

Institut Supérieur des Études Technologiques
ISET de Kairouan

Ce cours s'adresse à tous les étudiants en informatique

Cours Virtualisation et Cloud

Noureddine GRASSA

Kairouan, le 7 août 2014

Table des matières

1	La Virtualisation	4
1.1	Introduction	4
1.2	Les hyperviseurs	5
1.2.1	hyperviseur de type 1	5
1.2.2	hyperviseur de type 2	6
1.3	Les différents types de la Virtualisation	6
1.3.1	Virtualisation Complète	6
1.3.2	Para-Virtualisation	7
1.3.3	Les Isolateurs	8
1.4	Les principales solutions	9
1.4.1	XEN	9
1.4.2	KVM	10
1.4.3	VMware ESX	10
1.4.4	Hyper-V	10
1.4.5	OpenVZ	11
1.4.6	LXC	11
1.5	Domaines d'application	12
1.6	La Virtualisation de Stockage	13
1.6.1	Stockage en Réseau	13
1.7	Conclusion	16
2	Cloud Computing	17
2.1	Définition	17
2.2	Les différents types du Cloud	17
2.2.1	SAAS(Software as a Service)	17
2.2.2	PAAS(Plateform as a Service)	18
2.2.3	IAAS(Infrastructure as a Service)	19
2.3	Les architectures Cloud	19
2.3.1	Le Cloud Privé	19
2.3.2	Le Cloud Public	20
2.3.3	Le Cloud Hybride	21

2.3.4	Avantages et Bénéfices	22
2.4	Les différents acteurs du Cloud	23
2.4.1	Amazon	23
2.4.2	SalesForce	24
2.4.3	Microsoft	24
2.4.4	Google	24
2.4.5	OpenStack	25
2.5	Conclusion	25
	Bibliographie	25

Chapitre 1

La Virtualisation

1.1 Introduction

Un serveur est un ordinateur utilisé à distance depuis différents postes de travail, ou autres serveurs. Il possède des ressources matérielles, principalement CPU, mémoire, disques et interfaces réseau. Ces ressources sont utilisées par des applications, non pas de manière directe, mais en s'appuyant sur un système d'exploitation. La virtualisation de serveurs est un ensemble de techniques et d'outils permettant de faire tourner plusieurs systèmes d'exploitation sur un même serveur physique. Le principe de la virtualisation est donc un principe de partage : les différents systèmes d'exploitation se partagent les ressources du serveur. Pour être utile de manière opérationnelle, la virtualisation doit respecter deux principes fondamentaux :

Le cloisonnement : chaque système d'exploitation a un fonctionnement indépendant, et ne peut interférer avec les autres en aucune manière.

La transparence : le fait de fonctionner en mode virtualisé ne change rien au fonctionnement du système d'exploitation et a fortiori des applications.

La transparence implique la compatibilité : toutes les applications peuvent tourner sur un système virtualisé, et leur fonctionnement n'est en rien modifié. .

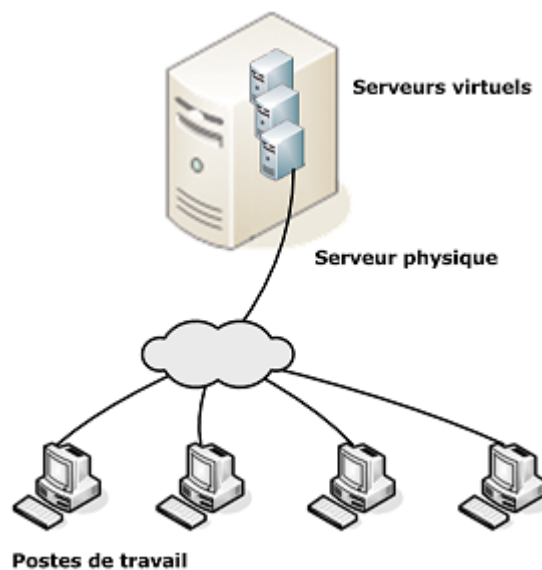


FIGURE 1.1 – Architecture Virtualisée

1.2 Les hyperviseurs

Un hyperviseur est une plate-forme de virtualisation qui permet à plusieurs systèmes d'exploitation de travailler sur une même machine physique en même temps.

1.2.1 hyperviseur de type 1

Un hyperviseur de Type 1, ou natif, voire "bare metal" (littéralement "métal nu"), est un logiciel qui s'exécute directement sur une plateforme matérielle; cette plateforme est alors considérée comme outil de contrôle de système d'exploitation. Un système d'exploitation secondaire peut, de ce fait, être exécuté au-dessus du matériel. L'hyperviseur type 1 est un noyau hôte allégé et optimisé pour ne faire tourner initialement que des noyaux de systèmes d'exploitation invités adaptés et optimisés à cette architecture spécifique, ces systèmes invités ayant "conscience" d'être virtualisés. Sur des processeurs ayant les instructions de virtualisation matérielle (AMD-V et Intel VT), le système d'exploitation invité n'a plus besoin d'être modifié pour pouvoir être exécuté dans un hyperviseur de type 1. Quelques exemples de tels hyperviseurs plus récents sont Xen, Oracle VM, ESX Server de VMware.

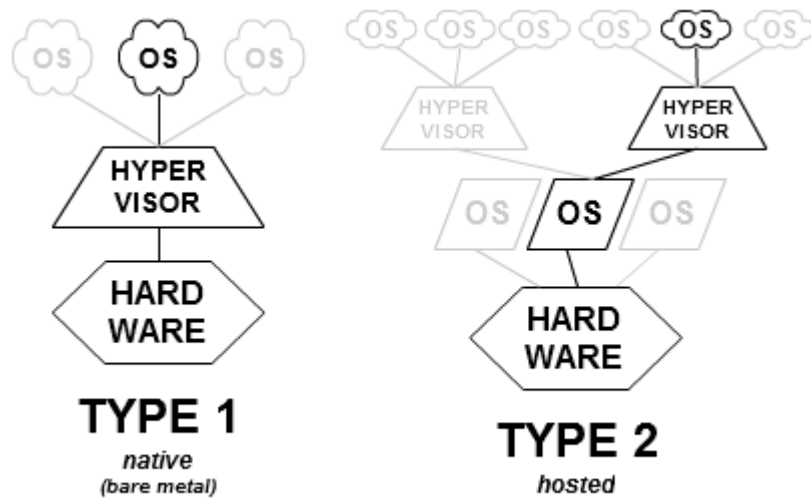


FIGURE 1.2 – Les types des hyperviseurs

1.2.2 hyperviseur de type 2

Un hyperviseur de Type 2 est un logiciel qui s'exécute à l'intérieur d'un autre système d'exploitation. Un système d'exploitation invité s'exécutera donc en troisième niveau au-dessus du matériel. Les systèmes d'exploitation invités n'ayant pas conscience d'être virtualisés, ils n'ont pas besoin d'être adaptés. Quelques exemples de tels hyperviseurs sont VMware Workstation, VMware Fusion, l'hyperviseur open source QEMU, les produits Microsoft Virtual PC et Virtual Server, VirtualBox d'Oracle, de même que Parallels Workstation de SWsoft et Parallels Desktop.

1.3 Les différents types de la Virtualisation

1.3.1 Virtualisation Complète

La virtualisation dite complète permet de faire fonctionner n'importe quel système d'exploitation en tant qu'invité dans une machine virtuelle. Pour l'utilisateur final, ce type de virtualisation est la plus simple à mettre en place et est la plus pratique.

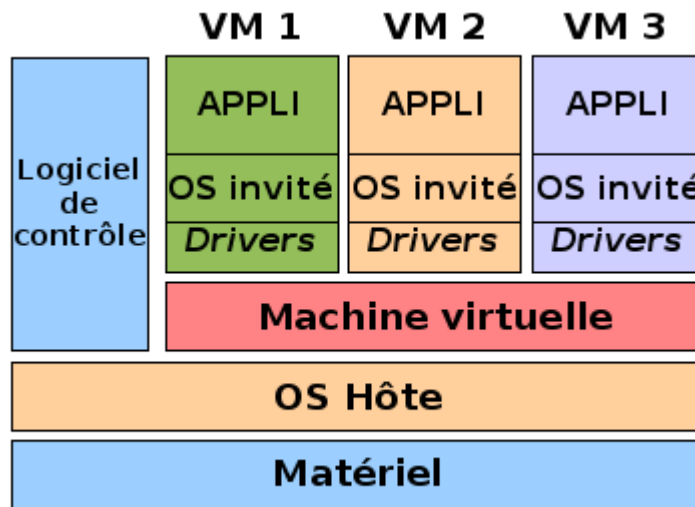


FIGURE 1.3 – Virtualisation Complète

Principe : L'hyperviseur crée un environnement virtuel complet simulant littéralement un nouvel ordinateur complet, avec du "faux matériel". À quelques rares exceptions, le système d'exploitation invité (installé dans la machine virtuelle) ne communique qu'avec ce faux matériel simulé, rendant étanche l'environnement virtualisé.

Limitations : Ce type de virtualisation ne permet de virtualiser que des systèmes d'exploitation prévus pour la même architecture matérielle que le processeur physique de l'ordinateur hôte. Par exemple, un ordinateur équipé d'un processeur Intel x86 sera incapable de virtualiser un système d'exploitation prévu pour fonctionner dans une architecture PowerPC.

Quelques hyperviseurs de virtualisation complète :

- VirtualBox
- VMWare Player, VMWare Workstation
- Parallels Desktop for Windows et Linux
- KVM

1.3.2 Para-Virtualisation

La paravirtualisation fait intervenir un hyperviseur. Il s'agit d'un noyau allégé au dessus duquel viendront se greffer les systèmes invités. Contrai-

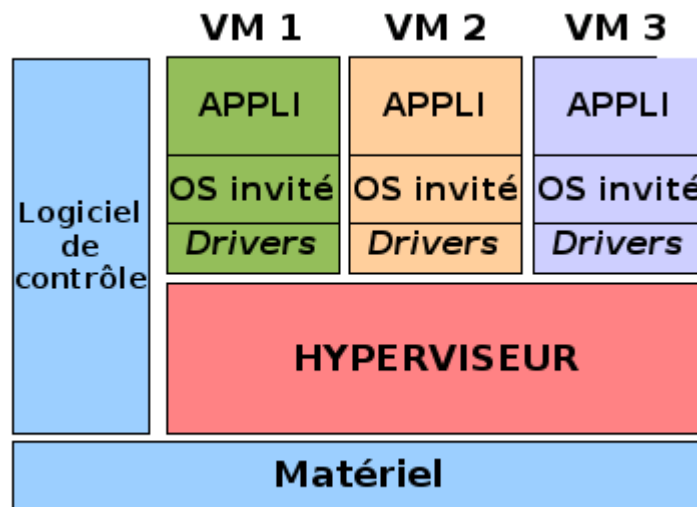


FIGURE 1.4 – Para-virtualisation

rement à un système traditionnel de machines virtuelles où la virtualisation est transparente, avec la paravirtualisation, le système invité doit avoir conscience qu'il tourne dans un environnement virtuel ce qui implique d'employer un noyau modifié. Ce type de virtualisation permet des performances bien plus importantes que la virtualisation totale (assistée par matériel, que nous avons vu plus haut) .

Quelques hyperviseurs de virtualisation assisté :

- XEN
- VMWare ESX/ESXi
- Hyper-V (Microsoft)
- xVM

1.3.3 Les Isolateurs

Un isolateur est un logiciel permettant d'isoler l'exécution des applications dans ce qui sont appelés des contextes, ou bien zones d'exécution. L'isolateur permet ainsi de faire tourner plusieurs fois la même application dans un mode multi-instance (plusieurs instances d'exécution) même si elle n'était pas conçue pour ça. Cette solution est très performante, du fait du peu d'overhead (temps passé par un système à ne rien faire d'autre que se gérer), mais les environnements virtualisés ne sont pas complètement isolés. La performance est donc au rendez-vous, cependant on ne peut pas vraiment

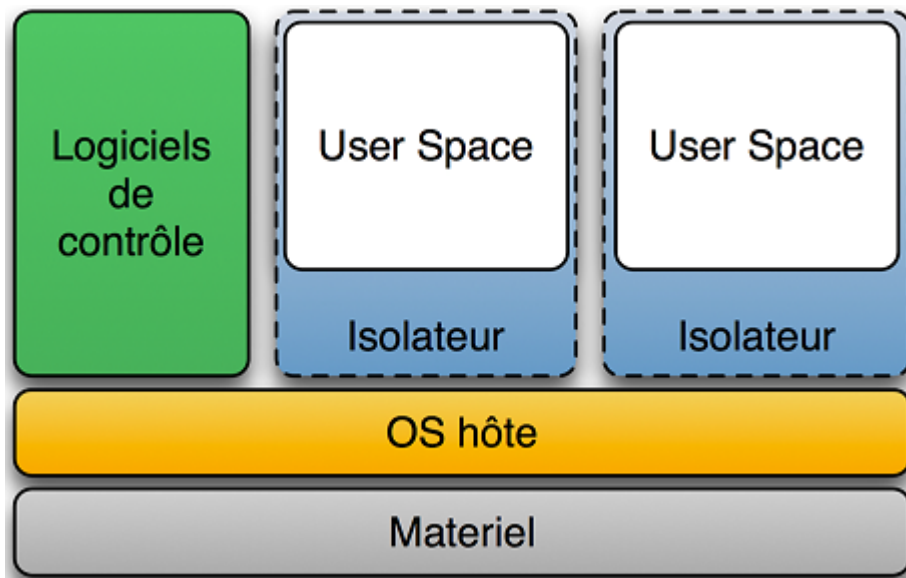


FIGURE 1.5 – Isolateur

parler de virtualisation de systèmes d’exploitation. Uniquement liés aux systèmes Linux, les isoleurs sont en fait composés de plusieurs éléments et peuvent prendre plusieurs formes.

Par exemple : Linux-VServer (isolation des processus en espace utilisateur) ; chroot (isolation changement de racine) ; BSD Jail (isolation en espace utilisateur) ; OpenVZ : libre, (partitionnement au niveau noyau sous Linux) ; LXC : libre, (usage des Cgroups du noyau Linux).

1.4 Les principales solutions

1.4.1 XEN

Xen est une solution de virtualisation open source développée initialement par le département informatique de l’Université de Cambridge. Son développement est aujourd’hui activement sponsorisé par Citrix, qui a racheté l’éditeur initial XenSource. Citrix distribue une version commerciale de Xen, nommée Citrix XenServer, particulièrement adaptée à la virtualisation des OS Microsoft Windows et Linux RHEL et SLES. Elle est dotée d’une interface d’administration avancée, et d’un accès au support technique. Quant aux fonctionnalités, elles sont les mêmes que dans la version distribuée librement. De grandes sociétés comme IBM ont contribué au développement de

Xen, et de gros efforts ont été faits par Citrix pour assurer une compatibilité parfaite avec Windows, compatibilité aujourd'hui reconnue par Microsoft.

1.4.2 KVM

KVM, Kernel Virtual Machine, est intégré depuis le noyau 2.6.20 et permet une virtualisation matérielle et donc une accélération de la virtualisation de système d'exploitation.

C'est un système optimisé pour la virtualisation de serveur. Pour virtualiser des systèmes de type desktop, on peut lui préférer virtualbox. KVM semble en effet plus performant en consommation de processeur mais plus lent pour l'émulation du périphérique graphique. L'utilisation d'un bureau virtualisé dans VirtualBox pourra donc laisser une meilleure impression à l'utilisateur. Vous pouvez tout de même préférer KVM pour sa meilleure compatibilité avec des systèmes d'exploitations anciens ou peu populaires.

Néanmoins, KVM est complètement libre, performant et très facile à installer et à utiliser. L'interface graphique virt-manager pourra aider à paramétrer KVM et pourra rendre la vie plus simple aux administrateurs réseaux. Mais Vous ne pouvez pas utiliser KVM en même temps que VirtualBox. Il faudra en effet fermer KVM pour utiliser VirtualBox et vice versa. Ou désactiver le support de la virtualisation processeur dans VirtualBox.

1.4.3 VMware ESX

VMware vSphere est un logiciel d'infrastructure de Cloud computing de l'éditeur VMware, c'est un hyperviseur de type 1 (Bare Metal), basé sur l'architecture VMware ESXi.

VMware vSphere nécessite une configuration matérielle restreinte précisée dans le guide de comptabilité VMware.

La gestion de ce serveur hôte peut se faire via plusieurs possibilités : par le navigateur Web avec une connexion directe, par une console cliente avec une connexion directe ou par un outil de gestion centralisée nommé VMware vCenter Server qui permet d'administrer l'ensemble des machines virtuelles, des hôtes physiques, de leurs ressources et des options de l'environnement (High Availability, vMotion, Storage vMotion, Distributed Resource Scheduler, Fault Tolerance) depuis une seule console.

1.4.4 Hyper-V

Hyper-V, également connu sous le nom de Windows Server Virtualization, est un système de virtualisation basé sur un hyperviseur 64 bits de la

version de Windows Server 2008.

Il est possible d'utiliser la console Hyper-V sur Windows 7. Dans le sens inverse, de nombreux systèmes d'exploitation peuvent tourner à l'intérieur de Hyper-V :

Bien évidemment pour les systèmes d'exploitation Microsoft Windows 8.1, Windows 8, Windows 7 (sauf édition familiale), Windows Vista SP1/SP2 (sauf édition familiale), Windows Server 2008 x64 SP1/SP2 et R2, Windows Server 2003 x64 SP2 et R2 SP2, Windows 2000 SP4, Windows XP Professionnel SP2/SP3 et x64 SP2

Pour les systèmes d'exploitation linux :

SUSE Linux Enterprise Server 10 SP1/SP2 et 11

Red Hat Enterprise Linux 5.2 x64 et versions ultérieures

Ubuntu 12.04 LTS et versions ultérieures

1.4.5 OpenVZ

Une des solutions les plus avancées et matures dans le domaine de l'isolation est OpenVZ. Ce produit se présente sous la forme d'un patch pour le noyau Linux, et d'un ensemble d'outils d'administration. Le patch du noyau permet à un système GNU/Linux de gérer des contextes virtualisés. Les outils d'administration permettent de créer, d'instancier, et de contrôler les environnements virtuels. Rappelons que la technologie d'isolation ne permet d'exécuter que des serveurs virtuels Linux sur un hôte OpenVZ, même si ces serveurs peuvent être de distributions différentes. Certaines distributions Linux proposent des versions packagées d'OpenVZ. En particulier, la distribution Debian GNU/Linux, dans les versions Lenny et Squeeze, permet dès l'installation du serveur physique de mettre en place cette solution en quelques secondes via son système de packages. Il faut cependant noter que OpenVZ a été remplacé par LXC dans la version Squeeze, paru en 2013.

Le projet OpenVZ fournit aux systèmes GNU/Linux une méthode de virtualisation. Cette virtualisation se situe au niveau du noyau de l'OS. Cela rend possible l'exécution de multiples instances d'OS GNU/Linux sur la même machine. Ces instances fonctionnant de façon complètement sécurisées et partageant intelligemment les ressources du serveur hôte.

1.4.6 LXC

LXC est une solution de virtualisation de type isolateur. Cette solution permet la virtualisation par container au niveau du noyau. LXC est très récent et remplace Linux-VServer et OpenVZ. Aussi, LXC est dès à présent

intégré au noyau, ce qui n'a jamais été le cas des deux solutions citées précédemment.

L'isolateur tire avantage de la possibilité, unique sous UNIX et Linux, de partager le noyau avec d'autres processus du système. Cette virtualisation à noyau partagé utilise une fonctionnalité nommée chroot. Cette fonctionnalité modifie le système de fichiers racine d'un processus pour l'isoler de manière à fournir une certaine sécurité. On parle alors de "jail" ou "chroot".

1.5 Domaines d'application

Voici quelques exemples d'application de ces techniques de virtualisation, dans les domaines où elles sont couramment mises en place.

Les offres d'hébergement étaient traditionnellement distinguées en deux catégories : hébergement dédié et hébergement mutualisé.

Dans un hébergement dédié, le fournisseur met à disposition de son client un ou plusieurs serveurs, configurés selon ses besoins. Selon les cas, le contrat peut prévoir une plus ou moins grande autonomie du client par rapport à la configuration et l'exploitation de son serveur, mais du moins au plan technique, rien ne s'oppose à ce que le contrôle soit total.

Avec un hébergement mutualisé, le fournisseur utilise un même serveur pour plusieurs de ses clients. Il utilise différentes solutions de cloisonnement¹ pour maintenir une certaine étanchéité entre ces environnements.

Le partage de la ressource serveur permet bien sûr un coût très inférieur, particulièrement attractif pour les sites à faible trafic. Mais l'hébergement mutualisé simple a plusieurs handicaps :

- L'allocation des ressources du serveur n'est pratiquement pas contrôlée, de sorte que la qualité de service de chaque site peut être pénalisée par un pic de trafic, ou par la boucle d'un programme sur un autre site.

- La configuration logicielle est unique, et dictée par l'hébergeur. Elle fait le choix, en général, d'un même serveur Http, mais aussi très souvent d'un même outil de gestion de contenus et de base de données. La simple installation de telle ou telle librairie spécifique nécessaire à l'un des clients n'est en général pas possible. Et a fortiori, des configurations globales sur mesure sont interdites.

1. séparation

- En termes d'exploitation, chaque client est extrêmement confiné², de peur qu'il ne perturbe la configuration. Il dispose le plus souvent d'un simple accès en transfert de fichier sur son répertoire privé, et dans tous les cas n'a jamais l'accès root (administrateur) sur le serveur.

Entre ces deux modes d'hébergement, la virtualisation a permis un mode combinant les bénéfices de l'un et de l'autre : le partage de ressources d'une part, l'autonomie et le contrôle d'autre part.

C'est le mode que l'on appelle **VDS** pour *Virtual Dedicated Server*, un serveur dédié virtuel.

1.6 La Virtualisation de Stockage

Dans tout projet de virtualisation se pose, à un moment ou un autre, la question du stockage. Quelque soit la technologie utilisée, une machine virtuelle se compose de deux éléments :

Des ressources : part de CPU alloués, mémoire vive autorisée, nombre de cartes réseau virtuelles...

Des données : comme un serveur normal, on doit disposer d'un système d'exploitation, de bibliothèques, d'outils, d'applications et de leurs données.

Le stockage dépend de la technologie de virtualisation utilisée, et surtout de sa profondeur.

Dans les technologies de machine virtuelle, l'hyperviseur ne fournit au système virtualisé qu'un espace de stockage. Il peut s'agir d'un volume, ou simplement d'un fichier, on peut placer l'intégralité de cet espace sur un disque local, un réseau de stockage ou un autre serveur...

L'utilisation d'un disque local est la solution la plus avantageuse en terme de performances et de facilité d'administration. Cependant, l'utilisation d'un stockage en réseau permet d'ouvrir la voie à de nouvelles fonctionnalités.

1.6.1 Stockage en Réseau

Les pleines capacités des hyperviseurs modernes ne peuvent s'exprimer qu'au travers d'un stockage en réseau, en effet les hyperviseurs sont généralement gérés sous forme de *pools*, formant une *force de travail* globale qui se partageront les machines virtuelles à exécuter.

Cette vision n'est possible que si le stockage est lui aussi unifié : sans cela chaque hyperviseur ne peut faire tourner que les serveurs virtuels présents sur son disque local, et n'est donc pas interchangeable.

2. Enfermer

Disposant d'un réseau de stockage, chaque hyperviseur a accès à toutes les machines virtuelles, et peut donc exécuter n'importe laquelle, et la transférer sans interruption à un autre hyperviseur en fonction de sa charge. Nous allons présenter quelques technologies permettant de mettre en place un réseau de stockage.

NAS et NFS

Un NAS, ou stockage réseau (Network-Attached Storage) est simplement un serveur fournissant leurs fichiers à d'autres serveurs par le réseau.

NFS (Network file storage) est le standard universel pour l'accès aux fichiers sur un réseau, c'est le protocole le plus utilisé dans les NAS.

Dans le cadre d'un isolateur, il permet de stocker l'arborescence du serveur virtuel à distance. Dans le cadre d'une solution de virtualisation complète il permet de stocker à distance les fichiers contenant les disques durs de la machine virtuelle. Ce dernier cas est déconseillé hors des environnements de test : NFS n'est pas adapté à la lecture aléatoire dans un seul fichier. En revanche pour un isolateur, stocker les données en NFS est intéressant, et le deviendra encore plus avec les systèmes de fichier de nouvelle génération tels que ZFS, HAMMER ou btrfs, qui permettent des snapshots instantanés, le versionnement des arborescences, et autres fonctionnalités pour l'instant réservées aux baies de stockage haut de gamme.

En plus des matériels dédiés, la plupart des systèmes d'exploitation proposent une implémentation serveur NFS, ce qui permet d'utiliser n'importe quel serveur comme serveur de stockage NFS. Ces derniers utilisent alors soit des disques locaux, soit leur propre réseau de stockage SAN

SAN

Un SAN, ou réseau de stockage (Storage Area Network), est un réseau sur lequel circulent les données entre un système et son stockage. Cette technique permet de déporter tout le stockage interne d'une machine vers un équipement dédié. Les SAN sont des équipements dédiés, qui ne travaillent qu'aux plus basses couches du stockage, la notion de fichier leur est inconnue ; ils travaillent simplement sur des blocs de données et les fournit par le réseau à des serveurs qui eux sauront les utiliser. Cependant, les SAN les plus hauts de gamme sont dotés de capacités avancées, tel que la prise de cliché, ou encore la copie rapide de volumes. Les deux principaux protocoles d'accès à un SAN sont iSCSI et Fibre Channel.

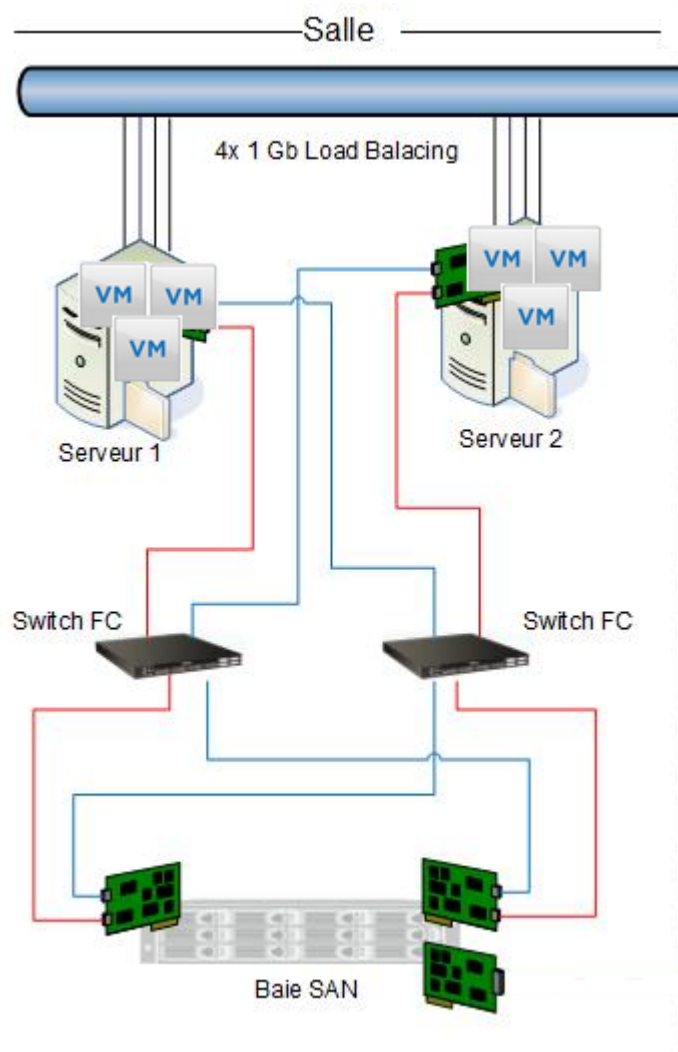


FIGURE 1.6 – Exemple de stockage SAN

iSCSI

iSCSI est un protocole d'accès disque fonctionnant sur un réseau Ethernet, il permet d'implémenter un réseau de stockage en profitant de la connectique et des équipements de commutation standards. Comme le NFS, il peut être soit implémenté par une baie de stockage dédiée, ce qui assure les meilleures performances, soit par un serveur classique disposant du logiciel adéquat, par exemple IET (iSCSI Enterprise Target) sous Linux. Voici un exemple de SAN : parmi les machines clientes du SAN, on retrouve un NAS : ces deux techniques peuvent être combinées car elles ne travaillent pas au même niveau.

Fibre Channel

La solution la plus haut-de-gamme pour implémenter un réseau de stockage est l'utilisation d'une baie dédiée et du protocole Fibre Channel. Basé sur des fibres optiques il assure une latence et un débit bien meilleurs que iSCSI, à un prix bien sûr plus élevé. Son principe d'utilisation est le même qu'un SAN iSCSI.

1.7 Conclusion

Dans le prochain chapitre on verra ce que c'est que Le Cloud Computing. Mais il faut savoir que celui-ci doit son apparition à la virtualisation. En effet la virtualisation est la pierre angulaire de plusieurs Clouds, puisque l'infrastructure virtualisée est indispensable pour assurer la flexibilité et l'évolutivité d'un Cloud.

Chapitre 2

Cloud Computing

2.1 Définition

Le Cloud Computing (informatique dans le nuage) est une technique permettant de gérer des ressources (serveurs) et d'adapter très rapidement une infrastructure à des variations de charge de manière totalement transparente pour l'administrateur et les utilisateurs.

Les applications proposées en mode Cloud Computing ne se trouvent plus forcément sur un serveur informatique hébergé chez l'utilisateur mais dans un nuage formé de l'interconnexion de serveurs géographiquement distincts réalisée au niveau de fermes de serveurs géantes (également appelées datacenters¹). Ceci est rendu possible par le procédé de virtualisation qui consiste à faire fonctionner plusieurs systèmes d'exploitation ainsi que leurs applications associées sur un seul serveur physique. La virtualisation permet ainsi de recréer plusieurs ordinateurs virtuels sur une seule et même machine physique.

2.2 Les différents types du Cloud

2.2.1 SAAS(Software as a Service)

SaaS est l'acronyme pour Software as a Service. Le mode SaaS est un mode d'utilisation d'une solution logicielle que se fait en utilisant l'application à distance qui est hébergée par l'éditeur. Le mode SaaS se rencontre couramment pour des applications logiciels relatives au CRM ou au web-marketing. La solution logicielle étant utilisée, le plus souvent, à partir d'un simple navigateur Internet, elle permet à l'entreprise d'être déchargée de toute

1. centre de traitement de données

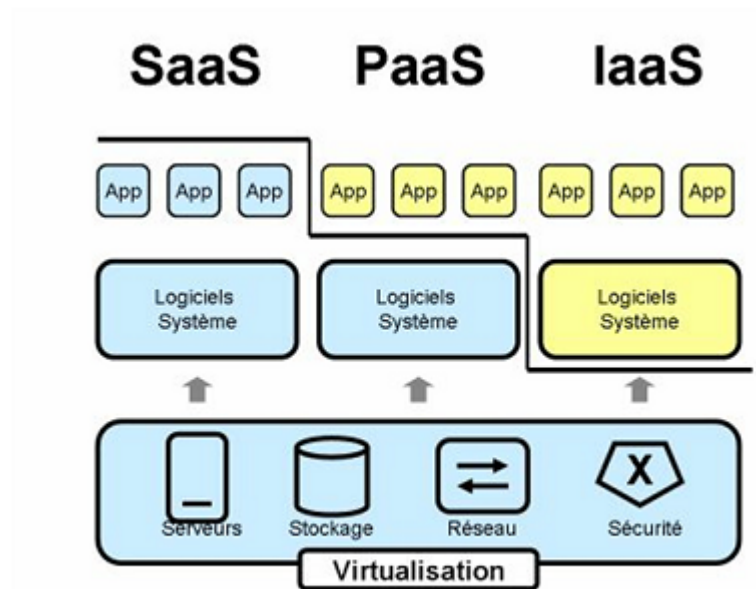


FIGURE 2.1 – Les différents types du Cloud

contrainte d'installation, de mise à jour ou de maintenance technique. Elle permet également d'être utilisée par les collaborateurs en situation de mobilité. La mise à disposition d'une solution SaaS peut être facturée par abonnement ou proportionnellement à l'usage et peut parfois comporter des frais de personnalisation et de mise à disposition du service. Dans le domaine du webmarketing, les plateformes de gestion des campagnes email, les outils de web analytique et les serveurs publicitaires sont généralement proposés en mode SaaS.

Avantage : plus d'installation, plus de mise à jour (elles sont continues chez le fournisseur), plus de migration de données etc. Paiement à l'usage. Test de nouveaux logiciels avec facilité.

Inconvénient : limitation par définition au logiciel proposé. Pas de contrôle sur le stockage et la sécurisation des données associées au logiciel. Réactivité des applications Web pas toujours idéale.

2.2.2 PAAS(Plateform as a Service)

Il s'agit des plateformes du nuage, regroupant principalement les serveurs mutualisés et leurs systèmes d'exploitation. En plus de pouvoir délivrer des logiciels en mode SaaS, le PaaS dispose d'environnements spécialisés au développement comprenant les langages, les outils et les modules nécessaires.

L'avantage est que ces environnements sont hébergés par un prestataire basé à l'extérieur de l'entreprise ce qui permet de ne disposer d'aucune infrastructure et de personnel de maintenance et donc de pouvoir se consacrer au développement.

Avantage : le déploiement est automatisé, pas de logiciel supplémentaire à acheter ou à installer.

Inconvénient : limitation à une ou deux technologies (ex. : Python ou Java pour Google AppEngine, .NET pour Microsoft Azure, propriétaire pour force.com). Pas de contrôle des machines virtuelles sous-jacentes. Convient uniquement aux applications Web. Les cibles sont les développeurs. Google App Engine est le principal acteur proposant ce genre d'infrastructures.

2.2.3 IAAS(Infrastructure as a Service)

Il s'agit de la mise à disposition, à la demande, de ressources d'infrastructures dont la plus grande partie est localisée à distance dans des Datacenters. L'IaaS permet l'accès aux serveurs et à leurs configurations pour les administrateurs de l'entreprise. Le client a la possibilité de louer des clusters, de la mémoire ou du stockage de données. Le coût est directement lié au taux d'occupation.

Avantage : grande flexibilité, contrôle total des systèmes (administration à distance par SSH ou Remote Desktop, RDP), qui permet d'installer tout type de logiciel métier.

Inconvénient : besoin d'administrateurs système comme pour les solutions de serveurs classiques sur site.

Les cibles sont les responsables d'infrastructures informatiques. Amazon EC2 est le principal qui propose ce genre d'infrastructures. **OpenStack** est un exemple d'infrastructure.

2.3 Les architectures Cloud

2.3.1 Le Cloud Privé

Le Cloud privé est prévu pour offrir aux entreprises, des services qui leur permettront de mieux maîtriser leurs ressources informatiques. L'entreprise peut gérer son infrastructure en solitaire, au rythme de ses besoins, ou passer par la mutualisation. Ce type de stockage peut être ainsi considéré d'une part, comme l'installation d'un réseau informatique propriétaire. D'autre part, on peut le désigner comme un centre de données qui fournit des services hébergés pour un certain nombre d'utilisateurs. Il leur délivre des accès,

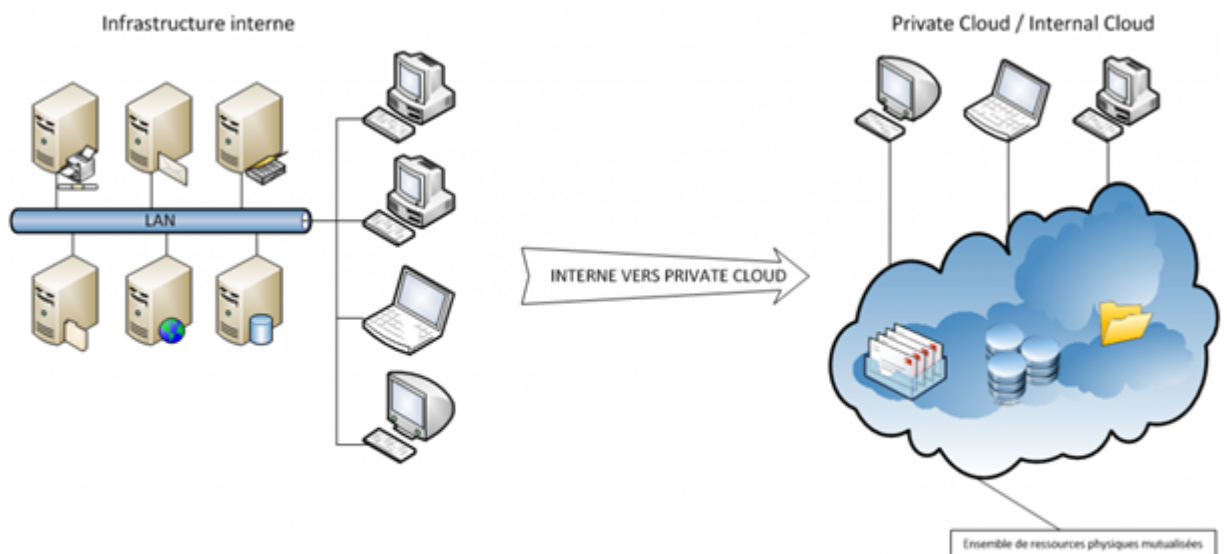


FIGURE 2.2 – Cloud privé

particulièrement favorables pour leurs applications hébergées.

On distingue les Cloud privés internes, utilisés par une entreprise pour satisfaire ses propres besoins. Ils sont administrés en interne par l'entreprise même. Il y a aussi les Cloud privés externes, destinés à satisfaire les besoins propres d'une entreprise cliente. Leurs gestions sont confiées à un prestataire extérieur. On parle alors de gestion externalisée. Dans ce cas, une partie de ses services externalisés, est prise en charge par un prestataire de confiance.

2.3.2 Le Cloud Public

Les ressources sont fournies par un prestataire et mutualisées pour un usage partagé par plusieurs clients. L'infrastructure est mise à la disposition du grand public (c'est-à-dire de plusieurs entreprises) mais elle appartient à un fournisseur de services informatiques, le niveau de service étant défini par le fournisseur et identique pour tous les utilisateurs.

Les exemples les plus évidents de Cloud computing ont tendance à tomber sous le modèle de Cloud public, car ils sont, par définition, accessibles au public. Les offres de Software as a Service (SaaS), comme le stockage Cloud et les applications office en ligne, sont peut-être les plus connues, mais les offres largement disponibles d'Infrastructure as a Service (IaaS) et de Platform as a

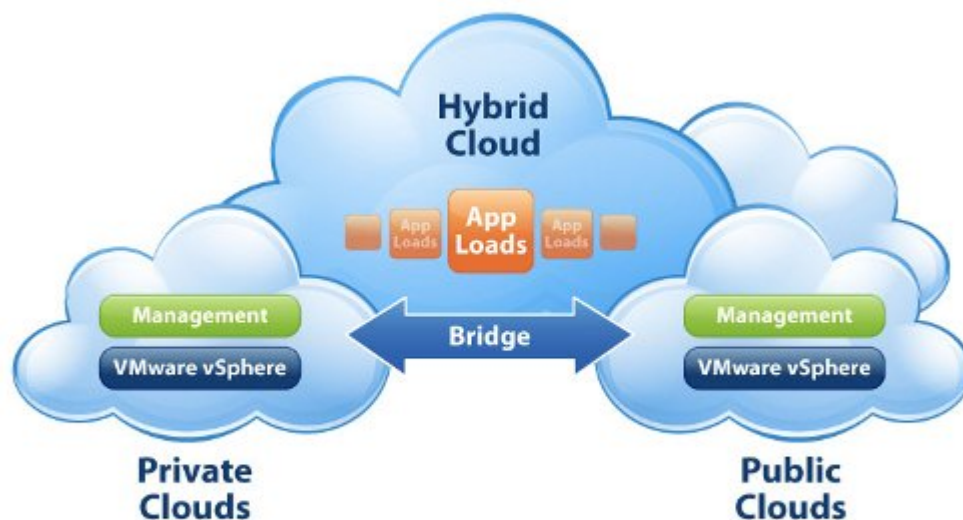


FIGURE 2.3 – Cloud hybride

Service (PaaS), qui incluent de l'hébergement web et des environnements de développement basés sur le Cloud, peuvent également correspondre à ce modèle (bien que toutes puissent exister au sein de Clouds privés). Les Clouds publics sont largement utilisés dans les offres adressées aux personnes privées, moins susceptibles d'avoir besoin de l'infrastructure et de la sécurité de Clouds privés. Toutefois, les entreprises peuvent toujours avoir recours au Cloud public pour rendre leurs opérations plus efficaces, par exemple pour le stockage de contenu non-sensible, la collaboration avec des documents en ligne et la messagerie web.

2.3.3 Le Cloud Hybride

Un Cloud hybride est un service Cloud intégré utilisant à la fois des Clouds privés et des Clouds publics pour remplir différentes fonctions au sein d'une même organisation. Si tous les types de services Cloud sont sensés offrir un certain niveau d'efficacité, à des degrés divers, les services de Cloud public sont susceptibles d'être plus avantageux au niveau des coûts et plus évolutifs

que des Clouds privés. C'est pourquoi une organisation peut maximiser son efficacité en utilisant des services de Cloud public pour ses opérations non-sensibles, et s'appuyer en revanche sur un Cloud privé lorsqu'elle en a besoin, faisant en sorte que toutes ses plateformes soient intégrées harmonieusement

2.3.4 Avantages et Bénéfices

Le Cloud Computing offre les fonctionnalités et avantages suivants :

Une grande extensibilité : les ressources Cloud sont disponibles à la demande depuis les vastes pools de ressources du Cloud public, afin que les applications qu'elles supportent puissent répondre de manière flexible aux fluctuations d'activités.

Avantages en termes de coûts : les Cloud publics rassemblent de plus grandes quantités de ressources afin de pouvoir bénéficier de vastes économies d'échelle. L'exploitation et la gestion centralisées des ressources sous-jacentes sont partagées entre tous les services Cloud alors que les composants, comme les serveurs, nécessitent moins de configuration personnalisée. Certaines offres figurant sur le marché sont même gratuites pour les clients, comptant sur la publicité pour alimenter leurs revenus.

Modèle de prix basé sur l'utilisation : les services de Cloud public utilisent souvent un modèle de tarification basée sur l'utilisation, selon lequel le consommateur peut accéder aux ressources dont il a besoin, quand il en a besoin, et ne paie que ce qu'il a effectivement utilisé, évitant ainsi un gaspillage de capacités.

Fiabilité : la grande quantité de serveurs et de réseaux impliqués pour créer un Cloud public et les configurations redondantes signifient qu'en cas de panne de l'un des composants, le service Cloud peut continuer à fonctionner sur les autres composants. Dans les cas où les Clouds tirent leurs ressources de nombreux data centres, un data centre entier peut tomber en panne et les services Cloud n'en seront pas affectés. En d'autres mots, il n'existe aucun point unique de défaillance susceptible de rendre un service de Cloud public vulnérable.

Flexibilité : il existe une multitude de services IaaS, PaaS et SaaS disponibles sur le marché et qui répondent au modèle de Cloud public, prêts à être utilisés en tant que services et accessibles depuis tout équipement disposant d'une connexion à Internet. Ces services peuvent remplir la plupart

des exigences en matière informatiques et fournir leurs avantages tant aux clients privés qu'aux entreprises. Les entreprises peuvent même ajouter des aspects de Cloud privé à leurs services de Cloud public en créant un Cloud hybride, notamment si elles doivent exécuter des fonctions sensibles.

Indépendance de l'emplacement : la disponibilité des services de Cloud public par connexion Internet assure que les services sont disponibles où que se trouve le client. Cela représente un nombre immense d'opportunités pour les entreprises, notamment l'accès distant à l'infrastructure IT (en cas d'urgence etc.) ou la collaboration avec des documents en ligne depuis de multiples emplacements.

2.4 Les différents acteurs du Cloud

Amazon, Microsoft, Google, Salesforce, Oracle, IBM, OpenStack : autant d'acteurs majeurs dont il faut parler pour comprendre les directions que prendra le Cloud Computing d'ici peu. Car, qu'on le veuille ou non, le Cloud dépend en grande partie de ces géants du web et de l'informatique, mais également de nos équipements de télécommunications.

2.4.1 Amazon

En 2002 Amazon, le site marchand de renommée internationale, lançait "Amazon Web Services". Le 'fait divers' relatant la location des ressources non-utilisées par le site-marchand à des entreprises, a donné l'idée à la société de Seattle de mettre en place un ensemble de services web destinés à plusieurs types de clients. Le catalogue de services s'est enrichi avec le temps, et on ne compte pas moins d'une quinzaine d'offres de services à l'heure actuelle. L'offre la plus connue est certainement Amazon Elastic Compute Cloud, ou EC2, qui permet un déploiement de machines virtuelles directement par le client, de manière automatisée, et en fonction du besoin, d'où le terme élastique. Les possibilités sont multiples et la technologie Xen, évoquée au chapitre sur la virtualisation, permet de définir chaque machine virtuelle comme un serveur virtuel privé. Les coûts sont calculés en fonction du temps d'utilisation. La localisation géographique semble également jouer un rôle dans la variation du prix à l'heure.

2.4.2 SalesForce

Le pionnier des offres Cloud Computing SalesForce, ou plutôt Salesforce.com est une société qui a lancé en 2003 des offres de Cloud Public. C'est donc officiellement le plus ancien prestataire dans ce domaine. Aujourd'hui encore, leurs offres sont uniquement composées de Cloud Public, et adressées aux entreprises (surtout les grands comptes). Sur leur site, on constate que les outils proposés sont résolument tournés vers le travail collaboratif, la gestion des ventes et le marketing relationnel. Certes, le grand public peut utiliser certains outils dans certaines conditions, mais l'offre gratuite est plutôt limitée, ce qui risque d'en arrêter plus d'un.

2.4.3 Microsoft

Office 365, c'est un peu la transformation de la suite bureautique Microsoft en version *à l'Online*. L'objectif est simple : concurrencer la suite bureautique en ligne de Google, et reconquérir les clients perdus au profit de Google. Microsoft entend adapter ses prix en fonction du client, à savoir que les petites structures paieront un prix plus adapté à leur structure. Là encore, on ne s'adresse pas vraiment au grand public.

Windows Azure, la seconde offre commerciale proposée par Microsoft, peut être vue comme la base d'une solution SaaS. Avec Windows Azure, Microsoft met en place un *à l'Os in the cloud*, et assure donc une couche qui va avoir comme rôle de structurer l'ensemble des technologies associées, que ce soit SQL, .NET, etc. Mais encore et toujours, Windows Azure reste *à l' Clairement destiné aux entreprises*, bien que les services Hotmail, MSN ou Xbox Live puissent être dans un avenir proche migrés sur la plate-forme.

Pour résumer, Microsoft met à disposition depuis très peu de temps des outils susceptibles de reprendre des parts de marché à Google (en adaptant ses offres aux petites entreprises), et de conquérir des entreprises désireuses d'opter pour une solution *à l' fraîche* de Cloud Computing, mais aucun véritable signe de changement majeur dans le quotidien du commun des utilisateurs grand public.

2.4.4 Google

Lancé en 2008, Google AppEngine est disponible uniquement en Cloud Public, sous la forme d'une offre gratuite. Sa grande force réside dans le fait qu'il soit justement gratuit, et propose des applications de qualité, malgré le très grand nombre de requêtes effectuées sur les serveurs.

Certes, des limitations techniques existent comme des restrictions sur les API (Application Programming Interface, ou Interface de programmation), la durée des requêtes, etc. mais que peut-on reprocher à des services ouverts au plus grand nombre, et gratuits. Notons tout de même que les GoogleApps sont un cran en dessous des "Office Web Apps" de Microsoft (sur lesquelles nous reviendrons), et que ces dernières bénéficient d'un fait considérable : l'habitude d'utilisation des consommateurs. Aussi, il est regrettable que Google ne propose pas d'offre adaptée aux "Grands Comptes", ceci permettant certainement de gagner en notoriété dans ce secteur, et de s'engager sur le terrain des entreprises, ce qui favoriserait la concurrence. Il convient toutefois de souligner que certaines start-ups utilisent ces outils.

Google semble donc vouloir mettre en place une stratégie "par le bas", en s'attaquant au grand public. Sur ce terrain, la firme de Mountain View semble bien partie, malgré des fonctionnalités trop basiques et en retrait par rapport aux Office Web Apps.

2.4.5 OpenStack

OpenStack est un logiciel gratuit beaucoup plus complet et ambitieux que les autres. Son but est de fournir toutes les briques nécessaires à la mise en place d'un service IaaS : Infrastructure as a Service. L'IaaS est l'une des facettes du Cloud Computing : il s'agit de fournir des ressources informatiques sous forme de machines virtuelles prêtes à l'emploi, et ce directement à la demande des utilisateurs, sans passer par la mise en place à la main des VM par un administrateur système. Le système IaaS gère la mise à disposition de la ressource en fonction de la capacité actuelle de la plateforme, et assure le suivi de consommation à des fins de facturation. OpenStack, développé initialement par Rackspace et la NASA, propose un regroupement de logiciels open source sous licence Apache. Ce regroupement permet de mettre à disposition des ressources de calcul (des machines virtuelles) ainsi que des espaces de stockage, en reprenant les principes de l'IaaS. Historiquement, Rackspace contribuait à la partie stockage d'OpenStack tandis que la partie calcul était développée par la NASA. À l'heure actuelle la fondation en charge du projet, nommé OpenStack Consortium, intègre plus de 150 entreprises, dont Canonical, Dell et Cisco.

2.5 Conclusion

Bibliographie

- [1] D.BRULEY. *Vmware : Vsphère et View*. ENI, 2ème Edition.
- [2] *Linux Magazine : Testez le Cloud*. GNU/Linux Magazine 142.