MYTHES ET LEGENDES DU CLOUD COMPUTING

18 décembre 2010

Ceci est un extrait du livre collectif "Mythes et légendes des TIC"

Ce livre collectif de l'association forum ATENA propose une approche originale pour expliquer différentes facettes des **T**echnologies de l'**I**nformation et de la **C**ommunication (**TIC**).

En soulignant les croyances largement partagées mais fausses, Ce livre s'adresse à tous les utilisateurs des **T**echnologies de l'**I**nformation et de la **C**ommunication qui désirent acquérir des connaissances et dominer les problèmes posés par l'Information, les systèmes d'information et les réseaux connectés à l'Internet. Ce livre apporte des réponses à leurs interrogations, leurs doutes, combat les idées fausses mais pourtant largement partagées, et donne des conseils pratiques basés sur nos expériences du terrain. Il aborde non seulement les aspects techniques mais aussi les aspects juridiques, humains et organisationnels qui se posent à tous.

La cible n'est pas une population d'experts. Notre livre s'adresse à un lectorat qui cherche des réponses pratiques à des problèmes concrets sans posséder la compétence pour bien appréhender les informations qu'on trouve dans des livres et des revues spécialisées.

Le fichier PDF de la version la plus récente du livre est en téléchargement libre à partir du Web de Forum ATENA en www.forumatena.org/?q=node/12, rubrique "Mythes et légendes des TIC", en laissant votre adresse e-mail.

Gérard Peliks

Président de l'atelier sécurité de Forum ATENA

Coordinateur de cet ouvrage

SOMMAIRE

MYTHES ET LEGENDES DES SYSTEMES DE CLOUD

Professeur Jean-Pierre Cabanel, INP / ENSEEIHT, membre de l'IRIT
Professeur Daniel Hagimont, INP / ENSEEIHT, membre de l'IRIT

Face à l'augmentation continuelle des coûts de mise en place et de maintenance des systèmes informatiques, les entreprises externalisent de plus en plus leurs services informatiques et confient leur gestion à des entreprises spécialisées (que nous appelons fournisseurs). L'intérêt principal réside dans le fait que le client de ces fournisseurs ne paie que pour les services effectivement consommés, alors qu'une gestion de ces services par le client ne serait pas complètement amortie, en particulier lorsque les besoins du client varient. Le « Cloud Computing » se situe dans cette orientation récente.

Devant le manque de consensus sur la définition de la notion de « Cloud Computing », reprenons celle de CISCO : "Cloud Computing is an IT ressources and services that are abstracted from the underlying infrastructure and provided on-demand and at scale in a multitenant environment".

Il s'agit donc de fournir aux clients (des entreprises) des services à la demande, illusion de l'infinité des ressources et enfin d'utiliser les mêmes ressources (mutualisation) pour tous les clients.

Cette stratégie offre plusieurs avantages parmi lesquels :

- Réduction des coûts pour le client. Il n'a plus besoin de gérer sa propre infrastructure et il est facturé en fonction de l'utilisation des services du Cloud.
- Flexibilité pour le client. Il peut augmenter la capacité de son infrastructure sans investissements majeurs, les ressources du Cloud étant allouées dynamiquement à la demande.
- Moins de gaspillage. Les infrastructures gérées chez les clients sont souvent sous-utilisées, alors que l'infrastructure d'un Cloud mutualise un ensemble de ressources pour un grand nombre de clients, ce qui permet d'augmenter le taux moyen d'utilisation des ressources.

Un exemple privilégié de mesure de ce gaspillage est la consommation électrique des infrastructures.

MYTHE N° 1 :
LE CLOUD EST JUSTE CE QU'ON APPELAIT AVANT LE "TIME SHARING" : LES APPLICATIONS NE SONT PLUS HEBERGEES CHEZ SOI ET ON NE PAYE QUE CE QUE L'ON CONSOMME

Le Cloud, c'est un peu plus compliqué. Les utilisateurs potentiels d'un Cloud se regroupent en 3 catégories : administrateur du Cloud, administrateur du client et utilisateur final.

L'administrateur du Cloud est responsable de l'administration des ressources matérielles et logicielles du Cloud. Il est notamment responsable de la gestion de la capacité d'hébergement du Cloud. Le Cloud doit donc fournir à son administrateur des services d'administration lui permettant de gérer les ressources matérielles et logicielles mises à disposition des clients.

Quant à l'administrateur du client, il utilise les ressources fournies par le « Cloud « pour gérer les applications finales du client. Il n'a pas une vue globale de l'environnement du Cloud,

mais seulement des ressources mises à la disposition du client et des applications gérées avec ces ressources.

En fonction du niveau de service fourni par le Cloud, on identifie 3 scénarios d'utilisation du Cloud :

- **Infrastructure as a Service (IaaS)** : Il s'agit du niveau le plus bas. Le Cloud fournit des ressources matérielles à ses clients (capacité de traitement, de stockage ...). Ces ressources matérielles peuvent être fournies directement au client (l'unité d'allocation est alors généralement une machine équipée d'un système d'exploitation) ou être virtualisées (l'unité d'allocation est alors généralement une machine virtuelle, plusieurs machines virtuelles pouvant s'exécuter sur une même machine physique) pour une gestion plus fine des ressources physiques. Pour ce niveau, le Cloud fournit un ensemble d'API permettant à l'administrateur du client d'utiliser un ensemble de ressources. L'administrateur du client a alors la responsabilité d'utiliser ces ressources (machines physiques ou virtuelles) pour y installer et gérer les applications utilisées par le client.
- **Platform as a Service (PaaS) :** Il s'agit d'un niveau intermédiaire dans lequel le Cloud ne fournit pas que des machines et leurs systèmes d'exploitation, mais également des logiciels appelés plateformes applicatives. Ces plateformes sont des environnements d'exécution pour les applications finales comme par exemple : les serveurs d'applications dans une architecture JEE. Ces plateformes applicatives sont maintenues par l'administrateur du Cloud, mais l'administrateur du client a la charge d'administrer les applications finales du client sur ces plateformes applicatives.
- **Software as a Service (SaaS)** : Il s'agit du niveau le plus haut dans lequel le Cloud fournit directement les applications finales à ses clients. L'administrateur du Cloud administre les applications finales et le rôle de l'administrateur du client est quasiment nul. Il est important de souligner qu'un Cloud de niveau SaaS peut être implanté par un acteur en s'appuyant sur un Cloud de niveau PaaS géré par un autre acteur, lui même implanté sur un Cloud IaaS.

MYTHE N° 2 : LE CLOUD COMPUTING EST UNE REVOLUTION TECHNOLOGIQUE

On peut penser que le « Cloud Computing » est une révolution technologique, mais non, c'est une orientation vers un mode de gestion des infrastructures informatiques des entreprises.

En adoptant cette orientation, on retrouve tout les problèmes classiquement adressés dans les infrastructures actuelles, et notamment :

- **La tolérance aux pannes**. Un service géré dans un Cloud soit tolérer les pannes dans le sens où il faut assurer la cohérence de l'état du service en cas de panne ainsi que sa disponibilité pour les usagers. La disponibilité peut être plus difficile à assurer du fait que les services sont déportés dans le Cloud et qu'une indisponibilité de la connexion entre le client et le Cloud peut lourdement affecter la disponibilité du service.
- **La sécurité.** Un service géré dans un Cloud doit résister à des utilisations malveillantes. La sécurité peut être délicate à assurer du fait que le Cloud peut héberger des applications pour le compte de différents utilisateurs (ce qui n'est pas le cas pour une infrastructure interne à l'entreprise cliente). De plus, l'utilisation d'un service nécessite une communication entre le client et le Cloud, ce qui peut constituer un talon d'Achille pour la sécurité.

- **L'interopérabilité et la portabilité.** Les clients des « Clouds » auront vite envie de pouvoir migrer des services d'un Cloud à un autre, ce qui nécessitera l'établissement de standards permettant de tels échanges.

Un problème apparaît toutefois plus crucial dans le domaine du Cloud Computing. Comme on l'a vu précédemment, l'organisation d'un Cloud implique deux administrateurs : l'administrateur du Cloud et l'administrateur du client. L'administrateur du Cloud doit déployer des logiciels (systèmes d'exploitation, machines virtuelles, plateformes applicatives ou logiciels pour l'utilisateur final) sur des machines physiques et les gérer à l'exécution (migration, répartition de la charge) afin d'assurer la qualité de service à ses clients.

L'administrateur du client doit effectuer les mêmes tâches d'administration dans le cas des scenarios PaaS et IaaS. Ces tâches d'administration ne peuvent être effectuées manuellement et une tendance générale est de fournir des environnements d'administration autonomes visant à automatiser au maximum ces tâches (on parle également plus généralement « d'autonomic computing ». Ces environnements d'administration autonome fournissent des formalismes permettant de décrire les actions à effectuer pour déployer des applications et les reconfigurer dynamiquement pour prendre en compte les conditions à l'exécution.

Il existe principalement trois types de système de Cloud et les problèmes de sécurité sont différents suivant la structure utilisée.

1. **Les systèmes privés** propres à un grand compte, avec si nécessaire quelques sous-traitants
2. **Les systèmes partagés** par plusieurs grands comptes
3. **Les systèmes publics**, ouverts à tout le monde

Un système de type Cloud se décompose en plusieurs parties :

- Des postes clients indépendants
- Un système de communication entre le poste client et le système.
- Des bâtiments qui abritent les ordinateurs Cloud
- Des ordinateurs, systèmes d'exploitation et logiciels du Cloud

Chacun des ces éléments est un des maillons de la chaine sécuritaire du système et impacte sur les paramètres suivants :

- Confidentialité
- Authentification
- Déni de service
- Pollution, destruction

La problématique de la sécurité d'un système de Cloud relève d'une tache ardue, et les protections envisagées vont diminuer la potentialité de généralisation d'utilisation de Cloud multiples pour un même client.

De manière induite, la problématique juridique est, elle aussi, très difficile : Qui va être responsable des aléas direct ou indirect qui surviendront ? Comment obtenir la réalité sur les causes des situations ?

Il y a quelques années, les constructeurs de « main frame », DEC, BULL, IBM etc., exploitaient des systèmes identiques au Cloud avec sur le plan sécuritaire plusieurs différences essentielles :

- Très souvent, les clients du point central, appartenaient à une même entité juridique: une banque, une industrie etc.

- Les systèmes de communications utilisés n'étaient pas l'Internet, ils permettaient un contrôle suffisant : lignes et réseaux spécifiques et propriétaires.
- La protection des ressources et la recherche des causes d'aléas étaient simplifiées, une seule entité juridique cliente et des systèmes de communication propriétaires des fournisseurs de « Main Frame » ou centre de ressources informatiques.

La nouvelle approche, modifie l'environnement précédemment présenté : clients avec des entités juridiques multiples, même si ces clients sont connus et identifiables à priori, et utilisation de moyens de communication ouverts et incontrôlables : l'Internet.

MYTHE N° 3 :
LE CLOUD PRIVE D'UN GRAND COMPTE EST COMPLETEMENT SECURISE

Dans les systèmes privés propriétaires d'un grand compte, ce type d'utilisation (très proche des PKI intra entreprise), le système est installé sur le site de l'entreprise et les risques sécuritaires sont minimisés. Ils relèvent de la protection des communications dans l'entreprise (internationales) et du contrôle des personnes et des systèmes dédiés au Cloud. Le responsable vis-à-vis des utilisateurs est alors le service informatique qui gère les services de Cloud. Sommes-nous face à un système qui possède un haut niveau de sécurité ?

Et bien cela n'est pas si clair, il est encore nécessaire de contrôler, les chemins utilisés par l'information afin que des copies illicites ne soient réalisées, de s'assurer de la pérennité du fournisseur du service, afin de ne pas perdre de l'information et ainsi désorganiser l'entreprise, contrôler les communications, etc.

Avec les systèmes réservés à plusieurs grands comptes, nous sommes en présence de la structure la plus exposée aux problèmes sécuritaires. En effet le site physique du Cloud n'est pas sous contrôle de l'entreprise mais contient des informations confidentielles de plusieurs entreprises.

MYTHE N° 4 :
LES INFORMATIONS STOCKEES SUR UN CLOUD PARTAGE SONT PROTEGEES, PAR CONTRAT, DES VIRUS, VERS ET AUTRES ATTAQUES

Les postes clients du système Cloud, utilisent sûrement des supports magnétiques amovibles, (il existe très peu d'application fermée) ou bien le poste client est utilisé pour d'autres travaux, ou dans le cas pire, le poste client est connecté à l'Internet de temps en temps.

Pensez-vous alors que les filtres anti virus du Cloud vont protéger les informations des entreprises clientes ? Et bien non ! En réalité ces filtres possèdent une efficacité toute relative et cela conduit au risque de pollution du Cloud par les virus et autres programmes malveillants positionnés par un client et ainsi polluer ou détruire des informations des entreprises clientes du Cloud

Vous imaginez peut-être, que les données des entreprises peuvent être séparées physiquement sur des machines différentes avec des accès réseaux différents ? Et bien non ! La réalité économique de ces systèmes oblige à mettre en commun les ressources afin de diminuer les coûts pour les clients.

Un fournisseur de systèmes de Cloud peut-il garantir par contrat la non destruction ou pollution des données stockées ? Les notions de virus et vers sont elles assimilées aux forces majeures : nature, guerre etc. ? La pérennité du fournisseur est-elle prise en compte par des clauses spécifiques ? Il semble que si l'on désire garder des coûts acceptables de service de Cloud, il soit très difficile de garantir de telles contraintes.

Pensez vous qu'il est possible, de détecter le client responsable d'une pollution ? Quelles sont les responsabilités partagées du Cloud et du client pollueur ?

Dans un environnement semi ouvert (les clients sont connus), la technique actuelle ne permet pas de protéger de la pollution un site de Cloud, de plus, cette dernière, peut être engendrée par un poste client, qui ne connait pas obligatoirement son propre état de pollution. Il est donc très difficile de remonter au client initial, et les autres clients du Cloud sont alors en droit de se retourner vers le propriétaire du Cloud dans le cas de pollution de leurs données.

De plus des postes clients peuvent eux-mêmes être pollués, par un Cloud pollué par un autre client. Cela montre l'interaction informatique entre des entreprises qui ne se connaissent peut être pas,

Peut être pensez vous que si vous participez à un Cloud, le fournisseur vous garantit un cloisonnement informatique étanche ? Et bien non ! Votre entreprise (vos postes connectés au Cloud) devient une partie de la toile tissée par le Cloud et votre informatique est alors assujettie aux aléas d'autres entreprises.

C'est un des problèmes très important lié au système de type Cloud.

MYTHE N° 5 :
SI VOUS QUITTEZ VOTRE FOURNISSEUR, VOTRE CONTRAT GARANTIT LA CONFIDENTIALITE ET LA RESTITUTION DE VOS INFORMATIONS ET LEUR DESTRUCTION

En dehors des problèmes de confidentialité et d'authentification relatifs aux communications électroniques entre plusieurs sites, les informations (confidentielles ou non) des clients sont stockées chez un tiers. Il se pose alors le problème de la confiance dans le tiers par rapport aux problèmes suivants :

- Accès à des informations de clients par des employés du tiers (espionnage).
- Pénétration du site par autrui qui est ou non un client. (usurpation d'identité)

Même si les informations sont chiffrées sur les machines du Cloud, le chiffrement est propriétaire (algorithme et clef) du Cloud et pas de chaque client. Un chiffrement propre à chaque client avec des clefs différentes pour chaque envoi, minimise les risques d'indiscrétion, mais complique la gestion du Cloud et ouvre la porte à d'autres problèmes.

Comment pensez-vous accorder votre confiance à un fournisseur de service de Cloud ? Quel niveau d'informations confidentielles êtes-vous prêt à confier à autrui ? Pensez vous que par contrat le fournisseur de Cloud va vous garantir la non divulgation en interne, ou par accès extérieur, des informations stockées ?

Ces questions montrent la difficulté d'accorder sa confiance à un fournisseur de service de Cloud, que vous ne contrôlez pas.

Vous pouvez aussi penser à changer de fournisseur de Cloud ou vous pouvez vous retrouver face à la disparition de votre fournisseur.

Alors se pose la question de la récupération de vos données et de l'effacement des informations des supports magnétiques utilisés. Pensez-vous que par contrat, votre fournisseur va vous garantir l'effacement de vos informations, c'est-à-dire la destruction des supports magnétiques ? Il semble peu vraisemblable que vous obteniez cette clause dans votre contrat.

Les systèmes ouverts au public ne peuvent correspondre au monde industriel, y compris aux PME/PMI. Les dangers sont très importants, ils correspondent à ceux relatifs au réseau internet. Aucun contrat ne pourra garantir la sécurité des informations, donc ils ne peuvent être utilisés que pour des informations ou traitement non confidentiels.

Comme les puissances de calcul, les volumes de stockage, les prix des logiciels continuent de s'améliorer, on peut se demander si le grand public nécessite ce type d'offre.

MYTHE N° 6 :

AVEC UN SERVICE DE CLOUD, JE N'AI PLUS BESOIN DE ME PREOCCUPER DE MA SECURITE ET DE LA DISPONIBILITE DES SERVICES, ET MON CONTRAT COUVRIRA LES RISQUES INFORMATIQUES ENGENDRES

Comme tout problème de sécurité, la problématique de l'utilisation de systèmes de type Cloud peut être formalisée par les deux idées antinomiques suivantes :

D'un coté les diminutions de coût engendrées par la mise en commun et la meilleure utilisation des ressources informatiques et, d'un autre coté une augmentation importante des risques d'espionnage, pollution etc. dans le monde informatique.

Avec un service de Cloud Computing, les problèmes de sécurité sont très fortement amplifiés : destruction, pollution, confidentialité etc., et la disponibilité des ressources est assujettie au fonctionnement du réseau. Il est plus facile de sécuriser des informations dans son entreprise que sur un système non propriétaire partagé et utilisable à travers un réseau.

Il est clair, que pour des entreprises stratégiques de taille importante, la notion de Cloud ne peut exister que dans le périmètre de l'entreprise, le Cloud est physiquement installé sur un des sites et les clients appartiennent à un même environnement. C'est une vieille utilisation, même si l'exploitation des calculateurs est confiée à un tiers qui peut être le fournisseur de système Cloud.

Pour des entreprises (PME-PMI) stratégiques, un vrai problème se pose, et l'analyse entre la perte financière engendrée par la copie (espionnage ou destruction) de document, et le gain obtenu par la diminution des coûts journaliers de l'informatique, est très difficile à évaluer et dépend de nombreux facteurs.

L'utilisation des systèmes de Cloud ouverts et gérés par des tiers devient alors limitée à des applications dont le niveau de confidentialité est faible, dans le monde industriel, la dernière molécule, le dernier programme etc. ne se partage pas, et les informations relatives à la comptabilité client sont protégées.

L'utilisation de système de type Cloud pose le problème de la confiance vis-à-vis du fournisseur du Cloud, mais aussi vis-à-vis de ses clients, il manque un gendarme.

Pensez-vous vraiment confier vos informations confidentielles à un tiers, et pensez vous que votre contrat couvrira les risques informatiques engendrés ? L'informatique évolue, les types d'attaques aussi, et un contrat signé à une date, ne peut envisager les évolutions dans les années suivantes.

QUELQUES PLATEFORMES EXISTANTES

Plusieurs plateformes ont émergé dans le domaine du Cloud Computing. Parmi les plus connues, nous pouvons citer :

- **Amazon Elastic Compute Cloud** (EC2) : il s'agit d'une plateforme de type IaaS basée sur les machines virtuelles Linux. EC2 fournit une plateforme de création de machines virtuelles personalisées (AMI pour Amazon Machine Image) et d'exécution de ces machines virtuelles.
- **Google App Engine** : il s'agit d'une plateforme de type PaaS de développement et d'exécution d'applications web. Une quantité de ressources minimum est allouée par la plateforme et peut évoluer en fonction des demandes de l'application.

- **Microsoft Live Mesh** : il s'agit d'une plateforme de type SaaS de stockage d'applications et de données. Elle assure la disponibilité et la synchronisation des données entre tous les équipements du client.

Ces quelques exemples montrent l'implication des grands acteurs. Si le Cloud Computing est plus une orientation stratégique et architecturale qu'une révolution technologique, il est clair que cette orientation risque de bouleverser les infrastructures informatiques de nos entreprises.

!

A PROPOS DES AUTEURS

Jean Pierre CABANEL est Professeur à l'Institut National Polytechnique (INP / ENSEEIHT) de Toulouse et membre du laboratoire IRIT (Institut de Recherche en Informatique de Toulouse), équipe Université. Il anime un groupe de recherche sur le futur des télécommunications. Ses travaux récents traitent de l'autonomie des vecteurs aériens et spatiaux.

Jean Pierre Cabanel est Docteur d'état de l'Université Paul Sabatier (Toulouse) en 1982. Il travaille en premier au sein du laboratoire IBM de Yorktown (USA), avant de retrouver les projets « pilotes »de l'INRIA dans le cadre du laboratoire de l'IRIT. Il anime avec le Professeur Guy Pujolle le « Working Group » 6.4 de l'IFIP sur les LAN et PABX et organise plusieurs congrès au sein de Sup Telecom. Paris.

Il travaille ensuite sur la problématique de la sécurité des systèmes de communication : PKI : Private Key Infrastructure, et TPC : Tierce Partie de Confiance. *jeanpierre.cabanel (at) free.fr*

Daniel HAGIMONT est Professeur à l'Institut National Polytechnique (INP / ENSEEIHT) de Toulouse et membre du laboratoire IRIT (Institut de Recherche en Informatique de Toulouse), où il anime un groupe de recherche autour des systèmes d'exploitation, des systèmes répartis et des intergiciels. Ses travaux plus récents concernent les systèmes d'administration autonomes.

Daniel Hagimont a obtenu un doctorat de l'Institut National Polytechnique de Grenoble en 1993. Après une année postdoctorale à l'Université de Colombie Britannique (Vancouver) en 1994, il a rejoint l'INRIA en 1995 comme Chargé de Recherche. Il a ensuite pris ses fonctions de Professeur en 2005. *daniel.hagimont (at) enseeiht.fr*