

4. Gestion des données urbaines dans les nuages informatiques

- Brève histoire des nuages informatiques
- Modèles de service et de déploiement
- **Technologie clé : la virtualisation**
- IaaS : les points de vue utilisateur et fournisseur
- PaaS : programmation et déploiement des applications
- Stockage de données
- Traitement de données
- Traitement de flux de données

Christine Morin

VILLES INTELLIGENTES : DÉFIS TECHNOLOGIQUES ET SOCIÉTAUX

Dans cette 3ème séquence, il m'a semblé utile de faire un focus technique sur la virtualisation qui est une technologie clé pour le cloud.

Naissance de la virtualisation à l'époque des mainframes (années 60)



2

La virtualisation est un concept qui est beaucoup plus ancien que le cloud puisque les premières techniques de virtualisation ont été implémentées à l'époque des mainframes, comme je vous l'ai dit en introduction.

Donc l'idée, c'était d'utiliser, en temps partagé, un ordinateur donc temps partagé entre plusieurs utilisateurs.

Avec l'émergence des ordinateurs individuels, la technologie était tombée en désuétude et c'est VMware qui l'a remise au goût du jour en proposant la technologie qui permet d'exécuter des systèmes d'exploitation distincts sur un même ordinateur individuel pour les processeurs X86.

Cependant certaines instructions du processeur X86 pouvaient, à l'époque, difficilement être virtualisées, ce qui entraînait des performances très moyennes. Et du coup, devant le regain d'intérêt pour la virtualisation, les constructeurs de processeurs n'ont pas tardé à faire évoluer les processeurs de manière à faciliter la virtualisation du système et ceci avec de bonnes performances.

Virtualisation : masquage d'une ressource physique derrière une ressource virtuelle

La **virtualisation d'un composant** (processeur, mémoire, périphérique d'E/S) à un **niveau d'abstraction donné** projette son interface et ses ressources visibles sur l'interface et les ressources d'un système réel sous-jacent.

3

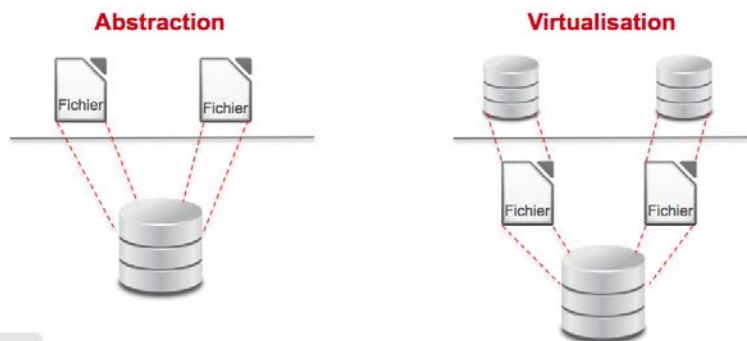
La virtualisation permet de **masquer une ressource physique derrière une ressource virtuelle**.

Les différents composants d'un ordinateur peuvent être virtualisés. Donc la virtualisation d'un composant tel que par exemple la mémoire, le processeur, un périphérique d'entrée sortie à un niveau d'abstraction donné, projette son interface et ses ressources visibles sur celles d'un système réel sous-jacent.

Différence entre abstraction et virtualisation

Abstraction et virtualisation sont deux notions différentes.

La virtualisation ne cherche pas à masquer les détails d'une interface complexe.



4

La virtualisation est une notion différente de l'abstraction. Contrairement à l'abstraction, la virtualisation ne cherche pas à masquer les détails d'une interface complexe.

Prenons un exemple pour bien comprendre. Considérons un disque de stockage. Les utilisateurs sont familiers avec la notion de fichier qui abstrait le disque en offrant une interface plus simple que l'interface du périphérique disque pour la manipulation des données qui sont à stocker sur le disque.

L'utilisateur ouvre le fichier écrit et lit des données dans le fichier et pour ce faire, il n'a pas à connaître les détails des disques magnétiques tels que le positionnement de la tête de lecture, la rotation des plateaux, la notion de secteur, etc.

Un disque virtualisé est toujours un disque, avec la même interface. Il est projeté sur un disque physique comme on le voit sur la partie droite du transparent.

Les fichiers stockés sur le disque virtuel sont stockés sur le disque physique qui correspond. Donc tout ce qu'on met dans le disque virtuel est stocké dans un fichier qui lui-même est stocké sur le disque physique.

Donc plusieurs disques virtuels peuvent être projetés sur un même disque physique. Par exemple, pour donner l'illusion à chaque utilisateur d'un gros disque physique qu'il dispose de son propre disque pour son utilisation exclusive.

Les différents types de virtualisation

1. Virtualisation du matériel (virtualisation système)
2. Virtualisation du système d'exploitation
3. Virtualisation de processus

5

Il existe principalement 3 types de virtualisations :

- La virtualisation du matériel, qu'on appelle aussi virtualisation système,
- La virtualisation du système d'exploitation,
- La virtualisation de processus.

Dans le contexte des clouds, seuls les 2 premiers types de virtualisation vont nous intéresser et nous allons les détailler dans la suite de cette séquence.

Nous n'aborderons pas le 3ème type de virtualisation, la virtualisation de processus dont un exemple est la machine virtuelle Java qui offre la portabilité des programmes Java sur différents systèmes.

1. Virtualisation du matériel : machines virtuelles

- Virtualisation de tous les composants d'un ordinateur



6

Commençons par la virtualisation du matériel qui est concrétisée par la notion de machine virtuelle. Une notion importante dans le cloud.

Une machine virtuelle virtualise tous les composants d'un ordinateur et il est donc possible d'exécuter un système d'exploitation sur une machine virtuelle puisqu'elle se présente pour le système d'exploitation exactement comme une machine physique.

Sur une machine physique, grâce au système de virtualisation, plusieurs machines virtuelles peuvent être exécutées chacune exécutant son propre système d'exploitation.

Sur l'exemple représenté ici, 3 machines virtuelles s'exécutent. L'une exécute le système Windows tandis que les 2 autres exécutent 2 distributions du système Linux. Red Hat et Debian ici sur l'exemple.

Le logiciel de gestion de machine virtuelle s'appelle un hyperviseur.

Son rôle est d'isoler et de gérer le partage des ressources physiques entre les différentes machines virtuelles qui s'exécutent donc sur un même serveur.

Hyperviseur

- Gestionnaire de machines virtuelles
- Deux types d'hyperviseur
 - Virtualisation de type 1 : exécution native sur la machine physique
 - Virtualisation de type 2 : exécution au-dessus du système d'exploitation

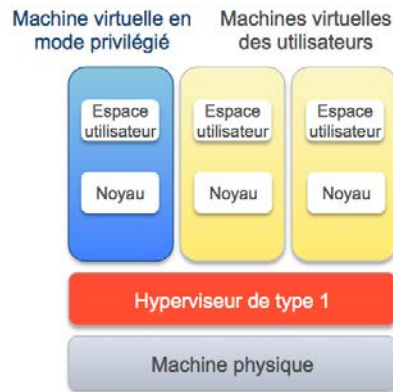
7

Il existe 2 types d'hyperviseurs qui correspondent respectivement à ce qu'on appelle la virtualisation de type 1 et la virtualisation de type 2.

Avec la virtualisation de type 1, l'hyperviseur s'exécute directement sur la machine physique.

Avec la virtualisation de type 2, l'hyperviseur s'exécute au-dessus du système d'exploitation qui s'exécute nativement sur une machine physique.

Virtualisation de type 1



8

On voit sur le schéma ici un **hyperviseur de type 1** qui s'exécute directement sur la machine physique sous-jacente et dont le rôle se limite à la gestion des ressources physiques de cette machine.

2 types de machines virtuelles s'exécutent au-dessus de cet hyperviseur :

- une **machine virtuelle particulière**, représentée ici en bleu qui est appelée **machine virtuelle privilégiée** et qui va exécuter certaines primitives, par exemple pour l'accès aux périphériques pour le compte des autres machines virtuelles.
- Les autres machines virtuelles, représentées en jaune ici, sont les machines virtuelles des utilisateurs.

Virtualisateur de type 2



9

Le schéma ici est représenté un **hyperviseur de type 2**. Ce type d'hyperviseur s'exécute au-dessus du noyau du système d'exploitation que l'on appelle le **système d'exploitation hôte** parce qu'il s'exécute sur la machine qui va héberger les machines virtuelles des utilisateurs. Les machines virtuelles des utilisateurs, représentées ici en jaune, vont de manière analogue à ce qu'on voyait avec la virtualisation du type 1, faire appel à des services fournis par les services de haut niveau du système d'exploitation hôte de la machine physique représenté à gauche en bleu.

2. Virtualisation du système d'exploitation : conteneurs



10

Par rapport à la virtualisation du matériel, la virtualisation du système d'exploitation est plus légère puisque le noyau est commun à tous les conteneurs.

Chaque conteneur est donc isolé et utilise une quantité de ressources bornées.

Dans les clouds, la virtualisation permet de partager un même serveur physique entre plusieurs machines virtuelles, ou plusieurs conteneurs selon la technologie, qui appartiennent donc à des clients différents.

Opérations sur les machines virtuelles

- Migration à chaud



- Sauvegarde d'état



11

2 fonctionnalités des machines virtuelles et des conteneurs sont exploitées dans les clouds : la **migration à chaud** et la **sauvegarde d'état**.

La **migration à chaud** de machine virtuelle permet de **déplacer une machine virtuelle en cours d'exécution** depuis une machine physique vers une autre de manière transparente aux applications qui s'exécutent au sein de cette machine virtuelle.

La migration à chaud permet de mettre en œuvre des stratégies de gestion de ressources au sein des data centers. Par exemple, à des fins d'économies d'énergie, on pourra faire ce qu'on appelle de la consolidation, c'est-à-dire regrouper plusieurs machines virtuelles sur un petit nombre de serveurs physiques de sorte à pouvoir éteindre les serveurs inutilisés de sorte à ce qu'ils consomment moins d'énergie, globalement, au niveau du data center.

La **sauvegarde d'état** consiste à prendre, en quelque sorte, une **photo de l'état de la machine virtuelle à un instant donné**, donc qui comporte l'état de ses registres, l'état de sa mémoire, l'état du disque. Cet état peut être conservé, par exemple sur disque, pour pouvoir redémarrer ultérieurement la machine virtuelle. Par exemple, pour des considérations de tolérance aux fautes ou pour la déplacer entre des data centers distants.

La réduction des émissions de carbone est une considération importante dans les villes intelligentes. Les clouds étant de très gros consommateurs d'énergie, il est essentiel de mettre en œuvre des politiques de gestion de ressources qui vont contribuer à réduire l'impact écologique des clouds.

Illustrations & photos : crédits

p. 2 : by Dave Mills, Wikimedia Commons

p. 11 : by Franz Pfluegl, Fotolia

p. 11 : by eralt, Pixabay, CC0 Public Domain