La virtualisation

Introduction

Qu'est-ce que la virtualisation ?

La virtualisation consiste, en informatique, à exécuter sur une machine hôte, dans un environnement isolé, des systèmes d'exploitation. On parle alors de virtualisation du système ou bien la virtualisation des applications, on parle alors de virtualisation applicative.

Ces ordinateurs virtuels sont appelés serveur privé virtuel (Virtual Private Server ou VPS) ou encore environnement virtuel (Virtual Environment ou VE). La virtualisation permet de faire fonctionner sur une seule machine physique plusieurs machines virtuelles avec des systèmes d'exploitation différents.

Une machine virtuelle permet de faire fonctionner des systèmes d'exploitation invités avec un logiciel qui fonctionne un système hôte. Ce procédé permet aux systèmes d'exploitation invités de dialoguer directement avec le matériel. Ainsi un système d'exploitation standard exécute une application en charge d'émuler une machine virtuelle au sein de laquelle il est possible de faire tourner un système d'exploitation invité. Cette machine émule un ou plusieurs processeurs, un espace de mémoire et les entrées-sorties.

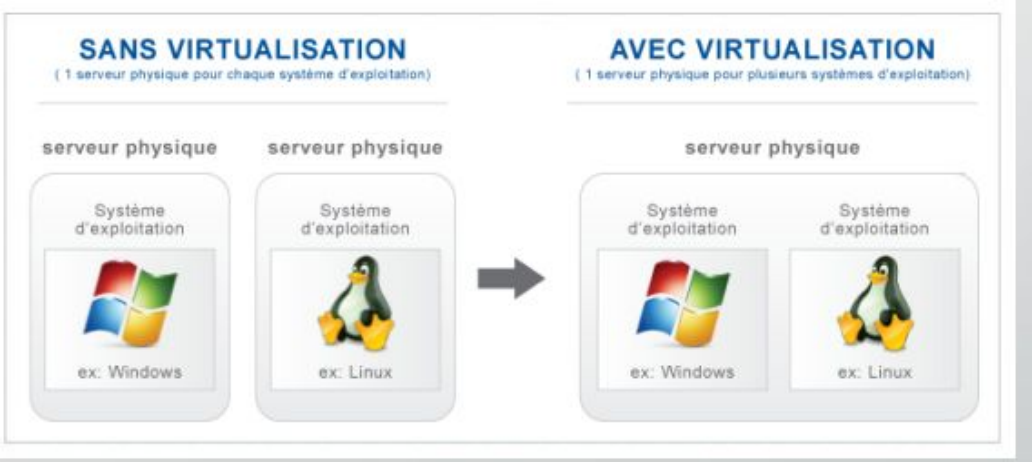

Figure 1 : Représentation schématique d'une machine virtualisée et une machine non virtualisée[1]

Introduction aux différentes techniques de virtualisation :

La virtualisation permet de faire fonctionner sur une seule machine physique plusieurs machines virtuelles avec des systèmes d'exploitation différents. On distingue plusieurs méthodes de virtualisation :

Noyau en mode utilisateur :

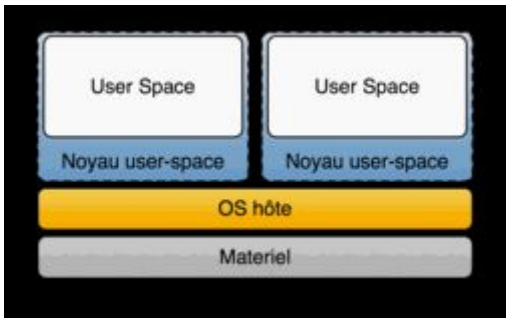

Figure 2 : Représentation schématique d'un noyau en mode utilisateur[2]

Le principe est de faire tourner le noyau d'un ou plusieurs systèmes d'exploitations invités dans un espace avec des accès restreints, appelé espace utilisateur. Le noyau est le logiciel cœur du système, qui permet notamment la communication entre logiciels et matériels, la gestion des différentes tâches et l'accès à la mémoire, le processeur et les périphériques. Les espaces utilisateurs quant à eux sont gérés par l'OS hôte et bornés à l'exécution de certaines instructions. De plus, l'accès à une partie de la mémoire lui est interdit. Cela s'oppose aux espaces noyau (ou privilégiés) qui ont accès à toutes les ressources disponibles (mémoire et calcul).

Cette solution est très peu performante, car deux noyaux sont empilés et l'isolation des environnements n'est pas gérée de même, l'indépendance par rapport au système hôte est inexistante. Elle sert surtout au développement du noyau.

Isolateur :

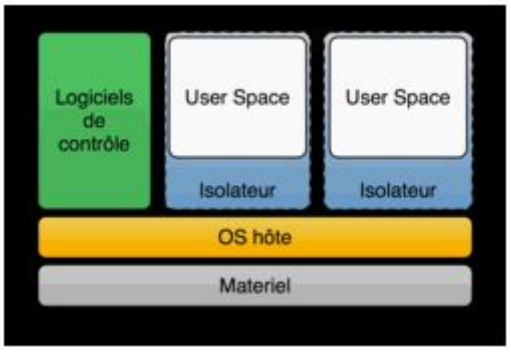

Figure 3 : Représentation schématique d'un isolateur [2]

Un isolateur permet l'exécution d'une application de façon isolée sur un système d'exploitation en créant des espaces-utilisateurs. En cloisonnant les programmes dans ces espaces, cela permet un système stable même en cas d'instabilité sur un programme. De plus, cela permet de limiter l'accès aux ressources des différents processus et applications. La consommation en ressources est faible ce qui permet d'avoir de bonnes performances avec cette solution.

Hyperviseur (para-virtualisation) :

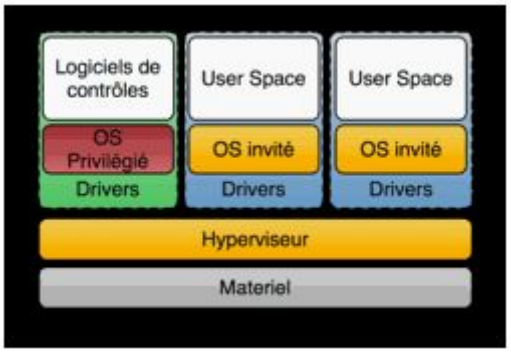

Figure 4 : Représentation schématique d'un hyperviseur[2]

Un hyperviseur est un noyau système très léger et optimisé pour gérer les accès des noyaux d'OS invités à l'architecture matérielle. Ce système a pour unique tâche de gérer ses systèmes invités.

Actuellement l'hyperviseur est la méthode de virtualisation d'infrastructure la plus performante du fait que les machines virtuelles communiquent directement sans passer par la couche matérielle mais elle a pour inconvénient d'être contraignante et onéreuse, bien que permettant plus de flexibilité dans le cas de la virtualisation d'un centre de traitement informatique.

Attention ne pas confondre machine virtuelle et conteneur ;

- Machine virtuelle (Virtualisation) : Imitation virtuelle d'un appareil informatique créé, à l'aide d'un logiciel hyperviseur et doté d'un système d'exploitation complet.

- Conteneurisation : Un conteneur est essentiellement une partition logique servant à isoler les applications sur un même serveur. Plutôt que de répliquer entièrement un système d'exploitation pour chaque application, comme dans une machine virtuelle, les conteneurs permettent aux applications d'un serveur de partager le même noyau du système d'exploitation. Ce système d'exploitation partagé est appelé OS hôte.

Notions fondamentales :

Chaque outil de virtualisation met en œuvre une ou plusieurs de ces notions :

- **Couche d'abstraction matérielle et/ou logicielle :**

En informatique, et plus particulièrement en architecture, une couche d'abstraction matérielle (HAL : *hardware abstraction layer*) est un logiciel intermédiaire entre le système d'exploitation et le matériel informatique. Il offre des fonctions standardisées de manipulation du matériel informatique tout en cachant les détails techniques de la mise en œuvre.

Les producteurs des systèmes d'exploitation incluent une couche d'abstraction matérielle dans leurs produits. C'est une pièce de logiciel importante dans les systèmes d'exploitation portables susceptibles d'être utilisés sur différents types de matériel : en cas de portage seule la couche d'abstraction matérielle nécessite une adaptation.

- **Système d'exploitation hôte (installé directement sur le matériel) :**

L'OS (système d'exploitation) hôte est le logiciel installé sur un ordinateur, qui interagit avec le matériel sous-jacent. Le terme désigne principalement un système d'exploitation qui exécute une machine virtuelle, par opposition au système d'exploitation « invité » à l'intérieur de cette MV.

Le plus souvent, on parle d'OS hôte pour le système d'exploitation qui interagit avec le matériel et exécute un hyperviseur de type 2. Egalement appelé « hyperviseur hébergé », celui-ci s'exécute au-dessus d'un système d'exploitation hôte au lieu d'interagir directement avec le matériel. L'hyperviseur de type 2 peut ainsi créer plusieurs machines virtuelles exécutant chacune un système d'exploitation invité. Ce dernier n'est pas nécessairement le même que l'OS hôte.

On prend l'exemple pratique d'un ordinateur exécutant le système d'exploitation OS X d'Apple. Si un utilisateur souhaite exécuter une application disponible uniquement pour les systèmes d'exploitation Windows, il peut avoir recours à la virtualisation et installer un hyperviseur de type 2, tel que VMware Fusion, sur l'ordinateur qui exécute OS X.

A l'aide de l'hyperviseur VMware Fusion, il peut ensuite créer une VM et y installer Windows 10 comme système d'exploitation. Il sera ainsi en mesure d'exécuter son application Windows dans la machine virtuelle. Dans ce cas, l'instance originale d'OS X installée sur l'ordinateur est considérée comme l'OS hôte, tandis que Windows 10 (exécuté sur la VM) représente le système d'exploitation invité.

Figure 5: Représentation du fonctionnement d'un hyperviseur[3]

- **Systèmes d'exploitation (ou applications) « virtualisé(s) » ou « invité(s) » ;**

Les OS invités sont hébergés par le système d'exploitation hôte, ce dernier orchestre les accès au matériel demandés par les OS invités.

- **Virtualisation des données :**

La virtualisation des données est une approche permettant d'unifier les données de plusieurs sources dans une même couche afin que les applications, les outils de génération de rapports et les utilisateurs finaux puissent accéder aux données sans avoir besoin de détails sur la source, l'emplacement et les structures de données d'origine.

Pourquoi virtualiser ?

Il peut sembler étrange de simuler d'autres machines sur une machine hôte : un système d'exploitation est conçu pour utiliser du matériel qui est entièrement sous son contrôle. La juxtaposition de plusieurs systèmes non conçus pour communiquer entre eux peut faire craindre des inefficiences auxquelles s'ajoute le fait que le processus de virtualisation consomme des ressources. Le tableau n'est pas aussi sombre. D'une part, on évite une grande partie de ces inefficiences juste en disposant de disques différents pour chaque système, et d'autre part le coût de la mémoire permet à chacun de ces systèmes de rester résident, et parfois avec de larges sections de code partagées.

De plus, il est courant pour une entreprise de disposer de plusieurs serveurs fonctionnant à 15 % de leur capacité, de façon à pouvoir faire face aux pointes de charge. Un serveur chargé à 15 % consomme autant d'énergie qu'un serveur chargé à 90 %, et regrouper plusieurs serveurs sur une même machine s'avère rentable si leurs pointes de charge ne coïncident pas systématiquement, même en incluant la charge de la virtualisation.

Concept général de la virtualisation

Historique : évolution de la virtualisation jusqu'à nos jours

Une bonne part des travaux sur la virtualisation fut développée au centre scientifique de Cambridge d'IBM en collaboration avec le MIT, où fut mis au point le système expérimental CP/CMS, devenant ensuite le produit (alors nommé hyperviseur) VM/CMS. Par la suite, les mainframes ont été capables de virtualiser leurs systèmes d'exploitation avec des technologies spécifiques et propriétaires, à la fois logicielles et matérielles. En 1979 fut annoncé par exemple sur les IBM 4331 et 4341 un accélérateur VM optionnel et microcode.

Dans la deuxième moitié des années 1980 et au début des années 1990, on a créé des embryons de virtualisation sur des ordinateurs personnels. Ces solutions pouvaient être soit purement logicielles, soit couplées à du matériel additionnel (ajout de processeur, carte réseau, etc.). Et c'est sur des ordinateurs Amiga équipés de processeurs hétérogènes comme le 80386 et 80486, 68xxx, et PPC qu'il était possible de lancer d'autres OS comme Windows, Mac OS, voire Linux, le tout en multitâche sous AmigaOS. C'est sur cette machine en avance sur son temps que la technologie de virtualisation a été pleinement exploitée et encore inégalée aujourd'hui. Pour les PC, il y avait des émulateurs comme le SideCar et PC Task. Sur Macintosh, Emplant et ShapeShifter. Les grands Unix ont suivi avec les architectures NUMA des Superdome d'HP (PA-RISC et IA-64) et des E10000/E15000 de Sun (UltraSparc).

Dans la seconde moitié des années 1990, les émulateurs sur x86 des vieilles machines des années 1980 ont connu un énorme succès, notamment les ordinateurs Atari, Amiga, Amstrad et les consoles NES, SNES, Neo-Geo AES. La société VMware développa et popularisa à la fin des années 1990 et au début des années 2000 un système propriétaire de virtualisation logicielle des architectures de type x86 pour les architectures de type x86. Les logiciels libres Xen, KVM, QEMU, Bochs, Linux-VServer, Oracle VM VirtualBox et les logiciels propriétaires mais gratuits VirtualPC, Virtual Server et VMware Server ont achevé la popularisation de la virtualisation dans le monde x86. VMware a dernièrement rendu gratuite une version allégée de son hyperviseur phare ESX3i. Les fabricants de processeurs x86 AMD et Intel ont mis en œuvre la virtualisation matérielle dans leurs gammes dans la seconde moitié de l'an 2000. [2[4]

L'ETSI a mis en place en novembre 2012 un groupe de travail afin de standardiser le NFV (en), une nouvelle approche de la virtualisation, auquel participent de nombreuses sociétés spécialisées dans les technologies d'intelligence réseau (Qosmos, Procera Networks ou encore Sandvine) et des laboratoires de recherche universitaire (University of Campinos, University Carlos III of Madrid, University of the Basque Country, University of Patras, etc.)

Les différents types de virtualisation

La virtualisation peut prendre plusieurs formes en fonction du domaine et de l'utilisation à laquelle on la destine. On dénombre actuellement 6 types de virtualisation différentes qui permettent chacune de jouer le rôle de plusieurs ressources.

Remarques

La virtualisation est possible grâce à l'utilisation d'un logiciel que l'on appelle hyperviseur et qui est relié directement au matériel afin de permettre une fragmentation unique du système en plusieurs environnements sécurisés distincts. On appelle cela des machines virtuelles et ces dernières exploitent la capacité de l'hyperviseur à séparer les ressources du matériel et à les distribuer de manière convenable contrairement au Cloud qui permet un accès à un même ensemble de ressources approvisionnée de manière informatique à un service ou une entreprise.

Figure 6: Représentation des différents types de virtualisation

Virtualisation des postes de travail ou *virtualization desktop*

Ce type de virtualisation est très apprécié et utilisé au sein des entreprises. Un système standard de bureau opérationnel est encapsulé dans une machine virtuelle accessible par les différents utilisateurs. Cette technique permet de faire tourner plusieurs environnements sur le même système physique en reproduisant l'environnement d'un ordinateur dans le but d'offrir aux professionnels la possibilité d'accéder à leur fichiers et à leurs applications personnelles depuis n'importe quel poste de l'entreprise.

Cette virtualisation permet donc d'exécuter sur plusieurs machines un serveur où le client obtient un affichage à distance sur son ordinateur. Au sein d'une entreprise, il y a donc un affichage sur une centaine de postes physiques d'une image virtuelle du poste utilisateur qui est en fait exécutée sur un serveur distant.

La *virtualisation desktop* est possible grâce à l'hébergement du poste de travail virtuel sur un serveur spécifique appelé VDI pour *Virtual Desktop Infrastructure*. Ce dernier va exécuter l'ensemble de l'environnement du poste c'est-à-dire le système d'exploitation et les applications. Ce type de virtualisation nécessite un hyperviseur dont nous détaillons l'intérêt et le fonctionnement plus en détail dans une autre partie de l'exposé. Cet hyperviseur va fonctionner sur un serveur de centre de données afin d'héberger la machine virtuelle du bureau. Plusieurs hyperviseurs peuvent être utilisés mais cela induit alors souvent l'utilisation de plusieurs logiciels de gestion bien qu'il existe des logiciels capables de gérer plusieurs hyperviseur à la fois.

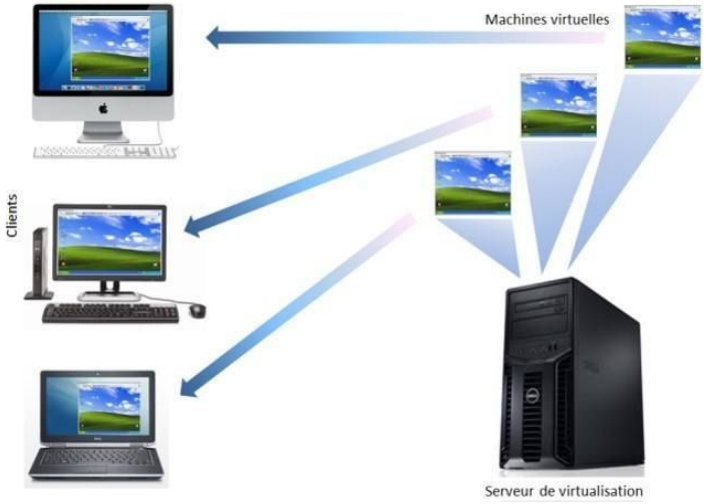

Figure 6 : Schéma expliquant le principe de la virtualisation des postes de travail

Exemple : TSlog, Citrix, Virtuel Bureau

Virtuel bureau permet de simplifier l'utilisation des postes de travail en proposant une mise à jour automatique des logiciels utilisés ainsi qu'une sauvegarde des données afin de permettre une utilisation mobile du système.
(commercialisé par Avenir numérique)

Avantages

Cette technologie entraîne beaucoup de flexibilité et notamment, lors des situations de mobilité tout en facilitant le transfert des environnements de travail aux équipes de sous-traitants. Elle induit aussi une réduction des coûts associée à la multiplications des postes de travail tels que ceux associés aux licences de systèmes par exemple (un PC consomme environ 100 Watts alors qu'un pont d'accès seulement 20 Watts). Cette méthode offre aussi une reprise en cas de sinistres ainsi qu'une économie d'énergie.

Inconvénients

Cette technologie peut entraîner une dégradation des méthodes et des performances (notamment sur les applications multimédia et 3D) et il n'existe pas de réel standard de stockage ce qui rend une utilisation multiple plus complexe. Cela signifie aussi qu'il est impossible de travailler en "offline", et un accès au réseau et aux serveurs est donc indispensable. Cette technique ne peut donc pas être utilisée dans les tunnels ou bien dans les avions et implique une défaillance totale de la production en cas de problème avec le réseau ou un problème matériel.

En effet, si un des serveurs tombent en panne, ce n'est plus un ordinateur qui sera hors-service mais des dizaines. Un des inconvénients majeurs reste aussi la complexité de la mise en place de cette technologie ainsi que le risque que cela implique pour la sécurité en cas de mauvaise configuration du réseau.

Virtualisation du stockage ou *Sotfware Defined storage*

La virtualisation du stockage est une technologie qui permet à plusieurs périphériques de stockage physique, de se regrouper au sein d'un seul et unique périphérique de stockage et ce dernier sera géré depuis une console centrale. Cet unique espace de stockage sera alors visible par les hôtes qui verront alors ce périphérique comme leur propre disque.

Un disque dur virtuel présent au sein d'une machine virtuelle permet de stocker les données et celui-ci se présente sous la forme de fichier dans le système de fichiers de l'hôte :

- VHD chez Microsoft
- VDI chez Oracle
- VMDK chez VMWare
- OVF format ouvert

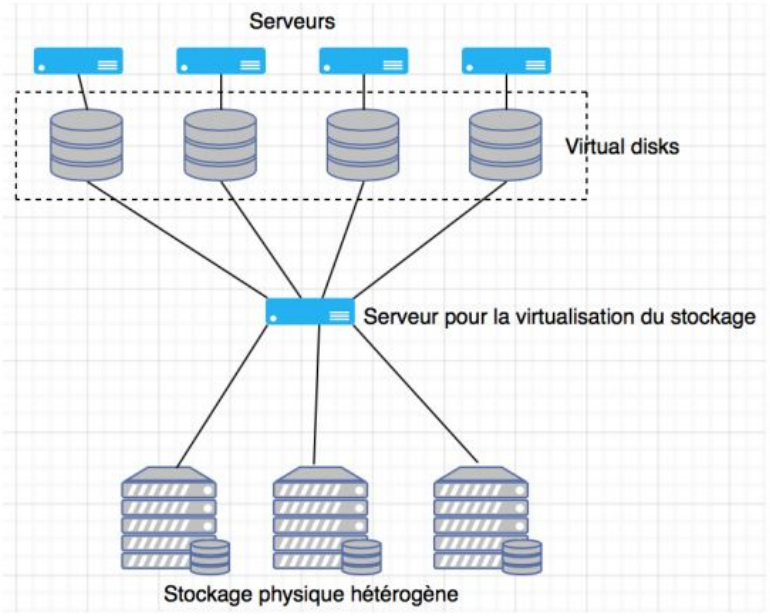

Figure 7 : Schéma représentatif de la virtualisation du stockage

Cette virtualisation peut être mise en oeuvre grâce à plusieurs architectures de stockages :

> DAS ou Direct Attached Storage

Cette technique fait référence à un système de stockage informatique qui va être lié de manière directe à un serveur ou à un ordinateur au lieu de transiter par un réseau. Aucun réseau de stockage n'est alors utilisé puisque le serveur est directement connecté par le biais d'un adaptateur de stockage (USB, micro USB pour relier un disque dur externe par exemple).

Exemple :

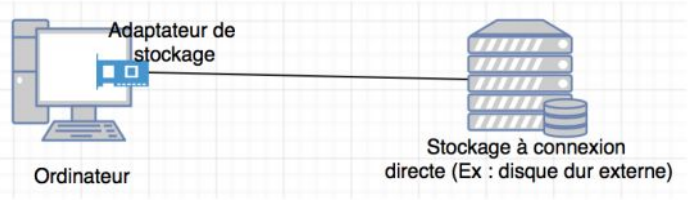

Figure 8 : Représentation du fonctionnement du DAS

> JBOD ou Just a bunch of Disks

Cette méthode permet d'augmenter la capacité de stockage en présentant directement les disques à un serveur comme s'il était liés physiquement. Généralement, ces boitiers sont des boîtiers externes connectés à un ou plusieurs serveurs via des connexions SAS ou eSATA permettant la combinaison de plusieurs disques physiques dans le but de créer un pool de stockage.

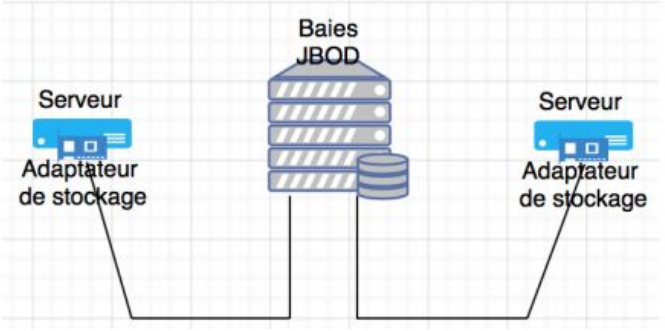

Figure 9 : Représentation du fonctionnement JBOD

> NAS ou Network Area Storage

Ce périphérique de stockage est un périphérique dédié fournissant aux nœuds de réseaux locaux ou LAN un stockage partagé basé sur des fichiers par le biais d'une connexion Ethernet.

Un des avantages de ce type de stockage est la possibilité de fournir à plusieurs clients sur le réseau l'accès à un même fichier où les dépendances vis-à-vis de l'emplacement sont cachés (c'est-à-dire où les données sont physiquement stockées). Les périphériques NAS sont généralement dépourvus de clavier ou d'écrans et leur gestion se fait par le biais d'un programme utilitaire. Chaque NAS correspond sur le réseau à un "nœud de réseau indépendant" qui va possède une adresse IP unique.

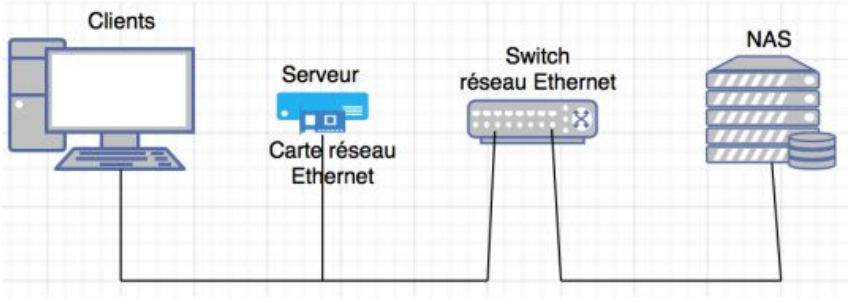

Figure 10 : Représentation du fonctionnement du NAS

> SAN ou Storage Area Network

Un réseau de stockage SAN est un réseau dédié qui va permettre une mutualisation des ressources de stockage. Ici, on a accès aux disques de stockage en mode bloc c'est-à-dire que le stockage est accessible en mode bloc grâce au système de fichier des serveurs. Cela va introduire un niveau d'abstraction entre les serveurs et le système de stockage permettant une plus grande flexibilité pour les administrateurs.

Cela permet donc une évolution de l'espace de stockage où l'espace du disque n'est plus limité par les serveurs et évolue en fonction de l'ajout ou non de disque de stockage.

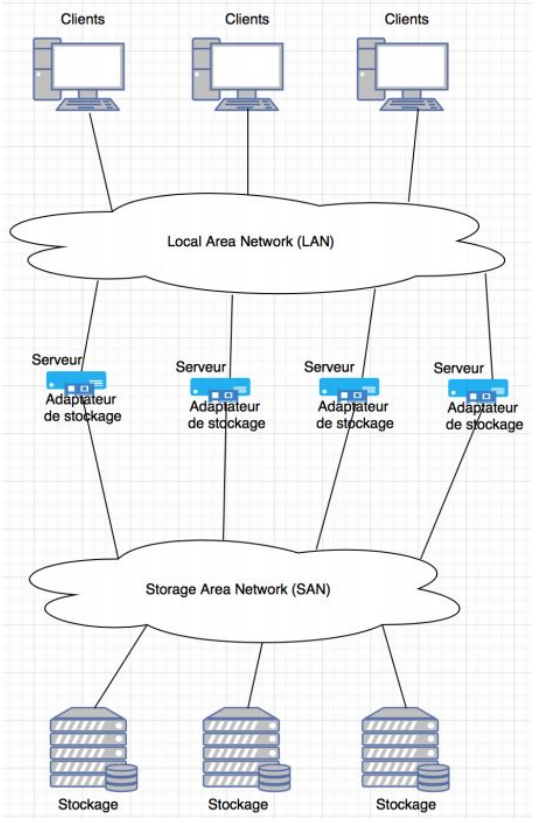

Figure 11 : Représentation du fonctionnement du SAN

Avantages

Ce genre de virtualisation permet de nombreuses fonctionnalités avantageuses telles que l'adjonction d'un périphérique de stockage supplémentaire sans interruption des services, un regroupement des unités de disques durs / technologie de stockage de différentes vitesses, de différentes tailles et de différents constructeurs ainsi qu'une réallocation dynamique de l'espace de stockage c'est-à-dire réajuster les capacités de stockages en fonction des besoins de l'utilisateur et/ ou de l'évolution structurelle de l'entreprise

Cette méthode induit aussi une homogénéisation du stockage associée à une meilleure optimisation des performances et de la vitesse.

Inconvénients

La mise en place de cette technologie impose un respect strict des matrices de compatibilités entre les constructeurs en plus de supposer que celles-ci soient supportés par la couche de virtualisation imposée. On introduit donc de nouvelles contraintes au niveau de la plateforme de virtualisation.

Exemple : SANSymphony développé par Datacore

Cette technologie permet de placer une couche de virtualisation évolutive sur les infrastructures de stockage. De la sorte, la cohabitation entre différents matériaux de stockage est possible.

Virtualisation du réseau ou *virtualisation network*

Lorsque l'on parle de *virtualisation network*, on entend par là, le processus qui reproduit un réseau physique et ses différents composants (les ports, les routeurs, pare-feu, VPN). La virtualisation des réseaux consiste à partager une même infrastructure physique (débit des liens, ressources CPU des routeurs) au profit de plusieurs réseaux virtuels isolés.

Ce réseau sera alors complètement défini par le logiciel. Cette virtualisation combine le matériel et le logiciel afin de créer un réseau entièrement défini par le logiciel. Les entités physique des périphériques deviennent alors des commutateurs virtuels qui vont exister en tant qu'entité pilotés par logiciels et non plus en tant qu'équipement avec des ports physiques.

Il existe deux notions importantes :

- **NFV pour Network Function Virtualization**
  Capacité à dissocier le matériel du logiciel pour les équipements réseaux : en effet, de nombreuses fonctions réseaux peuvent s'exécuter de manière indépendante sur un même matériel générique.
- **SDN ou Software Defined Networks**
  Capacité de configurer les équipements réseau à la volée en fonction des besoins de l'application/ service au moyen d'un contrôleur réseau

Ces deux notions sont indépendantes l'une de l'autre mais leur déploiement est souvent simultané.

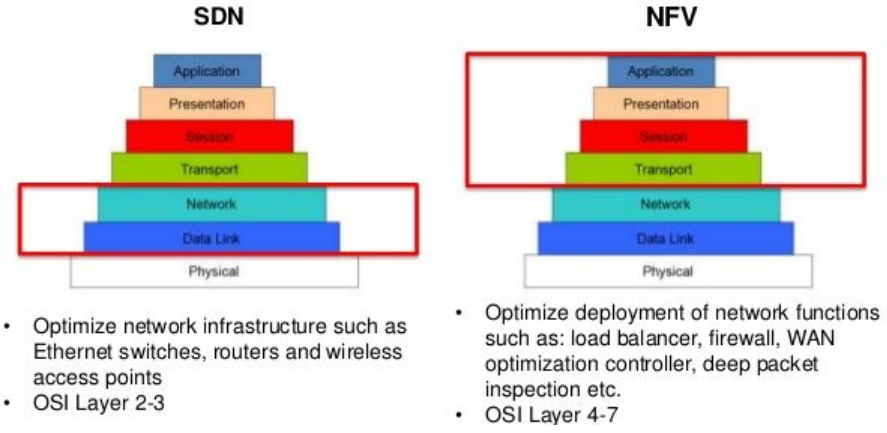

Figure 12 : Représentation d'une machine virtualisée et une machine non virtualisée

Cette méthode permet donc, par l'utilisation de logiciels de virtualisation, de mettre à disposition des réseaux isolés cloisonnés mais fonctionnant sur une infrastructure virtualisée.
On peut la diviser en deux sous-groupes : la virtualisation interne (au niveau du serveur) et la virtualisation externe (au niveau du datacenter).

> Interne

Les construction du réseau sont purement virtuelle et vont exister au sein d'un seul système. Elle se compose d'un système qui va utiliser des machines virtuelles ou des zones dont les interfaces réseaux sont configurées sur au moins une NIC physique aussi appelés Cartes d'interfaces réseau virtuelles ou NIC virtuelles (VNIC). Ces conteneurs communiquent les uns avec les autres en devenant un réseau virtuel sur un seul hôte.

Réseau virtuel privé

Ils sont différents des réseaux privés virtuels VPN car un logiciel VPN crée un liaison sécurisée entre deux systèmes d'extrémités, c'est donc un réseau virtuel privé sur un système qui n'est pas accessible par les systèmes externe. Cette isolation est possible qu'en configurant les VNIC sur les ethersubs.

Exemple : activation de la connectivité entre les différentes machines virtuelles sur le serveur

> Externe

Lorsque la configuration interne commence à intégrer des périphériques réseau externes physiques ou des serveurs externes supplémentaires, on parle de virtualisation de réseau externe. Ils sont composés de plusieurs réseaux locaux administré par le logiciel en tant que entités unique. Ce réseau est formé par des blocs de construction standard composé par le matériels de commutation et la technologie logicielle VLAN.

Exemple : les grand réseaux d'entreprises et les centres de données

Un VLAN ou *virtual local area network* est un réseau local qui regroupe un ensemble de machines de façon logique et non physique. On dénombre plusieurs niveaux de réseau virtuels :

- Les réseaux virtuels de niveau 1 que l'on appelle réseaux virtuels par port ou *port-based* VLAN
- Les réseaux virtuels de niveau qui sont aussi appelés réseaux virtuels par adresse MAC ou *MAC address-based VLAN*
- Les réseaux virtuels de niveau 3 regroupant les réseaux virtuels par adresse de sous réseau aussi connu sous le nom *Network address-based VLAN* et les réseaux virtuels par protocole

Exemple : VMWare NSX, OpenStack, Kubernetes

Composants de la virtualisation du réseau

> Cartes réseau VNIC

Les VNIC sont des périphériques réseaux virtuels avec les mêmes interfaces de liaison de données qu'une NIC physique et ils sont configurés sur une liaison de données sous-jacentes. Une fois configurées, elles vont se comporter comme des cartes d'interface réseau physique et les ressources du systèmes les traitent comme des cartes d'interfaces réseau physique.

> Commutateurs virtuels

Ces derniers sont créés automatiquement lors de la création d'une VNIC. Quand, un port de commutateur reçoit un paquet sortant depuis l'hôte connecté à ce port, le paquet ne peut pas accéder à une destination sur le même port. C'est un inconvénient pour les systèmes qui ont été configurés par des réseaux virtuels puisque ceux-ci partagent la même carte d'interface réseau.

Les paquets sortants passent par un port de commutateur sur le réseau externe et les paquets entrants ne peuvent pas atteindre leur zone de destination car les paquets ne peuvent pas revenir via le même port que celui par lequel ils ont été envoyés. Les commutateurs virtuels vont fournir à ces zones une méthode de transmission de paquets et va ouvrir un chemin de données pour que les réseaux virtuels communiquent les uns avec les autres.

> Ethersubs

Ces derniers sont des pseudo-cartes réseaux Ethernet et il est possible de créer des VNIC sur ces ethersubs. Ceux-ci deviennent alors indépendants des cartes réseaux virtuels sur le systèmes. On peut donc construire un réseau virtuel privé qui sera isolé des autres réseaux virtuels sur le système et sur le réseau externe. Ces outils permettent de limiter l'accès d'un environnement réseau à des personnes spécifiques et l'accès est donc impossible pour les utilisateurs du réseau global.

Ce type de virtualisation est un avantage considérable lors du déploiement d'infrastructures informatiques nécessitant une mise en place rapide en plus de faciliter leur reproduction.

Virtualisation des serveurs

C' est une technique permettant d'exécuter simultanément plusieurs systèmes d'exploitation isolés dans des machines virtuelles, sur un seul serveur physique. On entend par serveur l'ensemble système d'exploitation et applications. Cette technologie permet de créer plusieurs serveurs ou postes de travail informatiques sur une seule plateforme matérielle. À l'utilisation, les systèmes d'exploitation virtualisés ne se distinguent pas des systèmes d'exploitation classiques.

La virtualisation des serveurs permet une utilisation plus efficace des ressources informatiques car la plupart des serveurs utilisent moins de 15 % de leurs capacités, ce qui favorise leur prolifération et leur complexité. Avant la virtualisation des serveurs, il était courant d'avoir du matériel sur-utilisé ou sous-utilisé dans le même datacenter. Avec la virtualisation, il est possible de déplacer des charges de travail entre les machines virtuelles selon la charge. Le même serveur physique peut aussi fonctionner sur plusieurs systèmes d'exploitation de serveur et plusieurs configurations, ce qui augmente son efficacité.

Cette technique est également l'un des maillons de l'agilité des entreprises. Elle crée en effet un environnement hautement disponible qui s'assure que toutes vos applications sont accessibles à tout moment. Si l'un de vos serveurs tombe en panne ou plante, toutes les machines virtuelles peuvent être redémarrées automatiquement sur une autre machine, sans temps d'arrêt ni perte de données. Ce type de virtualisation permet de réaliser des économies substantielles sans entraîner de véritable changement dans le service informatique. Ici Quelques hôtes d'hyperviseurs et une console de gestion remplacent des dizaines, voire des centaines de serveurs physiques. Concrètement cette technique simule en effet les serveurs physiques en changeant leur identité, leurs numéros, leurs processeurs et leurs systèmes d'exploitation. Cela évite à l'utilisateur de gérer en permanence des ressources serveur complexes.

Quelques exemples de virtualisation de serveurs sont:

- ESXI de VMware.
- Hyper-V et Virtual Server de Microsoft.
- Xen/Citrix et Virtuozzo de SW soft.

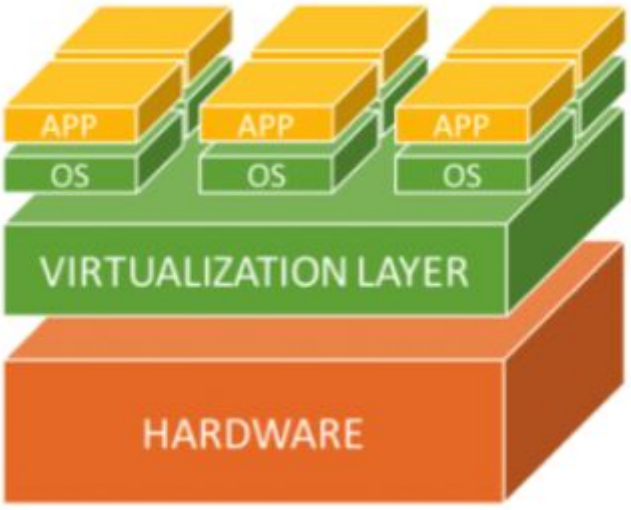

Figure 13 : Représentation de la virtualisation des serveurs

Sur cette figure nous observons comment un serveur physique fait fonctionner plusieurs serveurs virtuels, dans le sens qu'ils supportent des applications et/ou des services bien réels, s'exécutant au sein de machines virtuelles .

Avantages

- Une réduction des coûts opérationnels (matériel, énergie, espace) et une meilleure exploitation du serveur physique
- Une amélioration de la disponibilité des serveurs
- Une plus grande souplesse pour gérer l'évolution des besoins informatiques
- Suppression de la prolifération et de la complexité des serveurs

Inconvénient

L'indisponibilité du serveur physique provoque de fait l'indisponibilité de l'ensemble des serveurs supportés et donc des services et applications associés.

Virtualisation des applications

La virtualisation des applications permet aux applications d'être totalement dissociées du système d'exploitation. Son principe de fonctionnement est le suivant: un administrateur informatique implémente, des applications à distance sur un serveur au sein du datacenter de l'entreprise ou via un service d'hébergement. L'administrateur informatique utilise ensuite, le logiciel de virtualisation d'applications pour livrer ces applications sur le poste d'un utilisateur, ou tout autre appareil connecté. L'utilisateur peut alors accéder aux applications et les utiliser comme si elles étaient installées en local sur sa machine et ses actions sont renvoyées vers le serveur pour être exécutées.

De cette manière, même si une application est défectueuse, elle ne va pas amputer les autres applications ni le système d'exploitation. Les applications sont considérées comme des services virtuels et n'ont donc pas besoin d'être installées sur chaque ordinateur. En revanche, elles sont quand même exécutées depuis des postes de travail afin d'utiliser leurs ressources.
L'intérêt principal de ce type de solution est qu'il n'est pas utile d'installer les logiciels applicatifs physiquement sur les postes de travail. Cela évite également le déploiement des patchs directement sur les systèmes avec les risques de dysfonctionnement généralement associés.

Les apports fonctionnels de la virtualisation d'applications sont l'élimination des conflits entre applicatifs, la récupération rapide des applications en dysfonctionnement et la diminution importante de la durée des phases de test avant le déploiement. Un autre point fort pour les développeurs est qu'ils peuvent faire tourner plusieurs versions d'un même applicatif sur un seul OS. Il est également possible de diffuser les applications à la demande sur tous les types de dispositif en mode streaming (mode de diffusion et de lecture de contenus (son, vidéo) en flux continu).
exemple: ThinApp de VMware, XenApp de Citrix ou App-V de Microsoft , docker.

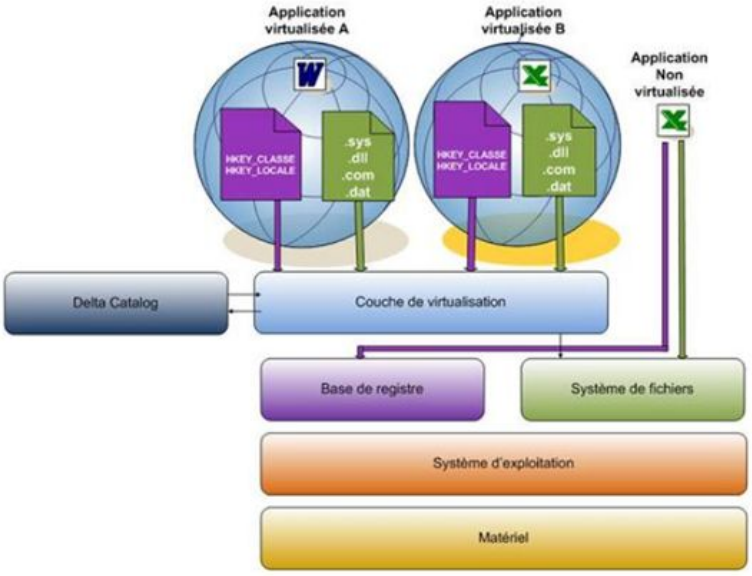

Figure 14 : Représentation de la virtualisation d'une application

Ce schéma présente un exemple de virtualisation d'application sous Windows. Une couche de virtualisation est ajoutée entre, les programmes virtualisés et le système d'exploitation qui intercepte les appels systèmes (base de registre, disque, applications).

Le système de fichiers et la base de registre virtuels ne sont pas des copies de ceux du système d'exploitation. Ils regroupent uniquement les modifications effectuées par l'application pour qu'elle puisse fonctionner. Si l'application veut altérer la configuration système, elle ne le fera que dans sa copie de la base de registre. De même, elle n'aura accès qu'à ses propres versions et de fichiers de configuration système. Il n'y a donc pas de conflit avec les autres applications, c'est ce que l'on appelle le concept de bulle. Ainsi,plusieurs applications peuvent opérer dans différentes bulles et elles demeurent indépendantes les unes des autres. Une telle garantie d'indépendance entre les applications limite fortement le volume des tests de régression nécessaires en cas de changement de système d'exploitation.

Avantages

- Gestion centralisée des droits d'accès aux applications
- Déploiement quasi instantané des applications et mise à jour centralisée
- Simplification de l'administration et du déploiement du parc informatique notamment lors de l'installation des nouvelles versions
- Isolement des applications permettant de pallier les incompatibilités.Créer pour chaque application des copies des ressources partagées

Inconvénients

- Le support du multimédia reste délicat

- Il est nécessaire d'investir dans des serveurs puissants
- Même si la redirection des périphériques est plutôt bien gérée, elle complique leur utilisation.

Les différentes techniques de virtualisation

La virtualisation est une technologie qui doit s'adapter aux différentes couches technologiques d'une infrastructure informatique. Il existe plusieurs variantes d'architecture de virtualisation aujourd'hui :

> L'isolateur
> L'émulateur
> L'hyperviseur et para-virtualisation
> La virtualisation complète

L'isolateur

Un isolateur est un logiciel permettant de confiner les autres applications que l'on souhaite virtualiser dans un contexte qui leur est propre. On parle aussi de "zone d'exécution" et le fait de dédier un espace mémoire spécifique à l'application virtualisée permet d'obtenir plusieurs instance d'une application initialement conçue pour une instance unique.

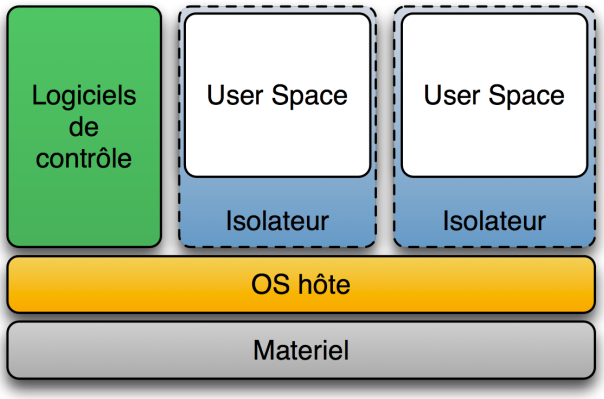

Figure 15 : Représentation du fonctionnement d'un isolateur

L'isolation ne peut pas virtualiser tout un système de d'exploitation mais permet d'offrir des performances supérieures. Ces isolateurs sont majoritairement disponible pour les systèmes Linux

Exemple : LInux-VServer, chroot, BSD Jail, OPenVZ

Cette solution est très performante et économique en mémoire mais cela nécessite un partage du code noyau.

L'émulateur

L'émulateur est un système informatique est considérée comme un système pouvant fonctionner comme un autre système à travers des fonctions pour lesquelles il n'avait pas été prévu à l'origine.

Exemple : Fonctionnement sur un PC des émulateurs de vieilles consoles de jeux tels que NES, QEmu qui est un hyperviseur de type VirtualBox

On doit respecter plusieurs principes :

- **Indépendance** : l'instance émulée est indépendante de la plateforme matérielle et l'émulateur joue le rôle d'une passerelle entre les deux machines
- **Isolation** : l'émulateur est isolée des autres instances ce qui permet de ne pas impacter la plateforme physique
- **Equivalence** : l'émulation consomme beaucoup de ressources et les performances entre une application exécutée sur un socle émulé et sur un environnement physique propriétaire ne seront pas forcément identiques

L'hyperviseur

L'hyperviseur est un gestionnaire de machine virtuelle ou VM qui va permettre la création des différentes versions virtuelles des ordinateurs et de leurs systèmes d'exploitation. Ce gestionnaire va les fusionner en un seul serveur physique afin de permettre une utilisation plus efficace des ressources matérielles. Il permet également aux utilisateurs d'exécuter simultanément différents systèmes d'exploitation sur le même ordinateur.

La virtualisation est mise en place grâce à l'utilisation d'un programme spécifique ou d'un ensemble de logiciels, matériels ou de micrologiciels que l'on appelle hyperviseur. Cet hyperviseur permet la réalisation et l'exécution de de diverses machines virtuelles. Il existe deux types d'hyperviseurs :

> L'hyperviseur de type I dit aussi hyperviseur VM natif
> L'hyperviseur de type II que l'on appelle aussi hyperviseur hôte VM

Hyperviseur de type I ou hyperviseur VM natifs

La création du premier hyperviseur de ce type remonte aux années 1960 avec IBM qui a développé ce que l'on appelle aujourd'hui les hyperviseurs natifs.
Dans ce type d'hyperviseur, l'exécution de l'hyperviseur permet de contrôler le matériel et le système d'exploitation invité et ce, directement sur la machine hôte. Les machines hôte font donc fonctionner la machine virtuelle sur le système d'exploitation en tant que processus invité.

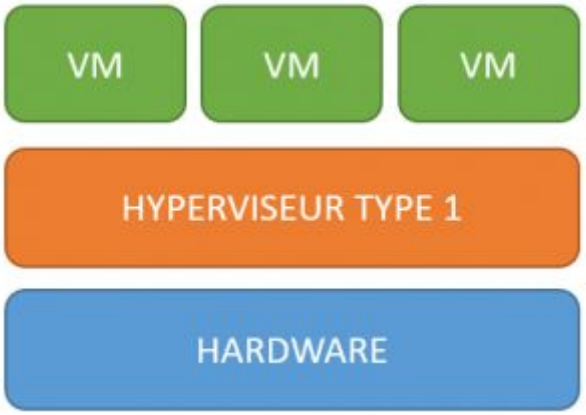

Figure 16 : Représentation du fonctionnement d'un hyperviseur de type I

Avantage

Un maximum de ressources peuvent être alloué aux machines virtuelles puisque ce type d'hyperviseur est directement lié à la couche matérielle.

Inconvénient

On ne peut exécuter qu'un seul hyperviseur à la fois.

Hyperviseur de type II ou hyperviseur hôte VM

Les hyperviseurs de type II nécessitent eux, le support d'un système d'exploitation opérationnel permettant de fournir des services de virtualisation. Ici, les hyperviseurs sont gérés sur un système d'exploitation conventionnel de manière similaire aux autres programmes informatiques exécutés. Ces hyperviseurs font donc abstraction des systèmes d'exploitation invités des systèmes d'exploitation invités au sein du système d'exploitation hôte.

C'est généralement un logiciel lourd qui tourne sur l'OS hôte permettant de lancer un ou plusieurs OS invités. Cette solution est très comparable à un émulateur sauf que dans ce cas, le microprocesseur, la RAM et la mémoire de stockage (via un fichier) sont directement disponibles par les machines virtuelles.

Exemple : Station et travail et boîte virtuelle de VMWare , VirtualBox

Cette solution possède un coût en performance mais elle permet la cohabitation de plusieurs OS hétérogènes sur une même machine. les échanges entre les machines se font donc alors par les canaux standards de communication entre systèmes d'exploitation (TCI /IP et les autres protocoles réseaux).

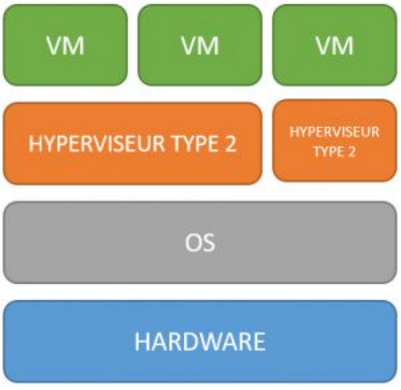

Figure 17 : Représentation du fonctionnement d'un hyperviseur de type II

Avantages

Ce type d'hyperviseur permet d'exécuter plusieurs hyperviseurs en même temps puisque ceux-ci ne s'installent pas directement sur la couche matérielle.

Inconvénient

Cet hyperviseur ne peut pas fournir autant de ressources matérielles que ceux du type 1.

Virtualisation complète

La virtualisation est dite complète lorsque le système d'exploitation invité n'a pas conscience d'être virtualisé et permet de faire fonctionner n'importe quel système d'exploitation en tant qu'invité dans une machine virtuelle. L'OS qui est virtualisé n'a aucun moyen de savoir qu'il partage le matériel avec d'autres OS. Ainsi, l'ensemble des systèmes d'exploitation virtualisés s'exécutant sur un unique ordinateur, peuvent fonctionner de manière totalement indépendante les uns des autres et être vu comme des ordinateurs à part entière sur un réseau. Sa caractéristique principale est que les système invités n'ont pas à être modifiés pour être utilisés dans une machine virtuelle utilisant une technologie de virtualisation.

En virtualisation complète, la machine physique qui va émuler le matériel pour le système invité doit être dotée d'un OS ainsi que d'une surcouche applicative. Un des gros intérêts de cette technique de virtualisation est de pouvoir émuler n'importe quelle architecture matérielle. On peut donc faire fonctionner les OS que l'on désire indépendamment de l'architecture du système hôte. On l'utilise en milieu industriel pour faire fonctionner des applications sur des architectures matérielles non encore commercialisées . Il faut savoir que ce type de virtualisation n'est que logiciel, aucune fonctionnalité matérielle de virtualisation n'est utilisée .

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

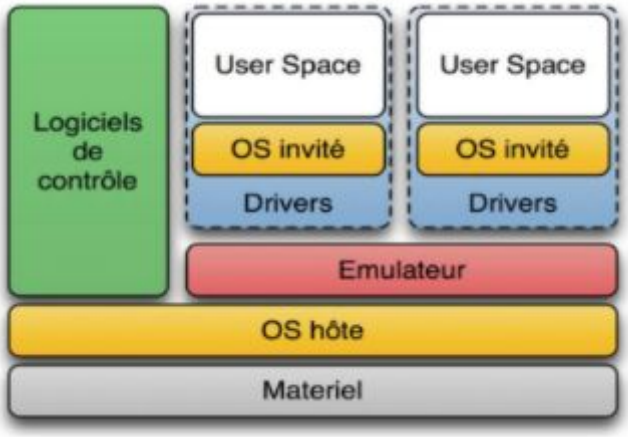

Figure 19 : Représentation schématique de la virtualisation complète

Comme présenté dans le schéma, toutes les requêtes de la machine virtuel sont émulées et cette couche d'émulation envoie l'élément demandé.

L'inconvénient de la virtualisation est la performance. Il est nécessaire d'émuler de nombreux composants physiques et de capter toutes les instructions émanant des machines virtuelles ce qui consomme beaucoup de ressources. Des tests montrent des pertes de performance de l'ordre de 20% à 80%. Il est possible d'améliorer significativement les performances en ajoutant des pilotes de paravirtualisation qui permettent un meilleur interfaçage avec la couche logicielle sous-jacente.

Avantages

- L'accès au ressources matérielles de chacune des VM est contrôlé ce qui empêche à une machine installée d'impacter l'autre

- Isolation des systèmes d'exploitations entre OS de la VM et l'OS de l'hôte
- Plus facile de porter la VM d'une machine hôte à une autre car les machines sont indépendantes
- Si un système d'exploitation tombe en panne, les autres continuent à fonctionner

Les avantages et les inconvénients de la virtualisation

La virtualisation comporte de nombreux avantages mais aussi des inconvénients.

Les avantages

La mise en place de la virtualisation permet de réduire les coûts puisqu'elle induit une réduction du nombre de serveurs physiques nécessitant alors moins de place pour les héberger associé un coût de maintenance plus faible. La diminution du nombre de serveurs physique implique aussi une diminution de la consommation énergétique allant de pair avec une diminution de la pollution numérique.

L'établissement d'une virtualisation implique aussi une meilleure exploitation des ressources qui sont souvent, jusqu'à la mise en œuvre de cette technologie, sous exploitée et cela permet une optimisation des capacités matérielles à disposition.

La virtualisation induit aussi une meilleure agilité en permettant une certaine libération des contraintes matérielles et encourage de ce fait, une flexibilité des processus et la mobilité des équipes.

Les inconvénients

L'introduction de la virtualisation entraîne aussi certains désagréments. En effet, les performances dépendent essentiellement de la puissance de traitement de l'objet et de la mémoire du système hôte. Elle peut aussi entraîner une diminution du niveau de sécurité avec l'apparition de failles de sécurité et de nouvelles menaces.

Paravirtualisation

On parle de paravirtualisation lorsque les systèmes d'exploitation doivent être modifiés pour fonctionner sur un hyperviseur de paravirtualisation. Les modifications sont en fait des insertions de drivers permettant de rediriger les appels système au lieu de les traduire.

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

Quelques exemples de paravirtualisation : XEN ,VMWare ESX/ESXi, Hyper-V (Microsoft) , xVM

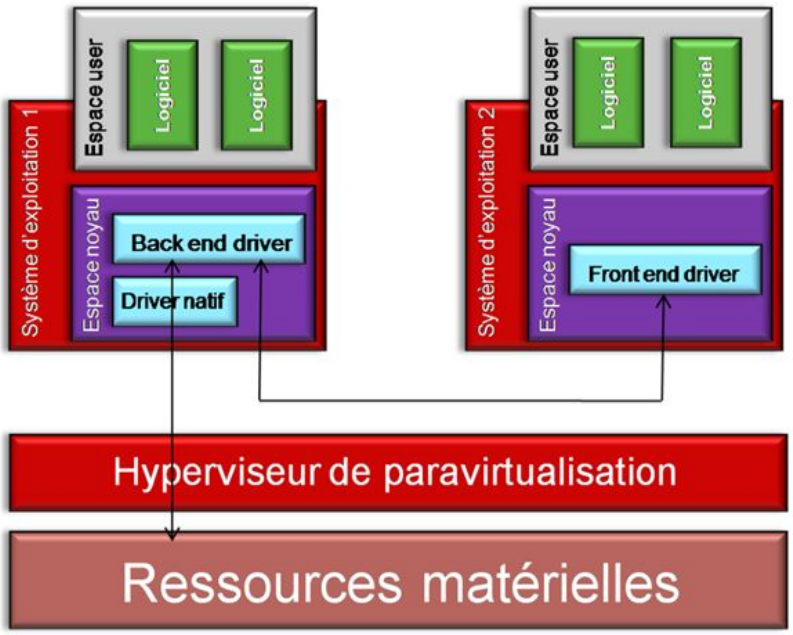

Figure 18 : Représentation de la paravirtualisation

Sur la figure ci-dessus des drivers backend et frontend sont installés dans les OS para-virtualisés. Ils permettent, au lieu de traduire les appels système comme cela est fait dans la virtualisation complète, de ne faire que de la redirection (ce qui est beaucoup plus rapide). Il est donc intelligent d'utiliser un tel mécanisme pour accéder à du matériel potentiellement très sollicité (disque dur, interface réseau...).

En fait, ce qu'il se passe dans une telle technologie, c'est que le contrôle d'un ou plusieurs matériel(s) est donné à un des OS virtualisé (celui qui contient le driver backend), ici le système d'exploitation 1. Une fois cela compris, il sera simple d'imaginer que l'OS 2, qui souhaite accéder au hardware, devra passer par son driver front end qui redirigera les appels système vers l'OS 1.

L'inconvénient de cette technique est donc la dépendance d'un OS virtualisé vis à vis d'un autre qui se crée par ce mécanisme de driver. En effet, si l'OS 1 tombe en panne, l'OS 2 ne pourra plus accéder au matériel.

inconvénient

- Il est nécessaire d'adapter les systèmes d'exploitation pour chaque couche de virtualisation

Informations supplémentaires

Certaines sources citent aussi l'existence d'un sixième type de virtualisation, la virtualisation des données.

Virtualisation des données ou Data virtualization

Cette technique de virtualisation repose sur l'abstraction des détails techniques traditionnels des données et du Data Management telles que la localisation, la performance ou bien encore le format.

Cela permet d'ouvrir l'accès aux données et d'accroître la résilience. De plus, la Data Virtualization permet aussi de consolider les données en une source unique afin de simplifier leur traitement, le but étant de permettre aux utilisateurs finaux d'accéder aux données sans se préoccuper de l'emplacement et de la structure d'origine de la source.

Les avantages de cette technique sont multiples puisqu'elle permet d'unifier la sécurité des données pour l'ensemble de l'entreprise grâce à la création de groupes en plus d'accroître l'agilité de l'équipe de développement au cours de lancement de projet d'intégration de données par exemple.
Cette méthode fournit en temps réel des données opérationnelles qui ont été traitées et nettoyées afin de répondre aux besoins les plus récents.

Application de la virtualisation : Docker

Docker est une solution Open Source qui s'appuie sur la notion de container au lieu de la notion de machine virtuelle. L'utilisation de cette méthode permet un gain considérable autant en termes de ressources matérielles que de performances. Cette technologie semble particulièrement adaptée pour les Clouds numériques.

Les avantages de Docker sont multiples : cette technologie permet d'isoler différents environnement sous un même OS tout en ne nécessitant pas l'utilisation de virtualisation par le biais de VMWare, Oracle VM.

Cette technologie comporte tout de même certains inconvénients notamment avec l'impossibilité d'utiliser des OS différents sur un seul serveur physique contrairement à la virtualisation ou bien encore avec les difficultés qu'elle engendre pour la définition d'une sécurité poussée entre les différents environnements mis en place au sein d'un même serveur physique.

⇒ Docker est une nouvelle tendance technologique