres Etats signataires de l'accord instituant l'Espace économique européen.

Dans des cas exceptionnels uniquement, votre adresse IP sera transmise intégralement à un serveur de Google aux USA, puis tronquée là-bas. Google utilisera ces informations dans le but d'évaluer votre utilisation du site, de compiler des rapports sur l'activité du site à destination de son éditeur et de fournir d'autres services relatifs à l'activité du site et à l'utilisation d'Internet. Google ne recoupera pas votre adresse IP transmise dans le cadre de Google Analytics avec toute autre donnée détenue par Google.

Vous pouvez désactiver l'utilisation de cookies en sélectionnant les paramètres appropriés de votre navigateur. Cependant, cette désactivation pourrait vous empêcher d'utiliser pleinement certaines fonctionnalités de ce site. Vous pouvez en outre empêcher que les informations collectées par le cookie et concernant votre utilisation du site (y-

compris votre adresse IP) ne soient transmises à Google et exploitées par Google, en téléchargeant et en installant le plug-in de navigateur.

Seule la connexion entre l'internaute et les serveurs de Google est sécurisée par défaut avec le protocole HTTPS via un tunnel SSL. Si la transmission des données est donc chiffrée, ce n'est pas le cas pour les serveurs de la firme.

2.1.6. CHIFFREMENT SUR GOOGLE

Oui. Les données sont chiffrées à plusieurs niveaux. Le protocole HTTPS (HyperText Transfer Protocol Secure) est appliqué systématiquement pour tous les échanges entre les utilisateurs et les services G Suite. En outre, tous les services utilisent la confidentialité de transmission parfaite (PFS, Perfect Forward Secrecy). Google chiffre également les messages échangés avec d'autres serveurs de messagerie qui utilisent le protocole TLS sur 256 bits et les clés de chiffrement 2048 RSA lors des phases de validation et d'échange de clés. Cette méthode permet de protéger les e-mails échangés entre les utilisateurs de nos clients et des tiers externes qui utilisent également TLS.

La confidentialité de transmission parfaite exige que les clés privées d'une connexion ne soient pas stockées. Quiconque déchiffre une clé unique ne peut pas déchiffrer les connexions des mois passés. Même l'opérateur du serveur ne peut pas déchiffrer les sessions HTTPS a posteriori.

Google s'efforce en permanence de renforcer le chiffrement et de l'appliquer à de nouveaux liens et services. G Suite inclut l'outil Mobile Device Management (MDM) pour Android et iOS qui permet notamment d'activer des appareils, de supprimer à distance les données d'un appareil et d'effectuer un chiffrement en fonction de règles. Le service MDM vous permet de contrôler vos données et de faciliter l'accès aux informations d'entreprise par les personnes qui utilisent leur appareil personnel, sans compromettre la sécurité.

2.1.7. PRESENTATION GMAIL

Gmail est un service de messagerie gratuit offert par Google. Il est en caractère commercial, publicitaire et multilingue permettant d'avoir un compte dont Google est propriétaire. Les messages reçus par Gmail peuvent être lus par les clients de messageries ou encore de navigateur web. Gmail est le seul ou un des seuls fournisseurs de boîte aux lettres électroniques au grand public qui chiffre le message d'un bout à l'autre de la chaîne. En effet la connexion est chiffrée par une clé de 128 bits entre l'utilisateur et les serveurs de Gmail que ce soit via l'interface Web sécurisée en HTTPS ou via un client de messagerie qu'il utilisera du SMTP. Pour transmettre le message dans le serveur SMTP du destinataire,

si celui-ci lui autoriser, Gmail utilise un chiffrement TLS. Cependant les messages sont en clair à l'endroit de leurs parcours, le temps de leur analyse statistique par les algorithmes de Google permettant de cerner les centres d'intérêt des correspondants.

A son lancement le 01 avril 2004 l'inscription nécessitait au préalable une invitation, mais deux ans plus tard la version beta fut ouverte au public avec la capacité initiale de 1 Go, le 13 mai 2013 il eut augmentation et Google décida d'unifier les espaces des stockages de ses différents services pourvu qui ait l'augmentation.

La messagerie fonctionnant en synergie avec le moteur de recherche Google qui a fait connaître la firme avec le soutien des fondateurs Larry Page et Sergueï Brin.

2.1.7.1. PROTOCOLE SUPPORTES PAR GMAIL

- POP3 : pour la récupération de ses mails sur un logiciel de courrier électronique de manière sécurisée grâce au protocole SSL ;
- IMAP : pour récupérer ses mails, les réorganiser ou en envoyer d'autres sur son compte Gmail avec un logiciel de courrier électronique ;
- SMTP : pour l'envoi des courriers depuis un logiciel de courrier électronique ;
- HTTPS : pour sécuriser la connexion à l'interface.

2.1.7.2. METADONNEES DE GMAIL

Une métadonnée, c'est en quelque sorte une donnée sur une donnée. On peut trouver dans ce tiroir sémantique (m.nouvelobs.com , s.d.) :

- Numéros de téléphone sortants et entrants ;
- Heure, date et localisation des appels (pas leur contenu) ;
- Émetteur, destinataire et date des courriels (pas leur contenu, encore une fois) ;
- Titre de l'e-mail ;
- Recherches internet ;
- Transactions bancaires ;
- Carnets d'adresses sur les réseaux sociaux.

Le programme baptisé « *Immersion* » permet de voir les métadonnées c.à.d demande à l'utilisateur d'entrer son adresse et son mot de passe Gmail puis balaye l'ensemble des courriels échangés depuis la création du compte. En quelques secondes, l'application crée un graphique interactif et dynamique constitué de petites bulles pour chaque correspondant, plus ou moins grosses selon le volume d'échanges, et les liants entre elles.

SECTION 2 : SERVICE DES RESEAUX SOCIAUX OU MEDIA SOCIAUX

2.2.1. DEFINITION DES RESEAUX SOCIAUX

Un réseau social « une structure définie par des relations entre des individus ». Concrètement, c'est l'ensemble des individus avec qui une personne est en contact. Il s'agit également de liens entre des personnes : les habitants d'un quartier, des passionnés d'art, une famille. Depuis le début des années 2000, l'apparition du web 2.0 a permis de développer les réseaux sociaux grâce aux différentes communautés en ligne. Le but des réseaux sociaux sur internet est tout d'abord de rencontrer des personnes qui ont des intérêts communs, garder le contact avec ces personnes et enfin de reprendre contact avec des personnes perdues de vue ainsi que de maintenir le lien avec des personnes distantes.

2.2.1.1. Le web 2.0

Le web est défini comme une toile virtuelle, c'est un réseau informatique mondial reliant des documents entre eux qui peuvent être exploités grâce à l'utilisation d'hyperliens permettant la navigation. Bien que l'aspect technique soit très compliqué, le web est démocratisé depuis 1995 et continue à se développer. Aujourd'hui, nous sommes dans l'ère du web 2.0. Le web 2.0 est une évolution de l'internet qui permet, grâce à une multitude de dispositifs (ordinateur, téléphone portable, tablette tactile), une interaction constante entre les internautes (blog, réseaux sociaux), qui peuvent eux-mêmes créer du contenu grâce aux évolutions des navigateurs.

2.2.1.2. Impact des réseaux sociaux

Plus d'un milliard d'utilisateurs sur 300 plateformes différentes, 70% des internautes sont présents sur un réseau social minimum. L'explosion des réseaux sociaux est bien réelle. De plus, les réseaux sociaux touchent tous les individus de tous âges.

2.2.1.3. Chaque réseau social a son intérêt

Chaque réseau social n'a pas la même cible et représente une offre différente pour ceux-ci. Il existe des réseaux sociaux pour être relié avec des amis, avec des collègues, avec des marques, ou même pour partager des vidéos, des photos. Selon les préférences de chacun, un réseau correspondra aux attentes d'un utilisateur qui souhaite intégrer une communauté.

2.2.1.4. Typologie des réseaux sociaux

A. Les réseaux sociaux

Un réseaux social est une plateforme qui permet le contact entre les internautes grâce à la création de profils personnels auxquels les amis, membres de la famille et connaissances peuvent avoir accès. Il prévoit des services de messagerie et de discussion instantanée on parle de « chat ». Les réseaux sociaux encouragent l'échange de contenus divers tels que des photos, des vidéos, des articles de presse, des sites internet, mais aussi des opinions, des statuts, etc. ils favorisent la communication en permettant aux membres de garder contact et alimentent un sentiment communautaire.

B. Les blogs

Un blog est une sorte de journal en ligne qui permet la publication de billets d'humeur affichés de façon antéchronologique sur des thèmes variés tels que la politique, la musique, le cinéma, la littérature, la mode etc.

C. Le microblogging

Est tout simplement un dérivé du blog et permet à ses utilisateurs de publier de courts messages à l'attention d'un cercle de « followers » du flux.

D. Sites de création et de partage de contenu

Ce sont des plateformes qui permettent aux utilisateurs de créer et de publier du contenu de toute sorte afin de le partager avec d'autres internautes et de recueillir leur avis.

E. Forums

C'est la manifestation la plus ancienne des médias sociaux. Les forums sont des plateformes sur lesquelles se déroulent des discussions thématiques entre les internautes, des sujets de conversation sont ouverts par les internautes qui viennent poser leurs questions, partager et recueillir des informations, des recommandations etc.

F. Les mondes virtuels

Plateformes qui répliquent un monde virtuel fantastique ou proche de la vie réelle, les mondes virtuels permettent aux joueurs d'interagir sous la forme d'un avatar. Ils sont comparables aux jeux vidéo mais nécessitent une connexion internet.

G. Les services de géolocalisation

L'essor du smartphone ces deux dernières années a permis le développement d'applications de géolocalisation sur téléphone mobile. Ainsi les utilisateurs peuvent signaler à leurs contacts où ils se trouvent, le but étant de créer un contact physique à travers la création d'un contact virtuel.

2.2.2. YOUTUBE

2.2.2.1. PRESENTATION

Les chaînes YouTube nous permettent de visualiser les vidéos par les internautes, ainsi seules les personnes inscrites ont ces accès d'envoyer des vidéos de façon illimitée et ces vidéos sont accessibles par catégorie à l'aide de mots-clés (tags) comme le fait Flickr ou Technocratie, ils peuvent être importées sur un blog personnel, tout internaute inscrit peut poster des commentaires et noter les vidéos en ligne.

La société a été créée en février 2005 par trois employés de l'entreprise Paypal par Chad Hurley, Steve Chen et Jawed Karim. La plateforme devait être à l'origine un site de rencontre basé sur l'utilisation de vidéos, mais les créateurs ont renoncé à cette idée quelques jours après son lancement public. La première vidéo de YouTube est d'ailleurs celle de Jawed Karim où il commente sa visite au Zoo de San Diego, la vidéo a été mise en ligne le 23 avril 2005. Le 9 octobre 2006, Google fait l'achat de YouTube ce qui constitue la quatrième plus grosse opération d'acquisition de Google.

2.2.2.2. PROTOCOLE SUPPORTES PAR YOUTUBE

- RTMP (Real Time Messaging Protocol) : est un protocole réseau propriétaire développé par Adobe Systems pour la diffusion de flux de données en Streaming (audio, vidéo ou autre) entre le serveur et le client généralement le lecteur flash.

Les variantes des protocoles RTMP sont :

- Protocole complet fonction sur TCP et exploite le port 1935
- RTMP (RTMP Tunneling) encapsule RTMP dans des requêtes http afin de passer les pare-feu
- RTMPS est similaire à RTMP mais il utilise la connexion sécurisée TLS/SSL
- RTMPE (Encrypted Real Time Messaging Protocol) : il permet le transfert sécurisé des données sans SSL
- RTP (Real Time Transport Protocol) sa version sécurisée SRTP : est protocole standard fondé sur l'ancien protocole propriétaire RDP (en voie d'obsolescence) il est utilisé pour ses possibilités de transmission Multicast,

pour les flux en direct et ses possibilités de resynchronisation et d'adaptation de débit. RTMP reste non standard et soumis à licence pour son implémentation complète, notamment les flux cryptent.

- HTTP (Hyper Text Transport Protocol) : un protocole internet de base permettant d'échanger des données entre un client (navigateur web) et serveur web.
- HTML5 (Hyper Text Markup Language 5) : dernière révision majeure du HTML format de données conçu pour représenter les pages web, il a été finalisé en octobre 2014 qui est spécifié de deux syntaxes d'un modèle abstrait défini en termes de DOM (HTML5 et XHTML5, il comprend une couche application avec nombreuses API, ainsi qu'un algorithme afin de pouvoir traiter les documents à la syntaxe non conforme.

2.2.2.3. METADONNEES DE YOUTUBE

Les métadonnées peuvent être liées à d'autres types de documents, tels que les fichiers PDF, les fichiers audio, des photos, des fichiers zip ou vidéo. Les métadonnées sont particulièrement utiles pour la vidéo. Contrairement à un document, les métadonnées ne sont pas intégrées à une vidéo. Les moteurs de recherche indexent le contenu d'une vidéo uniquement sur le titre donné à celle-ci. Il est donc nécessaire de lier des éléments textes descriptifs aux vidéos. Ces métadonnées sont répertoriées en tant qu'information sur un fichier. Il est joint au dossier et peut inclure des informations de copyright, auteur, mots-clés, taille du fichier, format du média, description du tournage, le lieu, le sous-titrage, le sous-titrage informatifs et bien plus. Si vous envoyez par réseau ou copiez votre fichier vidéo, toutes les informations des données se déplacent également (Formation.mdesigner.fr , s.d.).

Deux types de métadonnées vidéo existent :

- Les métadonnées peuvent être éditées au sein de votre caméra. Elles stockent les informations techniques de prises de vues telles que l'ouverture, la vitesse d'obturation, les coordonnées GPS et autres. Le format standard est interopérable. Le format de fichier de la collecte de ces informations est l'EXIF. La plupart des données EXIF ne sont pas modifiables après leurs écritures, à l'exception de la date, la durée et les données GPS.
- Écrire les métadonnées manuellement en fournissant toutes informations nécessaires sur le contenu vidéo. Cela peut être la transcription des dialogues, les conversations ou autres descriptions. Ces entrées de textes sont compréhensibles par les ordinateurs

et facilite leur recherche. Vous gagnez dès lors en visibilité et vous offre les possibilités de communication.

La plupart des métadonnées vidéo est automatiquement créé. Mais les données écrites par l'homme sont nettement plus importantes. Elles favorisent un plus grand potentiel de communication. YouTube et autres portails d'hébergement vidéo ont grandement besoin de ces descriptions. Elles donnent aux utilisateurs plus d'information sur votre vidéo.

Tous les logiciels de montage vidéo doivent intégrer les métadonnées. Ces métadonnées sont stockés sous forme XML. Adobe Bridge crée les données dans un fichier XMP. Plusieurs applications créent ces fichiers XMP : xmp manager, sous linux et autres. Le format XMP est devenu le standard de facto. Cependant, les métadonnées sur les plateformes de visualisation de YouTube ont leur propre étiquetage. Vos métadonnées seront donc non conservées lors de l'encodage et nécessitera de leur fournir les informations nécessaires en ligne.

2.2.3. FACEBOOK

2.2.3.1. PRESENTATION FACEBOOK

Facebook est un site Internet qualifié de réseau social qui permet à chaque membre d'interagir avec ses amis en publiant et partageant des contenus divers comme :

- Informations personnelles (état civil, parcours scolaire et professionnel) ;
- Fichiers multimédia (photos, musique, vidéo...) ;
- Liens internet, articles, humeur du moment,

Du point de vue technique, Facebook est une réussite vraie, que ce soit au niveau de l'utilisation ou de nombre des services proposés. Les services que les jeunes allaient piocher ici et là, sont désormais réunis en un seul endroit c.à.d. chat, messagerie électronique, partage de photos et vidéos... Facebook aura pu qu'à même réussit à démoder la pratique des blogs chez les ados et à les faire fuir des skyblogs.

En 2004, Mark Zuckerberg crée un site appelé " The Facebook" dans le but de mettre en relation les étudiants d'Harvard. Le site est d'abord conçu comme un trombinoscope en ligne. Rebaptisé Facebook en 2005.

2.2.3.2. PROTOCOLE SUPPORTE PAR FACEBOOK

- Tcp/Ip : c'est un ensemble des protocoles utilisés pour le transfert des données sur Internet.

- Zéro : pour accélérer la connexion sécurisée.

2.2.3.3. ALGORITHME DE FACEBOOK

L'algorithme de Facebook est le sujet de nombreuses conversations au sein de la communauté de social media manager et créateurs de contenus. Ce dernier évolue régulièrement et de manière soudaine au rythme de la vision stratégique de Facebook depuis maintenant 10 ans.

Mais pourquoi l'algorithme de Facebook est-il sujet à tant d'attention de la part des marketeurs ? Comme nous l'évoquons dans le préambule, l'algorithme va décider quelle publication sera mise en avant selon chaque utilisateur. Donc en somme de la portée totale du contenu et de l'engagement qui en découlera. Ces deux éléments (portée et engagement) sont le nerf de la guerre sur les réseaux sociaux et notamment Facebook qui est la première plateforme mondiale avec plus d'un milliard d'utilisateurs actifs...

- ✓ Quelle formule mathématique se cache derrière ?

La visibilité d'une publication se fait à travers 4 critères principaux auxquels s'ajoute de nombreux autres éléments de moindre importance. Ces principaux critères de visibilité sont :

- Le Créateur : l'intérêt qu'engendre l'utilisateur (profil ou page Facebook) qui publie le contenu,
- Le Post : la performance de la publication soit son engagement,
- Le Type : la publication postée que préfère l'utilisateur (statut, photographie, liens, vidéo),
- La Récence : le degré de « fraîcheur » de la publication en terme de date.

Ces quatre critères cumulés forment donc ce qu'on appelle l'algorithme de Facebook sous cette formule mathématique

- ✓ $VISIBILITÉ = CRÉATEUR \times POST \times TYPE \times RÉCENCE$

Le social media manager et brand manager se plaignent régulièrement du déclin de la portée organique sur Facebook. Cela est dû principalement à l'augmentation du nombre d'utilisateurs ces 10 dernières années. En plus du nombre d'utilisateurs, ces derniers sont plus enclins à partager et publier eux-mêmes des contenus (statuts et photos) qu'auparavant. Tout ceci limite l'affichage de contenu sur la timeline des utilisateurs (<https://siecledigital.fr/2016/09/08/fonctionne-reellement-algorithme-de-facebook/>, 2016).

Contrairement aux idées « conspirationnistes » dénonçant une manipulation de Facebook pour inciter à utiliser son offre publicitaire. L'algorithme est le résultat du succès de Facebook, ce qui fait que ses affaires soient dites en passant...

- ✓ Les principaux critères influençant l'affichage sur votre timeline
- Le Créateur ou l'utilisateur qui publie : Plus vous interagissez avec l'auteur de publications (profil ou page), plus l'algorithme de Facebook vous affichera sur votre timeline les contenus de cette personne ou marque. Ce qui explique que vous ne voyez que peu des posts d'amis de longues dates...

Les interactions qui ont un impact sont les mentions J'aime sur la publication, les commentaires, les partages, les clics pour afficher la publication ou encore le fait de ralentir le scroll sur le contenu.

- L'engagement sur la publication : Ce critère fonctionne sur le principe que « *le monde amène encore plus de monde* » : plus la publication aura d'engagements, plus l'algorithme de Facebook augmentera sa portée sur les timelines. Ce qui aura un impact direct sur ce même engagement.
- Le Type de publication : Plus vous êtes engageant sur certains types de publication, comme par exemple les vidéos et les statuts, plus Facebook vous proposera ce genre de contenus afin de toujours fournir les publications qui vous intéressent.
- La Récence ou la « *fraîcheur* » du contenu : L'algorithme de Facebook va privilégier les publications récentes au profit des anciennes. Ce qui est assez logique pour une plateforme sociale. Cependant Facebook va vérifier la dernière date à laquelle vous vous êtes connecté. Donc si votre dernière connexion remonte à une semaine, il est possible que l'algorithme vous affiche une publication d'un de vos amis remontant à 5 jours, car celle-ci aura reçu de nombreux engagements.
- La répétition du contenu : Si un contenu, par exemple un lien vers le même article, a été publié par plusieurs membres de votre communauté, Facebook va interpréter cela comme une publication suscitant un grand intérêt et donc afficher ce contenu en mentionnant que : « *X de vos amis ont partagé et commenté cette publication.* ».
- Les nouveaux produits de Facebook : La plateforme propose régulièrement de nouvelles fonctionnalités comme dernièrement les Facebook Live. Ainsi une publication utilisant une des dernières fonctionnalités aura plus de visibilité dans les timelines. Une manière pour Facebook d'avoir des retours sur les performances de ses nouveaux produits.

Enfin, l'algorithme va évoluer en fonction des comportements des utilisateurs et se caler sur ses intérêts. C'est là qu'entre en jeu l'indexation des contenus pour prioriser l'affichage sur la timeline. Comme le montre l'exemple ci-dessus, le score d'indexation et donc la visibilité sera propre à chaque utilisateur selon son comportement sur le réseau social.

✓ Qu'en est-il de la visibilité des publicités ?

Les Facebook ads ne viennent pas supplanter les publications naturelles de vos amis ou éditeurs, mais s'insèrent dans la timeline. L'algorithme de Facebook pour calculer la visibilité des publicités est similaire aux contenus naturels, mais avec quelques spécificités :

- Le nombre de publicités affichées est limité,
- Les publicités sont soumises aux centres d'intérêts de chaque utilisateur. Ces derniers peuvent être consultés et personnalisés.

2.2.3.4. FACEBOOK DANS LE CLOUD

La plupart des logiciels utilisés en interne chez Facebook sont en grande partie des solutions en mode Cloud à 70% des solutions utilisées en interne reposent sur le modèle SaaS. Le choix des offres SaaS et PaaS de Salesforce, le réseau social utilise donc les solutions SaaS de Salesforce pour gérer ses ventes, mais a également bâti une douzaine d'applications Cloud maisons, dédiées au RH ou à la maintenance des Datacenter entre autres. Elles reposent sur la plate-forme Force.com (le PaaS de Salesforce).

Facebook utilise également Rypple : une solution d'évaluation des salariés en mode Cloud qui ressemble justement à un réseau social d'entreprise spécialisé dans les RH et la productivité. Rypple a d'ailleurs été racheté par Salesforce. Le CRM de Salesforce intégré aux technologies Oracle, le membre de la DSI de Facebook justement chargé de l'architecture Cloud, le système d'information interne de Facebook combine cloud privé et SaaS (www.journaldunet.com/solutions/dsi/informatique-de-facebook/le-cloud-computing-chez-facebook.shtml, 2012).

2.2.3.5. ECHANGE DE DONNEES FACEBOOK

Pourquoi Facebook est important ? Parce que nous sommes en train de voir émerger le futur standard d'échange de données personnelles sur Internet.

Les analyses du phénomène Facebook se sont lourdement transformées depuis son émergence, en seulement quelques mois. Après avoir été une tendance du web 2.0, Facebook est devenu le site qui reliait l'identité numérique à l'identité réelle des gens, puis le business model du futur, et enfin le danger pour les libertés publiques et la vie privée des citoyens.

Le tout à travers une débauche sans cesse croissante d'articles et de commentaires dans des médias de plus en plus variés.

Il est clair que le potentiel de l'outil est de nature à justifier cet intérêt. Comme avec le mail ou la messagerie instantanée, il est aujourd'hui possible de passer sa journée sur Facebook sans aller sur le web. Le réseau social n'est plus seulement un service, il est devenu une application.

C'est parce que Facebook est devenu une application que les gens commencent à se poser des questions sur les usages qui en découlent. Ceux des autres d'abord, c'est-à-dire les différents modèles économiques qui permettent de rentabiliser le temps passé sur le site. Ils sont d'autant plus difficiles à lister que l'ouverture de l'infrastructure de Facebook aux développeurs tiers permet l'apparition continue de nouveaux applicatifs, et que la structuration du site en micro-communautés rend impossible la dissémination de ces applications à l'ensemble du réseau d'un bloc. Elle force les développeurs à opter pour des stratégies de marketing et à rentrer en concurrence les uns avec les autres.

On voit ainsi fleurir depuis quelques semaines des articles consacrés à la protection de la vie privée. Les analyses prospèrent personne ne semble se rappeler que les législations européenne et française interdisent la constitution de fichiers contenant des préférences politiques ou religieuses. Quel que soit l'intérêt des journalistes pour ce sujet, il reste encore superficiel. De la même façon que YouTube s'est construit en parallèle du droit d'auteur, il semble que Facebook soit en train de se bâtir à côté de la réglementation de la vie privée.

Enfin, c'est bien sûr parce que Facebook est une application que les milieux économiques commencent à s'intéresser à son modèle économique. Faire tourner un site de cette taille représente un coût. Les nombreux ralentissements qui affectent Facebook en sont la preuve. Les premières analyses mettaient l'accent sur le potentiel offert par la constitution d'une base de donnée relativement complète des usagers, sur les possibilités offertes par les différents mécanismes d'avatar (achats de cadeaux à 1\$, service premium, etc.), ou sur l'opportunité de faire payer les applications tierces pour améliorer leur qualité de service (afficher les photos du widget slide.com plus rapidement). Face à ces possibilités infinies, l'annonce du modèle économique prôné par Marc Zuckerberg n'a pas manqué de décevoir puisqu'il ne s'agissait ni plus ni moins que d'une mise en application du manuel du targeting marketing : enregistrer le comportement du consommateur, le revendre aux annonceurs et espérer que le consommateur ainsi bien "ciblé" endosserait les produits et irait en faire la promotion à ses amis. Quel que soit l'angle sous lequel on le prenne, le modèle semblait bien trop naïf pour convaincre : ce n'est pas parce que les joueurs de world of warcraft ont pu

élaborer des communautés qui leurs sont propres et qui participent à la valorisation du jeu vidéo de Blizzard qu'on doit croire que le modèle est répliquable si facilement et espérer que les Facebookeurs iront aveuglément faire la promotion du dernier produit de coca cola.

Facebook comme protocole, il ne semble pas vraiment pertinent de remettre en cause l'intérêt des utilisateurs pour les réseaux sociaux. Les problèmes rencontrés aujourd'hui ne fléchissent pas l'intérêt de ces applications. Les chiffres sont éloquentes : 114 millions de membres en juin 2007 pour Myspace, 40 millions de membres pour Facebook. En revanche, tout le monde semble vouloir déterminer la valorisation de ces sites par analogie avec les sites web, c'est-à-dire en faisant une projection de leur capacité à proposer du trafic, des services premium, et une audience la plus qualifiée possible pour les annonceurs.

Cette analogie me semble erronée pour une raison simple, mais difficile à se représenter : Facebook n'est pas un site internet, c'est un protocole. Facebook n'est pas un site web, c'est une infrastructure, l'équivalent d'un logiciel de mail comme Hotmail ou Gmail, sauf que sur Facebook on échange des données personnelles plutôt que des emails.

Autrement dit, il ne faut pas penser Facebook comme si c'était un site internet. Facebook est un standard d'échange de données personnelles entre des individus. Ce standard a même un nom : FBML pour Facebook Markup Language dont Facebook.com n'est que l'application principale. Et comme ils peuvent créer des sites web en HTML, d'autres développeurs peuvent utiliser le FBML pour créer des applications sur Facebook comme "iLike" ou "vampire."

Facebook était considéré comme le standard le plus ouvert, celui qui permettait aux plus de développeurs de construire des applications tierces interopérables entre elles et avec l'application principale. À ce jour, il n'existe pas de standard ouvert à proprement parler, et même si le W3C discute de "social Protocol" depuis 1998, il n'a jamais pu réellement aboutir à une norme qui soit globalement utilisée. Il existe bien des initiatives comme FOAF (friend of a friend) qui permet de décrire les relations entre individus et s'appuie sur RDF (ressource description format, un standard du W3C), mais il n'est pas beaucoup utilisé (à part sur quelques blogs).

Nous ne sommes donc pas face à une guerre entre des sites web sociaux qui essaient d'avoir le plus gros trafic possible. Nous sommes face à la guerre des nouveaux standards du web : qui contrôlera désormais les échanges de données personnelles entre internautes, FBML, FOAF ou Open Social.

La question est d'autant plus importante que le choix semble aujourd'hui se limiter aux standards définis par 3 grands acteurs privés : Myspace, Facebook et Open Social. Aucun de ces formats ne semble vraiment s'ouvrir au public ou à ses représentants. D'une certaine façon, ce qui se joue aujourd'hui, c'est un peu l'équivalent de ce qui se jouait entre AOL et CompuServe avant le croisement de TCP/IP et de HTML au tournant des années 90 : des formats de données concurrents essaient de se tailler un domaine réservé le plus vaste possible pour s'imposer comme standard universel.

A titre d'exemple, il était alors impossible à un utilisateur de AOL de se rendre sur un site CompuServe. De même qu'il est aujourd'hui impossible à un compte Myspace d'importer ses contacts sur Facebook.

Deux solutions sont envisageables. À l'image de ce qui s'est passé dans l'univers de la bureautique, il est possible qu'un standard s'impose grâce aux règles du marché et devienne la norme de fait que les usagers seront obligés d'utiliser. Ou bien, à l'image du web, des standards ouverts émergeront peu à peu pour permettre à chacun de créer des applications et de développer des services interopérables les uns avec les autres.

On ne comprendrait plus aujourd'hui qu'un abonné Orange ne puisse pas envoyer d'email à un abonné Neuf Telecom. Il ne faut pas oublier que beaucoup de gens utilisent maintenant plus leurs messageries Facebook ou Myspace que leur logiciel ou leur boîte mail.

2.2.3.6. PROTECTION DE DONNEES FACEBOOK

Déclaration de protection des données personnelles pour l'utilisation des fonctions de Facebook, lorsque vous visitez les sites, les fonctions établissent un lien direct entre votre navigateur et le serveur Facebook. Facebook est ainsi informé que vous avez visité son site à partir de votre adresse IP. Si vous cliquez sur le bouton « J'aime » pendant que vous êtes connecté à votre compte Facebook, vous pouvez associer les contenus du site à votre profil Facebook. Facebook peut ainsi associer la visite de son site à votre compte utilisateur. Si vous ne souhaitez pas que Facebook associe la visite de nos sites à votre compte utilisateur Facebook, veuillez-vous déconnecter de votre compte Facebook.

2.2.3.7. MEMOIRE CACHE DE FACEBOOK

L'appétit vorace de Facebook pour les données personnelles n'est un secret pour personne. Régulièrement, son estomac numérique gargouille auprès de ses membres pour réclamer son dû. « Donnez-moi votre numéro de mobile, que je puisse renvoyer votre mot de passe en cas d'oubli » « Expliquez-moi vos opinions politiques, que je puisse mieux cibler mes publicités » « Dites-moi ce que vous lisez, ce que vous écoutez, ce que vous regardez, ce que vous cuisinez, que vos amis puissent en profiter » Et la plupart des 800 millions

d'inscrits, dociles, jettent leurs gros steaks de vie privée dans la gueule béante du réseau social. Ce que personne ne semble savoir, en revanche, c'est que le monstre est dépourvu de système digestif. Il se contente de stocker les données, toujours plus de données, sans ne jamais effacer.

Les messages sont sans doute l'aspect le plus scandaleux du dossier. Courriers privés et tchats entre amis sont tous archivés dans le même système de messagerie sur Facebook et y restent pour toujours. Le bouton « Supprimer » n'a qu'une fonction cosmétique : il permet de masquer le message aux yeux de l'internaute, mais reste sur les serveurs de Facebook avec la mention « statut : effacé ».

CONCLUSION PARTIELLE

Dans ce chapitre, nous avons expliqué quelques points importants sur les technologies citées ci-haut et leurs fournisseurs Cloud Computing.

CHAPITRE III : ANALYSE FONCTIONNELLE ET COMPARATIVE SUR GMAIL, YOUTUBE ET FACEBOOK

INTRODUCTION

Dans ce chapitre nous allons présenter par analyse fonctionnelle des différents cas choisis sur les technologies Gmail, YouTube et Facebook, aussi bien présenter la comparaison sur base de différentes capacités de stockages sur le nombre de Datacenter ou serveur.

3.1. ANALYSE FONCTIONNELLE

L'analyse fonctionnelle est une démarche qui consiste à rechercher et à caractériser les fonctions offertes par un produit pour satisfaire les besoins de son utilisateur.

- ✓ L'objet visé par la démarche peut être un objet, un matériel, un processus matériel ou vivant, une organisation, un logiciel, etc.
- ✓ Les besoins sont de toute nature et sont exprimés de façon individuelle ou collective, objective ou subjective, avec des degrés de justification disparates.
- ✓ La ou les fonctions étudiées sont également diverses : fonctions de service, fonctions d'évaluation, fonctions de traitement.
- ✓ Le cadre de l'étude doit être aussi pris en compte : contraintes ou variables déduites de l'environnement, la réglementation, des usages, etc.

Les souci principale d'un internaute est de :

- ✓ Se connecter sur différente plateforme (Gmail, YouTube et Facebook).
- ✓ De se communiquer (partage de ressources qui seront stockées quelques part).

3.1.1. IDENTIFICATION DES COMPOSANTS FONCTIONNELS

Ces différents composants sont basés sur l'environnement client-serveur. Il s'agit d'un modèle logique d'architecture applicative qui vise à modéliser une application comme un empilement de trois couches logicielles (ou niveaux, étages, tiers) dont le rôle est clairement défini :

- ✓ La présentation des données : correspondant à l'affichage, la restitution sur le poste de travail le dialogue avec l'utilisateur ou l'internaute ;
- ✓ L'intergicielle : Le traitement métier des données, correspondant à la mise en œuvre de l'ensemble des règles de gestion et de la logique applicative ;

- ✓ Persistances : correspondant aux données qui sont destinées à être conservées sur la durée, voire de manière définitive.

Dans cette approche, les couches communiquent entre elles au travers d'un « *modèle d'échange* », et chacune d'entre elles propose un ensemble de services rendus. Les services d'une couche sont mis à disposition de la couche supérieure. On s'interdit par conséquent qu'une couche invoque les services d'une couche plus basse que la couche immédiatement inférieure ou plus haute que la couche immédiatement supérieure (chaque couche ne communique qu'avec ses voisins immédiats).

A. DIAGRAMME DE SEQUENCE DES COMPOSANTS D'UNE ARCHITECTURE 3 TIERS

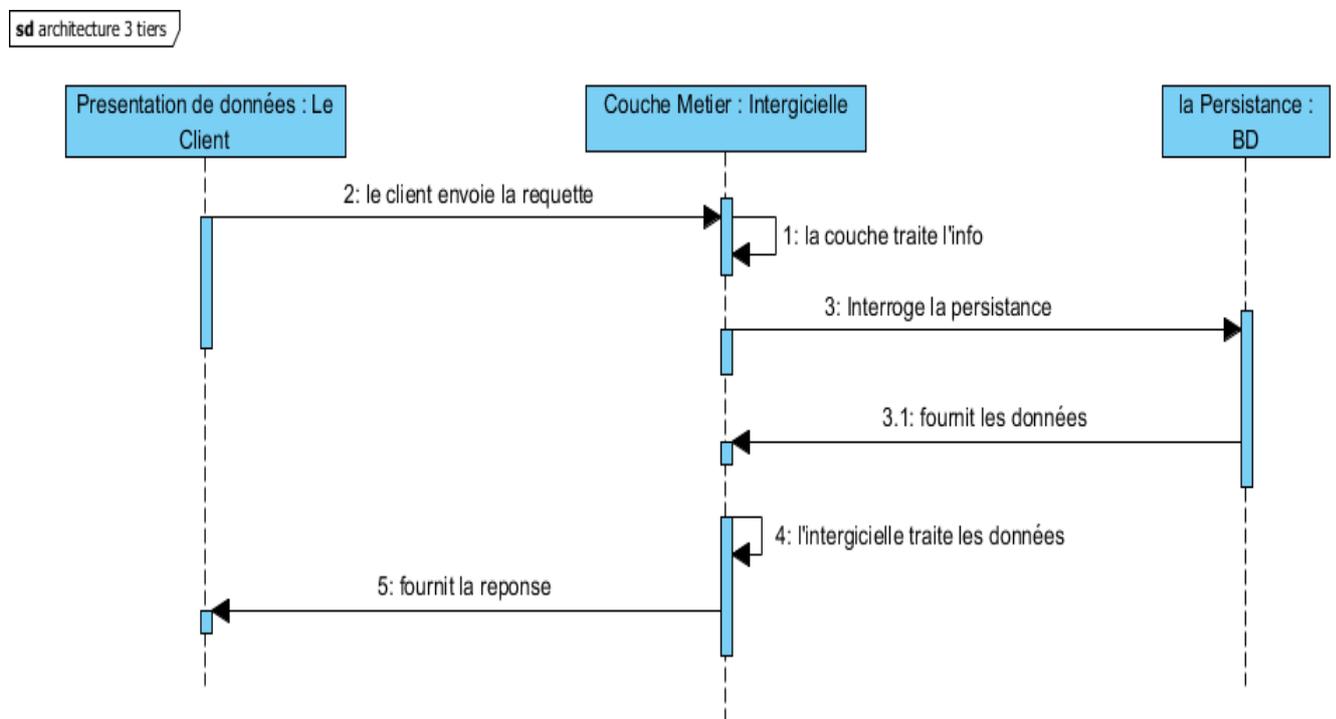


Figure 8 architecture 3 Tiers

3.1.2. FONCTIONNEMENT LOGIQUE DES DIFFERENTES TECHNOLOGIES

3.1.2.1. Serveur de Messagerie

Le fonctionnement du courrier électronique repose sur des ordinateurs puissants, reliés au réseau Internet en permanence, qui assurent les échanges d'informations. Ces machines sont appelées des serveurs, en l'occurrence des serveurs mails. Les courriers sont acheminés d'un poste vers un autre en suivant des protocoles bien précis. Pour mieux comprendre ce cheminement, tentons de présenter l'acheminement de courrier postal :

a. Cheminement du courrier postal :

- Rédiger un message et le glissez dans une enveloppe
- Poster votre courrier via votre bureau de poste
- La Poste envoie le courrier au destinataire
- Trouver votre courrier, à votre adresse, dans votre boîte aux lettres

b. Cheminement du courrier électronique (mail) :

- Rédigez un message sur votre ordinateur ;
- Le mail est envoyé grâce au protocole SMTP (La Poste joue le rôle du serveur SMTP) ;
- Le courrier électronique est envoyé au serveur POP ou IMAP du destinataire ;
- Consulter mail, en se connectant à la boîte aux lettres électronique.

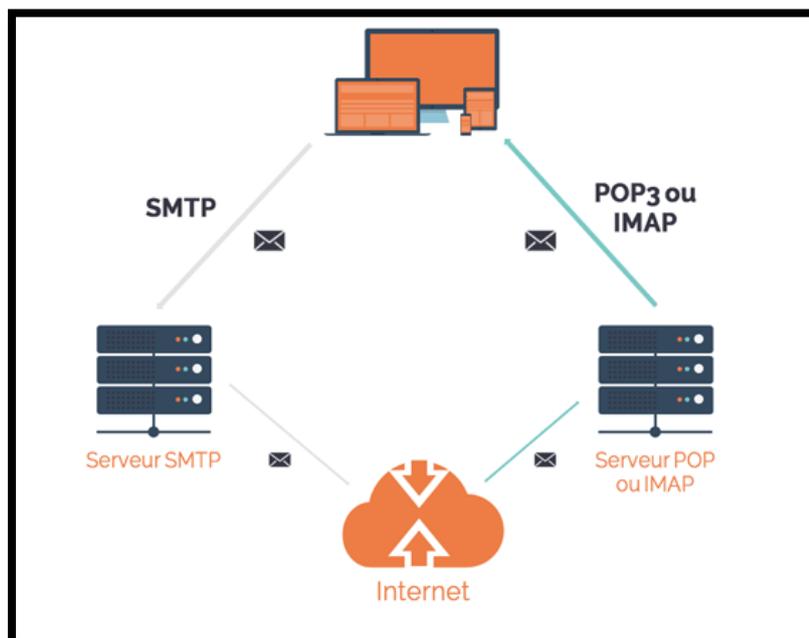


Figure 9 chemin de courrier électronique

A. Diagramme de séquence cheminement du courrier électronique

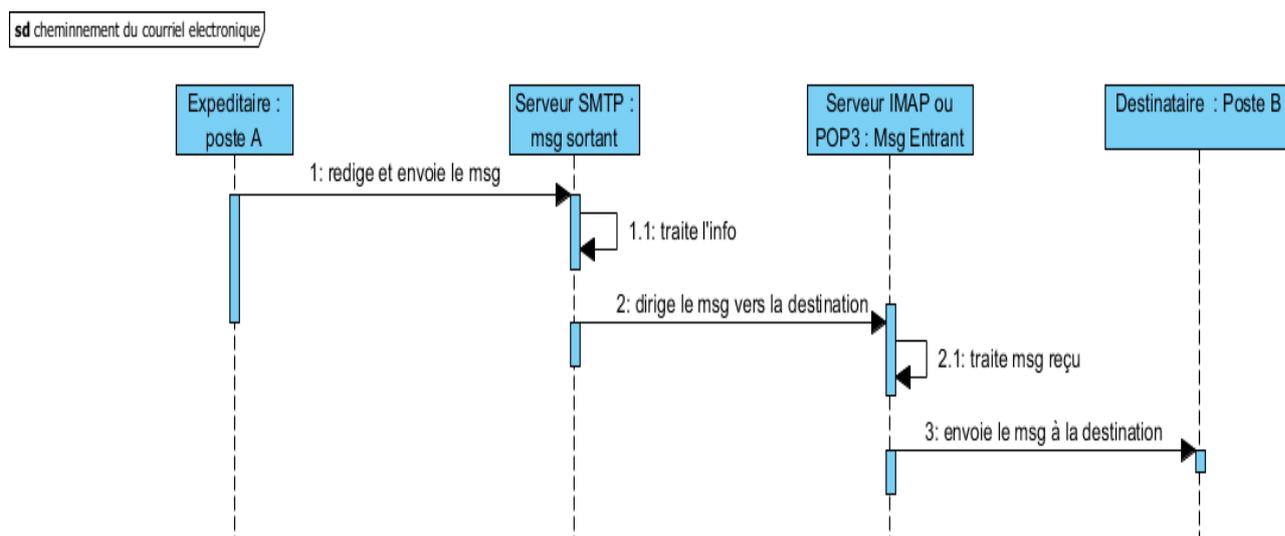


Figure 10 cheminement d'un courrier électronique

➤ Connexion au compte Gmail

Pour se connecter à Gmail, les différents composants qui interviennent :

PRESENTATION	INTERGICIELLE		PERSISTANCE
Interface homme-machine	web application	Interaction avec la base données	Base de données
	Servlet : un programme qui s'exécute étant qu'une extension du serveur	JDBC : permet l'accès à une base de données	
	Le JSP (java server page) : une technologie Java qui fournit les pages web dynamique. sépare la présentation sous forme de code	Hibernate : permet un développement facile de développement de la couche Persistance	

	HTML et les traitements écrit en Java sous forme de Servlet	d'une application	
--	---	-------------------	--

Tableau 2 connexion au compte Gmail

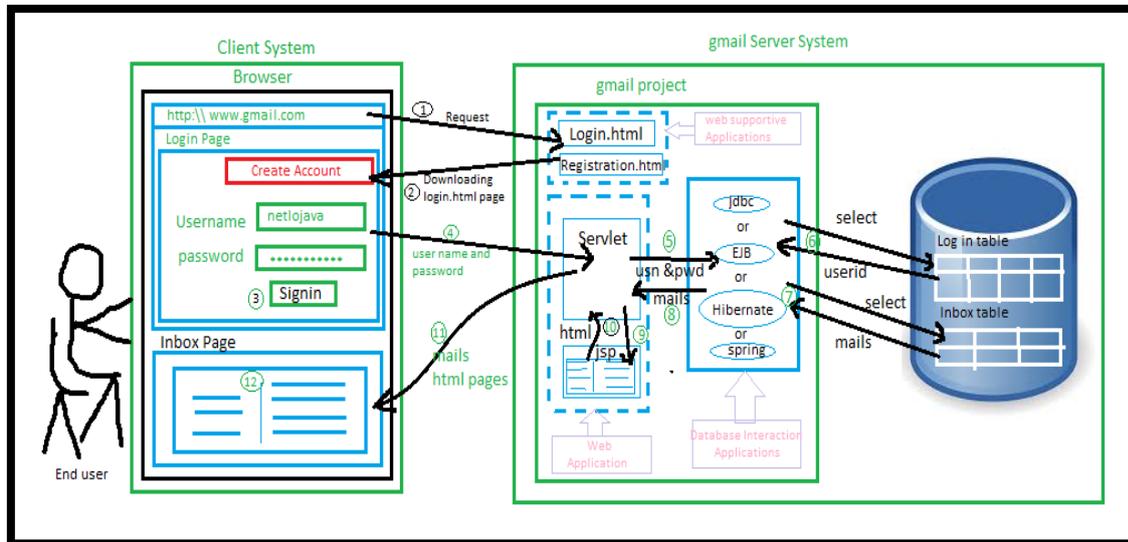


Figure 11 : Système Gmail

B. Diagramme de séquence de connexion au compte Gmail

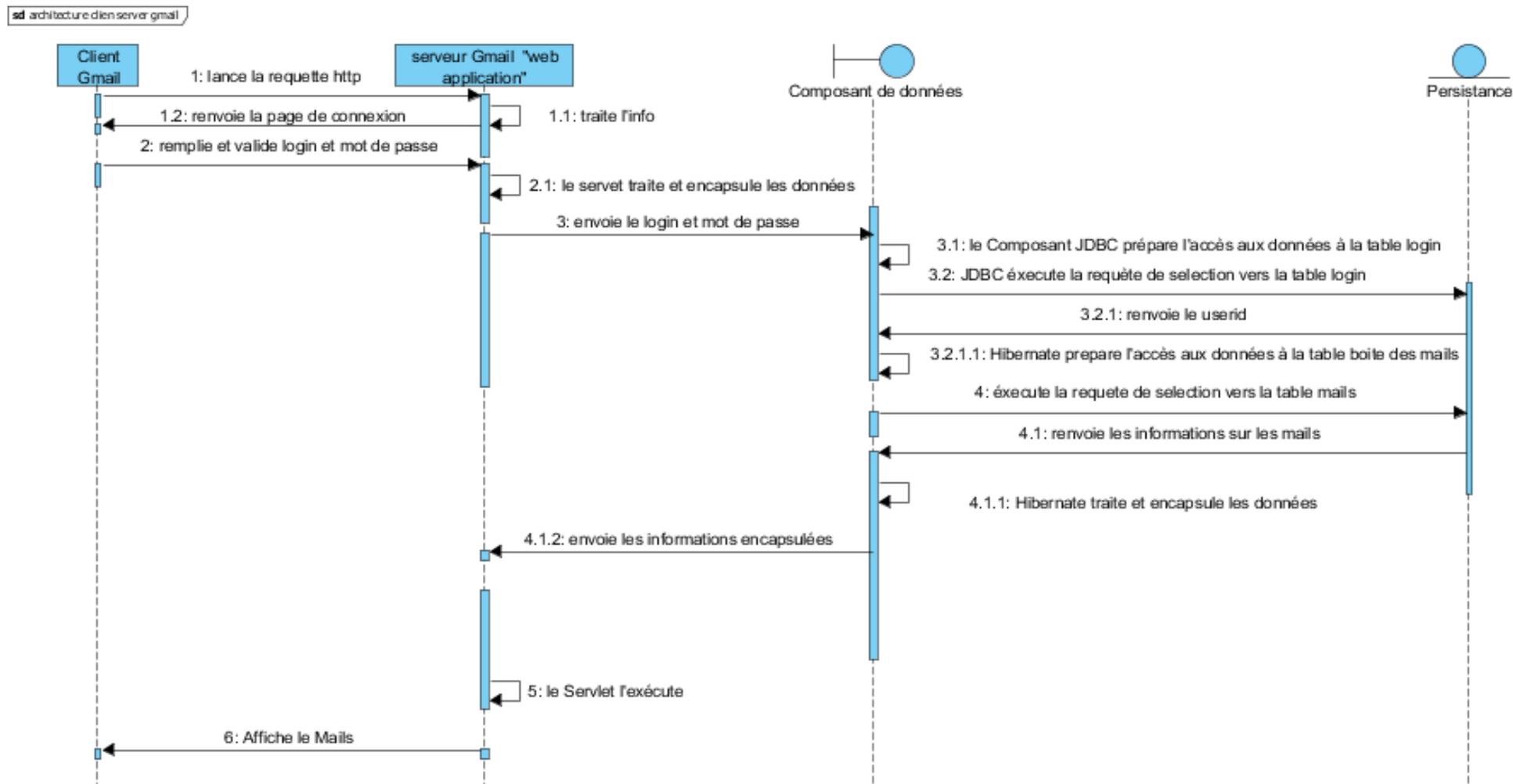


Figure 12 connexion Gmail

3.1.3. SYSTEME FACEBOOK

A. Diagramme de séquence de connexion sur Facebook

sd connexion sur facebook

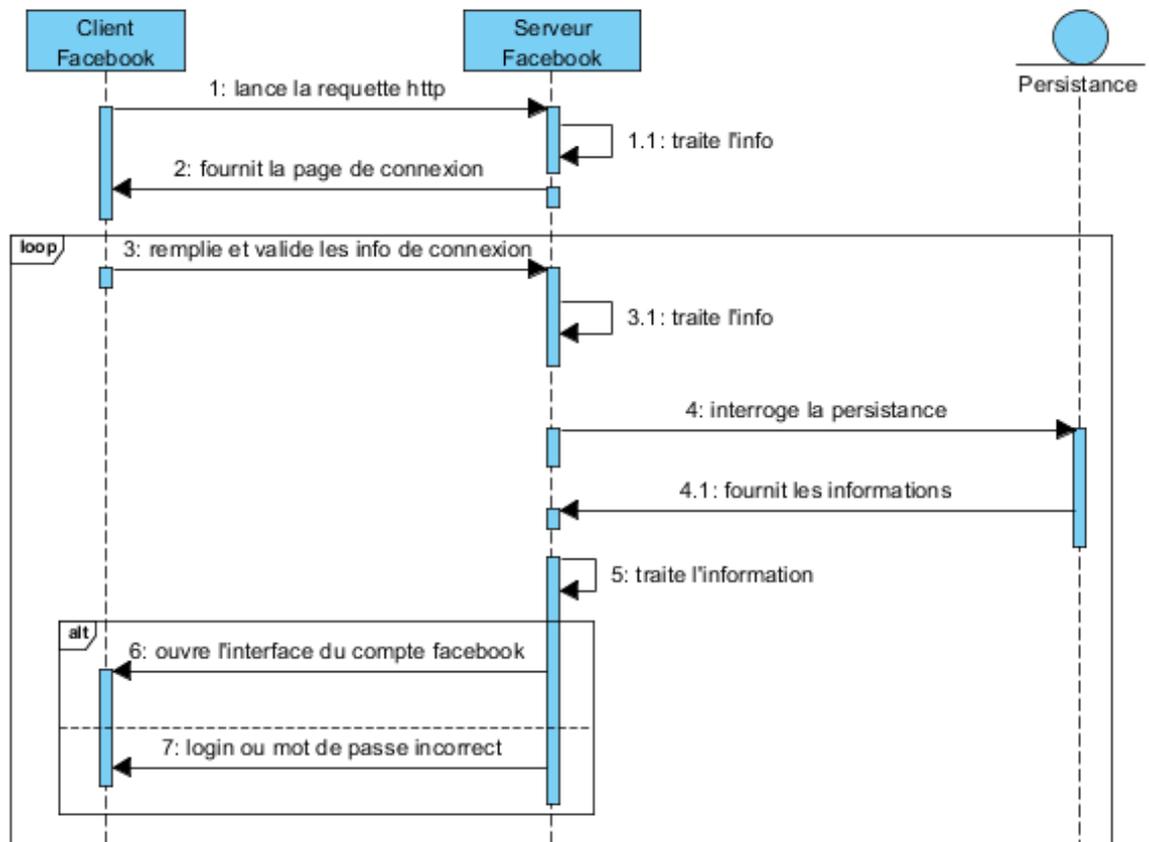


Figure 13 se connecter sur Facebook

3.1.4. YOUTUBE

A. Diagramme de séquence de connexion sur YouTube

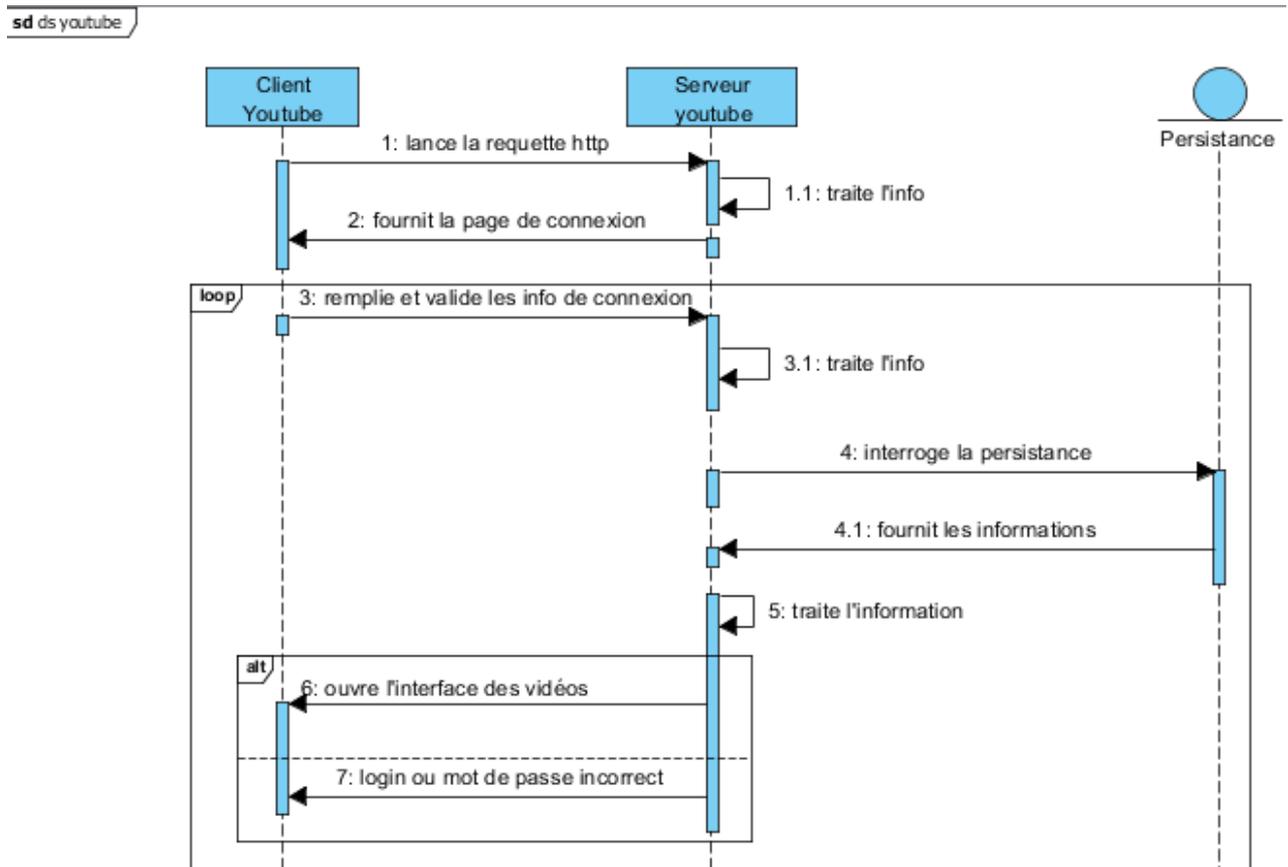


Figure 14 se connecter sur YouTube

3.2. ANALYSE COMPARATIVE

3.2.1. Du point de vue protocole

Le protocole est un ensemble de contraintes permettant d'établir une communication entre deux entités.

Les différents protocoles utilisés par Gmail, YouTube et Facebook sont :

Protocoles	Technologie	Commentaire
Http	YouTube ,Facebook, Gmail	Permet d'échanger les données entre client et serveur web

Zero	Facebook	Permet d'accélérer la connexion sécurisé
RTMP	YouTube	Permet la diffusion de flux de données en streaming (audio, vidéo ou autre)
IMAP,	Gmail	Récupérer l'email et le réorganiser ou en envoyer d'autre sur le compte Gmail avec un logiciel de courrier électronique
HTTPS	Gmail, Facebook, YouTube	Permet de sécuriser la connexion d'interface
TCP/IP	Gmail, Facebook, YouTube	Permet des Transfer des données sur internet
POP 3	Gmail	Récupération d'email sur un logiciel de courrier électronique d'une manière sécurisée au protocole
SMTP	Gmail	Permet d'envoyer les courriers depuis un logiciel de courrier électronique

Tableau 3 : Différentes protocoles

3.2.2. Du point de vue stockage

A. Datacenter Google

Google géant américain de la Mountain View possède plusieurs centres de stockage et de traitement des données.

La topologie du Google, est un réseau constitué de 34 centres de données réparties dans le monde, dont on retrouve 19 aux Etats-Unis (2 autres encours de construction), 11

autres en Europe, 3 autres en Asie ainsi 1 au Brésil (https://fr.m.wikipedia.org/wiki/Plateforme_Google, 2017).

Mettre ses données dans un data center externe permet de sécuriser ses données sans investir dans son propre matériel et sans assumer la maintenance de l'infrastructure. Google avec c'est multitude des services que l'on retrouve de plus en plus dans notre quotidien (Gmail, Google Drive, Android, YouTube, etc). Mais au cœur de Google, nous retrouvons toujours la même constante l'information. Et bien sûr, il faut stocker ces informations.

Sans révéler les informations exactes il semblerait que Google possède aujourd'hui près d'une quarantaine de Datacenter dans le monde entier. Environ une vingtaine sont situées aux USA, le reste étant principalement basé en Asie et en Europe. Bien que les chiffres ne soient pas publiés, certains estiment que Google maintient près de 2 millions de serveurs. Ils assurent en permanence, 7J/7 et 24h/24, le fonctionnement de Google. Son moteur de recherche indexe 20 milliards de pages par jour pour répondre à 3,3 milliards de requêtes quotidiennes et 40 000 par seconde. Mais ce n'est pas du tout Google c'est aussi 425 millions d'utilisateurs de sa messagerie Gmail disposant chacun d'un espace de stockage de 10 Go, soit 4,25 Eo (exa octets). La société compte aussi 250 millions de membres sur le réseau social Google+. Ajoutons aussi 500 millions d'utilisateurs du système d'exploitation mobile maison Android. Autant de données qui demandent une immense capacité de stockage et d'opérations qui nécessitent une grande puissance de calcul. Côté vidéo, YouTube recense 800 millions d'utilisateurs qui passent 4 milliards d'heures à regarder des vidéos chaque mois. Ils envoient aussi 72 heures de vidéos par minute pour 1,3 milliard de vidéos hébergées en tout. Selon WordStream la régie publicitaire Double Click que Google décrit comme sa « *pierre angulaire* », affichait 180 milliards de publicités par mois, ciblant 90% des internautes.

B. Datacenter Facebook

L'architecture réseau pour supporter l'évolutivité sans cesse croissante des besoins du média social. Quand on s'appelle Facebook la question de la topologie réseau prend un sens particulier lors de la construction d'un Datacenter. Les problématiques sont connues, un grand volume de données et de requêtes qui circulent tous les jours, un besoin de haute disponibilité en temps réel, la prise en charge des liens vers des clients et prestataires externes. Le réseau social a donc repensé son architecture traditionnelle de réseau avec l'extension de son Datacenter à Altoona dans l'Iowa.

Ce dernier alimenté par des énergies 100% renouvelables via un parc éolien, a construit son réseau sur la technologie dite « *data center fabric* ». Alexey Andreyev ingénieur réseau chez Facebook, explique dans un blog qu'il faut distinguer « *le trafic* » « *machine to user* » qui se définit comme les requêtes, ou la création de contenu d'applications est très important, il s'agit de la partie émergée de l'iceberg. Le trafic « *machine to machine* » au sein du Datacenter est en croissance exponentielle et le volume double en moins d'une année.

Un saucissonnage en pod de 48 racks serveurs, d'où l'idée d'avoir une approche désagrégée de l'architecture réseau, au lieu d'avoir des clusters avec beaucoup d'équipements réseaux, nous avons cassé le réseau en plusieurs petites unités identiques des serveurs pods et créer une connectivité haute performance entre les pods au sein du Datacenter. La taille de ces petites unités est de 48 racks de serveurs, chaque Pod est relié à 4 commutateurs fabric (spécialement élaboré par Facebook) avec des liens 40 G permettant d'atteindre une capacité de bande passante de 160 G pour un rack de serveurs connectés en 10 G. Une partie software adaptée à l'avantage de cette solution modulaire est de pouvoir la répliquer au sein du Datacenter et de requérir que des commutateurs basiques pour gérer l'agrégation top of rack, cette simplicité autorise plusieurs options de routage via le protocole BGP 4 (le seul retenu par l'entreprise). Dans le même temps, Facebook souligne avoir travaillé sur un contrôleur BGP centralisé qui permet de contourner les chemins de routage via un logiciel, cette approche flexible est dénommée « *DCCO (distributed control, centralized override)* », cela signifie aussi que le réseau social a élaboré son propre logiciel de gestion et de configuration du Datacenter fabric. Quand la firme veut intégrer un nouvel équipement, il est automatiquement reconnu et configuré en cas de problème, cela ressemble au même processus que le commissionnement d'une machine virtuelle. Facebook disposerait 30000 serveurs et 80 milliards d'images

Des chiffres toujours plus impressionnants publiés autour de l'infrastructure du mastodonte des réseaux sociaux qu'est devenu Facebook. Il doit stocker chaque jour 500 To de nouvelles données sur son infrastructure informatique. Un volume de données stocké via un cluster Hadoop d'une capacité de 100 petabytes (<http://www.zdnet.fr/actualites/facebook-500-to-de-donnees-enregistrees-quotidiennement-39775333.htm>, 2012).

3.3. ETUDE COMPARATIVE ENTRE GMAIL, YOUTUE ET FACEBOOK

Technologie	Point convergent	Point divergent
Gmail YouTube Facebook	Ces trois technologies de l'information et de la communication utilisent une architecture 3 tiers	<ul style="list-style-type: none"> ✓ GMAIL : est un service de messagerie électronique permettant d'envoyé les e-mails ✓ YOUTUBE : est un réseau social numérique de contenu c'est à dire partage de vidéos pour lesquels les fonctionnalités de réseau sont secondaire et basés sur une activé particulière. ✓ FACEBOOK : est un réseau social numérique de contact c'est à dire généralistes pour lesquels les fonctionnalités de mise en relation sont principales.

Tableau 4 : Comparaison du point de vue convergence et divergence

➤ Comparaison du point de vue capacité de stockage

TECHNOLOGIE	NOMBRE DE DATACENTER
Gmail et YouTube	17 Datacenter (1 millions de serveurs) soit 50% de Datacenter Google.
Facebook	30000 de serveur de stockage

Tableau 5 : Comparaison du point de vue capacité de stockage

CONCLUSION PARTIELLE

Dans ce chapitre nous avons pu comprendre quelques fonctionnements de différentes technologies de messagerie électronique et de réseaux sociaux. Nous avons présenté quelques diagrammes de séquence et les différents protocoles qui sont utilisé par Gmail, YouTube, Facebook ainsi que la comparaison dans le nombre de Datacenter de stockage de données tout en étudiant les points de convergence et de divergence sur ces différentes technologies d'internet.

CONCLUSION GENERALE

Ce présent mémoire aborde le Cloud Computing comme un modèle d'informatique dématérialisée permettant l'accès via un réseau et sur demande à un ensemble partagé des ressources configurables : réseaux, serveurs, stockage, applications, etc., réparties "en nuages" à divers endroits.

Cependant, un des avantages clés de Cloud Computing vient du fait qu'il peut être utilisé par des entreprises et des individus qui désirent externaliser certains de leurs besoins informatiques vers un fournisseur de nuage. Ce fournisseur peut ensuite fournir des services payant pour l'utilisation de ses clients. Du point de vue financier, cette relation est intéressante parce qu'elle permet à toutes les parties de réaliser des économies d'échelle. En effet, un fournisseur peut acquérir des ressources physiques en grande quantité pour servir ses nombreux clients.

L'objectif de notre travail de fin d'étude, à aborder principalement le fonctionnement sur les technologies d'Internet Gmail, YouTube, Facebook ainsi que la capacité de présenter leurs nombres de Datacenter ou serveurs qu'elles posséderont. Nous avons donné une idée générale sur le Cloud Computing, son architecture, ses avantages, ses caractéristiques principales et ses différents services. Nous avons fait par la suite un état de l'art pour les différentes technologies ci-hauts, en présentant leurs fournisseurs sur l'infonuagique, leur fonctionnement logique en présentant quelques diagrammes de séquence, comparant leurs différents protocoles dans un tableau.

Ce travail était pour nous une belle expérience et la formation sur les technologies d'internet d'en connaître sur le domaine informatique car il nous a ajouté des nouvelles connaissances.

C'est ainsi que nous demandons à tout chercheur scientifique enthousiasmé par cette œuvre de pousser encore les idées plus loin.

Références

- Formation.mdesigner.fr* . (s.d.). Récupéré sur Formation MDesigner: Formation.mdesigner.fr
- Fr.m.wikiversité.org*. (s.d.). Récupéré sur Wikiversité: Fr.m.wikiversité.org
- <http://mysocialselling.com/algorithm-google/>. (2016, 04 25). Récupéré sur Mysocialselling: mysocialselling.com
- <http://www.silicon.fr/facebook-ecarterait-le-cloud-dibm-pour-whatsapp-177305.html>. (2017, 06 08). Récupéré sur Silicon: www.silicon.fr
- <http://www.zdnet.fr/actualites/facebook-500-to-de-donnees-enregistrees-quotidiennement-39775333.htm>. (2012, 08 23). Récupéré sur zdnet: www.zdnet.fr
- https://fr.m.wikipedia.org/wiki/Plateforme_Google. (2017, 05 09). Récupéré sur m.wikipedia: fr.m.wikipedia.org
- https://fr.m.wikipedia.org/wiki/Plateforme_Google. (2017, 05 09). Récupéré sur wikipedia: fr.m.wikipedia.org
- <https://siecledigital.fr/2016/09/08/fonctionne-reellement-algorithme-de-facebook/>. (2016, 09 08). Récupéré sur Siecledigita: siecledigital.fr
- <https://www.developpez.com/actu/84212/-Bluemix-permet-au-developpeur-d-exprimer-toute-sa-creativite-et-construire-rapidement-une-application-entretien-avec-Vincent-Chartier-de-IBM-France/>. (2015, 10 28). Récupéré sur Développez: www.developpez.com
- <https://www.journalducm.com/2017/03/08/seo-algorithmes-google-panda-pinguin-colibri-15229/>. (2017, 03 08). Récupéré sur Journalducm: www.journalducm.com
- m.nouvelobs.com* . (s.d.). Récupéré sur Nouvelobs: m.nouvelobs.com
- www.androidpit.fr/*. (s.d.). Récupéré sur Android pit: www.androidpit.fr
- www.journaldunet.com/solutions/dsi/informatique-de-facebook/le-cloud-computing-chez-facebook.shtml. (2012, 01 12). Récupéré sur Journaldunet: www.journaldunet.com
- www.ouest-france.fr/*. (s.d.). Récupéré sur Ouest France: www.ouest-france.fr
- www.wikipédia.org*. (s.d.). Récupéré sur Wikipédia: www.wikipédia.org

TABLE DES MATIERS

<i>DEDICACE</i>	i
<i>IN MEMORIAM</i>	ii
<i>AVANT-PROPOS</i>	iii
LISTE DES FIGURES	v
LISTE DES TABLEAUX	v
LISTE DES SIGLES ET ABREVIATIONS	vi
INTRODUCTION GENERALE	1
CHAPITRE I : CONCEPTS ET TERMINOLOGIE DU CLOUD COMPUTING	6
INTRODUCTION	6
1.1. HISTORIQUE	6
1.2. DEFINITION	7
1.3. CONCEPTS DE BASE DU CLOUD COMPUTING	8
1.3.1. MODELES DES SERVICES DU CLOUD COMPUTING.....	8
1.3.2. MODELES DE DEPLOIEMENT DU CLOUD COMPUTING.....	11
1.3.2.2. CLOUD PUBLIC	12
1.3.2.3. CLOUD COMMUNAUTAIRE	13
1.3.2.4. CLOUD HYBRIDE	13
1.5. ELEMENTS CONSTITUTIFS DU CLOUD COMPUTING	15
1.5.1. INFRASTRUCTURE.....	15
1.5.2. VIRTUALISATION	15
1.5.3. INTERFACE DE SERVICE	15
1.5.4. DATACENTER (CENTRE DE DONNEES)	16
1.6. DIFFERENCE ENTRE VIRTUALISATION ET CLOUD COMPUTING	17
1.7. DIFFERENCE ENTRE GRID ET CLOUD COMPUTING	17
1.8. AVANTAGE DU CLOUD COMPUTING	18
1.9. INCONVENIENTS DU CLOUD COMPUTING	19
1.10. SECURITE DANS LE CLOUD COMPUTING	19
1.10.1. CONFIDENTIALITE	20
1.10.2. INTEGRALITE.....	20
1.10.3. DISPONIBILITE.....	20
1.11. TECHNOLOGIE DE STOCKAGE EN RESEAU	20
1.11.1. STOCKAGE SAN.....	20

1.11.2.	STOCKAGE NAS.....	22
1.11.3.	STOCKAGE DE DAS	22
1.12.	ARCHITECTURE DU CLOUD COMPUTING	23
1.13.	SERVEUR DU CLOUD COMPUTING	24
1.14.	BIGDATA	25
	CONCLUSION PARTIELLE.....	27
	CHAPITRE II : LES TECHNOLOGIES SUR INTERNET DANS LA MESSAGERIE ELECTRONIQUE ET RESEAUXSOCIAUX.....	28
	INTRODUCTION.....	28
	SECTION 1 : SERVICE DE MESSAGERIE GOOGLE MAIL	28
2.1.1.	PRESENTATION GOOGLE.....	28
2.1.2.	LES ALGORITHMES DE GOOGLE.....	28
2.1.3.	GOOGLE DANS LE CLOUD.....	30
2.1.3.1.	Stockage Gmail et YouTube dans le Cloud	30
2.1.3.2.	Fonctionnement de Google Drive	31
2.1.3.3.	Avantages de la technologie Google Drive	31
2.1.3.4.	Inconvénients de la technologie Google Drive.....	34
2.1.3.5.	Sécurité dans Google Drive.....	34
2.1.4.	ECHANGE DE DONNEES GOOGLE	34
2.1.5.	PROTECTE DE DONNEES GOOGLE	35
2.1.6.	CHIEFFREMENT SUR GOOGLE.....	36
2.1.7.	PRESENTATION GMAIL.....	36
2.2.2.1.	PRESENTATION	40
2.2.2.2.	PROTOCOLE SUPPORTES PAR YOUTUBE	40
2.2.2.3.	METADONNEES DE YOUTUBE	41
2.2.3.1.	PRESENTATION FACEBOOK.....	42
2.2.3.2.	PROTOCOLE SUPPORTE PAR FACEBOOK	42
2.2.3.3.	ALGORITHME DE FACEBOOK.....	43
2.2.3.4.	FACEBOOK DANS LE CLOUD	45
2.2.3.5.	ECHANGE DE DONNEES FACEBOOK	45
2.2.3.6.	PROTECTION DE DONNEES FACEBOOK.....	48
2.2.3.7.	MEMOIRE CACHE DE FACEBOOK.....	48
	CONCLUSION PARTIELLE.....	49

CHAPITRE III : ANALYSE FONCTIONNELLE ET COMPARATIVE SUR GMAIL, YOUTUBE ET FACEBOOK	50
INTRODUCTION	50
3.1. ANALYSE FONCTIONNELLE	50
3.1.1. IDENTIFICATION DES COMPOSANTS FONCTIONNELS.....	50
3.1.2. FONCTIONNEMENT LOGIQUE DES DIFFERENTES TECHNOLOGIES ...	52
3.1.3. SYSTEME FACEBOOK	56
3.1.4. YOUTUBE.....	57
3.2. ANALYSE COMPARATIVE	57
3.2.1. Du point de vue protocole	57
3.2.2. Du point de vue stockage	58
3.3. ETUDE COMPARATIVE ENTRE GMAIL, YOUTUE ET FACEBOOK	61
CONCLUSION PARTIELLE	62
CONCLUSION GENERALE	63
Références	64